

PETUA PENGELUARAN DAN POKOK RAWAK
UNTUK AUTOMASI DIAGNOSIS BERSEPADU
KANSER PROSTAT

MUHAMMAD UZAIR BIN SAHRIN

UNIVERSITI KEBANGSAAN MALAYSIA

PETUA PENGELUARAN DAN POKOK RAWAK UNTUK AUTOMASI
DIAGNOSIS BERSEPADU KANSER PROSTAT

MUHAMMAD UZAIR BIN SAHRIN

DISERASI YANG DIKEMUKAKAN UNTUK MEMENUHI
SEBAHAGIAN DARIPADA SYARAT MEMPEROLEHI
IJAZAH SARJANA TEKNOLOGI MAKLUMAT

FAKULTI TEKNOLOGI SAINS DAN MAKLUMAT
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2018

PENGAKUAN

Saya akui karya ini adalah hasil kerja saya sendiri kecuali nukilan dan ringkasan yang tiap-tiap satunya telah saya jelaskan sumbernya.

01 September 2018

MUHAMMAD UZAIR BIN
SAHRIN
P83911

PENGHARGAAN

Dengan lafaz Bismillahirrahmanirrahim, saya memanjatkan rasa kesyukuran ke hadrat Ilahi di atas limpah kurnia serta kasih sayangNya, saya telah berjaya menyiapkan projek serta laporan akhir sebagai salah satu kursus yang perlu dilengkapkan untuk memenuhi syarat bergraduasi serta sebagai pemegang Ijazah Sarjana Teknologi Maklumat. Walaupun menghadapi pelbagai cabaran dan rintangan di dalam penyediaan projek ini tetapi telah berjaya diatasi dengan berkat pertolongan daripada Allah, dorongan ibu bapa, serta rakan- rakan.

Pertama sekali saya ingin mengucapkan terima kasih kepada kedua ibu bapa saya kerana sentiasa mahu terbaik untuk anaknya, jasamu dikenang duhai keluargaku. Setinggi-tinggi penghargaan juga diberi kepada semua pensyarah dan kakitangan Fakulti Teknologi Dan Sains Maklumat yang sentiasa memberi bantuan dan kerjasama sepanjang tempoh pengajian penulis di Universiti Kebangsaan Malaysia. Ucapan terima kasih juga kepada keluarga dan rakan-rakan yang menjadi pembakar semangat dan tidak jemu memberi pendapat dan kritikan sepanjang projek ini dijalankan.

Terima kasih khas ditujukan kepada penyelia, Prof Madya Dr. Siti Norul Huda Sheikh Abdullah Abdul Shukor yang banyak memberi sokongan, dorongan, nasihat dan tunjuk ajar sepanjang projek ini dijalankan. Dengan tunjuk ajar serta dorongan yang telah beliau berikan banyak membantu dalam menyiapkan projek ini. Jutaan penghargaan dan ucapan terima kasih juga saya ucapkan kepada penyelia kedua saya iaitu Dr. Afzan Binti Adam yang banyak memberikan bimbingan kepada saya dalam menyiapkan projek ini.

Ucapan terima kasih tidak terhingga untuk kepada semua pihak yang terlibat secara langsung mahupun secara tidak langsung. Jasa kalian akan sentiasa saya kenang dan hanya Allah S.W.T. yang dapat membalas jasa kalian semua.

ABSTRAK

Jumlah penghidap kanser prostat merupakan keenam tertinggi di dunia dan merupakan penyakit yang rumit untuk dikenal pasti. Sehingga kini, masih tiada kaedah yang paling mudah dan pantas untuk diagnosis penyakit ini. Kebiasaannya kanser prostat didiagnosis secara manual oleh pakar berdasarkan pengetahuan dan pengalaman perubatan mereka. Sekumpulan pakar akan membincangkan status penyakit dalam sesi Risiko Penstrataan setelah melakukan beberapa siri ujian termasuk pemeriksaan fizikal, darah, biopsi dan imej-imej visual. Walaubagaimana pun, keputusan diagnosis imej visual terdedah kepada isu persetujuan bersama jika dinilai oleh beberapa pakar dalam satu-satu masa, terutamanya keputusan ahli histopathologi yang merupakan piawaian emas dalam pengredan kanser. Namun, proses ini mengambil masa yang agak lama. Oleh itu, objektif kajian ini adalah untuk (1) merekabentuk, membangun dan menilai sistem pakar kanser prostat berasaskan petua pengeluaran yang menggabungkan tiga jabatan iaitu urologi, radiologi dan histopatologi dan (2) mencadang dan membanding model diagnosis penyakit kanser prostat menggunakan kaedah pokok rawak dan petua pengeluaran untuk jabatan radiologi. Bagi objektif pertama, pengetahuan diagnosis daripada pakar dan rekod perubatan pesakit dikumpulkan bagi menjana petua pengeluaran dan disimpan dalam bentuk peraturan asas yang mudah untuk ditembak, difahami dan ditafsir sebagai enjin taabir. Bagi objektif kedua pula, perlombongan pengetahuan dengan teknik pokok rawak dipilih sebagai medium utama yang digunakan untuk menyelidik masalah bagi mengkaji keupayaan meramal tahap bahaya kanser prostat. Sebanyak 51 kes pesakit kanser prostat, 2 orang pakar urologi, seorang pakar radiologi dan seorang pakar patologi dari PPUKM terlibat secara aktif dalam pembangunan sistem pakar ini. Hasil penilaian dari kaedah kajian kes klinikal sebenar menunjukkan sistem pakar ini dapat menghasilkan ketepatan petua sehingga 76.4%, dan terbukti mampu menghasilkan model diagnosis yang boleh dipercayai. Setiap jabatan bersetuju bahawa sistem ini dapat membantu dalam memberi pandangan kedua kepada pakar perubatan daripada segi pengredan kanser prostat dengan skor 4.03/5 dengan skala Likert. Selain itu, sistem pakar yang menggunakan pengujian data bagi teknik pokok rawak pula mencapai 96% ketepatan. Secara keseluruhannya, sistem Automasi Diagnosis Bersepadu Kanser Prostat Secara Petua Pengeluaran Dan Pokok Rawak dapat dibangun mengikut objektif dan skop kajian yang telah ditetapkan.

PRODUCTION RULE AND RANDOM TREE FOR INTEGRATED DIAGNOSIS AUTOMATION OF PROSTATE CANCER

ABSTRACT

The number of patient with prostate cancer is currently the sixth highest in the world and it is a complicated disease to identified. Until now, there is still no method that most convenient and quick to diagnose this disease. Prostate cancer usually diagnosed manually by experts based on their medical knowledge and experiences. A group of experts will discuss the status of severity in a Risk Stratification session after doing a series of tests including physical examination, blood, biopsy and analyzing medical images. However, the results of visual image diagnosis are exposed to the issue of mutual consent if evaluated by specialists in one time, particularly the results of member histopathologi which is the gold standard in trading of cancer. But, this process takes quite some time. Therefore, the objectives of this study is to (1) design, develop and evaluate prostate cancer expert system based on the production rule combining three Departments, namely Urology, Radiology and Histopathology and (2) propose and compares model diagnosis of prostate cancer using random tree method and production rule for Department of Radiology. For the first objective, diagnosing the knowledge from expert as well as patient medical records is collected and used to generate production rule. These are stored in the form of rules to trigger the correct output in the inference engine. For the second objective, random tree was selected as the mining technique to model the knowledge in radiology department in predicting the level of prostate cancer severity. A total of 51 cases of prostate cancer patients with 2 urologists, a radiologist and an expert on pathology of PPUKM actively involved in the development of the expert system. Findings from the clinical case study method shows the real system experts to produce precision of rules up to 76.4%, and proved to be able to produce a reliable diagnosis model. Each department agree that this system could help in giving the second opinion to physician in terms of grading prostate cancer with a score of 4.03/5 on the Likert scale. Additionally, expert systems using data testing for random tree techniques achieve 96% accuracy. Overall, Integrated automation system in Diagnosis of Prostate Cancer Production Rule and Random Tree can be developed according to the objectives and scope of the study has been set.

KANDUNGAN

		Halaman
PENGAKUAN		ii
PENGHARGAAN		iii
ABSTRAK		iv
ABSTRACT		v
KANDUNGAN		vi
SENARAI JADUAL		x
SENARAI SINGKATAN		xvi
BAB I	PENGENALAN	
1.1	Pendahuluan	1
1.2	Permasalahan Kajian	3
1.3	Objektif Kajian	5
1.4	Skop Kajian	5
1.5	Pendekatan Kajian	6
1.6	Model Kerangka Penyelidikan	6
1.7	Organisasi Penulisan	6
1.8	Rumusan	8
BAB II	KAJIAN KESUSTERAAN	
2.1	Pengenalan	9
2.2	Peringkat Kanser Prostat Dan Rawatan	12
	2.2.1 Pembedahan (Urologi)	14
	2.2.2 Radiologi	14
	2.2.3 Histopatologi	14
	2.2.4 Onkologi	15
2.3	Pengesanan Dan Kaedah Diagnosis Kanser Prostat	15
	2.3.1 Pengimbasan Tulang	16
	2.3.2 SPECT / CT	18
	2.3.3 MRI	18
	2.3.4 PET	19
2.4	Etika Penyelidik (Perubatan)	20
2.5	Kepintaran Buatan Dalam Bidang Perubatan	20
	2.5.1 <i>Computer Aided Diagnosis (CAD)</i>	20

	2.5.2	Penggunaan Kepintaran Buatan Dalam Kanser Prostat	21
2.6		Perlombogan data	23
	2.6.1	Regrasi	26
	2.6.2	Pengelasan	26
	2.6.3	Ramalan	27
	2.6.4	Pengelompokan	27
	2.6.5	Pembelajaran Mesin	28
2.7		Rumusan	35
BAB III		KAEDAH METODOLOGI	
3.1		Pengenalan	37
3.2		Fasa 1: Pemahaman Teori dan Aplikasi	40
3.3		Fasa 2: Kajian Awal Dan Rangka Kerja	42
	3.3.1	Kaedah Pembangunan	43
	3.3.2	Perisian Bagi Perlombogan Data	44
3.4		Fasa 3: Fasa Perlaksanaan dan Uji Kaji	44
	3.4.1	Asas Logik Predikat	45
3.5		Fasa 4: Penghasilan Keputusan	46
	3.5.1	Pengujian Dan Penilaian Sistem	47
	3.5.2	Kaedah Penilaian Bagi Model Pokok Rawak	47
	3.5.3	Penilaian Pengelas Bagi Model Pokok Rawak	49
3.6		Rumusan	49
BAB IV		MODEL PETUA PENGELUARAN AUTOMASI DIAGNOSIS BERSEPADU KANSER PROSTAT	
4.1		Pengenalan	51
4.2		Rekabentuk Dan Model Pengetahuan	51
	4.2.1	Pengumpulan Pengetahuan	52
	4.2.2	Pangkalan Data	54
	4.2.3	Analisis Pengetahuan (Fakta dan Data)	65
	4.2.4	Enjin Taabir	65
	4.2.5	Ingatan Kerja	66
	4.2.6	Rekabentuk Antara muka	71
	4.2.7	Kemudahan Penerangan	79
4.3		Kaedah Ujian Kepenggunaan	81
4.4		Keputusan Dan Analisa Petua Pengeluaran	82
4.5		Rumusan	90

BAB V	MODEL POKOK RAWAK	
5.1	Pengenalan	91
5.2	Pengujian Set Data Bagi Menjana Petua	91
5.3	Model Petua Pokok Rawak Bagi Jabatan Radiologi	92
	5.3.1 Perolehan Data	92
	5.3.2 Pembahagian Set Data	93
	5.3.3 Pra-Promposesan Data Bagi Petua Pokok Rawak	94
5.4	Konfigurasi Ekperimen	98
	5.4.1 Eksperimen Pokok Rawak	98
	5.4.2 Pengujian Anova	105
5.5	Pengujian Model	106
5.6	Hasil Keputusan Petua Bagi Model Pokok Rawak	108
5.7	Rumusan	110
BAB VI	RUMUSAN DAN CADANGAN	
6.1	RINGKASAN KAJIAN	111
6.2	Rumusan kajian	112
6.3	Kelebihan Dan Kekurangan kajian	113
	6.3.1 Kelebihan	113
	6.3.2 Kekurangan	113
6.4	Persamaan Sistem	113
6.5	Sumbangan	114
6.6	Cadangan Masa Hadapan	114
6.7	Rumusan	115
RUJUKAN		116
Lampiran A	Perolehan Pengetahuan	123
Lampiran B	Eksperimen J48	131
Lampiran B.1	Pemilihan Bilangan Minimum Objek J48	132
Lampiran B.2	Pemilihan Bilangan Lipatan J48	133
Lampiran B.3	Pemilihan Benih J48	134
Lampiran B.4	Keputusan Dan Analisa J48	135
Lampiran C	Eksperimen Hutan Rawak	137

Lampiran C.1	Pemilihan Kedalaman Maksimum HUTAN RAWAK	138
Lampiran C.2	Pemilihan Bilangan Slot Perlakuan Hutan Rawak	139
Lampiran C.3	Pemilihan Bilangan Ciri Hutan Rawak	140
Lampiran C.4	Pemilihan Benih Hutan Rawak	141
Lampiran C.5	Keputusan Dan Analisa Hutan Rawak	142
Lampiran D	Surat Latihan Sangkutan	144
Lampiran E	Surat Kelulusan Etika	146
Lampiran F	Surat Minit Mesyuarat Etika	147
Lampiran G	Laporan Pesakit	148
Lampiran H	Manual Pengguna	149
Lampiran I	Contoh Borang Soal Selidik	164

SENARAI JADUAL

No. Jadual		Halaman
Jadual 3.1	Perjumpaan proses pengambilan data	40
Jadual 4.1	Metadata jadual <i>login</i>	62
Jadual 4.2	Metadata jadual <i>patient</i>	62
Jadual 4.3	Metadata jadual <i>general_medical_report</i>	62
Jadual 4.4	Metadata jadual <i>radiology_record</i>	64
Jadual 4.5	Metadata jadual <i>histopatologi_record</i>	64
Jadual 4.6	Keputusan markah purata bagi pengujian terhadap sistem bagi jabatan pembedahan (urologi)	83
Jadual 4.7	Keputusan markah purata bagi pengujian terhadap sistem bagi jabatan radiologi	84
Jadual 4.8	Keputusan markah purata bagi pengujian terhadap sistem bagi jabatan histopatologi	85
Jadual 4.9	Keputusan markah purata pengujian penerimaan petua penagluaran yang dikumpul bagi urologi	87
Jadual 4.10	Keputusan markah purata pengujian penerimaan petua penagluaran yang dikumpul bagi jabatan radiologi	87
Jadual 4.11	Keputusan markah purata penerimaan petua penagluaran yang dikumpul bagi jabatan histopatologi	88
Jadual 5.1	Nilai yang diperoleh mengikut nisbah pecahan peratusan	92
Jadual 5.2	Set data rekod pesakit dari Pusat Perubatan UKM	93
Jadual 5.3	Pengolahan atribut	96
Jadual 5.4	Pemilihan kedalaman maksimum pokok untuk latihan pokok rawak	99
Jadual 5.5	Pemilihan bilangan minimum untuk latihan pokok rawak	100
Jadual 5.6	Pemilihan bilangan lipatan pokok untuk latihan pokok rawak	101
Jadual 5.7	Pemilihan benih untuk latihan pokok rawak	102

Jadual 5.8	Prestasi antara pokok rawak berdasarkan peratus pecahan yang berbeza	103
Jadual 5.9	Prestasi pokok rawak berdasarkan keesahan silang yang berbeza	104
Jadual 5.10	Keputusan pengujian ANOVA pokok rawak berdasarkan peratus pecahan yang berbeza	105
Jadual 5.11	Keputusan pengujian ANOVA pokok rawak berdasarkan keesahan silang yang berbeza	106
Jadual 5.12	Keputusan pengujian set data ketepatan untuk teknik pembelajaran mesin pokok keputusan J48, pokok rawak dan hutan rawak	107
Jadual 6.1	Pengumpulan maklumat penemberengan CT	130
Jadual 6.2	Pemilihan bilangan minimum objek untuk latihan J48	132
Jadual 6.3	Pemilihan bilangan lipatan pokok untuk latihan J48	133
Jadual 6.4	Pemilihan benih untuk latihan J48	134
Jadual 6.5	Prestasi antara J48 berdasarkan peratus pecahan yang berbeza	135
Jadual 6.6	Prestasi J48 berdasarkan keesahan silang yang berbeza	136
Jadual 6.7	Pemilihan kedalaman maksimum pokok untuk latihan hutan rawak	138
Jadual 6.8	Pemilihan bilangan slot perlakuan untuk latihan hutan rawak	139
Jadual 6.9	Pemilihan bilangan ciri untuk latihan hutan rawak	140
Jadual 6.10	Pemilihan benih untuk latihan hutan rawak	141
Jadual 6.11	Prestasi antara hutan rawak berdasarkan peratus pecahan yang berbeza	142
Jadual 6.12	Prestasi pokok rawak berdasarkan keesahan silang yang berbeza	143

SENARAI ILUSTRASI

No. Rajah		Halaman
Rajah 2.1	(a) Kadar kematian dan bilangan bagi kanser prostat (b) Perbandingan kadar kematian bagi kanser prostat	10
Rajah 2.2	TNM Pengelasan karsinoma daripada prostat	13
Rajah 2.3	Kamera gamma dengan keupayaan untuk memperoleh satah seluruh badan dan tomografi imej	17
Rajah 2.4	(a) Normal imbasan tulang (pelvik) (b) Tidak normal imbasan tulang (pelvik)	18
Rajah 2.5	PET / CT untuk mengesan penyakit metastatic	19
Rajah 2.6	Langkah – langkah KDD (Knowledge Discovery in Database)	24
Rajah 2.7	Kaedah – kaedah dalam perlombongan data	25
Rajah 2.8	Teknik-teknik dalam perlombongan data	26
Rajah 2.9	Model pengelasan	27
Rajah 2.10	Pokok keputusan bagi konsep main_tenis	30
Rajah 2.11	Struktur asas bagi pokok pengelasan	34
Rajah 2.12	Seni bina umum hutan rawak	35
Rajah 3.1	Model kerangka kerja sistem	38
Rajah 3.2	Penghalusan kerangka kerja kajian	39
Rajah 4.1	Kerangka model pengetahuan	52
Rajah 4.2	Contoh skrip perbualan perolehan pengetahuan dari pakar urologi.	53
Rajah 4.3	Rajah konteks bagi jabatan pembedahan (urologi)	55
Rajah 4.4	Rajah konteks bagi jabatan radiologi	55
Rajah 4.5	Rajah konteks bagi histopatologi	56
Rajah 4.6	Rajah aliran data paras sifar jabatan pembedahan (urologi)	56
Rajah 4.7	Rajah aliran data paras sifar jabatan radiologi	57

Rajah 4.8	Rajah aliran data paras sifar jabatan histopatologi	57
Rajah 4.9	Sub modul bagi jabatan urologi	58
Rajah 4.10	Sub modul bagi jabatan radiologi	59
Rajah 4.11	Sub modul bagi jabatan histopathologi	59
Rajah 4.12	Carta alir proses diagnosis kanser prostat	60
Rajah 4.13	Rajah hubungan entiti sistem pakar	61
Rajah 4.14	Contoh laporan klinikal bagi kes di jabatan pembedahan (urologi)	67
Rajah 4.15	Pembuktian petua dengan klausa predikat	67
Rajah 4.16	Contoh input kemasukan pengetahuan bagi jabatan pembedahan (urologi)	68
Rajah 4.17	Contoh keputusan diganosis rekod pembedahan (urologi)	68
Rajah 4.18	Contoh input kemasukan pengetahuan bagi jabatan radiologi	69
Rajah 4.19	Contoh keputusan diganosis rekod radiologi	69
Rajah 4.20	Contoh input kemasukan pengetahuan bagi jabatan histopatologi	70
Rajah 4.21	Contoh keputusan diganosis rekod histopathologi	71
Rajah 4.22	Antara muka log pengguna	72
Rajah 4.23	Antara muka pengurusan kemasukan maklumat pesakit	73
Rajah 4.24	Antara muka pengurusan mengubahsuai dan memadam maklumat pesakit	73
Rajah 4.25	Antara muka pengurusan kemasukan rekod perubatan jabatan pembedahan (urologi)	74
Rajah 4.26	Antara muka pengurusan mengubahsuai rekod perubatan jabatan pembedahan (urologi)	74
Rajah 4.27	Antara muka pengurusan paparan rekod perubatan jabatan pembedahan (urologi)	75
Rajah 4.28	Antara muka pengurusan kemasukan rekod perubatan jabatan radiologi	76
Rajah 4.29	Antara muka pengurusan mengubahsuai rekod perubatan jabatan radiologi	76

Rajah 4.30	Antara muka pengurusan paparan rekod perubatan jabatan radiologi	77
Rajah 4.31	Antara muka pengurusan paparan rekod perubatan jabatan radiologi dalam meramal tahap kaser prostat	77
Rajah 4.32	Antara muka pengurusan rekod jabatan histopatologi	78
Rajah 4.33	Antara muka pengurusan mengubahsuai rekod jabatan histopatologi	78
Rajah 4.34	Antara muka pengurusan paparan rekod jabatan histopatologi	79
Rajah 4.35	Contoh kemudahan penerangan bagi jabatan pembedahan (urologi)	80
Rajah 4.36	Contoh kemudahan penerangan bagi jabatan radiologi	80
Rajah 4.37	Contoh kemudahan penerangan bagi jabatan histopatologi	81
Rajah 4.38	Keputusan markah purata pengujian terhadap sistem bagi semua jabatan	86
Rajah 4.39	Keputusan markah purata pengujian penerimaan ketua penagluaran bagi semua jabatan	89
Rajah 5.1	Pendiskretan data	97
Rajah 5.2	Bilangan slot perlakuan melawan kadar ketepatan untuk pokok rawak	99
Rajah 5.3	Bilangan minimum pokok melawan kadar ketepatan untuk pokok rawak	100
Rajah 5.4	Bilangan lipatan pokok melawan kadar ketepatan untuk pokok rawak	101
Rajah 5.5	Benih melawan kadar ketepatan untuk pokok rawak	102
Rajah 5.6	Prestasi antara pokok rawak berdasarkan peratus pecahan yang berbeza	104
Rajah 5.7	Prestasi antara Pokok Rawak berdasarkan keesahan silang yang berbeza.	105
Rajah 5.8	Model pokok rawak	108
Rajah 5.9	Perwakilan pengetahuan secara secara predikat logik tertib kedua	110

Rajah 6.1	(a) Petua-petua bagi Modul Urologi yang apabila dibincang dengan pakar Urologi, dan (b) perwakilan pengetahuan secara secara predikat logik tertib kedua	124
Rajah 6.2	(a) Petua-petua bagi Modul Radiologi yang apabila dibincang dengan pakar radiologi, dan (b) perwakilan pengetahuan secara secara predikat logik tertib kedua.	125
Rajah 6.3	(a) Petua-petua bagi Modul Histopatologi yang apabila dibincang dengan pakar histopatologi, dan (b) perwakilan pengetahuan secara secara predikat logik tertib kedua.	126
Rajah 6.4	Pseudokod diagnosis urologi	127
Rajah 6.5	Pseudokod diagnosis radiologi	128
Rajah 6.6	Pseudokod diagnosis histopatologi	129
Rajah 6.7	Bilangan minimum objek melawan kadar ketepatan untuk J48	132
Rajah 6.8	Bilangan lipatan pokok melawan kadar ketepatan untuk J48	133
Rajah 6.9	Benih melawan kadar ketepatan untuk J48	134
Rajah 6.10	Kedalaman maksimum pokok melawan kadar ketepatan untuk hutan rawak	138
Rajah 6.11	Bilangan slot perlakuan melawan kadar ketepatan untuk hutan rawak	139
Rajah 6.12	Bilangan ciri melawan kadar ketepatan untuk hutan rawak	140
Rajah 6.13	Benih melawan kadar ketepatan untuk hutan rawak	141

SENARAI SINGKATAN

ANN	Artificial Neural Networks
APCID	Automation Of Prostate Cancer Integrative Diagnosis
CAD	Computer Aided Diagnosis
CC	Correlation Coefficient
CT	Computed Tomography
ERD	Entity Relationship Diagram
FL	Fuzzy Logic
GA	Genetic Algorithm
HK2	Human Kallikrein-2
IDE	Integrated Development Environment
MCS	Multiple Classifier System
MSE	Mean Squared Error
MDP	Methylene Diphosphonate
MRI	Magnetic Resonance Imaging
NDA	Non-Parametric Discriminant Analysis
PAP	Prostatic Acid Phosphatase
PPUKM	Pusat Perubatan Universiti Kebangsaan Malaysia
PSA	Prostate-Specific Antigen
RMSE	Root Mean Square Error
RNN	Recurrent Neural Networks
ROC	Receiver Operating Characteristics
TNM	TNM Classification of Malignant Tumours
TNM	Classification of Malignant Tumours
UKM	Universiti Kebangsaan Malaysia

BAB I

PENGENALAN

1.1 PENDAHULUAN

Kanser prostat kini adalah kanser yang ke- 6 biasa di dunia (dari segi bilangan kes baru). Pada tahun 2000, anggaran bilangan kes adalah 513,000 yang mewakili 9.7% daripada kanser pada lelaki dengan 201,000 kematian (Eble 2004). National Cancer Registry Malaysia mempunyai 502 kes kanser prostat didiagnosis dan dilaporkan pada tahun 2007. Ia merupakan kanser keempat paling biasa untuk lelaki di Malaysia (Omar dan Tamin 2011).

Sehingga tahun 2012, pendekatan statistik yang tepat terhadap sistem pengkomputeran lembut adalah sebahagian daripada alat-alat yang telah dibangunkan untuk membantu dalam perubatan membuat keputusan terutama dalam meramal peringkat patologi akhir. Sebagai contoh, Jadual Partin et al. (2001) digunakan secara meluas di Eropah dan Amerika Utara sebagai panduan untuk memilih terapi muktamad yang menggunakan jadual kebarangkalian yang menggabungkan pra operasi serum Prostat Spesifik Antigen *Prostate-specific antigen* (PSA), peringkat klinikal TNM di mana sebuah sistem notasi yang menggambarkan peringkat kanser yang berasal dari tumor yang kukuh dengan kod abjad dan Gleason gred daripada biopsi prostat sebagai pembolehubah untuk mengira kemungkinan peringkat patologi akhir (positif penglibatan *vesicle* mani organ-terkurung, penembusan capsular, dan nod limfa penglibatan). Kemas kini terbaru jadual adalah pada tahun 2007 (Castanho et al. 2013). Seterusnya, kaedah pengkomputeran lembut seperti rangkaian neural buatan, logik kabur dan algoritma genetik pula mula digunakan dalam bidang perubatan untuk tujuan diagnosis dan prognosis dalam Sistem Pakar Kabur Bagi Meramalkan Patologi Peringkat Kanser Prostat dua puluh tahun yang lalu (Castanho et al. 2013).

Percubaan awal menggunakan Rangkaian Neural Buatan (RNN) oleh Snow et al. (1994) digunakan untuk meramalkan keputusan pembedahan dengan saringan lelaki yang tidak normal dan untuk meramalkan hasil rawatan selepas *prostatectomy*. Sistem berasaskan petua kabur Castanho et al. (2013) dibangunkan pada tahun 2008 adalah salah satu contoh yang menggunakan logik kabur untuk mengendalikan dengan ketidakpastian dan ketidak tentuan dalam meramalkan peringkat patologi kanser prostat dengan nilai *The Area Under Receiver Operating Characteristic Curve* (AUROC) untuk penyakit organ-terbatas dengan teknik yang tersebut adalah 0.76.

Algoritma genetik (GA) telah digunakan untuk mengenal pasti dan mengelaskan ovari, prostat dan kanser paru-paru dan perancangan rawatan kanser prostat (Castanho et al. 2013). Menurut Castanho et al. (2013), ramai penyelidik telah membuat penggunaan sifat pelengkap dari 3 teknik dalam *hybridising* ANN, FL dan GA untuk meningkatkan keupayaan sistem. Sistem Neuro-fuzz'y telah digunakan bagi mengelaskan kanser prostat dan mendiagnosis kanser payudara. Sistem Genetic-fuzzy telah digunakan dalam diagnosis kanser payudara dan telah berjaya mencapai prestasi pengelasan tinggi (Castanho et al. 2013).

Secara keseluruhan, perubatan dan sains komputer adalah dua silang disiplin yang boleh digunakan untuk meningkatkan gaya hidup dan pembangunan teknologi baru. Selain itu, sistem pengimejan berbantuan komputer dan sistem pemprosesan imej yang direka untuk membantu pengamal perubatan membuat keputusan berstruktur, separa berstruktur dan tidak berstruktur. Oleh itu, kajian ini pada *Automation Integrative Diagnosis of Prostate Cancer with Knowledge Rules and Random Tree* atau bersepadu Kanser Prostat Secara Petua Pengeluaran Dan Pokok Rawak adalah relevan untuk diterokai dan dibina.

1.2 PERMASALAHAN KAJIAN

Diagnosis kanser prostat adalah rumit dan terdapat beberapa kes di mana jenis kanser yang sukar atau mustahil untuk mengenal pasti (Cinar et al. 2009; Chen dan Lin 2011) mendakwa bahawa terdapat hanya beberapa alat pengimejan yang berjaya dan boleh diagnosis kanser prostat secara berkesan dalam memantau rawatan terapeutik untuk penyakit ini (Kim et al. 2010).

Keputusan akhir tentang status penyakit yang dihidap oleh pesakit dikeluarkan berdasarkan keputusan dari histopatologi. Kesemua maklumat yang diperolehi daripada pakar-pakar perubatan setiap jabatan akan dimasukkan ke dalam teknologi maklumat. Dengan menggunakan sistem kanser prostat ini, diharap ia dapat menjimatkan masa, beban kerja dan tenaga kerja yang diperlukan. Sistem kanser prostat ini bukanlah sebagai gantian kepada pakar perubatan tetapi bertujuan untuk membantu mereka. Begitu juga pesakit boleh mendapatkan keputusan yang lebih cepat. Jika berlaku kelewatan mendiagnosis akibat kekurangan doktor yang berkemampuan untuk mendiagnosis sampel tisu tersebut, risiko kanser menjadi semakin kritikal adalah tinggi. Masalah tersebut boleh dilihat seperti di bawah.

a. Pengetahuan Berasaskan Pemerhatian Pakar (*Interobserver Agreement*)

Kebiasaannya, doktor akan memaklumkan cabang yang betul dalam mengesan secara manual perkembangan ketumbuhan dan kesan sampingan lain terhadap rawatan. Preskripsi oleh pakar-pakar akan dibuat berdasarkan pengalaman lepas kerana rakaman manual yang sedia ada dan kurang memahami keadaan pesakit akibat gangguan komunikasi (Fang et al. 2016). Perbezaan penyiasatan dan teknik (J A. Fausto et. Al (2012)) pengimejan mengambil kira masa, kos dan kadang kala memberi kerja yang bahaya kepada pengamal perubatan (*Ractification*) (Penzkofer dan Tempny - Afdhal 2014) analisis imej dan rawatan tidak mudah kerana bahagian-bahagian badan yang berbeza memberikan imej dan maklum balas berbeza kepada rawatan. Dimana, tahap rendah integrasi keputusan dalam setiap Jabatan Radiologi, onkologi, patologi dan Urologi (Pillay et al. 2016). Tambahan lagi, sesetengah imej yang ditangkap mengandungi kesamaran yang tidak jelas dan separuh atau masalah-masalah yang tidak

berstruktur. Kebanyakan lebih suka untuk memindahkan masalah ketidakpastian ini kepada patologis untuk siasatan lanjut terperinci sebelum apa-apa keputusan rawatan yang dibuat preskripsi oleh pakar-pakar akan dibuat berdasarkan pengalaman lepas kerana rakaman manual yang sedia ada dan kurang memahami keadaan pesakit akibat gangguan komunikasi (Fang et al. 2016). Petua yang dikumpulkan mengambil kira faktor-faktor risiko dan gejala kanser prostat dan menyimpan mereka sebagai fakta masalah dalam ingatan kerja. Selain itu, pengetahuan pakar dikumpul untuk menjana peraturan dan disimpan dalam asas peraturan. Seterusnya ia akan menjadi asas kepada set peraturan yang berbeza yang berkaitan faktor dan tanda-tanda kanser prostat.

b. Membandingkan Petua Dalam Mengukur Ketepatan

Petua yang terkumpul dalam komputer (pangkalan data atau laporan pemeriksaan) adalah kurang efisien dalam menguruskan keputusan yang melibatkan ketidakpastian (Tu, M.J.P. Castanho et al (2013)). Sehubungan itu, petua yang telah dikumpulkan dari pakar dan petua yang dijana menggunakan pembelajaran mesin diambil untuk dibandingkan dalam menghasilkan petua yang lengkap dan tepat. Namun, kebanyakan pembolehubah dalam membuat petua terlalu banyak untuk dipertimbangkan dalam membuat keputusan penggredan kanser prostat. (Delahunt et al. 2016). Justeru, kajian ini bertujuan untuk membantu para doktor perubatan dalam membuat keputusan tepat dalam masa yang singkat bagi menentukan tahap kanser prostat pesakit dan rawatan selanjutnya.

1.3 OBJEKTIF KAJIAN

Objektif kajian ini adalah untuk membangunkan satu perisian mengesan dan mengklasifikasi kanser prostat berasaskan kepada pengetahuan pakar perubatan. Objektif khusus dalam mencapai objektif utama (melalui rundingan bersama pakar radiologi, pakar patologi dan pakar urologi dalam sesi pemerolehan pengetahuan). Manakala sub-sub objektif adalah seperti berikut;

- 1) **Sistem Pakar:** Untuk merekabentuk, membangun dan menilai sistem pakar kanser prostat yang terdiri daripada tiga modul: wad pembedahan (urologi), radiologi dan histopatologi.
- 2) **Pembelajaran Mesin:** Mencadangkan dan membandingkan model diagnosis penyakit kanser prostat menggunakan kaedah pokok rawak dan petua pengeluaran untuk jabatan radiologi.

1.4 SKOP KAJIAN

Kajian ini akan memberi tumpuan kepada menganalisis perwakilan pengetahuan dalam pengurusan kanser prostat (51 Laporan Pesakit) dan imej diagnostik imbasan CT dan MRI (22 Imej dan Laporan) dengan menggunakan kepintaran buatan dan teknik pengkomputeran lembut (statistik). Beberapa kes pesakit yang termasuk di dalam laporan perubatan bersama-sama dengan radiologi, histologi, patologi dan data farmaseutikal (urologi) diperlukan untuk pengetahuan data lebih baik. Seterusnya pengumpulan data, maklumat, imej-imej akan meliputi semua peringkat kanser prostat.

Pembangunan sistem ini bertujuan untuk membantu diagnosis kanser prostat dengan lebih cepat dan pantas. Sistem ini dibangunkan dengan menggunakan pemahaman dan pengetahuan mengenai kanser prostat. Dengan wujudnya sistem pakar kanser prostat, diharap mampu untuk membantu pakar serta pesakit dalam menyediakan pengetahuan dan maklumat yang pantas serta dapat menyumbang dalam kajian saintifik.

Selain itu, beberapa kaedah interaktif juga turut diimplementasikan supaya penggunaan sistem ini lebih mesra dan memenuhi kehendak pengguna. Selain itu,

sistem ini juga diharap dapat menjadi rujukan dalam penambaan sistem pakar kanser prostat dari semasa ke semasa yang mampu berdiri sama tinggi dengan sistem pakar yang ada. Pengetahuan yang diambil mungkin tidak boleh digunakan secara universal atau sah untuk semua kes. Maklumat yang diperoleh melalui saiz sampel daripada kes-kes perubatan pesakit kebanyakannya untuk lelaki yang telah teruk.

1.5 PENDEKATAN KAJIAN

Kajian ini boleh membantu kanser prostat pengamal perubatan seperti radiologi, pakar pembedahan (urologi) dan pakar histopatologi untuk membuat keputusan yang pantas dan tepat. Pengumpulan data pesakit dijalankan sepenuhnya di Pusat Perubatan Universiti Kebangsaan Malaysia (PPUKM).

Selain itu, kajian ini boleh digunakan secara pedagogi atau memberi inspirasi kepada masa depan penyelidikan kanser prostat.

1.6 MODEL KERANGKA PENYELIDIKAN

Model rangka kerja penyelidikan disediakan bagi membolehkan kajian yang berlansung mempunyai arah tuju untuk mencapai objektif yang telah ditetapkan. Kajian ini melibatkan dua komponen utama iaitu membangunkan sistem pakar dan membanding model diagnosis kanser prostat menggunakan kaedah pokok rawak dan petua pengeluaran disediakan bagi menggambarkan model rangka kerja ini seperti di Rajah 3.1.

1.7 ORGANISASI PENULISAN

Penulisan tesis ini adalah meliputi enam bab utama. Bab I menerangkan berkaitan pernyataan masalah, persoalan kajian, objektif kajian, skop kajian, metodologi kajian secara umum, kepentingan kajian dan model rangka kerja penyelidikan kajian.

Bab II Kajian Kesusasteraan yang menerangkan kajian yang dijalankan oleh penyelidik berkaitan sistem pakar, faktor pembangunan sistem pakar, model sistem pakar, teknik perlombongan data, aspek nilai cipta dan model nilai cipta yang serupa

seta boleh diguna pakai dalam kajian ini. Akhir sekali, bab ini menjelaskan mengapa teknik pokok keputusan pokok rawak dipilih dan dibandingkan dengan petua pengeluaran.

Bab III menjalankan Metodologi Kajian menerangkan secara terperinci berkaitan metodologi sistem pakar yang diguna pakai dalam kajian ini bagi mendapatkan hasil kajian yang memenuhi objektif kajian yang ditetapkan. Fasa dan aktiviti berdasarkan metodologi tersebut yang dijalankan sepanjang kajian juga diperjelaskan dalam bab ini.

Bab IV Model Petua Pengeluaran menerangkan hasil kajian yang diperoleh dan dijana melalui pengetahuan pakar dan faktor yang bersesuaian untuk pembangunan sistem pakar berdasarkan petua nilai cipta sama serta analisis terperinci yang dilaksanakan terhadap hasil kajian tersebut. Faktor tersebut kemudiannya digunakan sebagai input kepada membangunkan petua pakar berdasarkan pengetahuan yang diperolehi.

Bab V Model Pokok Rawak yang menerangkan secara terperinci berkaitan kaedah penilaian terhadap petua pokok rawak pakar yang dibangunkan. Hasil perbandingan ketepatan ramalan bagi teknik pembelajaran mesin pokok keputusan terbaik diuji J48, pokok rawak dan hutan rawak sebelum pokok rawak dipilih. Petua pakar ini disahkan secara pepadanan dengan sistem dan pengesahan pakar pada kajian kes yang telah dikenal pasti. Seterusnya, hasil daripada ujikaji yang dijalankan, petua pokok rawak dan petua pengeluaran akan dibandingkan untuk menentukan teknik yang paling baik dalam meramal tahap bahaya kanser prostat. Pokok rawak yang telah dijana turut disertakan di dalam bab ini dan juga petua yang telah diekstrak daripada pokok rawak juga dinyatakan.

Bab terakhir iaitu bab VI adalah Kesimpulan yang menerangkan dengan lebih lanjut tentang rumusan yang diperoleh daripada kajian yang dijalankan berdasarkan hasil dan dapatan kajian. Sumbangan kajian serta cadangan kepada penyelidikan akan datang juga turut dibincangkan dalam bab ini.

1.8 RUMUSAN

Dalam kajian ini, pembinaan petua adalah medium utama yang digunakan untuk meramal kanser prostat melalui teknik pembelajaran mesin pokok rawak dan pangkalan pengetahuan. Langkah-langkah yang paling penting dalam kajian ini adalah pengumpulan data dan pra-pemprosesan data kerana aspek ini perlu dititik beratkan untuk mendapatkan model peramalan tahap kanser prostat yang terbaik.

BAB II

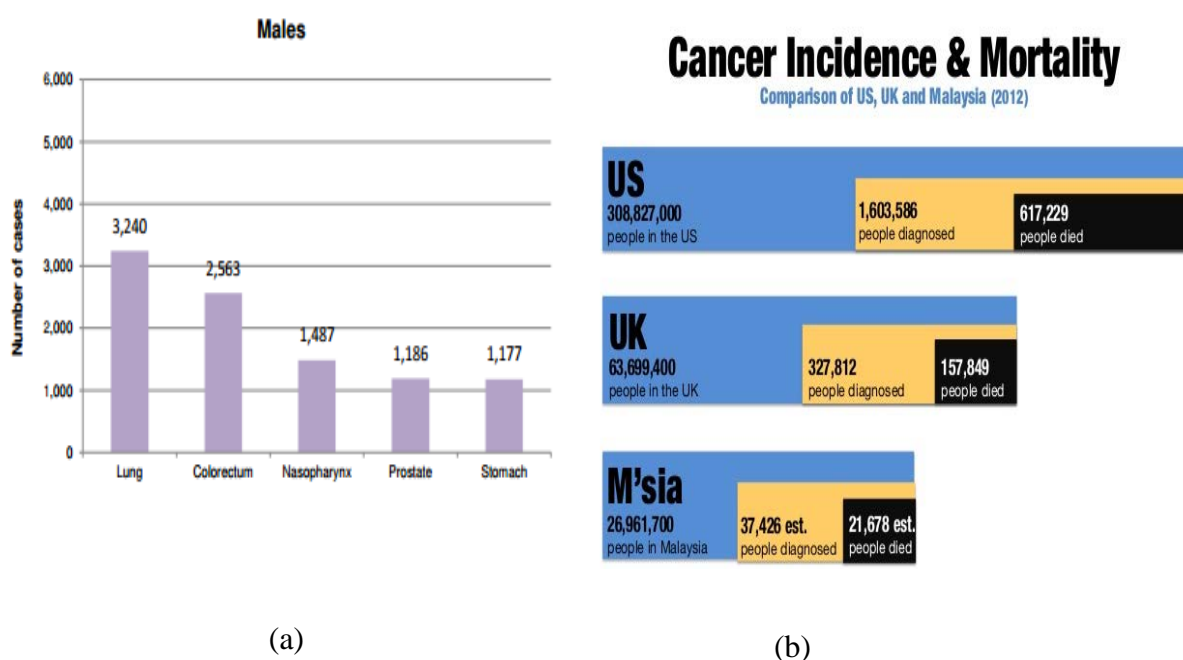
KAJIAN KESUSTERAAN

2.1 PENGENALAN

Tinjauan kesusteraan merupakan salah satu aspek yang penting dilakukan sebelum membangunkan sesuatu sistem. Dalam usaha membangunkan sistem beberapa perkara perlu dikenalpasti dan diambil kira untuk menentukan sistem yang terbaik yang dapat dibangunkan. Perkara seperti pernyataan masalah, cadangan penyelesaian, objektif kajian, skop kajian, metodologi kajian, jadual pembangunan sistem yang telah dibincangkan secara teliti mampu memberi asas yang kukuh untuk mendapatkan hasil sistem pakar yang terbaik dalam membantu pakar mendiagnosis kanser prostat.

Melalui kajian ini, pendekatan bagi teknik penyelesaian masalah, kekangan sistem dan pilihan teknologi yang boleh digunakan untuk menyelesaikan masalah agar dapat menangkap gambaran awal dan pengetahuan lebih mendalam berkenaan sistem yang akan dibangunkan. Segala informasi yang diperolehi perlu diteliti dan dikaji secara menyeluruh termasuk aspek kelebihan dan kelemahan. Kajian literatur diperlukan untuk meningkatkan kefahaman di dalam bidang yang dikaji. Tinjauan terhadap kajian terdahulu membantu penyelidik memikirkan hipotesis yang diperlukan bagi menjawab persoalan kajian dan memantapkan permasalahan kajian tersebut. Seterusnya, ia amat penting kerana dapat menonjolkan kebolehan tersendiri dalam mengenalpasti maklumat yang berkaitan dan seterusnya memaparkan tahap kefahaman kepada kajian yang dikaji seperti maklumat berkaitan penyakit kanser. Ia amat membantu dalam menambahbaik atau menukar sesuatu aspek dalam kajian yang dijalankan untuk membina sesuatu sistem yang lebih baik dan berkesan terhadap penggunaannya. Dalam setiap tinjauan ini terdapat beberapa cadangan dan cetusan idea yang diperolehi sabagai panduan untuk membangunkan automasi diagnosis bersepadu kanser prostat ini.

Kanser prostat adalah kanser yang paling kerap berlaku dan merupakan kanser keenam di dunia mengikut dari segi bilangan kes terbaru. American Cancer Society menganggarkan sebanyak 180,890 kes baru dan 26,120 kes kematian akibat kanser prostat pada tahun 2016. Kanser ini juga turut menjadi penyumbang kanser keempat di Malaysia yang mana berlaku terhadap golongan lelaki (Omar dan Tamin 2011). Antara pesakit yang didiagnos di Malaysia, sebanyak 15.8 peratus daripadanya lelaki berbangsa Cina, 14.8 peratus lelaki India dan 7.7 peratus lelaki Melayu menghidap penyakit berkenaan (Mohd Pilus 2015).”



Rajah 2.1 (a) Kadar kematian dan bilangan bagi kanser prostat (b) Perbandingan kadar kematian bagi kanser prostat

Sumber: Laporan World Health Organization, 2014 dan 2012

Rajah 2.1 (a), menunjukkan kes kanser prostat pada 2014 sebanyak lebih dari 3 kali ganda, yang dilaporkan pada tahun 2014 berbanding laporan dari Pendaftaran Kanser Kebangsaan sebanyak 502 kes kanser prostat didiagnosis pada tahun 2007. Manakala Rajah 2.1 (b) pula menunjukkan perbandingan kadar kematian akibat kanser antara Amerika Syarikat (34.5%), United Kingdom (48.2%) dan Malaysia (57.9%). Pengesanan awal mampu menyembuhkan pesakit dan seterusnya mengurangkan kadar kematian serta meningkat kadar hayat pesakit kanser (Campbell 2014). Tetapi,

kebanyakan kes kanser prostat di Malaysia hanya ditemui di peringkat yang agak parah di mana kanser telah mula merebak dan rawatan agak mustahil untuk berjaya (Cancercare 2016). Oleh itu, terdapat keperluan untuk meningkatkan kesedaran dan pengesanan awal kanser ini.

Beberapa kajian telah mula dijalankan menggunakan pendekatan pembelajaran mesin dan kecerdasan buatan (Castanho et al. 2013; Cinar et al. 2009). Beberapa teknik pengimejan boleh membantu untuk diagnosis dengan tepat dan berkesan bagi memantau kanser ini (Fausto et al. 2012). Antara yang terkini adalah pemprosesan imej histopatologi telah digunakan untuk mempercepatkan keputusan penggredan tisu displastik bagi pengesanan pra-kanser esofagus (Adam 2015; Adam et. al. 2012), kanser otak dan kanser sel darah, serta mengira anggaran percambahan kanser agresif (Alomari et al. 2016) dan penggredan kanser prostat (Albashish et al. 2015). Selain itu, diagnosis berdasarkan imej radiologi untuk kawasan kanser otak (Alomoush et al. 2014; Kaidar et al. 2013), kanser payu dara (Baharuddin et al. 2016) dan pengelasan kanser normal dan tidak normal (Sheikh Abdullah et al. 2016).

Diagnosis kanser prostat adalah sukar kerana terdapat ketumbuhan yang tidak dikenali (Murat et al. 2009). Keperluan untuk diagnosis kanser prostat adalah berdasarkan faktor gejala, pemeriksaan fizikal rektum dan peningkatan tahap prostatik antigen tertentu dalam darah (PSA). Sebagai contoh, Jadual Partin et al. (2001) digunakan secara meluas di Eropah dan Amerika Utara sebagai panduan untuk memilih terapi muktamad (Castanho et al. 2013) di mana ia menggabungkan kebarangkalian pra operasi serum PSA dengan penggredan TNM dan Gleason untuk mengira kemungkinan peringkat patologi akhir dengan kemas kini terbaru pada tahun 2007.

Tambahan lagi, kaedah-kaedah diagnosis yang berdasarkan faktor individu, pelbagai model regresi logistic (LR) dan rangkaian neural berulang (RNN) telah dibangunkan sejak 1994 (Murat et al. 2009). Menurut Castanho et al. (2013), ramai penyelidik telah mengguna sifat pelengkap daripada tiga jenis teknik dalam gabungan RNB, RL dan AG untuk meningkatkan prestasi sistem. Salah satu contoh adalah penggunaan Sistem Neuro-Fuzzy telah digunakan bagi mengelaskan kanser prostat dan mendiagnosis kanser payudara. Sistem Genetik Kabur pula telah digunakan dalam

diagnosis kanser payudara dan berjaya mencapai prestasi pengelasan tinggi (Castanho et al. 2013).

Selain itu, sistem pengimejan bantuan komputer dan sistem pemprosesan imej yang direka untuk membantu pengamal perubatan membuat keputusan berstruktur, separa berstruktur dan tidak berstruktur. Oleh itu, kajian-kajian lalu telah membuktikan bahawa Pengecaman Kanser Prostat adalah masih relevan untuk diterokai dan dibangunkan. Model pengecaman, pengelasan, peramalan serta kesan pelbagai pembolehubah masih luas untuk diterokai dan diperbaiki untuk mengurangkan statistik penghidap atau kematian akibat kanser prostat baik di Malaysia, mahu pun di peringkat dunia.

2.2 PERINGKAT KANSER PROSTAT DAN RAWATAN

Kanser prostat biasanya diperingkatkan dan berisiko strata oleh tiga sifat. Pertama adalah peringkat *tumor*, *lymph nodes* dan *metastasis* (TNM) yang memberitahu tahap kanser penyakit dan darjah rebakan ke nodus limfa yang jauh di dalam badan. Kedua dan akhir sekali, Gleason skor dan antigen khusus prostat (PSA) yang mengukur perbezaan sel tumor histologi dan jumlah PSA protein yang dilepaskan ke dalam aliran darah masing-masing (Smith et al. 2009).

Terdapat 4 peringkat kanser prostat dan penentuan untuk pesakit tertentu memerlukan 4 jenis data. Antaranya adalah saiz tumor primer (T), serantau nod limfa metastasis (N), metastasis jauh (M) dan penggredan histopatologi (G). Pengelasan terperinci daripada klasifikasi TNM untuk kanser prostat diambil daripada (Eble 2004) dan ditunjukkan dalam Rajah 2.2.

Klasifikasi kanker prostat menurut TNM
Tx: Tumor primer tidak dapat ditemukan
T1: Tumor tidak terdeteksi secara klinis, tidak dapat diraba atau terlihat dengan pencitraan <ul style="list-style-type: none"> • T1a: Tumor ditemukan secara kebetulan dari pemeriksaan histopatologis pada kurang dari atau sama dengan 5% dari jaringan yang direseksi • T1b: Tumor ditemukan secara kebetulan dari pemeriksaan histopatologis pada lebih dari 5% dari jaringan yang direseksi • T1c: Tumor ditemukan dari biopsi (karena ada peningkatan kadar PSA)
T2: Tumor terbatas pada prostat <ul style="list-style-type: none"> • T2a: Tumor mengenai kurang dari atau sama dengan setengah lobus prostat • T2b: Tumor mengenai lebih dari setengah lobus tetapi tidak pada kedua lobus prostat • T2c: Tumor mengenai kedua lobus prostat
T3: Tumor meluas melewati kapsul prostat <ul style="list-style-type: none"> • T3a: Tumor meluas di luar kapsul prostat (satu atau kedua sisi) termasuk penyebaran ke leher buli secara mikroskopik • T3b: Tumor mengenai kelenjar vesikula seminalis
T4: Tumor terfiksir atau mengenai struktur sekitar prostat selain vesikula seminalis: seperti katup luar buli, rektum, otot levator dan/atau dinding pelvis
Nx: Kelenjar getah bening regional tidak dapat dinilai
N1: Tidak ada penyebaran kelenjar getah bening N2: Penyebaran pada kelenjar getah bening regional
Mx: Penyebaran luas tidak dapat dinilai
M1: Penyebaran luas <ul style="list-style-type: none"> • M1a: Penyebaran kelenjar getah bening bukan regional • M1b: Penyebaran ke tulang • M1c: Penyebaran ke organ lain

Rajah 2.2 TNM Pengelasan karsinoma daripada prostat

Sumber: Ubahsuai Eble 2004

Pengredan Gleason masa ini merupakan sistem pengredan histopatologi yang utama untuk kanser prostat berdasarkan seni bina kelenjar. Terdapat 5 pola histologi atau gred dengan mengurangkan perbezaan yang ditakrifkan oleh sistem pengredan. Skor gleason adalah jumlah yang paling lazim (utama) dan tersebar luas kedua (sekunder) corak histologi yang terdapat dalam sel prostat biopsi. Sebagai contoh, jika corak utama ditemui adalah gred 4 dan corak gred sekunder adalah 3, skor Gleason adalah 7 (Eble 2004).

Penyelidikan perubatan telah mendapati bahawa tahap PSA biasa ada hubung kait dengan baik dengan peningkatan usia. Ini menunjukkan bahawa nilai-nilai yang berkaitan dengan usia (contohnya, lelaki di bawah umur 50 tahun mempunyai nilai PSA biasa kurang daripada *2.5 nanogram / mililiter*) boleh digunakan sebagai penanda kanser prostat. Ia telah kemudiannya didapati bahawa PSA serum tidak mempunyai sensitiviti yang tinggi dan kekhususan menjadi penunjuk kanser prostat yang baik. Masalah ini telah sebahagiannya diatasi dengan memperkenalkan pengiraan indeks PSA lain yang berkaitan dengan atau / dan penanda serum lain seperti halaju PSA, PSA menggandakan masa, *Prostatic Acid Phosphatase (PAP)* dan *Human Kallikrein-2 (HK2)* (Eble 2004).

Seterusnya, *prostatectomy*, *hormone terapi*, *cryosurgery*, intensiti tinggi tertumpu ultrasound atau kombinasi semua mereka adalah rawatan biasa selepas kanser prostat itu sedang didiagnosis. Setiap rawatan mempunyai kesan sampingan yang berbeza dan risiko, pada masa ini, tidak ada kata sepakat wujud pada rawatan yang terbaik untuk kanser prostat tetapi rawatan ini biasanya bergantung kepada keadaan kesihatan prostat kanser pesakit (Kim et al. 2010).

2.2.1 Pembedahan (Urologi)

Jabatan ini merupakan jabatan pertama yang perlu dikunjungi oleh pesakit sekiranya disyaki menghidapi kanser prostat atau masalah kencing. Doktor dari jabatan ini akan bertanyakan soalan kepada pesakit untuk memastikan tahap risiko pesakit menghidapi kanser prostat dan simpton kencing (Cancer.Org 2015). Sekiranya pesakit dikenalpasti mempunyai tahap risiko yang tinggi, pesakit adalah digalakkan untuk meneruskan pemeriksaan di bahagian radiologi ataupun mendapat rawatan selanjutnya.

2.2.2 Radiologi

Seterusnya, pesakit akan melakukan pemeriksaan pengimejan di jabatan radiologi untuk mengetahui prostat yang disyaki menghidapi kanser. Pemeriksaan pengimejan dilakukan dengan menggunakan sinar x, medan magnet, gelombang bunyi atau bahan radioaktif untuk menghasilkan gambar pada bahagian dalaman badan seseorang. Tujuan melakukan pemeriksaan pengimejan adalah untuk mengenalpasti bahagian yang disyaki ada sel kanser, mengesan kedudukannya dan pada tahap mana kanser telah merebak.

2.2.3 Histopatologi

Histopatologi merupakan satu bidang perubatan yang melakukan pemeriksaan mikroskopik terhadap sampel tisu untuk menentukan jenis dan grad sesuatu penyakit. Dalam proses diagnosis kanser prostat, pakar patologi akan menentukan sama ada sel kanser wujud atau tidak dalam sampel tisu yang diambil dari prostat pesakit melalui proses biopsi. Teknik-teknik biopsi yang digunakan merangkumi biopsi aspirasi jarum halus, biopsi jarum teras dan pembedahan biopsi. Sampel-sampel tisu akan diperiksa

oleh pakar patologi dengan menggunakan mikroskop untuk melihat sebarang ketidaknormalan yang ada.

2.2.4 Onkologi

Onkologi merupakan bidang perubatan yang melakukan kajian tentang tumor (Cancer.Net 2013). Seseorang doktor yang mempunyai pengkhususan dalam rawatan pesakit kanser dikenali sebagai pakar onkologi. Tiga bidang utama yang terdapat dalam onkologi adalah onkologi perubatan yang fokus dalam merawat kanser dengan ubat seperti kemoterapi, onkologi pembedahan yang merawat kanser dengan pembedahan dan juga onkologi radiologi yang menggunakan terapi sinaran untuk menghapuskan sel kanser.

2.3 PENGESANAN DAN KAEDAH DIAGNOSIS KANSER PROSTAT

Kanser prostat biasanya tidak menunjukkan tanda atau gejala serius terutama di peringkat awal. PSA dan *digital rectal examination* (DRE) adalah kaedah saringan untuk mengenalpasti kanser prostat. Jika tahap PSA yang tinggi dan pemeriksaan rektum menemui kawasan firma atau nodular, ia boleh menunjukkan satu petunjuk yang mencurigakan untuk kanser prostat (Ellsworth 2012).

Terdapat prosedur bukan invasif dan invasif dalam mendiagnosis kanser prostat teknik pengimejan seperti imbasan pengimejan resonan magnet (MRI), imbasan tomografi berkomputer (CT), imbasan tulang dan antigen (PSA). Tahap prostat khusus dalam darah yang merupakan bekas manakala transrectal ultrasonography, pemeriksaan rektum dan biopsi prostat adalah yang terakhir (Hoffelt et al. 2003; Eble 2004; Ou et al. 2009; Saritas et al. 2010; Kim et al. 2010). Imbasan (CT) kerap digunakan untuk mengukur isipadu prostat melalui sinaran *external beam radiotherapy* (EBRT) dan ruang-antara implantasi sering digunakan untuk mengukur isipadu prostat dan *interstitial implantation* (Hoffelt et al. 2003). Tulang scintigraphy (imbasan tulang radionuclide) adalah satu kaedah yang sensitif boleh digunakan untuk menentukan sama ada kanser prostat sudah menular ke bahagian tulang (Eble 2004).

MRI sering digunakan dalam perubatan, tetapi ia mampu menghasilkan kepekaan dan spesifikasi yang diperlukan untuk menentukan kanser prostat. Piawaian emas untuk diagnosis kanser prostat yang sedang digunapakai ialah biopsi prostat (Ou et al. 2009). Walau bagaimanapun, keputusan standard prosedur emas ini tidak dilakukan kepada pesakit risiko kanser yang rendah mungkin perlu mengelakkan proses ini kerana kos yang tinggi dan komplikasi yang mungkin timbul. Sebagai contoh, risiko mukosa rektum yang rosak (Saritas et al. 2010).

2.3.1 Pengimbasan Tulang

Pengimejan untuk mengesan perubahan patologi dalam tulang adalah scintigraphy tulang. Menurut pakar klinikal pengimejan imbasan tulang adalah penilaian penyakit di mana keumpulan pesakit yang dirujuk untuk imbasan tulang adalah pesakit yang mempunyai metastatik. Rujukan kebiasaannya bagi mereka yang berisiko tinggi dan untuk membuat tindak balas rawatan penilaian. Kanser prostat mempunyai kecenderungan untuk menular ke kelenjar limfa dan rangka sebagai organ pilihan.

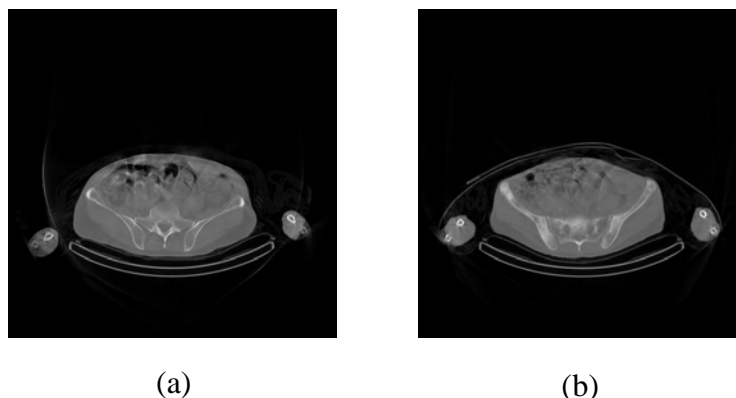
Pemeriksaan pengimejan dilakukan menggunakan kamera gamma Rajah 2.3. Imbasan tulang seluruh badan diperolehi 3 hingga 4 jam selepas pentadbiran 600 MBq ^{99m}- *Methylene Diphosphonate* (MDP). Prosedur imbasan mengambil masa lebih kurang 25 minit dan hasilnya adalah dua imej dua dimensi, satu anterior dan posterior imej. Imej-imej dua dimensi biasanya cukup untuk menunjukkan sama ada terdapat apa-apa perubahan patologi dalam rangka.



Rajah 2.3 Kamera gamma dengan keupayaan untuk memperoleh satah seluruh badan dan tomografi imej

Sumber: https://sc02.alicdn.com/kf/HTB1y1QsKpXXXXcHXFXXq6xXFXXu/NM-Gamma-Camera.jpg_220x220.jpg

Tafsiran imbasan tulang seluruh badan adalah satu penilaian visual, dan amat bergantung kepada tahap pengalaman doktor perubatan nuklear. Kawasan yang mempunyai pengambilan pengesanan tinggi dalam tulang boleh menjadi varian normal, tetapi pengambilan fokus dalam tulang sentiasa perlu dipertimbangkan dengan teliti. Pengambilan yang tinggi boleh disebabkan oleh trauma, dengan itu sejarah pesakit adalah penting. Pembaca imbasan tulang perlu mengambil kira penyetempatan, simetri, bentuk dan intensiti setiap luka untuk membezakan perubahan patologi akibat dari aktiviti metastatic atau daripada perubahan degeneratif dalam rangka. Normal dan tidak normal imbasan tulang (pelvik) digambarkan dalam Rajah 2.4.



Rajah 2.4 (a) Normal imbasan tulang (pelvik) (b) Tidak normal imbasan tulang (pelvik)

Sumber: Hospital Universiti Kebangsaan Malaysia

2.3.2 SPECT / CT

Maklumat fungsi dan struktur tentang kanser organ yang berbeza boleh diperolehi dengan menggunakan pengimejan dua modaliti, contohnya tomografi berkomputer *Computed tomography* (CT) digabungkan dengan satu foton pelepasan dikira tomografi *Single photon emission computed tomography* (SPECT). Bekas menyediakan maklumat struktur ke atas badan dan SPECT yang menyediakan data yang berfungsi. Tiga dimensi SPECT dan CT imej boleh digabungkan atau yang ditunjukkan dalam imej yang sama, supaya analisis serentak dijalankan. CT juga boleh menunjukkan sama ada terdapat apa-apa perubahan degeneratif pada sendi yang boleh menjelaskan peningkatan pengambilan pengesanan dalam imbasan tulang.

2.3.3 MRI

Satu lagi modaliti pengimejan yang digunakan pada pesakit *Prostate-specific antigen* (PCA) adalah pengimejan resonan magnet *Magnetic resonance imaging* (MRI). Modaliti ini boleh digunakan untuk peringkat tempatan kanser prostat awal. MRI membezakan tisu-tisu lembut dengan ciri-ciri yang berbeza. Kanser dalam prostat boleh diterangkan dengan lebih tepat dari segi lokasi dan berhubung dengan organ-organ lain, terutamanya yang berkaitan dengan kehadiran lanjutan Extracapsular, tetapi penggunaan MRI untuk PCA masih terhad.

2.3.4 PET

Positron emission tomography (PET) adalah satu lagi modaliti pengimejan nuklear, dan ia dapat memberikan resolusi tinggi imej tiga dimensi tubuh manusia. Kaedah ini, bersama-sama dengan CT, telah terbukti menjadi satu modaliti pengimejan yang berjaya yang menawarkan resolusi imej yang lebih tinggi dan spesifikasi yang lebih besar, membolehkan diagnosis metastasis dalam tulang dan tumor pada tisu lembut. Dalam berisiko tinggi pesakit prostat kanser menunjukkan sensitiviti dan spesifisiti yang lebih besar dalam pengesanan metastasis tulang menggunakan ^{18}F -Fluorida PET / CT daripada menggunakan imbasan tulang seluruh badan. Walau bagaimanapun, penglibatan PET dan MRI tidak terdapat banyak di pusat penjagaan kesihatan di dunia, dan kos ini modaliti pengimejan adalah lebih tinggi daripada imbasan tulang.

PET / MRI masih dalam penyelidikan, tetapi ia boleh menjadi penting untuk mengesan malignan dalam pengimejan Onkologi. PET / MRI boleh memainkan peranan kanser untuk pengurusan utama dalam bidang pengimejan molekul, dengan cara yang menjalankan angiogenesis tumor dan mekanisme sistem pengisyaratan sel dapat dijelaskan.



Rajah 2.5 PET / CT untuk mengesan penyakit metastatic

Sumber: <http://mrifresno.com/wp-content/uploads/2013/12/ct-scanner1.jpg>

2.4 ETIKA PENYELIDIK (PERUBATAN)

Seseorang penyelidik yang ingin melakukan kajian yang berkenaan dengan manusia mesti mengikut prosedur etika penyelidikan yang ditetapkan. Ini termasuk kajian yang berkenaan dengan orang, data mereka ataupun spesimen biologi. Tujuan mempraktikkan etika penyelidikan adalah untuk menjaga kesejahteraan sejagat bagi semua pihak yang terlibat dalam kajian. Carta alir proses pengurusan projek yang berkenaan dengan manusia bersama Research Ethics Committee (REC), Universiti Kebangsaan Malaysia adalah seperti yang dilampirkan di Lampiran E.

2.5 KEPINTARAN BUATAN DALAM BIDANG PERUBATAN

Perkembangan teknologi dalam bidang perubatan telah memberi harapan kepada jutaan pesakit yang sedang menderita. Teknologi seperti mammogram, ultrabunyi dan MRI telah membantu dalam mengesan kanser prostat. Selain itu, sistem pakar juga turut memainkan peranan yang penting dalam diagnosis kanser prostat.

2.5.1 *Computer Aided Diagnosis (CAD)*

Konsep asas CAD telah dicadangkan pada pertengahan 1980-an oleh University of Chicago. Tujuan utama mereka bentuk CAD adalah untuk menyediakan output komputer sebagai pandangan yang kedua untuk membantu pakar radiologi untuk menganalisis imej. Oleh itu, ketepatan dan konsistensi diagnosis radiologi boleh diperbaiki dan juga mengurangkan masa menganalisis imej.

Sejak sistem CAD digunakan sebagai pandangan yang kedua untuk membantu pakar radiologi, sistem CAD mempunyai kebolehan sama seperti radiologi dari segi mesin pembelajaran dan pengecaman pola. Oleh itu, pembelajaran mesin dan teknik pengecaman pola adalah penting dalam pembangunan sistem CAD.

Pengecaman pola adalah proses mengenalpasti ciri-ciri dari beberapa objek dalam data mentah dan membuat keputusan berdasarkan pengelasan output seperti mengklasifikasikan setiap objek kepada satu kategori yang mempunyai pelbagai corak (Garcia-Uribe et al. 2004).

Pada dasarnya, terdapat dua jenis sistem CAD. Sistem CAD yang pertama adalah untuk mengklasifikasikan semua imej CT kepada abnormal dan normal. Keduanya pula adalah sistem yang digunakan untuk mengklasifikasikan kes yang tidak diketahui untuk mengenalpasti bilangan jenis keabnormalan iaitu melebihi dua kelas. (Arimura et al. 2009).

Sebuah badan organisasi yang ditubuhkan dengan tujuan mengurangkan impak kanser di kalangan komuniti telah ditubuhkan di bawah nama The Cancer Australia team. Antara usaha yang dilakukan adalah pendedahan kepada risiko (*Risk Factor*) menghadapi kanser yang boleh dilayari di laman web rasmi mereka (canceraustralia.gov.au 2015). Faktor risiko ini boleh didapati di pautan <https://prostate-cancer.canceraustralia.gov.au/risk-factors>. Secara ringkasnya risiko ini ialah sebuah aplikasi web yang membenarkan orang umum mengenalpasti tahap risiko dengan menjawab soalan-soalan bersiri dipautan tersebut. Aplikasi web ini kemudiannya akan memberi cadangan berkenaan dengan tindakan yang harus dilakukan oleh responden berdasarkan tahap risiko yang dikenalpasti.

Artikel yang bertajuk *Cancer Detection Using Artificial Neural Network and Support Vector Machine: A Comparative Study* (2013) menyatakan bahawa penggunaan sistem pembelajaran Kecerdasan Buatan (AI) seperti Rangkaian Neural Buatan (ANN) dan Mesin Vektor Sokongan (SVM) boleh membantu doktor dalam mendiagnosis kanser dengan lebih tepat. Keupayaan kedua-dua algoritma ini telah terbukti dalam menghasilkan prestasi yang baik bagi pengelasan kanser.

2.5.2 Penggunaan Kepintaran Buatan Dalam Kanser Prostat

Pada tahun-tahun kebelakangan ini penyelidikan kecerdasan buatan yang dilakukan ke atas kanser prostat, banyak perhatian telah dibelanjakan untuk model tugas pengelasan.

Kaedah automatik sepenuhnya bagi segmentasi tumor prostat di mana mengintegrasikan pengelasan hutan rawak dengan hierarki yang bersyarat rombakan rawak dalam skim pengurangan tenaga (Horning 2010).

Penyelidik menangani masalah set data jujukan mikro berasaskan pengelasan kanser menggunakan cadangan yang baru Sistem Pengelas Pelbagai (MCS), yang dirujuk sebagai Hutan Putaran (*Rotation Forest*). Hutan putaran merupakan kaedah pertama yang dilakukan ke atas pengelasan set data *Microarray*. *Principal Component Analysis* (PCA), *Non-Parametric Discriminant Analysis* (NDA) and *Random Projections* (RP) digunakan untuk ciri transformasi di Hutan Putaran yang asal (Ho 1995).

Penyelidik ini menggunakan hutan rawak (Ye et al 2013; Bauer et al. 2012) sebagai pengelas dan bukannya mesin sokongan vektor (SVM). Hutan rawak adalah kumpulan pokok keputusan di mana dipilih secara rawak. Latihan dalam setiap pokok keputusan dilakukan dengan mengoptimumkan parameter fungsi berpecah di setiap pokok nod dengan memaksimumkan maklumat yang memisahkan data latihan. Bagi ujian pula, vektor ciri ini ditolak melalui setiap pokok, menggunakan ujian di setiap nod yang berpecah sehingga nod daun dicapai

Artikel yang bertajuk *Prostate Cancer Classification Using Logistic Regression and Random Forest On Pelvis Ct Images* menerangkan bahawa terdapat beberapa Metodologi prapemprosesan imej yang digunakan dalam kajian ini meningkatkan peluang untuk mengesan tumor dengan lebih tepat pada kebanyakan imej CT (Uzair et al. 2015). Ini membawa kepada kesahihan pengelasan kepada normal dan abnormal. Objektif eksperimen ini adalah untuk membandingkan prestasi Hutan Rawak dan Regresi Logistik (LG) ini terpilih kerana tiada perbezaan pada kesemua min pengelas ini yang telah dibuktikan dalam kajian ujian Anova (Nayef et al. 2013). Hutan Rawak menggunakan kedalaman maksimum pokok, bilangan slot perlakuan, bilangan ciri, bilangan pokok dan benih dengan nilai 0,1,10,20 dan 1, masing-masing. Manakala Regresi Logistik (LG) mempunyai kadar pembelajaran, bilangan lelaran dan bilangan vektor yang sama iaitu -1 dan $1.0E-8$, masing-masing. Keputusan perbandingan prestasi Hutan Rawak dan logistic regression (LG) mengikut keesahan silang yang berbeza. Keesahan silang yang digunakan adalah 10 lipatan dan 25 lipatan. Hutan Rawak memperoleh ketepatan tertinggi iaitu 88% pada keesahan silang adalah 10 lipatan. Walau bagaimanapun, Hutan Rawak memperoleh nilai terendah pada 10 dan 25 lipatan

berbanding Regresi Logistik (LG) yang mempunyai ketepatan 90% yang sama pada 10 dan 25 lipatan.

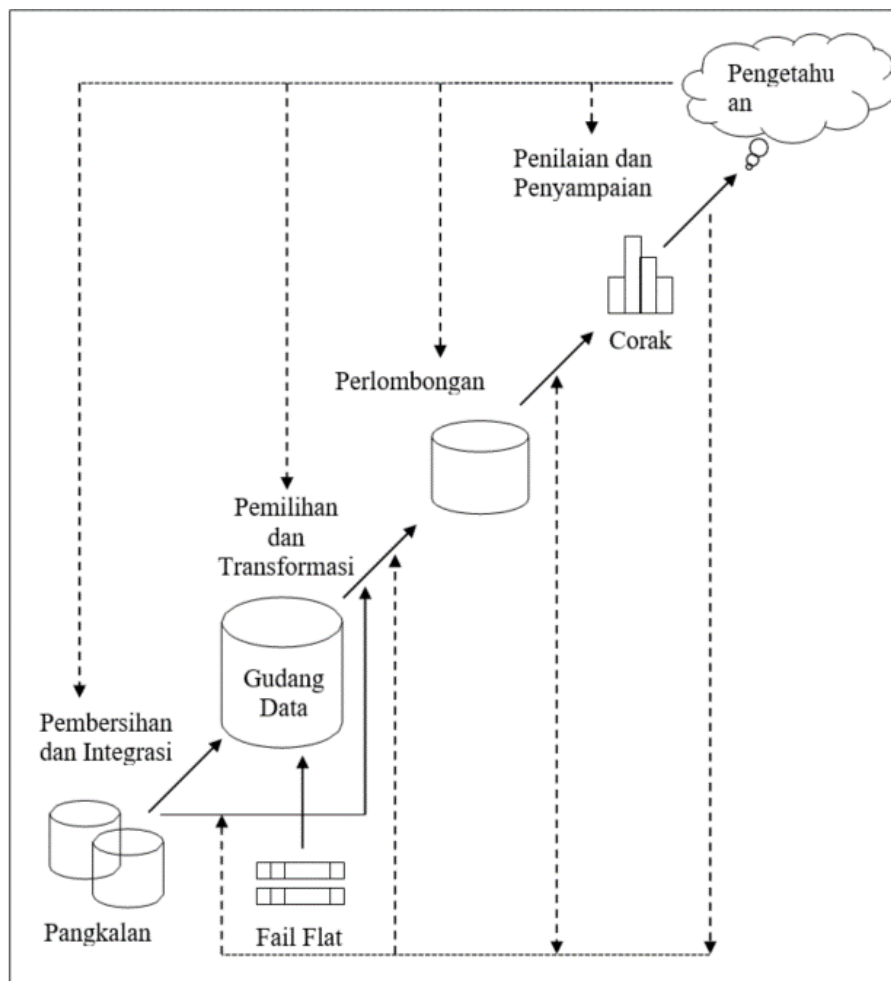
Sementara itu, pengimejan berkomputer dan sistem pemrosesan imej telah juga diterokai untuk membantu pengamal perubatan dalam membuat struktur, separuh berstruktur atau keputusan tidak berstruktur. Semua ini membuktikan bahawa pengesanan kanser prostat adalah masih relevan untuk siasatan lanjut dan pembangunan. Terdapat banyak ruang kajian pengesanan yang lebih baik, klasifikasi serta ramalan pemodelan untuk menurunkan statistik angka kematian disebabkan oleh kanser prostat di Malaysia serta dunia.

2.6 PERLOMBOGAN DATA

Perlombongan data sememangnya merupakan satu tugas yang sangat penting dalam membangunkan sebuah sistem pakar. Aktiviti perlombongan data ini biasanya dijalankan pada data-data yang banyak dan besar. Perlombongan data merupakan satu bidang yang telah berkembang hasil daripada maklumat dan data yang semakin bertambah. Perlombongan data juga melibatkan proses pembinaan model secara berulang menggunakan kaedah tertentu. Kaedah tersebut dianggap sebagai kaedah pintar iaitu melibatkan algoritma pembelajaran terhadap data. Ia adalah salah satu daripada cabang kecerdasan buatan yang bertujuan untuk mendapatkan pengetahuan tertentu daripada maklumat yang ada. Perlombongan data ialah kaedah yang digunakan untuk mengekstrak atau mendapatkan corak tertentu daripada data yang boleh dijadikan pengetahuan. Secara ringkasnya, perlombongan data adalah satu proses yang menggunakan peralatan penganalisis data untuk menemui corak dan hubungan dalam data yang mungkin dapat digunakan untuk membuat petua yang sahih (Mannila 1996).

Terdapat banyak bidang yang mengaplikasi perlombongan data iaitu dalam bidang seperti pemasaran, kewangan, perubatan, keselamatan dan kejuruteraan. Perlombongan data adalah proses penemuan pengetahuan dengan mengekstrak corak menarik iaitu tidak lazim, tidak diketahui sebelumnya dan berpotensi digunakan, daripada jumlah data yang besar yang disimpan dalam pangkalan data atau gudang data. Selain daripada itu, definasi-definasi lain perlombongan data ialah perlombongan data merupakan satu bahagian penting dalam KDD (*Knowledge Discovery in Databases*), di

mana KDD merupakan proses keseluruhan yang menukarkan data mentah kepada maklumat yang berguna, yang ditunjukkan dalam Rajah 2.6. Proses ini melibatkan langkah-langkah berikut:



Rajah 2.6 Langkah – langkah KDD (Knowledge Discovery in Database)

Seterusnya, merujuk kepada Hand et al. (2001), perlombongan data ialah analisis tentang set-set data (seringkali besar) untuk mencari hubungan yang boleh dipercayai dan data diringkaskan kepada bentuk baru yang lebih mudah difahami dan berguna kepada pemilik data. Menurut Michalski (2000), dalam tahap aplikasi pengetahuan, pengetahuan yang ditemukan akan digunakan atau digabungkan dengan sistem keputusan atau sistem pakar. Oleh itu, kaedah perlombongan data dapat membantu manusia untuk melakukan analisis, menginterpretasi dan memahami data dalam jumlah yang banyak serta menyokong proses pembuatan keputusan.

Han dan Kamber (2000) mendefinisikan perlombongan data kepada proses mengekstrak pengetahuan daripada jumlah data yang besar. Terdapat beberapa langkah dalam proses penerokaan pengetahuan dan perlombongan data merupakan salah satu langkah dalam proses penerokaan pengetahuan. Antara langkah lain ialah pembersihan data di mana data yang tidak konsisten akan disingkirkan, integrasi data iaitu dengan menggabungkan pelbagai data, pemilihan data di mana hanya data yang relevan akan dipilih daripada pangkalan data, pengubahan data iaitu dengan mengubah data kepada bentuk yang bersesuaian untuk perlombongan, penilaian corak untuk mengenalpasti corak menarik yang mewakili data dan seterusnya persembahan pengetahuan menggunakan visualisasi dan teknik pengetahuan. Dalam proses ini, corak yang menarik akan diambil dan disimpan dalam pangkalan pengetahuan.

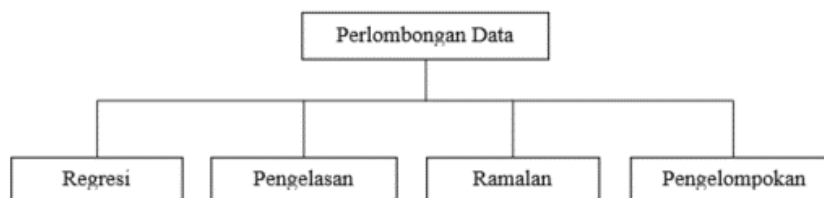
Rajah 2.7 menunjukkan kaedah-kaedah yang terdapat dalam perlombongan data. Proses perlombongan data sebenarnya menggunakan konsep pembelajaran mesin, statistik dan teknik persembahan keputusan supaya hasil yang diperolehi dapat difahami dengan mudah oleh manusia (Palous 2002). Perlombongan data juga ditakrifkan sebagai suatu proses mengekstrak maklumat yang tersirat, tidak diketahui dan berpotensi serta berguna daripada pangkalan data yang besar untuk membuat keputusan yang kritikal. Ia merupakan bidang yang merangkumi pembelajaran mesin, pengecaman corak, statistik, pangkalan data dan visualisasi untuk mengenalpasti pengeksktrakan maklumat daripada pangkalan data yang besar.



Rajah 2.7 Kaedah – kaedah dalam perlombongan data

Sumber: Ubahsuai Palous 2002

Beberapa teknik telah digunakan dalam kaedah perlombongan data untuk mengekstrakkan pengetahuan yang berguna kepada pengguna. Teknik-teknik tersebut adalah seperti regresi, pengelasan, ramalan dan pengelompokan seperti yang ditunjukkan dalam Rajah 2.8.



Rajah 2.8 Teknik-teknik dalam perlombongan data

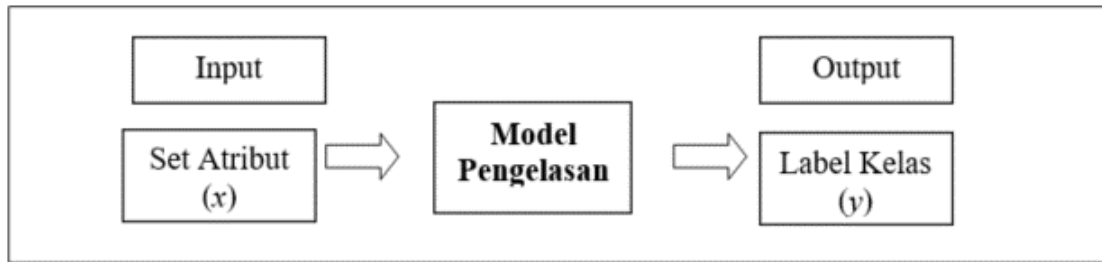
Sumber: Palous 2002

2.6.1 Regrasi

Regrasi merupakan teknik statistik yang tertua dan paling terkenal yang digunakan dalam perlombongan data. Regresi boleh digunakan untuk memodelkan hubungan di antara satu atau lebih daripada satu pembolehubah merdeka dan pembolehubah bersandar di mana pembolehubah-pembolehubah tersebut mesti mempunyai nilai yang berterusan. Regresi boleh dibahagi kepada regresi linear, regresi linear berganda, regresi tak linear, regresi logistik, regresi poisson, model log-linear dan pokok regresi.

2.6.2 Pengelasan

Pengelasan merupakan proses untuk menerbitkan fungsi atau model yang menentukan kelas sesuatu objek berdasarkan atributnya. Model yang terhasil daripada proses pengelasan dapat digunakan untuk meramal kelas objek yang tidak diketahui. Daripada pengelasan, pengetahuan yang berguna dapat dihasilkan. Menurut Zhang, pengelasan adalah sebahagian proses pembuatan petua yang paling kerap digunakan oleh sistem pakar. Ini disebabkan oleh model pengelasan mudah difahami dan mampu untuk meramalkan keputusan dengan tepat dan mudah. Pengelasan juga boleh meramal kategori seperti mengelaskan e-mel kepada label kelas spam atau non-spam mengikut tajuk dan kandungan mesej seperti yang ditunjukkan dalam Rajah 2.9.



Rajah 2.9 Model pengelasan

Sumber: Zhang 2000

Pelbagai aplikasi yang menggunakan pengelasan seperti dalam bidang kesihatan yang menggunakan pengelasan untuk mengenalpasti tahap penyakit kanser, mengelaskan tahap kelajuan angin untuk bidang meteorologi dan sebagainya kerana mampu untuk mengekstrak kelas data yang berguna. Selain itu, pengelasan juga menjana petua untuk ditafsirkan kepada pengetahuan. Antara teknik-teknik pengelasan ialah teknik pokok keputusan, set kasar, set kabur dan algoritma genetik (Han dan Kamber 2000).

2.6.3 Ramalan

Ramalan adalah serupa dengan pengelasan kecuali sasaran pembolehubah ramalan ialah nilai berterusan berbanding dengan kategori label. Sebagai contohnya, ramalan boleh digunakan untuk meramalkan harga saham tiga bulan depan, peratusan peningkatan kematian dalam lalu lintas tahun depan jika had laju dinaikkan dan kemenangan bola sepak berdasarkan perbandingan antara statistik pasukan.

2.6.4 Pengelompokan

Pengelompokan ialah membahagikan satu pangkalan data kepada beberapa kelompok yang mempunyai ciri-ciri yang sama. Setiap elemen yang terdapat di dalam kelompok mempunyai persamaan antara satu sama lain. Tujuan utama proses pengelompokan adalah untuk mengenalpasti corak sesebuah kumpulan, di mana membolehkan kita melihat persamaan serta perbezaan yang wujud antara kumpulan. Ini membolehkan andaian dapat dibuat berdasarkan kelompok yang telah dikelompokkan ini. Terdapat pelbagai kaedah di dalam pengelompokan data yang setiap satunya berfungsi mengikut

cara tersendiri dan mengeluarkan petua yang berbeza (Zait dan Mestaffa 1997). Menurut Fayyad (1997) perangkaian ialah proses mengumpulkan objek fizikal atau abstrak yang sama ke dalam kelas-kelas untuk mengetahui corak agihan keseluruhan set-set data.

2.6.5 Pembelajaran Mesin

Pembelajaran mesin berkait rapat dengan kajian, reka bentuk dan pembangunan algoritma yang memberikan keupayaan kepada komputer untuk belajar tanpa diprogramkan (Samuel 2000). Fungsi pembelajaran mesin ini adalah untuk mempelajari pengalaman sebelumnya yang diambil dari perlombongan data dan kebiasaanya data tersebut akan dibahagikan kepada dua set iaitu set data latihan dan set data ujian.

Pembelajaran mesin ini amat berguna kepada manusia kerana ianya tidak melibatkan campur tangan daripada manusia kerana program ini dibangunkan bagi membuat keputusan sendiri setelah menempuhi pelbagai pengalaman yang dihadapi. Pembelajaran mesin ini juga perlu melalui proses yang sama sebelum membuat keputusan sendiri dan ini amat sesuai bagi mendiagnosis kanser prostat kerana ianya dapat mentafsir simpton yang terdapat pada pesakit.

Terdapat dua pendekatan di dalam pembelajaran mesin antaranya ialah Pembelajaran Mesin Berpenyelia (PMB) dan Pembelajaran Mesin Tanpa Penyelia (PMTP). PMB ini adalah kaedah bagi mengelas serta mengagak sesuatu peristiwa dengan syarat set data latihan perlu dilabelkan terlebih dahulu. Model perlu disediakan bagi melalui proses untuk membuat petua dan dibetulkan apabila tersebut adalah salah. Manakala PMTP pula adalah kaedah bagi menentukan keberadaan data dalam kelas tertentu tanpa dilabelkan set data latihan tersebut.

a. Pengelasan Berasaskan Petua Pengeluaran

Pengelasan berasaskan petua merupakan teknik bagi mengkategorikan data kepada set masing-masing diberikan kepada salah satu kelas dan menjana petua serta mewakilkannya kepada pengetahuan untuk meramal keputusan dan menghasilkan pengetahuan atau maklumat baru. Pengelas akan mengkategorikan data menggunakan

petua $X \Rightarrow Y$ di mana X merupakan hubungan antara atribut dan Y merupakan hasil ataupun kelas. Kaedah petua digunakan kerana petua merupakan salah satu perwakilan ekspresif dan mampu dibaca oleh manusia untuk hipotesis dengan set petua (Lanzi 2007). Petua JIKA – MAKA digunakan sebagai tafsiran di mana: -

JIKA < syarat > ATAU < syarat > DAN < syarat > MAKA <kesimpulan>

Contoh petua ialah,

Petua: JIKA umur = 50 DAN sejarah_keluarga = ada ATAU risiko=sederhana DAN sejarah=ada Maka risiko_tinggi =pembedahan.

JIKA di sini merujuk kepada prasyarat atau sesuatu yang terdahulu manakala MAKA pula merujuk kepada akibat atau kesimpulan. Dalam bahagian JIKA, syarat boleh mengandungi satu atau lebih atribut (seperti umur = 50 DAN sejarah_keluarga = ada).

b. Pokok Keputusan

Pengelasan pohon keputusan *Decision tree* adalah menarik kerana terdapat banyak kelebihan. Ideanya ialah suatu pengelasan yang menarik, latihan yang dijalankan sentiasa secara lansung, terbaik dari semua dan pengelasannya yang sangat cepat. Penyelidik telah mengkaji tentang pengelasan ini pada dua dekad yang lalu dan banyak digunakan dalam aplikasi praktikal (Jordan dan Jacobs 1994) dan baru-baru ini, hubungan dengan pengelas lain seperti HMM kaedah (Park 1994) perceptron dan pelbagai aras (Schuermann dan Doster 1984).

Pokok keputusan merupakan carta aliran berbentuk struktur pokok di mana setiap nod dalaman ataupun nod bukan daun mewakili ujian bagi atribut, setiap cabangan mewakili hasil bagi ujian dan setiap nod daun atau nod terminal mewakili label kelas. Nod paling atas merupakan nod akar. Rajah 2.10 menunjukkan contoh pokok keputusan bagi konsep main_tenis.



Rajah 2.10 Pokok keputusan bagi konsep main_tenis

Kebanyakan kajian mencadangkan heuristik untuk membina pokok untuk ketepatan klasifikasi optimum atau bagi mengurangkan saiznya. Namun pokok yang dibina dengan data latihan yang tetap cenderung untuk menyesuaikan dengan data latihan. Pemangkasan balik pokok yang membesar sepenuhnya boleh meningkatkan generalisasi ketepatan data ghaib di mana selalunya pada perbelanjaan ketepatan pada data latihan. Kaedah kebarangkalian yang membolehkan keturunan melalui pelbagai cawangan dengan ukuran keyakinan yang berbeza juga tidak menjamin pengoptimuman ketepatan set latihan (Srimani dan Koti 2011). Kekurangan bagi induksi pokok keputusan ialah kemungkinan untuk menghadapi suboptima, kurang sesuai bagi tugas-tugas meramalkan keputusan yang bernilai berterusan dan adalah kompleks dan memerlukan masa yang banyak semasa mencarikan pembelahan yang paling bagus bagi setiap nod. Namun begitu, kelebihan kaedah pokok pokok ialah ia mudah untuk difahami dan ditafsirkan oleh manusia, informasi dalam pokok boleh ditukar kepada petua dan boleh digunakan bagi atribut-atribut kategori atau berterusan.

Kelebihan dari pokok keputusan adalah data yang sebelumnya kompleks dan sangat global dapat diubah menjadi lebih mudah dan spesifik serta ianya fleksibel untuk memilih data daripada nod yang berbeza. Data yang terpilih akan membezakan suatu kriteria dengan kriteria yang lain dalam nod yang sama. Kefleksibelan algoritma pokok keputusan ini meningkatkan kualiti keputusan yang dihasilkan jika dibandingkan ketika menggunakan cara penghitungan satu tahap yang lebih konvensional. Seterusnya, kelemahan kaedah ini berlaku apabila satu set data yang mengandungi ciri-ciri yang banyak. Dalam kes ini, pokok keputusan itu mungkin terlalu kompleks untuk difahami

dengan mudah. Untuk menyelesaikan masalah itu, pendekatan Cantasan pokok dikenakan ke atas pokok keputusan tersebut. Pendekatan menyelesaikan masalah *overfitting dataset* (dalam erti kata lain, pokok-pokok keputusan terlalu tidak sesuai) dengan menggunakan kaedah statistik untuk mencantas dahan yang kurang penting (nod) supaya pengguna mudah boleh menguasai keseluruhan struktur data. Pencantasan ialah teknologi yang berguna walaupun mungkin terdapat beberapa kesilapan kecil dalam pokok-pokok yang dihasilkan di bawah kaedah ini.

Kelebihan induksi pokok keputusan ialah mudah untuk difahami oleh manusia, informasi dalam pokok boleh ditukar kepada petua dan boleh digunakan bagi atribut-atribut kategori atau atribut-atribut yang berterusan. Keburukan bagi induksi pokok keputusan ialah kemungkinan untuk menghadapi suboptima, kurang sesuai bagi tugas-tugas meramalkan keputusan yang nilainya berterusan dan kompleks serta memerlukan masa yang banyak semasa mencari pembelahan yang paling bagus bagi setiap nod. Di bawah terdapat beberapa jenis pokok keputusan yang dipilih mengikut kesesuaian data iaitu J48, Random Forest dan pokok rawak. Walaupun ID3 juga merupakan teknik pokok keputusan, namun teknik ID3 tidak sesuai untuk digunakan kerana data sedia ada mempunyai data yang hilang dan tidak semua atribut berjenis nominal.

i. J48

Teknik J48 merupakan implementasi daripada C4.5 dalam WEKA®. Algoritma C4.5 menjana pokok keputusan berdasarkan satu set data input yang di label dan menjadi teknik yang sangat popular dan hebat untuk proses klasifikasi dan ramalan Quinlan. Pokok keputusan adalah satu struktur yang mudah di mana nod bukan terminal mewakili ujian ke atas satu atau lebih atribut dan nod terminal akan menghasilkan output. Ketika pembangunan pokok keputusan, cabang yang banyak mungkin akan mengakibatkan noise atau outlier pada set latihan. Pemangkasan pohon perlu dilakukan untuk menghapuskan cabang-cabang tersebut. Pohon yang melalui proses pemangkasan akan menjadi lebih kecil dan lebih mudah difahami. Pemangkasan pohon juga dapat digunakan untuk mengatasi masalah ketidaksesuaian. ketidaksesuaian terjadi kerana

terdapat noise pada data latihan, iaitu data yang tidak relevan sehingga mengakibatkan pohon memiliki subtree yang panjang dan tidak seimbang.

ii. Pokok Rawak

Pokok rawak adalah Pengelas yang diselia, Ia adalah suatu algoritma pembelajaran yang menjana banyak individu yang belajar. Ia menggunakan idea yang memenangi untuk membina set data untuk membina pokok keputusan secara rawak. Bagi setiap pokok nod terbahagi menggunakan pembahagi terbaik antara semua pembolehubah. Di dalam suasana hutan rawak, setiap nod menggunakan pembahagi terbaik di kalangan subset daripada jangkauan yang dipilih secara rawak pada nod itu. Pokok rawak telah diperkenalkan oleh Leo Breiman dan algoritma Adele Cutler yang terdiri daripada hutan rawak. Algoritma pokok rawak boleh menangani masalah kedua-dua klasifikasi dan regresi. Pokok rawak adalah merupakan kumpulan gabungan ramalan hutan rawak.

Mekanisme klasifikasi seperti berikut: Pengelas rawak pokok mendapat vektor mempunyai input, mengklasifikasikan setiap pokok di hutan, dan output label kelas yang mendapat majoriti "undi". Dalam kes yang regresi, jawapan Pengelas adalah purata bagi jawapan ke atas semua pokok-pokok di hutan. Pokok-pokok secara rawak adalah pada dasarnya gabungan dua algoritma yang sedia ada dalam pembelajaran mesin: satu model pokok akan bergabung dengan idea hutan secara rawak.

Model pokok rawak adalah pokok keputusan di mana setiap daun tunggal memiliki model linear yang dioptimum untuk ruang sub tempatan yang dijelaskan oleh daun ini. Hutan secara rawak telah menunjukkan untuk meningkatkan prestasi pokok keputusan yang ketara. kepelbagaian pokok dicipta oleh dua cara untuk rawak (Breiman 2001; Liaw 2013). Mula-mula data latihan dicuba dengan penggantian bagi setiap pokok tunggal seperti dalam memenangi. Kedua, apabila penanaman pokok, dan bukannya sentiasa pengkomputeran berpecah yang mungkin terbaik bagi setiap nod hanya subset rawak semua atribut dikira pada setiap nod, dan pecahan terbaik untuk subset yang dikira. Pokok tersebut telah bagi klasifikasi model pokok rawak untuk pertama kali menggabungkan model pokok dan hutan rawak. Kegunaan pokok rawak ini menghasilkan untuk berpecah pilihan dan dengan demikian mendorong pokok

seimbang, di mana satu persekitaran global untuk nilai *ridge value* merentasi semua daun, sekali gus memudahkan prosedur pengoptimuman (Liaw 2013).

iii. Hutan Rawak

Hutan rawak merupakan kombinasi ramalan pokok di mana setiap pokok bergantung kepada nilai-nilai vektor rawak sampel bebas dengan pengagihan yang sama untuk semua pokok di dalam hutan. Kesilapan umum bagi hutan daripada pengelasan pokok bergantung pada kekuatan setiap pokok dalam hutan dan hubungan antara mereka. Peningkatan ketepatan pengelasan adalah keputusan daripada pembesaran ensembel pokok dan membiar undian dengan mengambil kira kelas yang paling terkenal. Bagi membesarkan ensembel ini, vektor rawak sering dijanakan untuk mengawal tumbesaran setiap pokok (Ye et al. 2013).

Penganggaran ralat ujian dalam Hutan Rawak ditentukan daripada sampel latihan semasa pembesaran hutan. Bagi setiap sampel yang tidak terpilih dalam ikat but dan dipanggil sebagai sampel OOB (Breiman 2001). Dengan menggunakan sampel OOB sebagai input untuk pokok sepadan, ramalan dilakukan seolah-olah sampel ujian novel. Melalui simpan kira, undian yang tertinggi (pengelasan), purata (regresi) adalah dikira untuk semua sampel OOB daripada semua pokok. Ralat ujian yang telah dianggarkan adalah sangat tepat dalam latihan dengan mengambil kira bilangan pokok yang munasabah.

Beberapa persamaan perlu diambil kira untuk mengkategorikan ketepatan Hutan Rawak. Contohnya, diberi pengelasan ensembel $h_1(x), h_2(x), \dots, h_K(x)$, dan set latihan yang diperolehi secara rawak daripada taburan vektor rawak (X, Y) , Breiman mentakrifkan bahawa fungsi margin, mg adalah seperti :

$$mg(X, Y) = p_k \ln(h_k(X) = Y) - \max_{\{j \neq Y\}} p_k \ln(h_k(X) = j) \quad \dots(2.1)$$

yang mana $\ln(*)$ adalah fungsi indikator yang mengukur purata bilangan undian, p pada X, Y untuk kelas betul yang melebihi purata undian daripada kelas-kelas yang

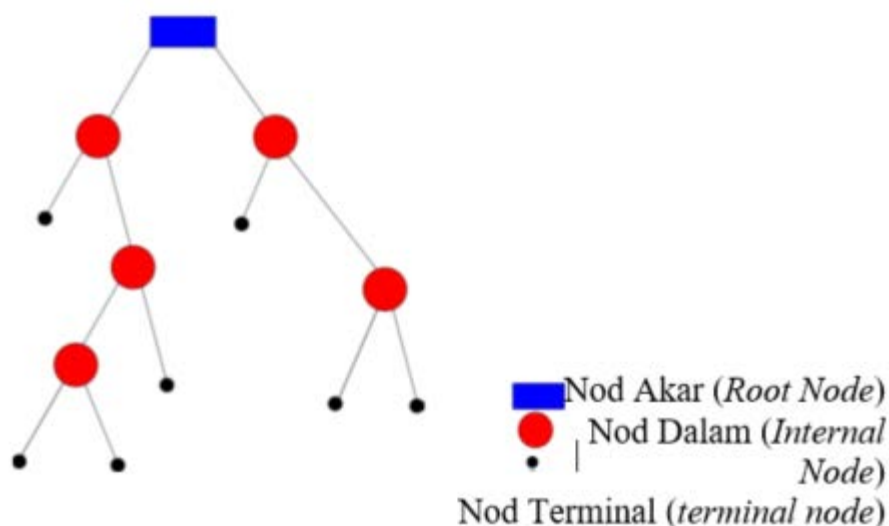
lain. Margin, mg yang paling besar bermaksud keyakinan pengelasan juga lebih tinggi. Menurut Breiman persamaan ralat generalisasi, RG adalah seperti:

$$RG^* = P_{X,Y}(mg(X,Y) < 0) \quad \dots(2.2)$$

di mana X, Y mewakili kebarangkalian bagi pembawa ruang X, Y .

1. Struktur Hutan Rawak

Struktur asas bagi pokok pengelasan adalah seperti Rajah 2.11.

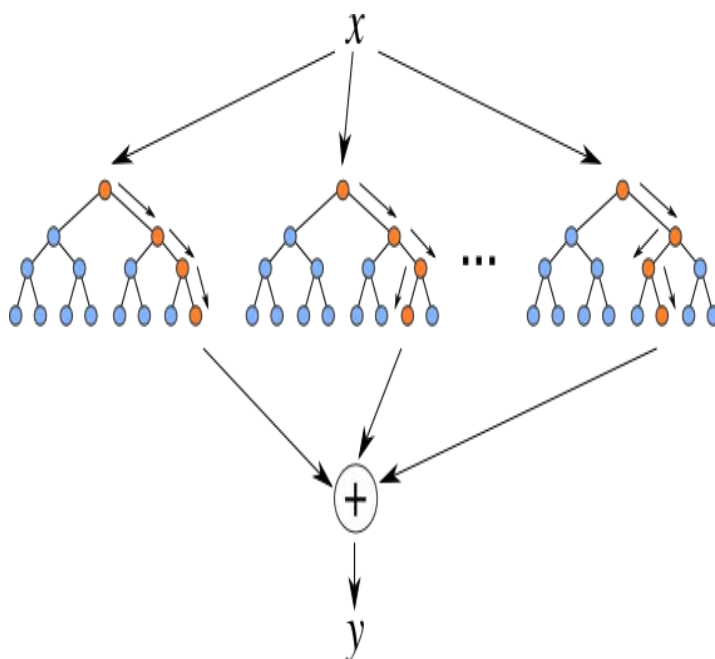


Rajah 2.11 Struktur asas bagi pokok pengelasan

Data latihan digunakan untuk membina model. Penjana pokok menentu ukur pemboleh ubah yang mana satu untuk pisahkan nod dan nilai untuk dipisahkan, membuat keputusan untuk berhenti (membuat nod terminal) atau pisahkan lagi semula dan memberikan nod terminal pada kelas.

Manakala model ensembel ialah menggabungkan keputusan yang diperolehi dengan model yang berlainan. Model ini mungkin boleh menjadi jenis yang berbeza atau sama. Keputusan model ensembel biasanya adalah lebih baik daripada model individu.

Rajah 2.12 menunjukkan seni bina umum bagi hutan rawak di mana B ialah bilangan pokok dalam Hutan Rawak dan k_1 , k_2 dan k_B dan k adalah label kelas (Razi dan Kuriakose 2005). Sebagai bilangan pokok yang semakin meningkat, kadar ralat set ujian berkumpul pada had yang bermaksud tiada yang lebih sesuai dalam Hutan Rawak yang besar (Ye et al. 2013).



Rajah 2.12 Seni bina umum hutan rawak

2.7 RUMUSAN

Proses diagnosis kanser prostat merupakan satu proses yang melibatkan pelbagai bidang perubatan. Pembangunan sistem pakar merupakan salah satu cara yang boleh membantu dan mempercepat proses diagnosis kanser prostat. Pelbagai kajian dan usaha dari bidang sains komputer telah dilakukan supaya dapat membantu pesakit kanser prostat dan memberi sumbangan kepada bidang perubatan. Tesis ini telah menggunakan petua-petua yang telah disahkan oleh pakar-pakar dalam bidang perubatan dalam bidang tertentu.

Selain itu, kebaikan dan keburukan teknik pokok keputusan dalam mengelas kanser prostat disamping mampu menjana petua-petua yang boleh difahami. Oleh itu, teknik pengelasan dan peramalan bagi pokok keputusan J48, pokok rawak dan hutan rawak telah dipilih disebabkan kesesuaian teknik ini kepada set data yang akan

digunakan dalam kajian. Kajian-kajian lepas yang berkenaan dengan klasifikasi kanser prostat juga dibincangkan supaya pemahaman terhadap skop kajian yang dijalankan lebih jelas. Tambahan lagi, sistem yang dibangunkan adalah sistem pelanggan-pelayan seperti yang dinyatakan oleh pihak HUKM. Perbandingan platform dan metodologi pembangunan perisian telah dibuat untuk mendapatkan kaedah yang terbaik dalam pembangunan sistem.

BAB III

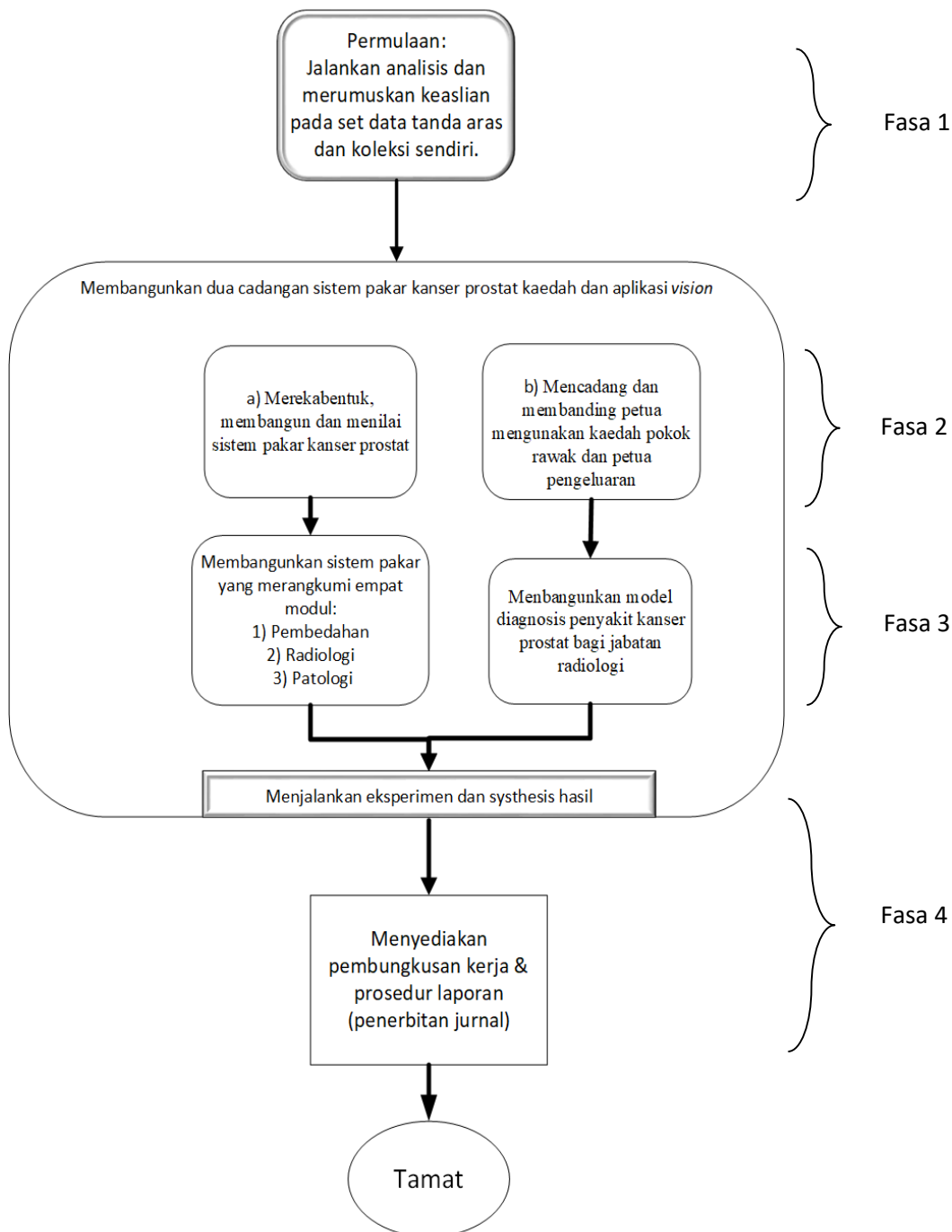
KAEDAH METODOLOGI

3.1 PENGENALAN

Setiap fasa semasa perlu diselesaikan sebelum fasa seterusnya sedia untuk dilaksanakan. Selepas fasa pengesanan dan pengesahan, jika prestasi sistem atau kebolehgunaan, tidak diselesaikan oleh pakar-pakar, prosedur daripada fasa pertama ke fasa terakhir akan dijalankan lagi. Proses pengulangan adalah untuk mengenal pasti isu-isu prestasi sistem atau kebolehgunaan disebabkan oleh satu atau lebih fasa dan kemudian menambah baik sistem sehingga pakar-pakar berpuas hati.

Kajian ini adalah susulan daripada paradigma kualitatif tafsiran kerana dua sebab. (1) Kajian ini bertujuan untuk memahami isu-isu diagnostik dalam pakar-pakar. Integrasi ini berlaku dalam konteks sistem sosial, struktur dan proses diagnosis perubatan organisasi. (2) Kajian ini akan menjadi penerokaan dalam fasa pertama untuk menetapkan tempat kejadian, meneroka, memahami dan menjadi biasa dengan fenomena perubatan dan diagnosis. Oleh yang demikian, ini kajian menggunakan kombinasi temubual exploratif diikuti dengan lebih mendalam kajian kes pelbagai rekod pesakit untuk pengumpulan data (i) meneroka dan memahami masalah kajian dengan lebih mendalam dan (ii) mencapai penyegitigaan di peringkat data dan kaedah, dijelaskan Creswell (2003), penyelidik boleh menggabungkan kaedah dalam satu turutan prosedur di mana niat penyiasat adalah untuk melebarkan atau menghuraikan penemuan-penemuan satu kaedah dengan kaedah lain. Selain itu, rekebentuk serta maklumat adalah mengikut spesifikasi oleh pakar. Parameter yang digunakan dalam ujikaji ini juga diterangkan. Hasil keputusan yang diperoleh selepas proses latihan dan pengujian data juga dicatatkan.

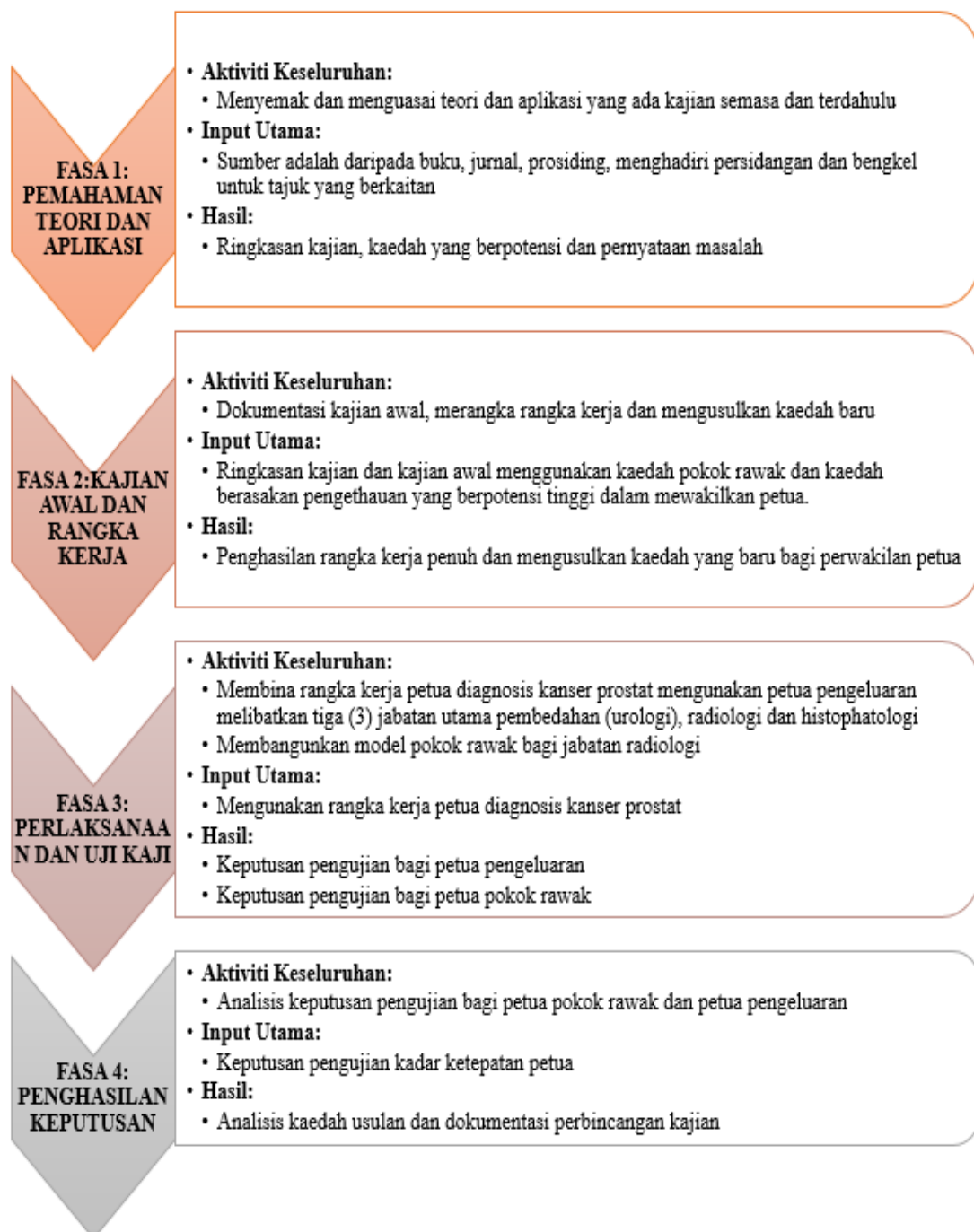
Seterusnya, metodologi yang akan diterangkan dalam fasa-fasa berikut melalui model kerangka pada Rajah 3.1:-



Rajah 3.1 Model kerangka kerja sistem

Sasaran kajian ini merangka kerja yang telah dicadangkan berdasarkan sasaran kajian yang telah dicadangkan. Ia merupakan satu sistem pakar diagnosis awal yang boleh menentukan tahap kanser prostat. Metodologi kajian ini diterangkan dengan lebih

perinci dalam Bab 3 ini. Metodologi kajian ini dibangunkan dalam empat fasa iaitu, fasa pemahaman teori dan aplikasi, fasa kajian awal dan rangka kerja, fasa pelaksanaan dan uji kaji dan akhir sekali ialah fasa pelaksanaan keputusan. Rajah 3.2 menerangkan setiap fasa yang dinyatakan.



Rajah 3.2 Penghalusan kerangka kerja kajian

3.2 FASA 1: PEMAHAMAN TEORI DAN APLIKASI

Dalam fasa ini, pemahaman teori melalui teknik temu bual dijalankan untuk memahami fenomena. Matlamat menjalankan temu bual adalah (1) untuk mendapatkan pandangan dan pengalaman daripada pakar-pakar perubatan kanser prostat masalah dan (2) untuk memberikan fleksibiliti penemuduga mengulas lanjut dan susulan kepada maklumat yang berharga yang tidak dijangka, dan siasatan untuk penjelasan lanjut. Dalam kajian ini, persoalan kajian, konteks penyelidikan dan kajian literatur adalah input utama untuk fasa temuduga di mana terlebih dahulu perlu meminta kebenaran rujuk Lampiran D. Menunjukkan proses pengambilan data yang dijalankan seperti Jadual 3.1:

Jadual 3.1 Perjumpaan proses pengambilan data

No	Tarikh	Tempat	Pakar	Penerangan
1	13/05/2016	Jabatan Urologi, HUKM	Dr Praveen Singam	Melawat wad pembedahan (urologi)
2	17/05/2016	Jabatan Patologi, HUKM	Dr Suria Hayati Bt. Md Pauzi	Proses temubual dan perolehan pengetahuan tentang sel kanser prostat yang diambil
3	09/06/2016	Jabatan, Radiologi, HUKM	Dr Syazarina Sharis Binti Osaman	Menpelajari sambil melawat jabatan radiologi serta melihat kaedah untuk menghasilkan gambar mealalui pancaran X-ray.
4	11/09/2016	Jabatan Urologi, HUKM	Dr Praveeb Singam	Melihat pembedahan kecil di mana Mereka mengambil sampel tisu daripada kelenjar prostat untuk mencari sel-sel

						kanser prostat
5	02/12/2016	Jabatan, Radiologi, HUKM	Dr	Syazarina Sharis Binti Osaman	Proses	pengambilan pengetahuan tentang gambar organ dalam yang diambil. Antara contoh data yang diambil adalah laporan pesakit seperti Lampiran G.
6	08/03/2017	Jabatan Urologi, HUKM	Dr	Praveen Singam	Perjumpaan	terakhir dengan Dr. Praveen mengenai petua yang dihasilkan melalui perolehan pengetahuan
7	16/05/2017	Jabatan Urologi, HUKM	Dr	Guan Hee Tan	Proses	pengesahan data petua yang diperolehi.
8	26/05/2017	Jabatan Patologi, HUKM	Dr	Suria Hayati Bt. Md Pauzi	Proses	pengesahan data petua yang diperolehi.
10	26/05/2017	Jabatan, Radiologi, HUKM	Dr	Syazarina Sharis Binti Osaman	Proses	pengesahan data petua yang diperolehi.
11	19/12/2017	Jabatan Urologi, HUKM	Dr	Guan Hee Tan	Kaji	selidik untuk mengetahui penilaian dan ketepatan sistem dan petua
12	29/12/2017	Jabatan Patologi, HUKM	Dr	Suria Hayati Bt. Md Pauzi	Kaji	selidik untuk mengetahui penilaian dan ketepatan sistem dan petua

13	04/01/2018	Jabatan, Radiologi, HUKM	Dr Syazarina Sharis Binti Osaman	Kaji selidik untuk mengetahui penilaian dan ketepatan sistem dan petua
----	------------	--------------------------------	--	---

Kajian kes adalah suatu siasatan empirikal yang menyiasat dan menjelaskan suatu fenomena kontemporari dalam konteks semula jadi. Ini adalah untuk menyiasat sesuatu fenomena yang baru muncul di mana kajian sebelum ini telah dijalankan. Kajian kes meneliti satu fenomena dalam persekitaran yang semula jadi dan menggunakan pelbagai kaedah pengumpulan data seperti temubual, analisis dokumen dan pemerhatian untuk mengumpul maklumat daripada rekod pesakit. Para pesakit yang dipilih adalah orang-orang yang hadir dengan kanser yang disyaki yang dikenal pasti oleh ujian urologi dan penyiasatan Radiologi. Di sini, kajian lokasi akan melibatkan Radiologi, pembedahan prostat dan histopatologi Pusat Perubatan Universiti Kebangsaan Malaysia.

Set data petua diagnosis kanser prostat yang dikumpulkan daripada 3 buah jabatan pembedahan (urologi), radiologi dan histologi diperolehi daripada Pusat Perubatan Universiti Kebangsaan Malaysia. Data diperolehi adalah dari bulan Mac 2016 sehingga Julai 2017, ia mengandungi 29 petua pengeluaran dan 9 petua pokok rawak secara keseluruhannya. Dalam fasa ini juga, mengenalpasti masalah yang dihadapi dalam proses pengumpulan data serta mencari jalan penyelesaian bagi menyelesaikan masalah tersebut. Bagi kajian ini, masalah yang dihadapi ialah data yang diperolehi mempunyai banyak kesamaran serta data-data yang diperolehi tidak beberapa cukup.

3.3 FASA 2: KAJIAN AWAL DAN RANGKA KERJA

Dalam fasa ini, satu rangka kerja untuk mengautomatiskan proses diagnosis akan dicadangkan berdasarkan kepada maklumat yang dikumpul dari fasa 1. Diagnosis rangka kerja secara automatik dapat memberi tumpuan kepada penggunaan teknik-teknik kepintaran buatan. Pemahaman secara menyeluruh dan jelas tentang masalah yang dihadapi dan proses penyelesaian dilakukan bagi mengatasi masalah tersebut adalah diperlukan. Oleh itu, fasa ini adalah penting dalam menjalankan kajian. Antara langkah-langkah untuk mengenalpasti masalah adalah apakah masalah yang hendak

dikaji atau dilaksanakan, bagaimana melaksanakan penyelesaian masalah tersebut dan mengapa pentingnya untuk melaksanakan masalah tersebut diulas oleh (Zuriyah 2011).

Kajian ini memerlukan model yang berasaskan petua untuk menjanakan pengetahuan yang mudah difahami, keseragaman data, data yang sederhana tetapi tidak besar dan banyak dengan pemisahan pengetahuan iaitu fakta- fakta dan bertaakulan untuk membangunkan sistem pakar diagnosis kanser prostat suhu yang lebih tepat, berkualiti dan mudah digunakan.

3.3.1 Kaedah Pembangunan

Untuk membangunkan sistem diagnosis kanser prostat, perkakasan dan perisian yang diperlukan adalah seperti yang berikut:

A. Java SE Development Kit 9 (JDK 9)

JDK 9 merupakan satu komponen yang diperlukan untuk membangunkan aplikasi Java. JDK membenarkan pembangun aplikasi Java untuk kompil dan nyahpepijat aplikasi Java pada fasa pembangunan system. Bahasa pengaturcaraan Java dipilih kerana mempunyai sifat "*Write Once Run Anywhere*" yang membenarkan aplikasi Java boleh dilakukan di semua platform seperti Windows, Linux dan Mac OS dengan syarat JRE dipasang di platform masing- masing. Bahasa pengaturcaraan Java juga membenarkan pembangun aplikasi Java melakukan sambungan pangkalan data dengan mudah.

B. Eclipse Oxygen

Eclipse merupakan alat yang boleh diguna oleh pembangun aplikasi Java untuk membuat pengaturcaraan dengan lebih cepat berbanding dengan penyunting teks biasa seperti Notepad.

C. Microsoft Access

Microsoft Access diperlukan dalam fasa pembangunan untuk membangunkan pangkalan data. Microsoft Access digunakan untuk membina jadual dan mengurus rekod dalam pangkalan data. Pangkalan data yang dibangunkan merupakan pangkalan

data mudah alir yang boleh diakses di mana-mana tanpa perlu dihoskan melalui pelayan pangkalan data.

3.3.2 Perisian Bagi Perlombongan Data

Perisian bagi perlombongan data yang diguna dalam kajian projek penyelidikan ialah perisian Weka versi 3.8. Weka (*Waikato Environment for Knowledge Analysis*) merupakan satu perisian pembelajaran mesin yang popular di mana bahasa pengaturcaraan ditulis dalam bahasa Java. Weka dibangunkan di Universiti Waikato, New Zealand dan merupakan perisian yang percuma yang boleh didapati di bawah *GNU General Public License*.

Antara kelebihan menggunakan perisian Weka ialah ia merupakan satu perisian mudah alih yang senang dilaksanakan dalam bahasa pengaturcaraan Java serta mampu beroperasi hampir kesemua platform dalam pengkomputeran moden. Selain itu, perisian Weka ini mudah untuk diguna kerana ia mempunyai antara muka pengguna grafik (*Graphical User Interface*) yang mesra pengguna.

Weka mampu membantu mempermudah tugas dalam perlombongan data yang khusus kepada prapemproses data, pengelompokan, pengelasan, regresi, visualisasi dan pemilihan fitur. Namun, perisian Weka ini juga mempunyai kelemahan iaitu ia tidak boleh melakukan perlombongan data yang jenis data pelbagai hubungan.

3.4 FASA 3: FASA PERLAKSANAAN DAN UJI KAJI

Semasa fasa ini, spesifikasi keperluan sistem yang dikenalpasti di peringkat sebelumnya akan ditukarkan kepada model yang sesuai bagi sistem yang dicadangkan. Sistem ini akan mempunyai antaramuka interaktif dan terdiri daripada modul iaitu pakar pembedahan (urologi). Radiologi modul, Patologi modul, Tamabahan lagi, seni bina sistem yang dicadangkan diberikan dalam Rajah 3.1.

3.4.1 Asas Logik Predikat

Dalam logik predikat formalisme logik cadangan dilanjutkan dan dibuatnya lebih halus daripada logik cadangan. Oleh itu, adalah mungkin untuk membentangkan ungkapan yang lebih rumit dari bahasa semula jadi dan menggunakannya dalam kesimpulan formal. Sebagai contoh:

“kanser yang hanya untuk lelaki. Roslan menghadapi masalah kencing. Oleh itu, Roslan disyaki menghidap penyakit kanser prostat”

Dalam pandangan logik cadangan, ayat adalah atom yang sama sekali berbeza dan tidak mempunyai bahagian yang diperlukan dalam analisis kesimpulan. Walau bagaimanapun, terdapat bahagian yang jelas dalam ayat tersebut. Dalam ayat pertama bercakap tentang kanser dan lelaki, dalam ayat kedua tentang roslan dan masalah kencing, dan dalam ayat terakhir ada yang diucapkan tentang roslan dan kanser prostat. Jadi, titik hubungan boleh dijumpai.

a. Komponen Sintaktik Logik Predikat

Predikat logik mengandungi semua komponen logik cadangan, termasuk cadangan pembolehubah dan pemalar. Di samping itu, logik predikat mengandungi terma-terma, predikat dan *quantifiers*. Terma-terma biasanya digunakan bagi menggantikan kata nama dan bukan kata nama. Mereka akan digabungkan menjadi ayat melalui predikat. Sebagai contoh, dalam ayat 'Amirul ada kanser', perkataan itu adalah 'Amirul' dan 'Kanser', dan predikat itu adalah 'ada'. Begitu juga jika ayat ini diterjemahkan ke dalam logik cadangan, bahawa 'Syahid' dan 'Kanser' kini dipanggil istilah. Logik predikat menggunakan *quantifiers* untuk menunjukkan jika pernyataan selalu benar, jika kadang-kadang benar, atau tidak benar. Dalam pengertian ini, *quantifiers* digunakan untuk menyesuaikan kata-kata seperti 'semua', 'beberapa', 'tidak pernah', dan ungkapan yang berkaitan.

Abjad p terdiri daripada sambungan dari (dari logik cadangan) \neg , \wedge , \vee , \rightarrow , $\$$ dan kurungan (,) dan dot (.) Simbol-simbol lain yang dimiliki oleh abjad adalah:

1. Pemalar individu a, b, c, \dots
2. Pembolehubah x, y, z, \dots
3. Simbol predikat P, Q, R, \dots
4. Simbol fungsi f, g, \dots
5. Simbol identity = *quantifiers* \forall dan \exists

3.5 FASA 4: PENGHASILAN KEPUTUSAN

Fasa ini melibatkan pelaksanaan pendekatan untuk modul-modul yang pelbagai telah dikenalpasti dalam fasa 2 dan juga menggunakan teknologi dan teknik yang dibangunkan dalam fasa 4. Sebuah sistem terdiri dua tugas (a) berasaskan pengetahuan (b) pembinaan petua akan berdasarkan pokok rawak kepada jabatan radiologi yang dipilih dalam fasa I kajian. Pengujian dan pembaikan sistem akan dilaksanakan. Akhirnya penilaian terhadap sistem akan dijalankan di Pusat Perubatan Universiti Kebangsaan Malaysia. Maklumbalas mengenai sistem akan digunakan untuk memperbaiki lagi sistem yang dicadangkan.

Pembangunan model merupakan satu proses yang kukuh dan model yang dibina akan digunakan sebagai pangkalan pengetahuan kepada sistem yang bakal dibangunkan ini. Dalam pembangunan model untuk diagnosis kanser prostat di Malaysia, pengelasan berasaskan petua dijalankan pada set data kanser prostat yang telah disediakan semasa pra-pemprosesan data. Teknik pengelasan berasaskan petua yang digunakan ialah pokok keputusan. Pokok keputusan merupakan algoritma yang biasanya dapat menghasilkan set petua yang berketepatan tinggi berbanding algoritma lain. Pengelas pokok rawak yang merupakan salah satu pengelas pokok keputusan digunakan dalam kajian ini untuk menjana petua. Seterusnya, data latihan akan digunakan untuk membina model menggunakan parameter-parameter yang telah dipilih dengan nilai yang berbeza untuk hasilkan model-model yang berbeza bagi setiap teknik pokok keputusan.

Petua yang dihasilkan semasa proses pengelasan akan diimplementasi ke dalam sistem. Sistem ini akan dibangunkan dengan menggunakan bahasa pengaturcaraan Java melalui *Eclipse Oxygen* sebagai platform untuk pembangunan sistem pakar. Sistem ini dibangunkan untuk memberi keputusan terhadap tahap kanser pesakit. Sistem ini memerlukan pengguna untuk memasukkan input yang diperlukan yang berkaitan dengan simpton pesakit. Pakar akan memberikan input pada paparan sistem dan sistem akan mengira dengan membandingkan petua dan memaparkan hasil keputusan. Data input merujuk kepada data pesakit yang dimasukkan ke dalam sistem untuk diproses. Kemudian, nilai- nilai input akan dibandingkan dengan petua yang sepadan yang dihasilkan semasa proses permodelan data. Proses pengiraan nilai yang padan dengan petua akan dikira melalui pengundian.

3.5.1 Pengujian Dan Penilaian Sistem

Setelah sistem siap dibangunkan, fasa pelaksanaan dan uji kaji sistem akan dijalankan. Pengujian dan penilaian sistem adalah penting bagi memastikan sistem berjalan dengan lancar di samping untuk menguji ketepatan sistem yang telah dibangunkan melalui perbandingan hasil petua keputusan sistem dengan output daripada data sebenar. Sebanyak 29 petua pengeluaran dan 9 petua pokok rawak telah validasi dan diuji yang telah diasingkan untuk tujuan pengujian sistem. Data yang digunakan untuk pengujian mengandungi 3 modul iaitu urologi, radiologi dan histopatologi. Selepas itu, data-data tersebut akan dimasukkan untuk menguji dan menilai sistem petua diagnosis kanser prostat yang dibangunkan. Kebanyakan hasil daripada petua menunjukkan data – data yang diuji menghasilkan output yang sama seperti data yang sebenar selepas dibuat perbandingan dengan keputusan asal. Walau bagaimanapun, terdapat sedikit perbezaan terhadap beberapa output yang dihasilkan.

3.5.2 Kaedah Penilaian Bagi Model Pokok Rawak

Klasifikasi adalah perkara yang sangat biasa bagi data perubatan. Terdapat dua idea yang sangat penting bagi tahap ini iaitu kepekaan dan kekhususan. Apabila mengambil kira keputusan ujian tertentu dalam dua populasi, satu penduduk positif, penduduk lain negatif, jarang memerhatikan pemisahan yang sempurna antara kedua-dua kumpulan. Malah, pengagihan keputusan ujian akan bertindih.

Untuk setiap titik potong yang mungkin atau nilai kriteria anda pilih untuk membezakan antara kedua-dua populasi, akan ada beberapa kes positif diklasifikasikan sebagai positif ($TP = \textit{True Positive Fraction}$), tetapi beberapa kes positif akan diklasifikasikan negatif ($FN = \textit{False Negative Fraction}$). Sebaliknya, beberapa kes negatif akan betul diklasifikasikan sebagai negatif ($TN = \textit{True Negative Fraction}$), tetapi beberapa kes negatif akan diklasifikasikan sebagai positif ($FP = \textit{False Positive Fraction}$). Penilaian akan dilakukan mengikut di persamaan bawah:

Kadar sebenar positif (TP) adalah peratusan dari petua dianogsis yang telah dikelaskan betul dan nombor kejadian sebenar positif dan palsu negatif nilai, menggunakan persamaan:

$$TP = \frac{TP}{TP + FN} \quad \dots(3.1)$$

Kadar sebenar negatif (TN) adalah peratusan petua dianogsis yang telah dikelaskan betul dan bilangan petua dianogsis sebenar negatif dan palsu yang mengandungi, seperti rumus berikut:

$$TN = \frac{TN}{TN + FN} \quad \dots(3.2)$$

Kepekaan kebarangkalian bahawa hasil ujian adalah positif apabila kategori adalah positif (kadar sebenar positif, dinyatakan sebagai peratusan). Kekhususan kebarangkalian bahawa hasil ujian akan menjadi negatif apabila kategori negatif (kadar sebenar negatif, dinyatakan sebagai peratusan). Tujuan adalah 100% sensitiviti dan spesifisiti, tetapi perkara ini tidak ada dalam pangkalan data sebenar, perlu mencari kombinasi terbaik yang boleh menjadi serasi.

Sensitiviti adalah mengukur perkadaran bahagian sebenar positif yang betul seperti contohnya (peratusan orang-orang sakit yang dikenalpasti dengan betul sebagai mempunyai keadaan).

$$TPR = \frac{TP}{P} = \frac{TP}{(TP + FN)} \quad \dots(3.3)$$

3.5.3 Penilaian Pengelas Bagi Model Pokok Rawak

Menurut Ching-Lai et al. (2007), ramalan yang salah akan menyebabkan terjadinya pengelasan yang salah dan seterusnya akan menghasilkan maklumat yang salah. Justeru, terdapat rumus yang digunakan untuk mendapatkan nilai ketepatan bagi model pengelasan.

Penilaian bagi kaedah pokok rawak dalam kajian projek ini dinilai menggunakan dua cara penilaian iaitu peratusan peramalan yang betul (*Percentage of Correctly Classified Instances*) dan ketepatan (*Accuracy*). Perisian Weka 3.8 berfungsi memastikan peratus data yang betul dan memeriksa semua nilai atribut untuk direkod serta membandingkan nilai-nilainya dengan nilai-nilai yang pemilihan fitur semasa proses *Selection Attributes*. Penilaian keberkesanan ini telah dibincangkan dengan lebih lanjut dari hasil dapatan bagi kajian ini. Dalam memilih algoritma model terbaik dan parameter pokok rawak iaitu *Correlation Coefficient (CC)*, *Root Mean Square Error (RMSE)*, dan *Receiver Operating Characteristics (ROC)* diuji dalam hasil implementasi.

Ketepatan peratus telah digunakan untuk menialai prestasi pencarian bacaan peratusan tertinggi set data. Penilaian akan dilakukan mengikut di persamaan bawah:

$$AC = \frac{TN + TP}{\#TN + \#TP + FN + FP} \quad \dots(3.4)$$

3.6 RUMUSAN

Bab ini keseluruhannya membincangkan mengenai fungsi sistem dan pembangunan sistem. Sistem ini dapat dibangunkan dengan menggunakan perisian dan keperluan lain yang dibincangkan dalam bab ini. Segala fungsi yang diperlukan dalam sistem akan dibangunkan dan dimasukkan ke dalam sistem ini. Seterusnya ialah objektif yang telah dihuraikan secara singkat agar setiap tahap dalam fasa pembangunan sistem ini,

mengikuti objektif yang telah diberikan. Begitu juga untuk skop kajian sistem pakar dalam membentuk pembangunan sistem pakar agar semasa pembangunan tidak lari dari skop yang ditentukan mengikuti perancangan yang telah dibuat. Penyataan masalah yang telah dinyatakan akan dapat diselesaikan mengikuti objektif yang telah dinyatakan sebelum ini. Ia bertujuan untuk membantu pakar dalam berurusan kanser prostat dan mengesan serta-merta tahap kanser sebenar pada pesakit prostat melalui beberapa saringan yang dijalankan seperti imej imbasan MRI, imej imbasan CT dan imej imbasan tulang. Dengan menggunakan teknologi ini dalam sistem diagnosis kanser prostat, pakar boleh mendapat maklumat pesakit dan membuat laporan keputusan diagnosis dengan menggunakan sama ada pelayanweb atau aplikasi mudah alih. Cara ini bukan sahaja boleh mempercepat proses diagnosis malah dapat mengurangkan penggunaan kertas dalam bidang perubatan. Sistem atau komponennya (contohnya data dan kaedah pemerolehan pengetahuan atau perisian yang digunakan untuk pembangunan) boleh digunakan untuk tujuan pembelajaran sepanjang hayat (pedagogi) atau memberi inspirasi kepada Penyelidikan Kanser prostat pada masa hadapan.

BAB IV

MODEL PETUA PENGELUARAN AUTOMASI DIAGNOSIS BERSEPADU KANSER PROSTAT

4.1 PENGENALAN

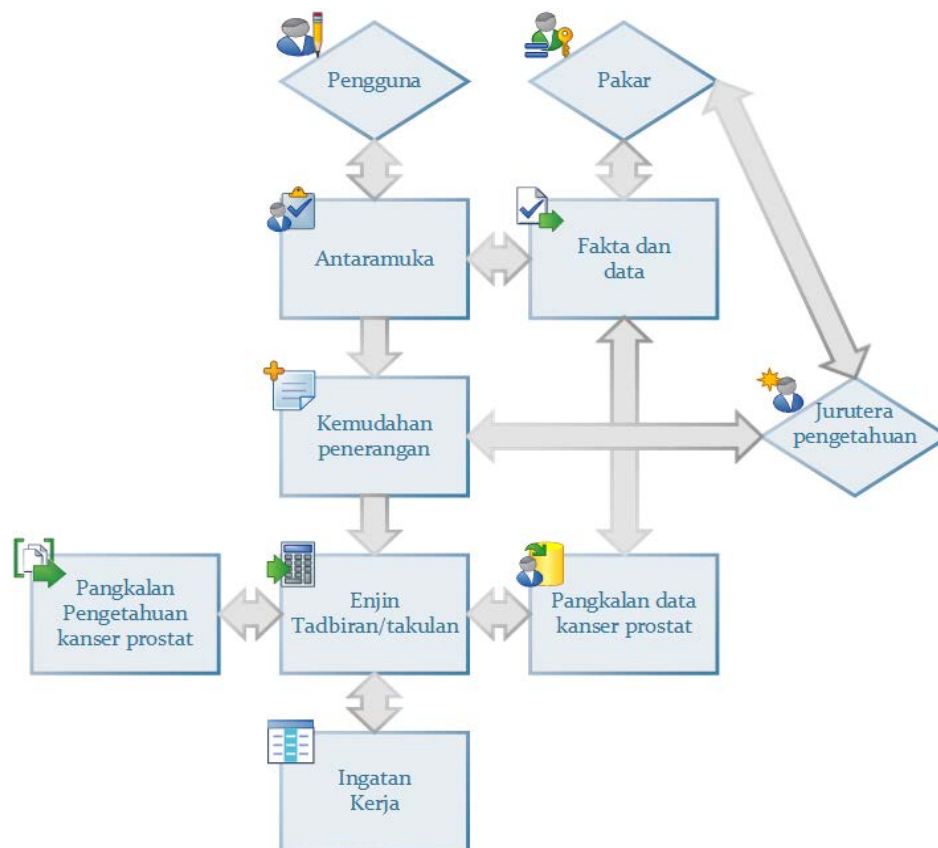
Dalam kajian ini, Sistem Automasi Diagnosis Bersepadu Kanser Prostat Secara Petua Pengeluaran Dan Pokok Rawak ialah petua sistem yang pakar yang dilatih untuk memudahkan mendiagnosis pesakit. Justeru, bab ini merumuskan hasil kajian serta membuat penilaian ke atas sistem yang dibangunkan menggunakan petua pengeluaran.

Setelah data perubatan pesakit dan pengetahuan yang diperolehi, kemudian semua disusun dan dianjurkan dengan mengkategorikan pengetahuan ke dalam kaedah-kaedah atau heuristik formal. Ia adalah semasa pra-pemprosesan dan menyusun data dan pengetahuan, model pengetahuan akan direka bentuk dan dibina untuk dimuatkan dalam kemungkinan pelaksanaan dan uji kaji dalam fasa 3.

4.2 REKABENTUK DAN MODEL PENGETAHUAN

Rajah 4.1 menunjukkan kerangka model pengetahuan yang dibangunkan untuk sistem pakar ini. Terdapat beberapa tugas perlu disiapkan dalam membangunkan model pengetahuan sistem pakar ini: (1) Antara muka, (2) pengetahuan fakta dan data, (4) kemudahan penerangan data (4) enjin taabir. Antara muka membantu pengguna untuk memasukkan data input pertanyaan, dan menggambarkan keputusan dalam bentuk laporan dan peta. Asas pengetahuan fakta dan data memperoleh domain pengetahuan, pengalaman, dan peraturan-peraturan daripada pakar-pakar dengan menggunakan teknik-teknik pengambilalihan dan perwakilan pengetahuan yang sesuai. Enjin taabir adalah perantaraan penterjemah berkenaan dengan asas pengetahuan pengalaman pakar-pakar yang dijadikan petua, pangkalan data dan ingatan kerja. Kemudahan

penerangan yang menganggap struktur mekanisme taabir dan asas pengetahuan di mana menjelaskan bagaimana pakar membuat keputusan tertentu.



Rajah 4.1 Kerangka model pengetahuan

4.2.1 Pengumpulan Pengetahuan

Mengumpul perolehan pengetahuan dari pakar kanser prostat merangkumi empat peringkat iaitu temu bual dengan pakar-pakar untuk kriteria pemilihan, reka bentuk keputusan dan pokok rawak untuk manipulasi pengetahuan, pengesahan dari pakar, pembangunan pangkalan pengetahuan dalam bentuk yang berasaskan petua. Empat pakar telah ditemubual dalam kemahiran bidang masing-masing dalam pembedahan (urologi), radiologi dan histopatologi. Selepas beberapa siri perbincangan, modul bagi setiap bidang diperolehi dan dijadikan sebagai atribut untuk menjana petua. Rajah 4.2 menunjukkan contoh skrip perbualan perolehan pengetahuan dari pakar urologi.

Saya	: Selamat sejahtera Doktor, saya gembira dapat berjumpa Doktor.
Pakar	: Selamat sejahtera, jemput duduk dan mari kita turun kepada tujuan asal kita.
Saya	: Saya ingin mengambil waktu doktor tentang masalah yang dihadapi oleh penghidap kanser prostat.
Pakar	: Masalah apa yang saya boleh bantu anda dalam memahami kanser prostat?
Saya	: Saya ingin bertanya jika pesakit didiagnosis dengan kanser prostat. Apa tindakan yang perlu dilakukan?
Pakar	: Tugas yang paling penting selepas didiagnosis adalah mendapatkan maklumat tentang keadaan anda. Kemudian anda dan doktor anda boleh bercakap mengenai tindakan terbaik. Oleh kerana terdapat pelbagai pilihan rawatan, membuat keputusan itu boleh menjadi rumit.
Saya	: Berapa banyak masa yang diperlukan untuk membuat keputusan?
Pakar	: Sebahagian besar orang-orang yang didiagnosis dengan kanser prostat, dalam erti kata lain, mempunyai banyak masa untuk membuat keputusan yang dianggap tentang rawatan."Mungkin satu perkara yang paling penting saya memberitahu pesakit adalah untuk tidak tergesa-gesa untuk keputusan, Ambil masa yang diperlukan".
Saya	: Berapa agresif kanser prostat boleh merebak?
Pakar	: Walaupun doktor sangat bagus dalam mengesan awal kanser prostat, mereka mempunyai kesukaran masa yang membezakan antara jenis penyakit itu berkembang perlahan dan agresif. Kebanyakan lelaki berusia 50 tahun yang lepas mempunyai sel-sel kanser prostat mereka yang tidak mempunyai simpton yang jelas bahawa mereka mempunyai kanser prostat. Hari ini, dengan menggabungkan maklumat dari ujian PSA, peperiksaan digital rektum, pengimejan kajian seperti ultrasound, dan keputusan biopsy, doktor boleh membuat penilaian berdasarkan tahap yang menimbulkan kanser prostat seseorang individu. Dalam menilai risiko, doktor juga berpendapat umur, etnik, dan sejarah keluarga.
Saya	: Kita kesuntukan masa doktor, kita akan sambung di pertemuan akan datang.
Pakar	: Baiklah saya jumpa kamu pada masa dan tarikh yang ditetapkan nanti.
Saya	: Terima kasih Doktor, sehingga ketemu lagi.

Rajah 4.2 Contoh skrip perbualan perolehan pengetahuan dari pakar urologi.

Petua – petua ini kemudiannya ditafsir dan dianalisa sebelum ditukar ke dalam bentuk pengetahuan. Berdasarkan 29 petua yang telah diperolehi dari pakar setiap bidang, beberapa analisa tentang corak petua yang dijana telah dilakukan di mana wad

pembedahan (urologi) merupakan tempat pertama pesakit perlu mendaftar dan mengetahui gejala serta simpton yang dihadapi, wad radiologi merupakan langkah di mana pakar mengesan kanser pada tubuh pesakit dan jabatan histopatologi untuk menentukan tahap kanser prostat.

Seterusnya, petua diterjemahkan dalam perwakilan pengetahuan dan pertimbangan untuk memberi maklumat tentang rekebentuk dalam sebuah sistem komputer rujuk 4.2.2f.

4.2.2 Pangkalan Data

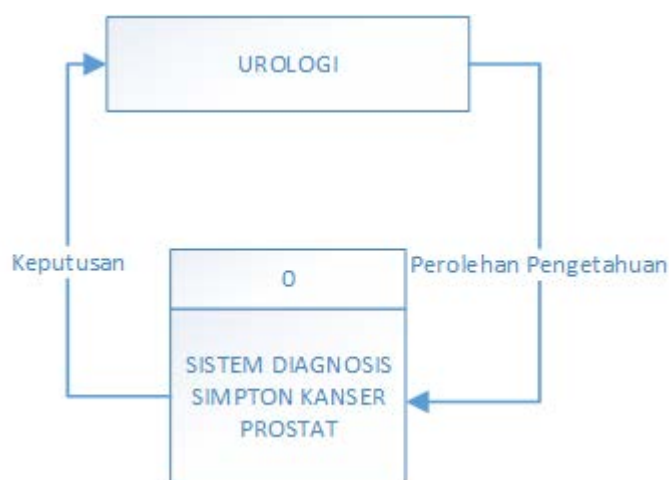
Pustaka pihak ketiga UCanAccess telah digunakan untuk membuat sambungan pangkalan data antara aplikasi sistem dengan pangkalan data Microsoft Access kerana pemacu JDBC tidak dibekalkan daripada pihak Microsoft untuk tujuan tersebut. UCanAccess merupakan pemacu JDBC sumber terbuka yang membenarkan aplikasi Java untuk melakukan transaksi SQL ke atas pangkalan data Microsoft Access yang diisukan bawah GNU Lesser General Public License 2.1.

a. Rajah Aliran Data (DFD)

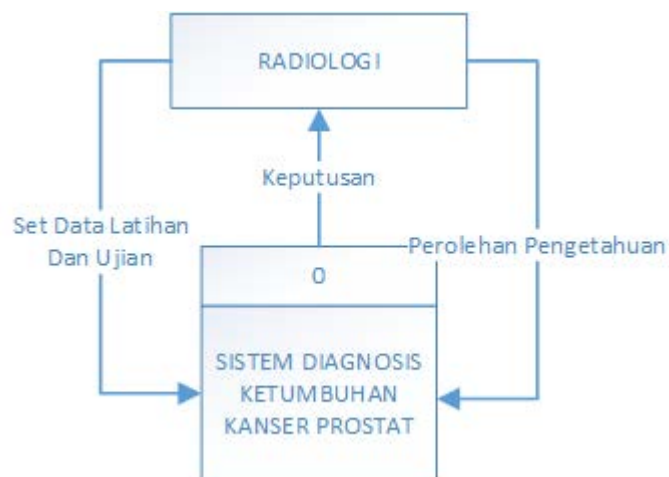
Gambar rajah aliran data (DFD) menggunakan pelbagai simbol untuk menunjukkan aliran sistem dalam mengubah data input ke dalam maklumat yang berguna. Rajah aliran data ialah kaedah grafik yang menjelaskan mengenai sesuatu sistem maklumat. Ia mewakili model logikal yang menunjukkan apa yang dilaksanakan oleh sistem, dan juga model fizikal bagaimana ia melaksanakannya. Terdapat empat (4) simbol asas dalam rajah aliran data. Ianya adalah simbol proses, aliran data, storan data dan juga entiti luaran. Dua (2) set model yang dipopular oleh Demacro dan Yourdan (1979) dan Gene dan Sarson (1979) untuk rajah aliran data bagi model Gane and Sarson dan model Yourdan.

b. Rajah Konteks

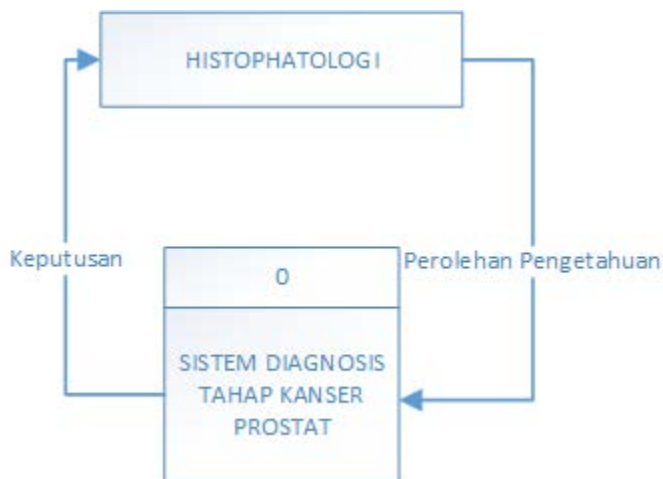
Rajah konteks adalah tahap paling atas berada dalam rajah aliran data. Ia merupakan gambaran umum bagi aliran model yang dibangunkan dan ia hanya mengandungi satu proses iaitu proses 0 yang menghubungkan fungsi keseluruhan model. Rajah di bawah menunjukkan rajah konteks untuk model jabatan pembedahan (urologi), radiologi dan histopatologi seperti di Rajah 4.3, Rajah 4.4 dan Rajah 4.5.



Rajah 4.3 Rajah konteks bagi jabatan pembedahan (urologi)



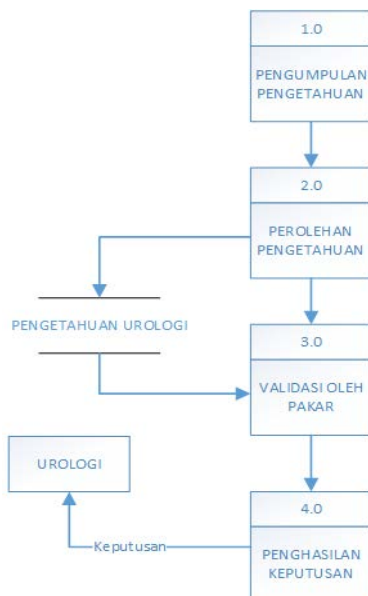
Rajah 4.4 Rajah konteks bagi jabatan radiologi



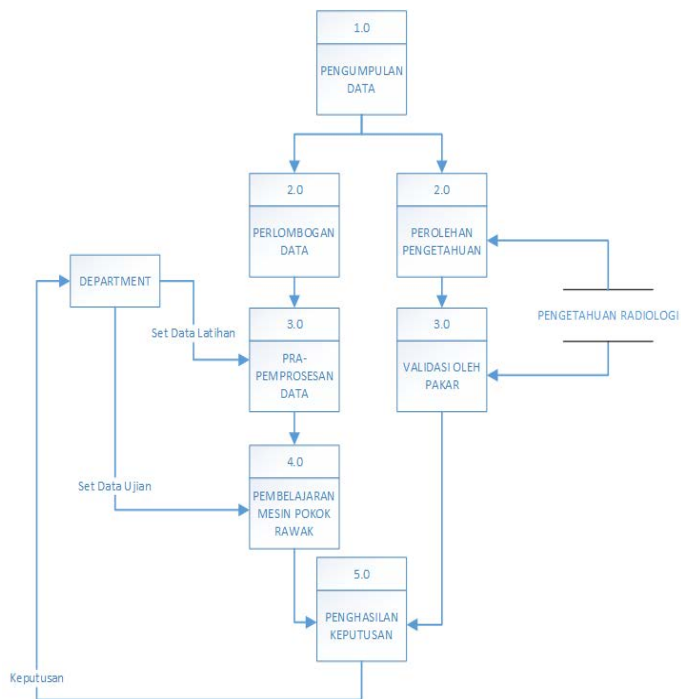
Rajah 4.5 Rajah konteks bagi histopatologi

c. Rajah Aliran Data Paras Sifar

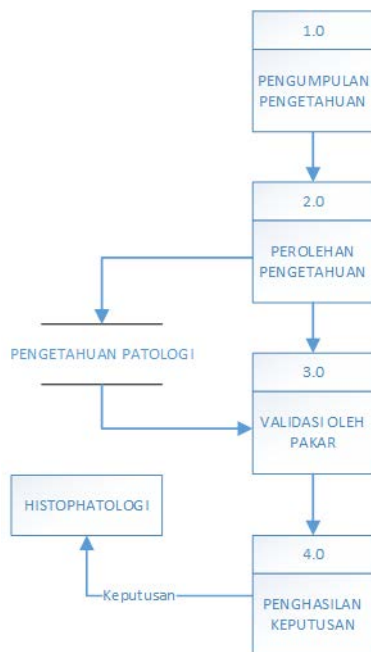
Rajah Aliran Data Paras Sifar atau 0 memberikan gambaran keseluruhan semua komponen yang berinteraksi untuk membentuk keseluruhan model dan mempunyai proses dalaman dan juga aliran data. Rajah di bawah menunjukkan rajah aliran data paras sifar bagi model jabatan pembedahan (urologi), radiologi dan histopatologi seperti di Rajah 4.6, Rajah 4.7 dan Rajah 4.8.



Rajah 4.6 Rajah aliran data paras sifar jabatan pembedahan (urologi)



Rajah 4.7 Rajah aliran data paras sifar jabatan radiologi

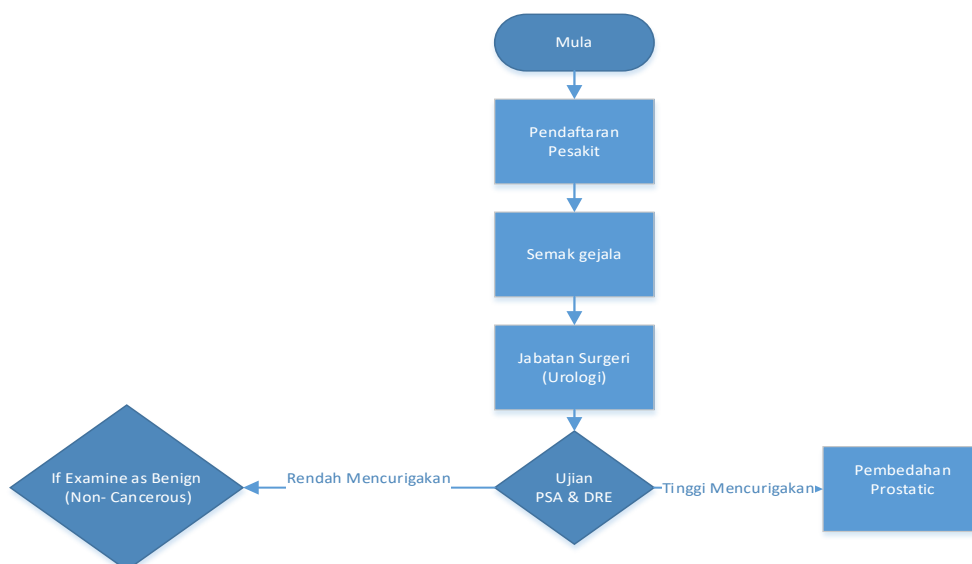


Rajah 4.8 Rajah aliran data paras sifar jabatan histopatologi

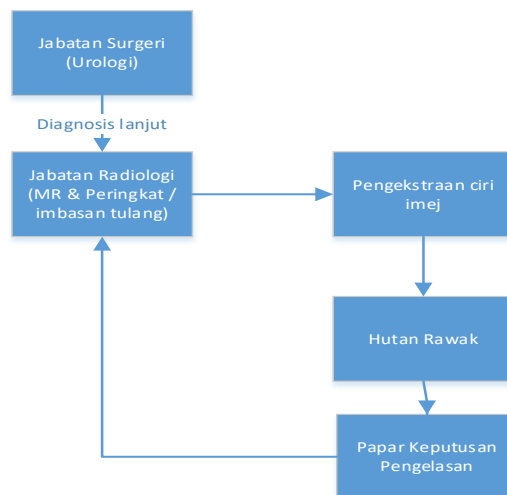
d. Carta Alir Proses Diagnosis Kanser Prostat

Carta aliran menjelaskan aliran proses-proses yang terlibat dalam sistem secara terperinci. Carta aliran ditunjukkan melalui pelbagai jenis kotak yang mewakili operasi proses dan susunannya dihubungkan dengan anak panah yang mewakili aliran proses. Carta aliran berperanan penting dalam proses analisis, rekabentuk, dokumentasi dan pengurusan proses sistem.

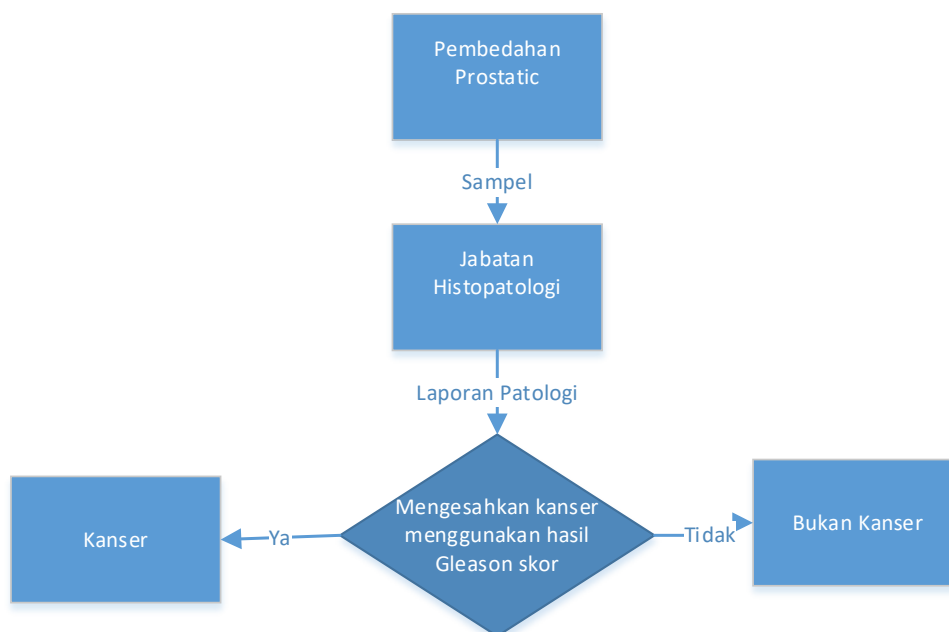
Carta alir proses diagnosis kanser prostat telah dibina seperti yang ditunjukkan dalam Rajah 4.12 pada halaman yang seterusnya sebelum pembangunan sistem. Carta alir yang lebih terperinci untuk proses mendiagnosis kanser prostat adalah seperti di Rajah 4.9, Rajah 4.10 dan Rajah 4.11.



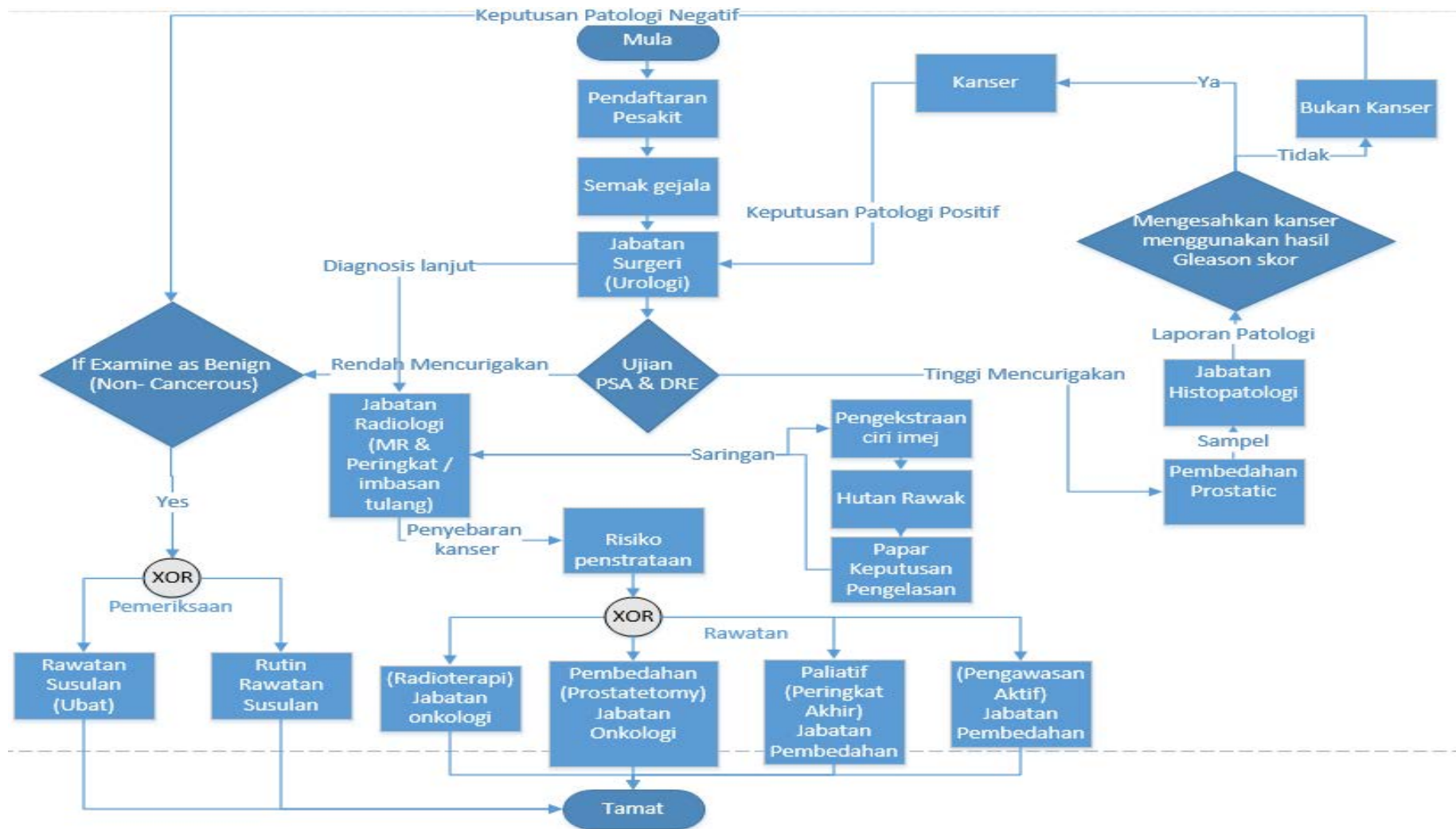
Rajah 4.9 Sub modul bagi jabatan urologi



Rajah 4.10 Sub modul bagi jabatan radiologi



Rajah 4.11 Sub modul bagi jabatan histopathologi

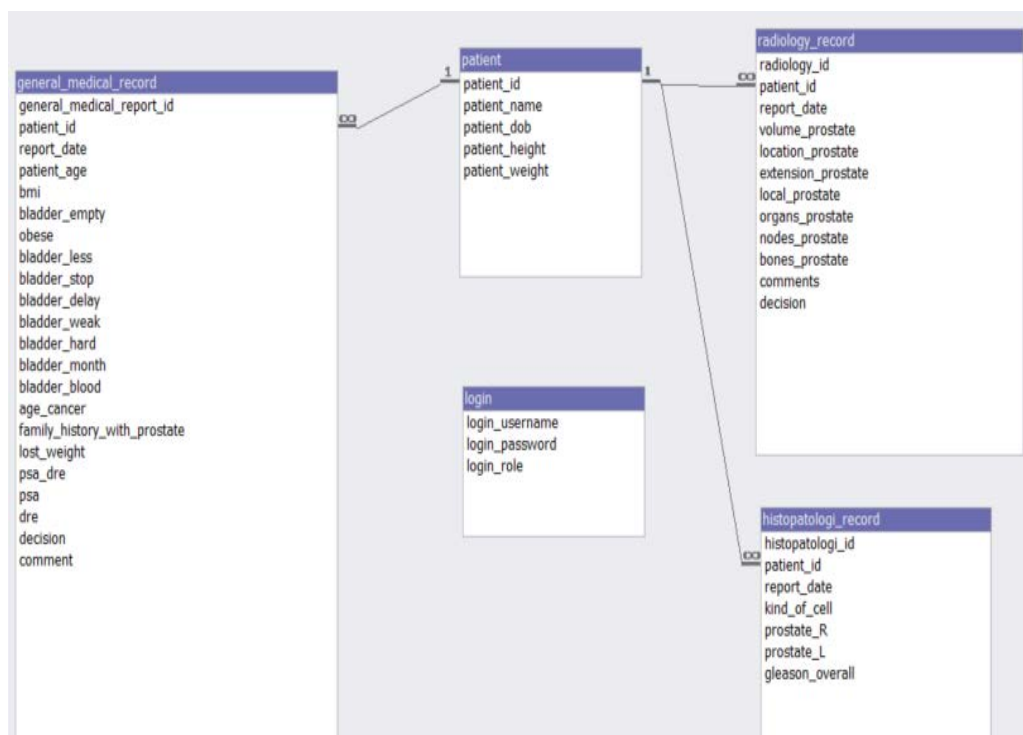


Rajah 4.12 Carta alir proses diagnosis kanser prostat

e. Rajah Hubungan Entiti

Rekabentuk data menjelaskan struktur pangkalan data sistem. Bagi memberikan gambaran yang lebih terperinci, rajah hubungan entiti (ERD) dan kamus data digunakan untuk menerangkan struktur pangkalan data bagi sistem pakar. Manakala, rajah aliran data digunakan untuk menggambarkan pemrosesan data antara sistem dengan entiti lain.

Rajah hubungan entiti (ERD) merupakan grafik khusus yang menggambarkan struktur logikal pangkalan data. Rajah ini digunakan untuk menjelaskan hubungan antara entiti. Selepas mengenal pasti entiti-entiti yang terlibat dalam sistem pakar, rajah ini disediakan untuk menunjukkan hubungan antara entiti-entiti. Selain itu, senarai atribut bagi setiap entiti juga dikenal pasti dan disenaraikan dalam rajah. Rajah 4.13 menunjukkan rajah hubungan entiti (ERD) sistem pakar yang terdiri daripada pelbagai entiti, seperti Log masuk (*login*), Pesakit (*Patient*), rekod pesakit umum (*General Medical*), rekod radiologi (*Radiology Record*), rekod hispatologi (*Hispatology Record*).



Rajah 4.13 Rajah hubungan entiti sistem pakar

f. Perwakilan pengetahuan

Jadual yang digunakan dalam pangkalan sistem kanser prostat adalah jadual jadual *patient* dan jadual *login*. Jadual *patient* mempunyai hubungan satu kepada banyak (1:M) dengan jadual *general_medical_record*, jadual *histopatologi_record* dan jadual *radiologi_record* kerana seorang pesakit boleh mempunyai banyak rekod perubatan tetapi satu rekod perubatan hanya untuk seorang pesakit sahaja. Jadual-jadual di bawah menjelaskan nama, jenis data dan penerangan untuk setiap atribut yang terdapat dalam setiap jadual. PK merupakan kunci primer manakala FK merupakan kunci asing yang terdapat dalam jadual masing-masing. Rujuk Jadual 4.1, Jadual 4.2, Jadual 4.3, Jadual 4.4 dan Jadual 4.5.

Jadual 4.1 Metadata jadual *login*

Sumber: Hospital Universiti Kebangsaan Malaysia

Nama atribut	Jenis data	Penerangan
▪ login_username	Short Text	PK. Nama pengguna yang digunakan untuk log masuk sistem.
▪ login_password	Short Text	Kata kekunci yang digunakan untuk log masuk sistem.
▪ login_role	Short Text	Atribut yang digunakan untuk menentu peranan pengguna.

Jadual 4.2 Metadata jadual *patient*

Sumber: Hospital Universiti Kebangsaan Malaysia

Nama atribut	Jenis data	Penerangan
▪ patient_id	AutoNumber	PK. Atribut yang menentukan pesakit. Dihasilkan secara automatik semasa pesakit dicipta.
▪ patient_name	Short Text	Name pesakit.
▪ patient_dob	Date/Time	Tarikh lahir pesakit.
▪ patient_height	Number	Ketinggian pesakit.
▪ patient_weight	Number	Keberatan pesakit.

Jadual 4.3 Metadata jadual *general_medical_report*

Sumber: Hospital Universiti Kebangsaan Malaysia

Nama atribut	Jenis data	Penerangan
▪ general_medical_report_id	AutoNumber	PK. Atribut yang menentukan rekod pembedahan (urologi) pesakit. Dihasilkan secara automatik semasa mencipta rekod baru.
▪ patient_id	Number	FK. Pemilik rekod perubatan ini.

▪ report_date	Date/Time	Tarikh rekod ini diambil.
▪ patient_age	Number	Umur pesakit apabila rekod ini diambil.
▪ Bmi	Number	Indeks jisim tubuh pesakit apabila rekod ini diambil.
▪ bladder_empty	Number	Kerap kali mempunyai sensasi tidak mengosongkan pundi kencing selepas selesai membuang air kecil. 0 - "Tiada Lansung", 1 - "Kurang daripada 1 kali hingga 5", 2 - "Kurang Dari separuh masanya", 3 - "Kira-kira separuh daripada masanya", 4 - "Lebih daripada separuh masanya", 5 - "Hampir selalu".
▪ bladder_less	Number	Selalunya terpaksa membuang air kecil kurang dari dua jam. 0 - "Tiada Lansung", 1 - "Kurang daripada 1 kali hingga 5", 2 - "Kurang Dari separuh masanya", 3 - "Kira-kira separuh daripada masanya", 4 - "Lebih daripada separuh masanya", 5 - "Hampir selalu".
▪ bladder_stop	Number	Sering mendapati telah berhenti dan mula semula untuk kencing. 0 - "Tiada Lansung", 1 - "Kurang daripada 1 kali hingga 5", 2 - "Kurang Dari separuh masanya", 3 - "Kira-kira separuh daripada masanya", 4 - "Lebih daripada separuh masanya", 5 - "Hampir selalu".
▪ bladder_delay	Number	Sering mendapati sukar untuk menanggungkan membuang air kencing. 0 - "Tiada Lansung", 1 - "Kurang daripada 1 kali hingga 5", 2 - "Kurang Dari separuh masanya", 3 - "Kira-kira separuh daripada masanya", 4 - "Lebih daripada separuh masanya", 5 - "Hampir selalu".
▪ bladder_weak	Number	Kerap kali mempunyai aliran kencing yang lemah. 0 - "Tiada Lansung", 1 - "Kurang daripada 1 kali hingga 5", 2 - "Kurang Dari separuh masanya", 3 - "Kira-kira separuh daripada masanya", 4 - "Lebih daripada separuh masanya", 5 - "Hampir selalu".
▪ bladder_hard	Number	Selalunya untuk meneran atau membebaskan untuk memulakan kencing. 0 - "Tiada Lansung", 1 - "Kurang daripada 1 kali hingga 5", 2 - "Kurang Dari separuh masanya", 3 - "Kira-kira separuh daripada masanya", 4 - "Lebih daripada separuh masanya", 5 - "Hampir selalu".
▪ bladder_month	Number	Berapa kali sebulan mempunyai kebiasanya dapat untuk membuang air kecil, dari Malam hingga masa bangun Pagi. 0 - "Tiada Lansung", 1 - "Kurang daripada 1 kali hingga 5", 2 - "Kurang Dari separuh masanya", 3 - "Kira-kira separuh daripada masanya", 4 - "Lebih daripada separuh masanya", 5 - "Hampir selalu".
▪ bladder_blood	Number	Air kencing didatangi dengan darah. 0 - "Tidak", 1 - "Ya".
▪ lost_weight	Number	Pesakit turun berat badan. 0 - "Tidak", 1

▪ family_history	Number	- "Ya".
▪ _with_prostate_cancer	Number	Pesakit ada saudara dekat yang menghidapi kanser prostat. 0 - "Tidak", 1 - "Ya".
▪ age_cancer	Number	Pesakit berumur 50 tahun keatas. 0 - "Tidak", 1 - "Ya".
▪ Psa	Number	Psa level. 0 - "Tidak", 1- "Ya".
▪ Dre	Number	Dre positif. 0 - "Tidak", 1- "Ya".
▪ Decision	Number	Keputusan pakar berdasarkan indeks pilihan jenis keputusan di antara muka. 0 - "Biopsy", 1 - "Follow -up", 2 - "Routine follow -up".
▪ Comment	Long Text	Komen doktor dan maklumat tambahan ke atas keputusan diagnosis.

Jadual 4.4 Metadata jadual *radiology_record*

Sumber: Hospital Universiti Kebangsaan Malaysia

Nama atribut	Jenis data	Penerangan
▪ prostate_id	AutoNumber	PK. Atribut yang menentukan rekod perubatan prosate pesakit. Dihasilkan secara automatik semasa mencipta rekod baru.
▪ patient_id	Number	FK. Pemilik rekod perubatan ini.
▪ report_date	Date/Time	Tarikh rekod ini diambil.
▪ volume_prostate	Text	Isipadu prostat (cm).
▪ location_prostate	Number	Lokasi lesion. 0 - "Apex right", 1 - "Apex left", 2 - "Medium grid right", 3- " Medium grid left", 4 - "Base right", 5 - " Base left".
▪ extension_prostate	Number	Extracapsular extension. 0 - "Tidak", 1 - "Ya".
▪ organs_prostate	Number	Organ yang terlibat. 0 - "Seminal vesicles", 1 - "External sphincter", 2 - "Rectum", 3- "Bladder", 4 - "Levator muscles", 5 - " Pelvic wall".
▪ nodes_prostate	Number	Nodes pesakit. 0 - "Non - suspicious", 1 - "Suspicious".
▪ bones_prostate	Number	Tulang pesakit. 0 - "Normal/Benign", 1 - "Suspicious of metasis".
▪ PIRADS	Number	Tahap PIRADS. 0- "1", 1- "2", 2- "3", 3="4", 4- "5"
▪ Comments	Long Text	Komen doktor dan maklumat tambahan ke atas keputusan diagnosis.
▪ Decision	Number	Keputusan pakar mengikut indeks pilihan keputusan di antara muka. 0 - "Follow-up urology", 1 - " Follow-up radiology".

Jadual 4.5 Metadata jadual *histopatologi_record*

Sumber: Hospital Universiti Kebangsaan Malaysia

Nama atribut	Jenis data	Penerangan
▪ histo_id	AutoNumber	PK. Atribut yang menentukan rekod perubatan histopatologi pesakit. Dihasilkan secara automatik semasa mencipta rekod baru.

▪ patient_id	Number	FK. Pemilik rekod perubatan ini.
▪ report_date	Date/Time	Tarikh rekod ini diambil.
▪ prostate_R	Number	Keputusan “Gleason score most predominant”.
▪ prostate_L	Number	Keputusan “Gleason score second predominant”.
▪ gleason_overall	Number	Keputusan keseluruhan “Overall Gleason Score”. 0 - "Low risk =<6 (Surveillance,Radiotherphy, Surgery)", 1 - "Intermediate risk =7 (Radiotherphy, Surgery)", 2 - "High risk =>8 (Radiotherphy, Surgery, Hormontherphy”.
▪ Decision	Number	Keputusan pakar mengikut indeks pilihan keputusan di antara muka. 0 - "Follow-Up urology", 1 - “Follow-up radiology”.
▪ Comments	Long Text	Komen doktor dan maklumat tambahan ke atas keputusan diagnosis.

4.2.3 Analisis Pengetahuan (Fakta dan Data)

Setelah aktiviti pengumpulan data dijalankan melalui kaedah temubual dan pengumpulan data, data tersebut telah diajara menggunakan petua pengeluaran dan pokok rawak. Petua – petua ini kemudiannya ditafsir dan dianalisa sebelum ditukar ke dalam bentuk pengetahuan. Berdasarkan petua yang telah diperoleh, beberapa analisa tentang corak petua yang dijana telah dilakukan di jabatan yang terlibat seperti Urologi, Hispatologi dan Radiologi.

4.2.4 Enjin Taabir

Enjin taabir ialah otak sistem pakar. Ia menggunakan struktur kawalan (peraturan penterjemah) dan menyediakan metodologi untuk pertimbangan. Ia bertindak sebagai jurubahasa yang menganalisis dan proses peraturan. Ia digunakan untuk melakukan tugas sepadan dari maklum balas yang diberikan oleh pengguna dan peraturan menembak. Tugas utama enjin taabir adalah untuk mengesan dengan cara melalui hutan petua untuk tiba pada kesimpulan.

Mekanisme taabir daripada sistem pakar diagnosis kanser prostat. Antara proses taabir adlah seperti berikut: (1) pangkalan data untuk mendapat maklumat diagnosis kanser prostat, (2) melakukan taabir berdasarkan asas pengetahuan, (3) memasukkan maklumat yang diminta dan disimpan oleh ingatan kerja. Sistem pakar menggunakan tiga teknik engine taabir untuk diagnosis kanser prostat. Enjin taabir akan menyemak

data yang diperolehi dan maklumat mengikut petua yang disimpan di dalam pangkalan pengetahuan dan menjana petua dalam memberikan penilaian dan cadangan.

4.2.5 Ingatan Kerja

Ingatan bekerja adalah satu sistem kognitif dengan kapasiti terhad yang bertanggungjawab bagi mengadakan maklumat yang ada untuk pemprosesan buat sementara waktu. kerja memori adalah penting bagi pertimbangan dan bimbingan keputusan dan tingkah laku dalam meperolehi kewibawaan data (data integriti). Dalam kata lain, ingatan kerja mempengaruhi tahap kesahihan, ketepatan, keselajaran, dan kesempurnaan data yang disimpan, dihantar, atau diproses oleh sistem. Untuk mewujudkan tahap kewibawaan data yang tinggi dalam sistem diagnosis kanser prostat, jenis input yang dipilih untuk interaksi antara pengguna dan sistem adalah amat penting dalam pemprosesan data buat sementara waktu.

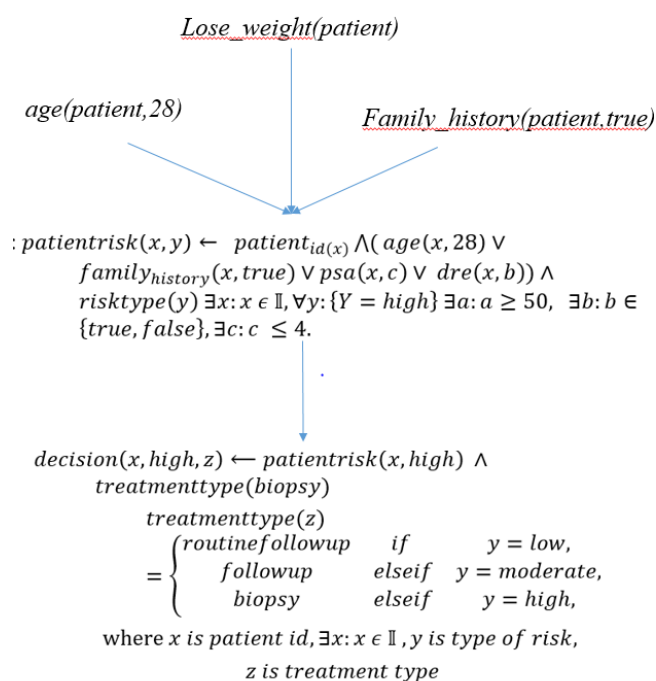
Sistem ini telah mengelakkan pengguna daripada menaip dengan menyediakan komponen dropdown di antara muka sistem. Pendekatan ini dapat mengelakkan pengguna daripada memasukkan data kosong ataupun data yang tidak mengikut format yang dijangkakan ke dalam sistem. Untuk input yang memerlukan pengguna untuk memasukkan nilai rentetan (*string*), pengesahan input (*input validation*) telah dilakukan sebelum diproses atau disimpan dalam sistem. Atribut-atribut pangkalan data yang digunakan untuk menyimpan data daripada input dropdown adalah dalam jenis data Number berdasarkan indeks pilihan dropdown tersebut. Sebagai contoh, pangkalan data akan menyimpan nilai 0 untuk pilihan “Tidak” yang mempunyai indeks 0 atau 1 untuk pilihan “Ya” yang mempunyai indeks 1 pada dalam komponen dropdown. Pendekatan ini dapat mengelakkan overhead yang diperlukan untuk memproses data sebelum dipaparkan atau disimpan dalam pangkalan data kerana pemprosesan rentetan adalah lebih mengambil masa berbanding dengan pemprosesan integer.

Selain daripada ini, operasi Boolean adalah mudah untuk dilaksanakan dengan operasi penambahan dengan menggantikan betul (TRUE) dengan 1, salah (FALSE) dengan 0, atau (OR) dengan operator tambah (+) dan dan (AND) dengan operator darab (\times). Sebagai contoh, operasi Boolean TRUE AND FALSE OR FALSE boleh dilaksanakan dengan operasi matematik $1 \times 0 + 0$. Keputusan operasi Boolean tersebut

ialah TRUE sekiranya jawapan adalah lebih besar atau sama dengan 1 manakala salah (FALSE) untuk jawapan 0. Kesimpulannya ialah pendekatan mengguna ingatan bukan sahaja dapat menjaga prestasi sistem malah dapat membantu dalam membuat diagnosis kanser prostat dengan mudah. Antara contoh penggunaan ingatan kerja dalam sistem diagnosis kanser prostat adalah:

28-year-old patient losing weight and has close relative with prostate CA. Less than half of the time need to push or strain to begin urination.

Rajah 4.14 Contoh laporan klinikal bagi kes di jabatan pembedahan (urologi)



Rajah 4.15 Pembuktian petua dengan klausa predikat

Rajah 4.16 menunjukkan antara muka yang digunakan oleh doktor perubatan jabatan pembedahan (urologi) untuk mengdiagnosis pesakit berdasarkan laporan klinikal dalam rajah Rajah 4.14 yang telah dibuktikan pada Rajah 4.15. Dalam antara muka ini butang *Diagnose* boleh ditekan setelah data yang pengguna telah dimasukkan untuk mengdiagnosis pesakit. Cadangan rawatan akan dipaparkan dan dipilih secara automatik pada *Decision dropdown* seperti yang ditunjukkan dalam Rajah 4.17. Doktor perubatan jabatan pembedahan (urologi) boleh melakukan keputusan akhir dengan

memilih pilihan yang disediakan dalam *Decision dropdown* sebelum menyimpan maklumat pesakit dalam sistem.

The screenshot shows the 'GENERAL MEDICAL (Adding)' window. It features a 'Home' button and a 'PATIENT INFORMATION' section with a dropdown for 'Select patient' (ID: 24, Name: Sulaiman) and input fields for Patient ID (24), Age (28), Height (178.0), Patient Name (Sulaiman), Weight (67.0), and BMI (21.15). An 'Update' button is present. Below is the 'IPSS SCORE' section with six questions about urinary symptoms, each with a dropdown menu. The score is 2/35. The 'FACTOR' section includes checkboxes for 'Patients aged 50 years and older?' (Yes), 'Patients lose weight?' (Yes), 'Patients have a close relative with cancer of the prostate?' (Yes), and 'Urine comes with blood?' (No). A 'Doctor's comment' text area is on the left. On the right, there is a 'Diagnose' button, a 'Decision' dropdown menu (set to 'Biopsy'), and a 'Save' button.

Rajah 4.16 Contoh input kemasukan pengetahuan bagi jabatan pembedahan (urologi)

This screenshot shows the same interface as Rajah 4.16, but with a 'Results Of General Medical Record' dialog box open. The dialog box contains an information icon, a title bar, and the following text: 'The risk level to have prostate cancer is : HIGH'. Under 'Factors contributing to the result:', there is a bulleted list: 'Patients lose weight.', 'Patients have a close relative with cancer of the prostate.', and 'Over the past month, how often have you had to push or strain to begin urination.'. The 'Suggestion' section states: 'Patients are recommended to perform the test PCA / DRE for further diagnosis.'. An 'OK' button is at the bottom of the dialog. The background interface is partially visible and dimmed.

Rajah 4.17 Contoh keputusan diagnosis rekod pembedahan (urologi)

Rajah 4.18 menunjukkan antara muka yang digunakan oleh pakar radiologi untuk mengdiagnosis rekod perubahan radiologi. Setelah memasukkan data yang diperlukan, cadangan rawatan boleh didapati dengan menekan butang *Diagnose*.

PIRADS dan *Decision dropdown* akan dipilih secara automatik selepas rekod perubahan radiologi diproses oleh sistem seperti yang ditunjukkan dalam Rajah 4.19. Pakar radiologi juga boleh melakukan keputusan terakhir dengan membuat pilihan secara manual sebelum menyimpan rekod tersebut dalam sistem.

The screenshot shows the 'Add Radiology Report' window. The 'RADIOLOGY ASSESSMENT' section is active, displaying various input fields and checkboxes. The 'Patient' information is ID: 24, Name: Sulaiman. The 'Prostate volume (cm)' is 23. The 'Lesion location' has 'Medium grid left' and 'Base right' selected. 'Extracapsular extension' is 'No'. 'Adjacent local organ involvement' is 'No'. 'Organ involved' includes 'Seminal vesicles', 'Levator muscles', and 'Pelvic wall'. 'Patient nodes' is 'Suspicious' and 'Patient bone' is 'Suspicious of metasis'. The 'Doctor's comment' field contains 'danger'. At the bottom, 'PIRADS' is set to 0 and 'Decision' is 'Follow_up'. There are buttons for 'Diagnose', 'Diagnose AI', 'Save', and 'Clear'.

Rajah 4.18 Contoh input kemasukan pengetahuan bagi jabatan radiologi

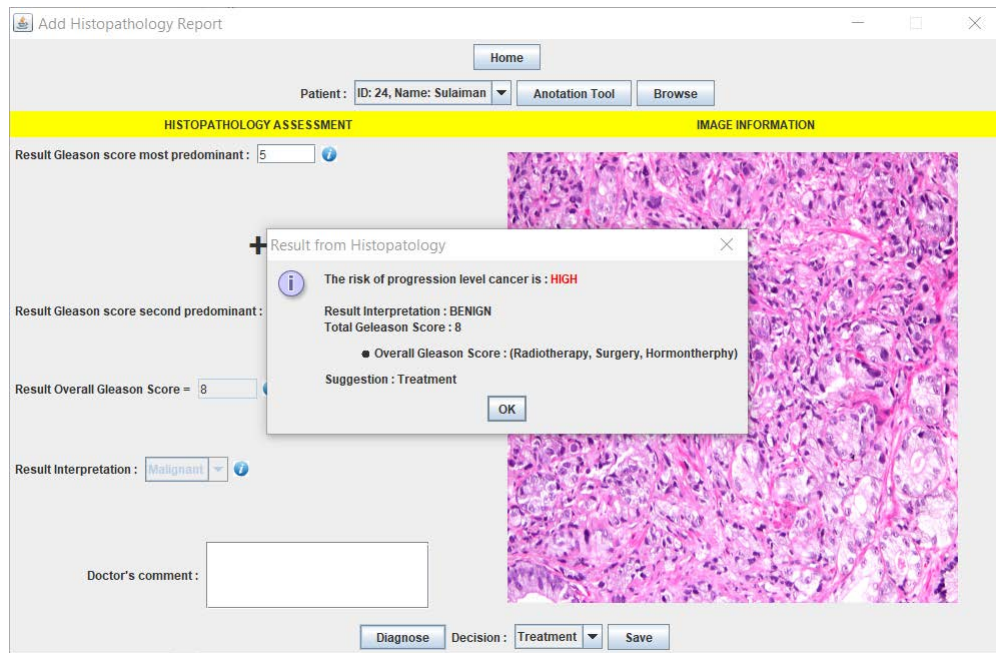
This screenshot is similar to the previous one but includes a 'Result from radiology' dialog box. The dialog box contains an information icon, the text 'The risk level to have prostate cancer is : HIGH', and a list of reasons: 'Volume prostate:23', 'Assymmetrical densities.', 'Suspicious nodes involve.', and 'Suspicious of metastasis.'. Below the reasons, it says 'Suggestion : Treatment' and has an 'OK' button.

Rajah 4.19 Contoh keputusan diganosis rekod radiologi

Diagnosis rekod perubatan histopatologi boleh dilakukan oleh pakar histopatologi dengan memasukkan data yang diperlukan dalam antara muka seperti yang ditunjukkan dalam Rajah 4.20. Cadangan rawatan akan dipilih secara automatik selepas menekan butang *Diagnose* akan melihat paparan input yang dipilih seperti yang ditunjukkan dalam Rajah 4.21 Seterusnya, pakar histopatologi melakukan keputusan akhir dengan memilih pilihan dalam *Decision dropdown* sebelum menyimpan maklumat rekod perubatan dalam sistem

The screenshot shows a web application window titled "Add Histopathology Report". At the top, there is a "Home" button and a "Patient" dropdown menu displaying "ID: 24, Name: Sulaiman". To the right of the patient information are "Anotation Tool" and "Browse" buttons. The main content area is split into two columns: "HISTOPATHOLOGY ASSESSMENT" on the left and "IMAGE INFORMATION" on the right. The assessment section includes input fields for "Result Gleason score most predominant" (5), "Result Gleason score second predominant" (3), "Result Overall Gleason Score =" (8), and "Result Interpretation" (Malignant). A plus sign (+) is visible between the first two Gleason score fields. Below these is a "Doctor's comment" text area. The image section shows a histopathology image of prostate tissue. At the bottom of the window, there are "Diagnose", "Decision" (Treatment), and "Save" buttons.

Rajah 4.20 Contoh input kemasukan pengetahuan bagi jabatan histopatologi



Rajah 4.21 Contoh keputusan diganosis rekod histopathologi

4.2.6 Rekabentuk Antara muka

Antara muka memainkan peranan yang penting di mana ia berfungsi sebagai medium interaksi antara pengguna dan sistem. Justeru, antara muka yang mesra pengguna amat penting demi memastikan pengguna dapat menggunakan sistem dengan mudah dan berkesan. Tambahan pula, antara muka yang dibangunkan hams memenuhi keperluan pelanggan dan sistem. Dalam pembangunan antara muka pengguna sistem diagnosis kanser prostat, *plugin WindowBuilder* dalam *Eclipse Oxygen IDE* telah digunakan untuk mengatasi kekangan masa dalam kajian. *WindowBuilder* dapat membantu pembangun aplikasi Java menjimatkan masa dalam pembangun antara muka dengan membekalkan fungsi tarik dan lepas untuk meletakkan komponen antara muka ke atas aplikasi dengan senang. Antara muka pengguna membentangkan maklumat penting kemudahan seperti sistem Penerangan, konsep flow Rajah, Glosari dan sistem panduan dan bantuan. Ia membolehkan pengguna untuk memasukkan maklumat yang dikehendaki dan sistem kemudian menilai data untuk mekanisme inferens untuk memperoleh sistem bantuan dan keputusan.

Antara muka yang dihasilkan merupakan antara muka kesetiaan tinggi yang boleh diguna semula semasa membangunkan sistem. Antara dua kriteria yang

difokuskan untuk membangunkan antara muka pengguna sistem diagnosis kanser prostat adalah:

- Mengandungi semua input yang diperlukan untuk tujuan disimpan sebagai rekod dalam pangkalan data.
- Mengurangkan penaipan pengguna untuk mengurangkan kadar ralat.

Seksyen ini akan menerangkan antara muka yang penting dalam sistem serta aliran bagaimana sistem ini berfungsi.

Rajah 4.22 menunjukkan antara muka log masuk yang merupakan halaman yang pertama pengguna akan nampak semasa melakukan sistem. Nama pengguna dan kata kunci adalah diperlukan untuk log masuk dan mengguna fungsi yang disediakan oleh sistem.



Rajah 4.22 Antara muka log pengguna

Rajah 4.23 dan Rajah 4.24 menunjukkan antara muka yang digunakan untuk mencipta (atas) dan mengemaskini (bawah) data pesakit. Pengguna perlu log masuk sebagai doktor pembedahan (urologi) umum sekiranya ingin menggunakan fungsi mengurus maklumat pesakit. Butang Save adalah digunakan untuk menyimpan maklumat pesakit dalam sistem setelah habis mengisi borang yang disediakan manakala butang Update dan Delete adalah digunakan untuk mengemaskini dan menghapuskan maklumat pesakit.

Patient Information

Home

Name :

Date of Birth eg.(30/11/1991):

Height :

Weight :

Clear Save

Rajah 4.23 Antara muka pengurusan kemasukan maklumat pesakit

Update Patient Information

Home

ID: 24, Name: Sulaiman

ID :

Name :

Date of Birth :

Height :

Weight :

Reset Update

Rajah 4.24 Antara muka pengurusan mengubahsuai dan memadam maklumat pesakit

Antara muka untuk pengurusan rekod perubatan jabatan pembedahan (urologi) adalah seperti yang ditunjukkan di Rajah 4.25. Dari kiri atas merupakan antara muka yang digunakan oleh doktor pembedahan (urologi) untuk mendiagnosis tahap risiko pesakit. Dropdown digunakan dalam sistem untuk mengelakkan pengguna daripada banyak menaip dan mengelakkan kesilapan kemasukan data. Butang Diagnose digunakan untuk mendapat cadangan keputusan daripada sistem berdasarkan input yang diberi. Ruang Komen Doktor adalah disediakan untuk pengguna sekiranya ada apa-apa maklumat tambahan yang ingin disimpan sekali dengan rekod perubatan. Doktor pembedahan (urologi) boleh mengemaskini dan menghapus rekod perubatan pesakit dalam halaman yang ditunjukkan di bahagian atas kanan Rajah 4.26. Halaman yang ditunjukkan di bahagian bawah Rajah 4.27 merupakan halaman yang membenarkan

pengguna untuk melihat sahaja rekod perubahan dan menyimpan sekiranya perlu dalam bentuk PDF dengan menekan butang Save PDF.

Rajah 4.25 Antara muka pengurusan kemasukan rekod perubahan jabatan pembedahan (urologi)

Rajah 4.26 Antara muka pengurusan mengubahsuaikan rekod perubahan jabatan pembedahan (urologi)

The screenshot displays a web-based medical record interface titled "General Medical Record (View)". It features a "Home" button at the top. The "PATIENT INFORMATION" section includes a dropdown for "Select patient" (ID: 24, Name: Sulaiman), a "Patient ID" field (24), an "Age" field (28), a "Select report" dropdown (Report ID: 52, Date: 2018-01-01), a "Patient Name" field (Sulaiman), and a "BMI" field (21.15). The "IPSS SCORE" section contains six questions about urinary symptoms with radio button options and a "Score: 18" field. The "FACTOR" section includes questions about age, weight loss, family history, and blood in urine, with "Yes" and "No" options. The "PSA/DRE" section shows "PSA Score Test: 0" and "DRE Examination: Negative". At the bottom, there is a "Doctor's comment" text area, a "Decision" dropdown (Biopsy), and a "Save PDF" button.

Rajah 4.27 Antara muka pengurusan paparan rekod perubatan jabatan pembedahan (urologi)

Pengguna sistem yang ingin mengurus rekod perubatan pesakit harus log masuk ke dalam sistem sebagai pakar radiologi. Fungsi mengurus rekod perubatan radiologi pesakit yang disediakan dalam antara muka seperti yang ditunjukkan di Rajah 4.28, Rajah 4.29 dan Rajah 4.30. adalah lebih kurang sama dengan mengurus rekod pembedahan (urologi) pesakit yang dibincangkan sebentar tadi.

Add Radiology Report

Home

Patient: ID: 25, Name: Faisal

Annotation Tool Detection Tool Browse

RADIOLOGY ASSESSMENT **IMAGE INFORMATION**

Prostate volume (cm):

Lesion location: Apex right Apex left Medium grid right

Medium grid left Base right Base Left

Extracapsular extension: No

Organ involved: Seminal vesicles External sphincter Rectum

Bladder Levator muscles Pelvic wall

Patient nodes: Non - suspicious

Patient bone: Normal/Benign

Doctor's comment:

PIRADS: 0 Diagnose Diagnose AI Decision: Follow_up Save Clear

Rajah 4.28 Antara muka pengurusan kemasukan rekod perubatan jabatan radiologi

Update Radiology Report

Home

Patient: ID: 25, Name: Faisal

Report: Report ID: 40, Date: 2018-05-06

Annotation Tool Detection Tool Browse

RADIOLOGY ASSESSMENT **IMAGE INFORMATION**

Prostate volume (cm): 23

Lesion location: Apex right Apex left Medium grid right

Medium grid left Base right Base Left

Extracapsular extension: No

Organ involved: Seminal vesicles External sphincter Rectum

Bladder Levator muscles Pelvic wall

Patient nodes: Non - suspicious

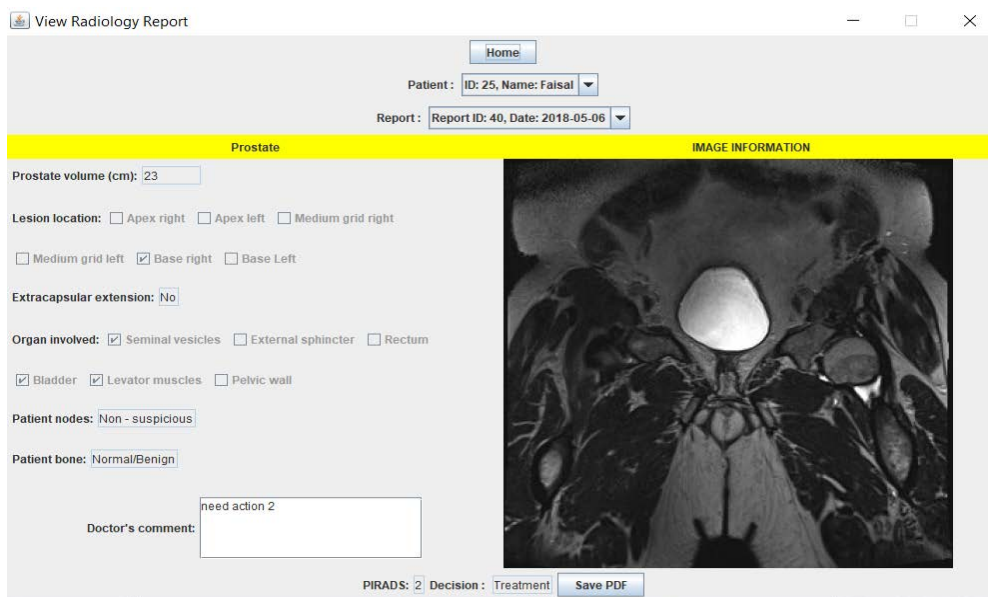
Patient bone: Normal/Benign

Doctor's comment: need action 2

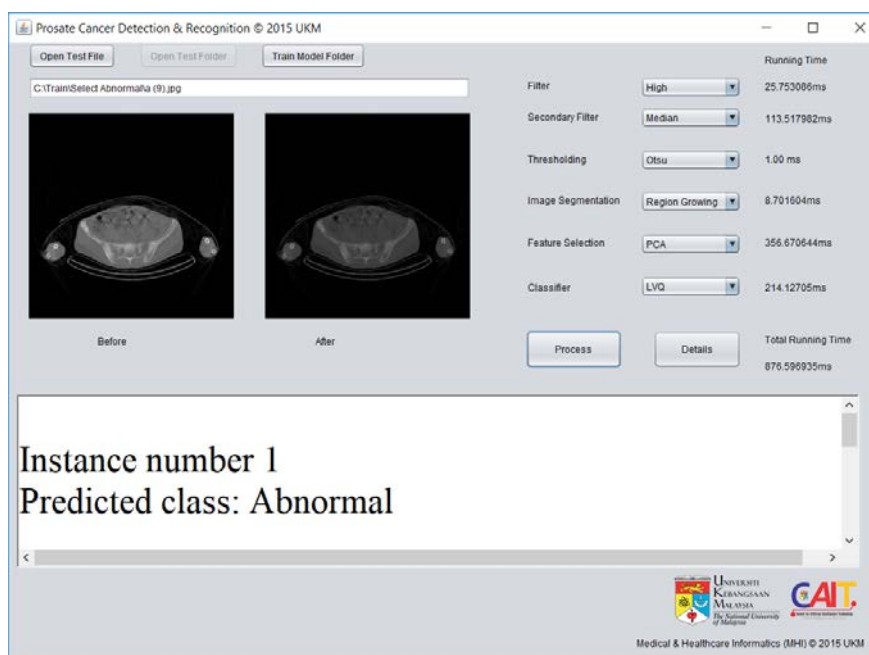
PIRADS: 2 Diagnose Diagnose AI Decision: Treatment Update Delete

Rajah 4.29 Antara muka pengurusan mengubahsuai rekod perubatan jabatan radiologi

Rajah 4.31 menunjukkan antara muka tambahan bagi meramal tahap kanser prostat bagi jabatan radiologi. Melalui antara muka utama ini, pengguna boleh memilih untuk mengetahui samada pesakit adalah normal *Benign* ataupun tidak normal *Malignant*. Selain itu, antara muka utama juga dimuatkan dengan maklumat tentang ketepatan ramalan dalam mengenalpasti kanser. Kesukaran dalam pengumpulan data menyebabkan kadangkala ramalan yang dibuat kurang tepat.



Rajah 4.30 Antara muka pengurusan paparan rekod perubatan jabatan radiologi



Rajah 4.31 Antara muka pengurusan paparan rekod perubatan jabatan radiologi dalam meramal tahap kaser prostat

Pengguna yang log in sebagai pakar histopatologi akan dapat menggunakan fungsi mencipta, mengemaskini, menghapus dan menyimpan rekod perubatan histopatologi pesakit dalam antara muka histopatologi seperti yang ditunjukkan di Rajah 4.32, Rajah 4.33 dan Rajah 4.34. Selain daripada itu, pakar histopatologi boleh menggunakan alat bantuan yang disediakan untuk menggedkan tahap malignan tumor pesakit semasa mencipta rekod perubatan pesakit.

The screenshot shows a web application window titled "Add Histopathology Report". At the top, there is a "Home" button and a patient selection dropdown menu showing "Patient: ID: 24, Name: Sulaiman". To the right are "Anotation Tool" and "Browse" buttons. Below this is a yellow header bar with "HISTOPATHOLOGY ASSESSMENT" on the left and "IMAGE INFORMATION" on the right. The main content area contains several input fields: "Result Gleason score most predominant:" with an empty text box and a help icon; a large "+" sign; "Result Gleason score second predominant:" with an empty text box and a help icon; "Result Overall Gleason Score =" with an empty text box and a help icon; "Result Interpretation:" with a dropdown menu set to "Benign" and a help icon; and "Doctor's comment:" with a large empty text area. At the bottom, there are "Diagnose" and "Save" buttons, and a "Decision:" dropdown menu set to "Follow - up".

Rajah 4.32 Antara muka pengurusan rekod jabatan histopatologi

The screenshot shows a web application window titled "Update Histopathology Report". It features a "Home" button and a patient selection dropdown menu showing "Patient: ID: 24, Name: Sulaiman". Below this is a "Report:" dropdown menu showing "Report ID: 9, Date: 2018-01-01", followed by "Anotation Tool" and "Browse" buttons. A yellow header bar contains "HISTOPATHOLOGY ASSESSMENT" and "IMAGE PROCESSING". The left side of the form has input fields: "Result Gleason score most predominant:" with the value "2" and a help icon; a large "+" sign; "Result Gleason score second predominant:" with the value "3" and a help icon; "Result Overall Gleason Score:" with the value "5" and a help icon; "Result Interpretation:" with a dropdown menu set to "Benign" and a help icon; and "Doctor's comment:" with a large empty text area. On the right side, there is a large image of a histopathology slide. At the bottom, there are "Diagnose", "Update", and "Delete" buttons, and a "Decision:" dropdown menu set to "Treatment".

Rajah 4.33 Antara muka pengurusan mengubahsuai rekod jabatan histopatologi

Rajah 4.34 Antara muka pengurusan paparan rekod jabatan histopatologi

4.2.7 Kemudahan Penerangan

Kemudahan penerangan dalam sistem pakar diagnosis kanser prostat adalah untuk menerangkan bagaimana sistem dapat membuat keputusan tertentu dan pengesahkan kaedah petua padanan yang berdasarkan ketepatan petua yang dijana. Dalam erti kata lain, Bilakah perlunya petua dipadankan.

Penjelasan ini boleh terdiri daripada bagaimana penyelesaian akhir atau pertengahan telah tiba untuk menjustifikasikan keperluan data tambahan. Di sini pengguna ingin bertanya soalan mengapa dan bagaimana sistem berfungsi sebagai pembelajaran dan perkongsian pengetahuan sistem dengan pengguna. Antara contoh penggunaan kemudahan penerangan dalam sistem diagnosis kanser prostat adalah:

GENERAL MEDICAL (Adding)

Home

PATIENT INFORMATION

Select patient: ID: 24, Name: Sulaiman

Patient ID: 24 Patient Name: Sulaiman
 Age: 28 Weight: 67.0
 Height: 178.0 BMI: 21.15

Update

IPSS SCORE

The International Prostate Symptom Score (IPSS) is an 8 question (7 symptom questions + 1 life question) written screening tool used to screen for, rapidly diagnose, track the symptoms of the symptoms of the disease benign prostatic hyperplasia (BPH)

Over the past month, how often have you had a sensation of not emptying your bladder completely after you finished urinating? Not at all

Over the past month, how often have you had to urinate again less than two hours after you finished urinating? Not at all

Over the past month, how often have you had a weak urinary stream? Not at all

Over the past month, how many times did you most typically get up to urinate, from the time you went to bed at night until the time you got up in the morning? None

Score: /35

FACTOR

Patients aged 50 years and older? No

Patients lose weight? No

Patients have a close relative with cancer of the prostate? No

Urine comes with blood? No

Doctor's comment:

Diagnose

Decision: Follow_up

Save

Rajah 4.35 Contoh kemudahan penerangan bagi jabatan pembedahan (urologi)

Add Radiology Report

Home

Patient: ID: 24, Name: Sulaiman

Annotation Tool Browse

RADIOLOGY ASSESSMENT **IMAGE INFORMATION**

Prostate volume (cm):

Lesion location: Apex right Apex left Medium grid right

Medium grid left Base right Base Left

Extracapsular extension: No

Adjacent local organ involvement: No

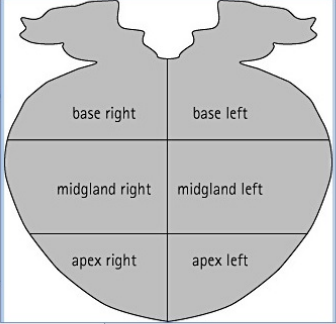
Organ involved: Seminal vesicles External sp

Levator muscles Pelvic wall

Patient nodes: Non - suspicious

Patient bone: Normal/Benign

Doctor's comment:



PIRADS: 0

Diagnose Diagnose AI

Decision: Follow_up

Save Clear

Rajah 4.36 Contoh kemudahan penerangan bagi jabatan radiologi

Rajah 4.37 Contoh kemudahan penerangan bagi jabatan histopatologi

4.3 KAEDAH UJIAN KEPENGGUNAAN

Fasa ini akan membincangkan tentang pengujian dan penilaian yang dilakukan terhadap hasil yang dikeluarkan oleh sistem. Dalam fasa ini, set data petua diagnosis kanser prostat akan digunakan untuk tujuan menguji dan menilai sistem yang dibangunkan. Setiap data akan dimasukkan ke dalam sistem dan hasil keputusan ramalan akan dibandingkan dengan kelas dari data set ujian. Penilaian sistem bergantung pada peratusan perbandingan keputusan petua tersebut. Matlamat pengujian sistem adalah untuk memastikan sistem dapat dijalankan dengan lancar.

a. Pengujian Unit Sistem

Pengujian unit adalah ujian yang dijalankan terhadap unit terkecil dalam sistem. Tujuan pengujian ini adalah untuk mengesan keralatan sistem dengan lebih awal dengan menfokuskan setiap kod yang diimplementasi ke dalam sistem sebelum digabungkan. Selain itu, pengujian ini juga untuk memastikan output yang dikeluarkan adalah konsisten dan tidak mengambil masa yang terlalu panjang untuk menghasilkan output.

b. Pengujian Integerasi Sistem

Pengujian ini adalah untuk memastikan bahagian-bahagian dalam antara muka sistem berfungsi dengan sistematik. Komponen sistem difokuskan dalam pengujian ini di mana unit kecil dalam sistem digabungkan. Dalam kajian ini, pengujian ini dijalankan melibatkan setiap modul antara muka untuk mengenalpasti ralat dalam sistem. Fungsi validasi data input disediakan supaya pengguna memasukkan kesemua input sebelum dapat menekan butang “Diagnosis” untuk diproses. Sekiranya pengguna tidak memasukkan kesemua input, maka sistem tidak dapat menjalankan proses dan input tidak lengkap.

c. Pengujian Sistem

Pengujian sistem adalah fasa pengujian yang melibatkan ujian keseluruhan sistem. Tujuan pengujian ini adalah untuk memastikan sistem tersebut memenuhi tujuan dan objektif sebenar sistem dibangunkan. Dalam kajian ini, set data akan diuji dengan membandingkan output yang dihasilkan oleh sistem dengan output yang diperoleh daripada data sebenar yang divalidasi oleh pakar.

4.4 KEPUTUSAN DAN ANALISA PETUA PENGELUARAN

Pengujian sistem diagnosis kanser prostat telah dilakukan dengan cara pengujian tertutup bersama dengan pakar dan pelajar dalam bidang pembedahan (urologi), radiologi dan histopatologi di Pusat Perubatan Universiti Kebangsaan Malaysia. Manual pengguna telah disediakan seperti yang dilampirkan di Lampiran H untuk penguji sistem semasa pengujian sistem dilakukan. Untuk mendapatkan maklum balas daripada penguji sistem, kaedah skala Likert telah digunakan untuk mengukur tahap keberkesanan sistem. Contoh borang pengujian penerimaan pakar yang digunakan sepanjang pengujian adalah seperti yang dilampirkan di Lampiran I.

Jadual 4.6 Keputusan markah purata bagi pengujian terhadap sistem bagi jabatan pembedahan (urologi)

Seksyen	N	1	2	3	4	5	Markah purata	Perbincangan
Kemudahan penggunaannya	1				2	3	4.33	Sistem mudah digunakan tetapi kekurangan maklumat yang diberikan.
	2				4	1		
	3				4	1		
	4			1	3	1		
Kejelasan input	5				5		4	Input yang diperlukan adalah lengkap tetapi tidak
	6				5			
Penerangan kenapa	7				5		4	Penjelasan keputusan sistem adalah memadai tetapi tidak mempunyai maklumat
	8				4	1		
Persembahan keputusan	9				5		4.06	Keputusan sistem adalah lengkap dan senang difahami.
	10			1	4			
Pertimbangan umum	11				5		3.86	Sistem adalah cepat dan berguna.
	12			1	4			

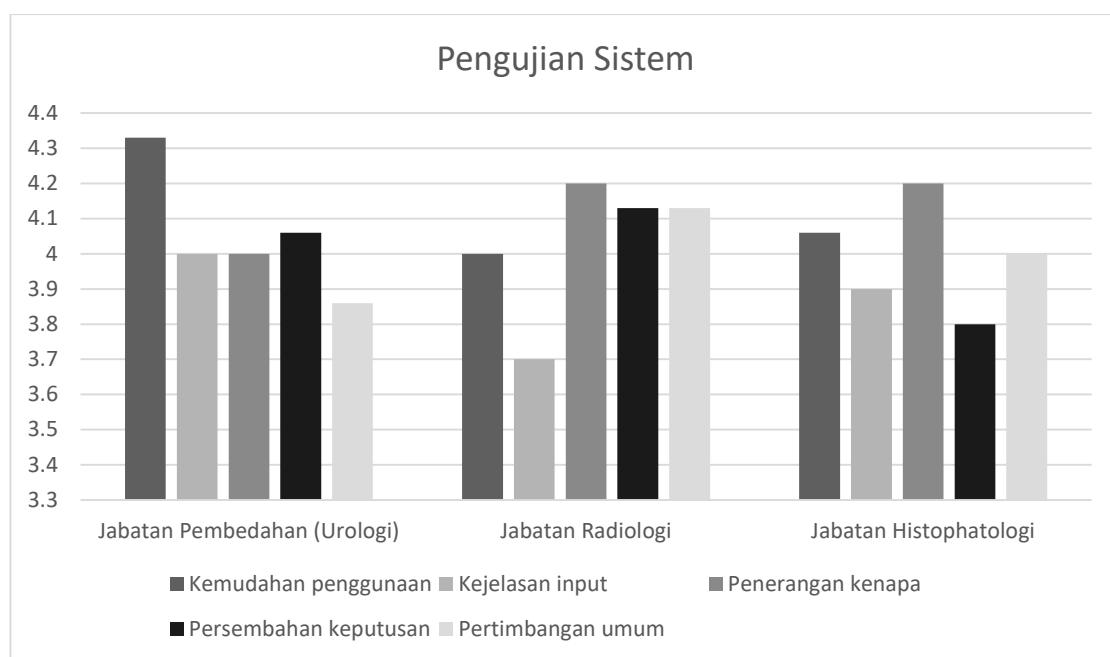
Jadual 4.7 Keputusan markah purata bagi pengujian terhadap sistem bagi jabatan radiologi

Seksyen	No	1	2	3	4	5	Markah purata	Perbincangan
Kemudahan pengguna an	1				4	1	4	Sistem mudah digunakan tetapi kekurangan maklumat yang diberikan.
	2				4	1		
	3		1		4			
	4			2	2	1		
Kejelasan input	5			2	3		3.7	Input yang diperlukan adalah lengkap tetapi tidak
	6				4	1		
Penerangan kenapa	7				3	2	4.2	Penjelasan keputusan sistem adalah memadai tetapi tidak mempunyai maklumat
Persempahan keputusan	8			1	3	1	4.13	Keputusan sistem adalah lengkap dan senang difahami.
	9			1	3	1		
	10			2	3			
Pertimbangan umum	11				4	1	4.13	Sistem adalah cepat dan berguna.
	12				2	3		

Jadual 4.8 Keputusan markah purata bagi pengujian terhadap sistem bagi jabatan histopatologi

Seksyen	N	1	2	3	4	5	Markah purata	Perbincangan
Kemudahan penggunaan	1			1	3	1	4.06	Sistem mudah digunakan tetapi kekurangan maklumat yang diberikan.
	2				3	2		
	3			1	4			
	4			3	1	1		
Kejelasan input	5			1	2	2	3.9	Input yang diperlukan adalah lengkap tetapi tidak
	6			1	2	2		
Penerangan kenapa	7			1	4		4.2	Penjelasan keputusan sistem adalah memadai tetapi tidak mempunyai maklumat
Persembahan keputusan	8			1	4		3.8	Keputusan sistem adalah lengkap dan senang difahami.
	9			1	4			
	10			1	4			
Pertimbangan umum	11				4	1	4	Sistem adalah cepat dan berguna.
	12				4	1		

Sebelum menganalisa maklum balas dalam borang pengujian penerima pengguna dengan mengambil markah purata untuk setiap seksyen pengujian bagi terhadap sistem semua jabatan dikira dan hasilnya adalah 4.03/5, keputusan yang didapati adalah seperti yang ditunjukkan dalam Rajah 4.38. Sistem yang dibangunkan menggunakan petua pengeluaran mampu membantu pakar dalam mendiagnosis kanser prostat dengan lebih mudah dan pantas. Ini disebabkan kemungkinan model yang menggunakan petua pakar dapat mengagak output bagi input data yang diberikan kerana petua yang diperolehi sesuai bagi menganalisis input yang diberikan disamping penerangan yang jelas yang diberikan sistem. Selain itu, petua pengeluaran yang mempunyai banyak nilai linguistik yang akan memberi keputusan yang lebih tepat bagi penentuan kanser prostat.



Rajah 4.38 Keputusan markah purata pengujian terhadap sistem bagi semua jabatan

Seterusnya, untuk mengukur tahap penerimaan pakar kepada data petua pengeluaran diagnosis kanser prostat yang dikumpul terhadap 15 pakar dalam penggunaan sistem bagi semua jabatan dikira dan hasil keseluruhannya adalah 76.4%, keputusan yang didapati adalah seperti yang ditunjukkan dalam Rajah 4.39 di mana temubual telah dijalankan dengan kehadiran 15 pakar-pakar. Komen dan cadangan yang telah dikumpul secara ringkas. Menurut Durkin (1994), pakar-pakar luar boleh menilai ketepatan keputusan dalam pelbagai jenis peratusan kes, atau sistem pula serta, atau

lebih baik daripada pakar-pakar. Borang cadangan pakar dan maklumbalas bagi setiap modul digambarkan dalam Jadual 4.9, Jadual 4.10 dan Jadual 4.11.

Jadual 4.9 Keputusan markah purata pengujian penerimaan petua penagluaran yang dikumpul bagi urologi

Seksyen	Markah purata (%)						Perbincangan
	Pakar 1	Pakar 2	Pakar 3	Pakar 4	Pakar 5	Keseluruhan	
Perolehan pengetahuan	80	80	80	80	80	4	Kaedah petua diperolehi lengkap.
Penerangan kemudahan	80	80	80	80	80	4	Pengetahuan yang diperolehi mudah untuk dijelaskan.
Bantuan kemudahan	80	80	80	80	60	3.8	Menpunyai kemudahan bantuan dalam menerangkan pengetahuan.
Keputusan sistem	80	80	80	80	80	4	Petua yang diperoleh jelas dan tepat.
Pengesyoran kepada orang lain	80	80	80	60	80	3.8	Kegunaan menyarankan sistem digunakan

Jadual 4.10 Keputusan markah purata pengujian penerimaan petua penagluaran yang dikumpul bagi jabatan radiologi

Seksyen	Markah purata (%)						Perbincangan
	Pakar 1	Pakar 2	Pakar 3	Pakar 4	Pakar 5	Keseluruhan	
Perolehan pengetahuan	60	60	80	80	80	3.6	Kaedah petua diperolehi lengkap.
Penerangan kemudahan	60	80	80	100	80	4	Pengetahuan yang diperolehi mudah untuk dijelaskan.
Bantuan	60	80	80	80	60	4.4	Menpunyai

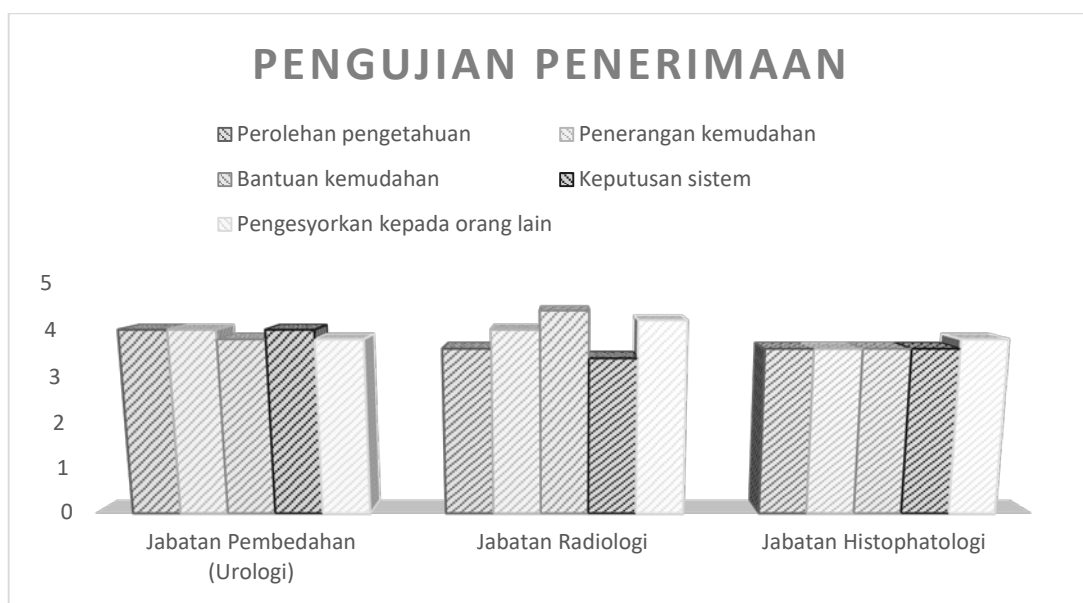
kemudahan								kemudahan bantuan dalam menerangkan pengetahuan.
Keputusan sistem	60	80	60	80	60	3.4		Petua yang diperoleh jelas dan tepat.
Pengesyorkan kepada orang lain	80	80	100	100	60	4.2		Kegunaan menyarankan sistem digunakan

Jadual 4.11 Keputusan markah purata penerimaan petua penegluaran yang dikumpul bagi jabatan histopatologi

Seksyen	Markah purata (%)					Keseluruhan	Perbincangan
	Pakar 1	Pakar 2	Pakar 3	Pakar 4	Pakar 5		
Perolehan pengetahuan	80	80	80	60	60	3.6	Kaedah petua diperolehi lengkap.
Penerangan kemudahan	80	80	80	60	60	3.6	Pengetahuan yang diperolehi mudah untuk dijelaskan.
Bantuan kemudahan	80	80	80	60	60	3.6	Menpunyai kemudahan bantuan dalam menerangkan pengetahuan.
Keputusan sistem	80	80	80	60	60	3.6	Petua yang diperoleh jelas dan tepat.
Pengesyorkan kepada orang lain	100	80	80	60	60	3.8	Kegunaan menyarankan sistem digunakan

Hasil dari pengujian menunjukkan jabatan urologi dan jabatan radiologi mendapat penerimaan pakar yang tinggi sama banyak sebanyak 3.92/5 skala Likert berbanding jabatan histopatologi sebanyak 3.64/5 skala Likert. Pakar-pakar mendapati terdapat beberapa komponen yang masih perlu dikemaskini terutama dari segi penggunaan istilah yang bersesuaian dengan pengguna dari bidang perubatan. Setelah kesemua pengujian dilakukan oleh pakar dalam setiap jabatan, sistem pakar yang

dibangunkan diperbaiki mengikut perkara yang telah dicadangkan. Istilah juga telah diubah supaya lebih bersesuaian dengan pengguna dari bidang perubatan. Walau bagaimanapun, masih terdapat beberapa perkara yang sukar untuk diperbaiki memandangkan kesukaran untuk mendapatkan imej dan data yang dikehendaki dari PPUKM. Berdasarkan maklum balas komen yang diberikan, secara keseluruhan, sistem pakar diagnosis kanser prostat dapat diterima dan 15 orang pakar bersetuju sistem ini mampu memudahkan mendiagnosis kanser prostat dengan lebih senang dan tepat.



Rajah 4.39 Keputusan markah purata pengujian penerimaan petua penegluaran bagi semua jabatan

Bagi sistem penilaian, 15 orang pakar yang mahir dalam bidang masing-masing dalam kanser prostat. Borang soal selidik komen diwujudkan untuk membantu mendapatkan cadangan mereka. Ia meliputi beberapa bahagian dan topik seperti teknik-teknik pengambilan ilmu, pengetahuan kandungan, Kemudahan penjelasan, pantas membuat keputusan, Kemudahan bantuan dan keyakinan mengenai cadangan. Maklum balas atau komen mereka yang diberi oleh kumpulan penguji adalah seperti yang berikut:

1. Mendapat maklumat lebih terperinci supaya dapat meningkatkan ketepatan petua.
2. Menambah fungsi integrasi dengan telefon pintar.

3. Menambah fungsi untuk memaparkan rekod keseluruhan jabatan.
4. Menambah fungsi gambar dalam pelayan sever.
5. Menambah fungsi anotasi gambar.
6. Menaik taraf antara muka supaya lebih mesra pengguna.
7. Mendapatkan data yang lebih terkini mengikut kes yang sebenar.

Selepas membuat analisis kuantitatif, soal selidik yang dijalankan bersama 15 pakar mengenai kanser prostat peringkat awal mencadangkan bahawa pelan rawatan ini hanya akan membantu pakar dan kebanyakannya diputuskan sendiri oleh pakar. Walaupun pada mulanya menerima keputusan akhir masih ditangan pakar, pihak yang di soal selidik dari masa ke semasa ingin mengkaji semula keputusan tersebut. Keputusan yang dihasilkan masih lagi dapat diterima oleh pakar dan tetap boleh diterima pakai dalam situasi ini.

4.5 RUMUSAN

Secara kesimpulannya, kumpulan penguji adalah berpuas hati dengan sistem diagnosis kanser prosat yang dibangunkan. Ralat yang dikenalpasti semasa pengujian juga dileraikan untuk meningkatkan tahap kepuasan pengguna. Komen- komen konstruktif daripada kumpulan penguji telah dicatat dan bakal digunakan untuk menambahbaik sistem untuk versi yang akan datang. Data yang baik dan berkualiti juga sangat penting untuk memastikan hasil petua yang diperoleh juga baik.

BAB V

MODEL POKOK RAWAK

5.1 PENGENALAN

Setelah teknik atau algoritma penemuan pengetahuan dipilih, perkara seterusnya ialah menggunakan teknik tersebut kepada data. Pemilihan parameter yang sesuai merupakan aspek penting untuk mendapatkan model yang baik. Justeru, kajian ini membandingkan beberapa parameter untuk diuji keberkesannya terhadap petua diagnosis kanser prostat. Antara parameter-parameter yang akan digunakan dalam kajian ini ialah Pokok Rawal di mana parameter yang akan digunakan ialah maxDepth, minNum, numFolds dan seed. Setiap parameter akan diubah-ubah nilainya

5.2 PENGUJIAN SET DATA BAGI MENJANA PETUA

Dalam permodelan petua pokok rawak dan petua pengeluaran, terdapat parameter yang boleh diubah nilainya iaitu dalam pecahan peratusan (*Percentage Splits*) untuk mendapatkan keputusan optimum berdasarkan nilai pembolehubah yang boleh ditentukan. Melalui kaedah pecahan peratusan, set data akan dibahagikan kepada dua set iaitu set data latihan dan set data pengujian bagi mendapatkan model yang terbaik. Kajian untuk mendapatkan model yang terbaik akan dijalankan menggunakan platform Weka 3.8. Pecahan peratusan ini akan menghasilkan sembilan set model yang dibahagikan dengan nisbah latihan ujian, yang merujuk kepada peratusan yang diambil dari data untuk latihan dan peratusan yang diambil dari data untuk pengujian.

Berdasarkan nilai pecahan peratusan yang ditunjukkan pada Jadual 5.1, model yang sesuai akan diambil dengan mencari nilai ketepatan tertinggi berdasarkan dua faktor utama iaitu jumlah petua yang dihasilkan dan peratusan ketepatan data (*Correctly Classified Instances*) Namun, data ujian akan diambil mengikut pengukuran nilai pecahan peratusan data ujian yang tinggi, mempunyai peratusan ketepatan data yang tertinggi dan nilai ROC yang tertinggi.

Jadual 5.1 Nilai yang diperoleh mengikut nisbah pecahan peratusan

Model	Data Latihan (%)	Data Ujian (%)	Ketepatan Data (%)	Bilangan Petua
1	10	90	34.78	10
2	20	80	87.80	10
3	30	70	86.11	10
4	40	60	93.54	10
5	50	50	92	10
6	60	40	95	10
7	70	30	53.33	10
8	80	20	70	10
9	90	10	100	10

5.3 MODEL PETUA POKOK RAWAK BAGI JABATAN RADIOLOGI

Data yang diambil oleh pakar radiologi dan laporan pesakit dari Pusat Perubatan UKM, Malaysia. Berikut adalah data sampel rekod pesakit di Jabatan Radiologi.

5.3.1 Perolehan Data

Semua pesakit dengan diagnosis diketahui adenokarsinoma prostat dan menjalani *resection* pembedahan (Urologi). Pengumpulan data akan diambil dari pejabat rekod perubatan, IRIS / OMS dan rekod pesakit dan keputusan H&E dari pelayan Medweb / Pusat Perubatan Universiti Kebangsaan Malaysia (PPUKM) PACS. Rujuk Jadual 5.2.

Jadual 5.2 Set data rekod pesakit dari Pusat Perubatan UKM

Nama Set	Jumlah	Attribut	Nombor Kelas	Sumber
Data	Kes			
Kanser Prostat	51	8	Tidak Normal	PPUKM
Jenis Kelas			Keseluruhan	
Tidak Normal			51	

Set data rekod pesakit yang dikumpulkan terdiri daripada 51 rekod laporan pesakit yang tidak normal Lampiran G. Bagi setiap rekod pesakit dipilih 9 atribut dan 1 kelas bagi menentukan tahap kanser prostat dan rawatan seterusnya.

5.3.2 Pembahagian Set Data

Perlombongan data yang merupakan data yang sudah diproses terbahagi kepada dua iaitu set data latihan dan set data pengujian. Kedua-dua set data ini digunakan untuk membina satu model pengelasan yang baik. Terdapat dua proses utama dalam pengelasan data. Pertama, model daripada set data latihan yang dibina digunakan dalam menentukan atribut kelas. Kedua, ketepatan menggunakan set data akan diuji mengikut model tersebut (Han dan Kamber 2011; Hor et al. 2007). Terdapat pelbagai kaedah yang boleh diguna pakai dalam pembahagian data. Antara contohnya ialah kaedah pengesahan silang lipatan-k.

Dalam kajian ini, terdapat kaedah pengujian silang lipatan-k yang digunakan kerana semua set data mempunyai peluang untuk dilatih dan diuji selain ia turut memberikan prestasi yang bagus (Guttierez-Osuna 2002). Dalam kaedah ini, data awal dibahagikan secara rawak kepada k subset atau lipatan, L_1, L_2, \dots, L_k , di mana saiz bagi setiap satu lipatan tersebut adalah sama. $k-1$ lipatan akan digunakan untuk latihan manakala yang selebihnya akan digunakan untuk pengujian. Sebagai contoh,

pembahagian data latihan dan data pengujian yang berbeza sebanyak sembilan kali seperti 90:10, 80:20, 70:30, 60:40, 50:50, 40:60, 30:70, 20:10, dan 10:90.

5.3.3 Pra-Promprosesan Data Bagi Petua Pokok Rawak

Rangka kerja ini akan kemudiannya disahkan oleh pakar-pakar pembedahan (urologi), Radiologi dan Patologi Pusat Perubatan Universiti Kebangsaan Malaysia.

Pembangunan model merupakan satu proses yang kukuh dan model yang dibina akan digunakan sebagai pangkalan pengetahuan kepada sistem yang bakal dibangunkan ini. Dalam pembangunan model untuk mendiagnosis kanser prostat, pengelasan berasaskan petua dijalankan pada set data pesakit yang telah disediakan semasa pra-pemprosesan data. Teknik pengelasan berasaskan petua yang digunakan ialah pokok keputusan. Pokok keputusan merupakan algoritma yang biasanya dapat menghasilkan set petua yang berketepatan tinggi berbanding algoritma lain. Pengelas pokok rawak yang merupakan salah satu pengelas pokok keputusan digunakan dalam kajian ini.

Sebelum proses perlombongan data hendak dilaksanakan, pra-pemprosesan data akan dilaksanakan bagi membersihkan data yang tidak diperlukan di mana boleh mempengaruhi petua. Menurut Han dan Kamber (2011), data yang tidak lengkap merupakan data yang negatif. Contohnya data yang tidak lengkap disebabkan oleh faktor-faktor seperti masalah daripada manusia, perkakasan atau perisian. Tambahan lagi, data hingar seperti ketika memasukkan data ataupun ketika transmisi data boleh mengakibatkan kepada data bising. Beberapa faktor seperti data yang berpunca daripada pelbagai sumber juga boleh membawa kepada data yang tidak konsisten seterusnya menjadi data kotor.

Pra-pemprosesan data yang akan dijalankan terbahagi kepada tiga peringkat iaitu, pembersihan data, pengurangan data, dan transformasi data. Penggunaan data yang tidak berkualiti secara tidak langsung boleh mengakibatkan keputusan yang diperolehi daripada perlombongan data tidak menarik dan boleh dikategorikan gagal. Antara faktor penyebab data gagal ataupun tidak berkualiti disebabkan, data yang bertindan dan tidak lengkap boleh mengakibatkan laporan statistik yang diperolehi tidak betul dan tidak menunjukkan angka yang sebenar, data yang hilang semasa proses

mengisi pangkalan data, data yang tidak konsisten dan data yang salah semasa kemasukan data. Oleh sebab itu, pra-pemprosesan dianggap proses yang sangat penting dalam perlombongan data kerana jika data tidak berkualiti maka pengetahuan yang akan diperolehi juga tidak berkualiti.

Terdapat beberapa aktiviti utama yang boleh dilaksanakan semasa proses pra-pemprosesan. Di antaranya ialah pembersihan data, integrasi data, transformasi data, pengurangan dan pendiskretan data. Berikut merupakan beberapa langkah yang dilaksanakan semasa pra- pemprosesan data.

a) Pembersihan Data

Tujuan pembersihan data adalah untuk mengisi data yang hilang atau tidak lengkap atau membuang data yang tidak berkualiti yang terdapat pada set data dalam meningkatkan kualiti data. Antara kaedah yang digunakan untuk mengisi data yang hilang adalah dengan mengira nilai majoriti bagi bilangan atribut tersebut. Proses ini penting untuk penambahbaikan kualiti data. Proses pembersihan data yang boleh dijalankan adalah seperti mengambil nilai majoriti, menggunakan nilai purata atribut, menggunakan pembolehubah global atau mengagihkan nilai atribut secara sama rata. Data yang hilang bermaksud set data tersebut adalah tidak lengkap. Hal ini boleh terjadi semasa proses cerapan disebabkan oleh masalah teknikal pada alat- alat cerapan mengakibatkan data yang direkod tidak lengkap. Proses pembersihan yang dijalankan adalah dengan membuang data dari set data yang mempunyai dua atau lebih data yang hilang. Hasilnya, tiada data yang disingkirkan daripada data asal yang mengandungi 51 data. Bilangan set data tidak berubah.

b) Pengurangan Data

Atribut data yang tidak digunakan dalam kajian kes ini akan dikeluarkan daripada set data. Pengeluaran atribut ini mampu memberikan peningkatan dalam keberkesanan perlombongan data dan petua yang akan dilaksanakan. Terdapat 22 atribut yang berlainan dalam semua set data dari laporan pesakit di jabatan radiologi. Dalam kajian kes ini, 8 atribut dipilih dan dikekalkan dalam set data

dan atribut yang lain dibuang kerana didapati tidak memberi impak dalam membina petua. Atribut yang dikekalkan adalah *Volume*, *Lesion Location*, *Extracapsular Extension*, *Organ Involved*, *Suspicious Patient Nodes*, *Abnormal Patient Bones*, *PIRADS* dan *DECISION*. Atribut tersebut dipilih kerana relevan kerana menjadi tunjang utama yang berkaitan dengan petua. Selain itu, atribut yang dikeluarkan adalah attribute pecahan dari *Lesion Location* dan *Organ Involve* terdapat 13 kesemuanya. Pada peringkat awal, antara atribut yang dikeluarkan hendak dipilih dalam kajian ini, tetapi didapati atribut tersebut tidak perlu untuk memperoleh perwakilan data yang lebih kecil dengan tidak mengganggu keaslian data supaya proses permodelan berjalan dengan lebih lancar. Pengeluaran atribut ini juga bertujuan bagi mengelakkan daripada berlakunya penghasilan keputusan yang tidak tepat bagi petua yang akan dijana. Atribut yang kekal dan yang telah diolah adalah seperti dalam Jadual 5.3 di bawah:-

Jadual 5.3 Pengolahan atribut

KEKAL		BUANG	
Atribut	Bentuk	Atribut	Bentuk
<i>Volume</i>	Numerik	<i>Local Adjacent</i>	Numerik
<i>Lesion Location</i>	Numerik	Apex right	Numerik
<i>Extracapsular Extension</i>	Numerik	Apex left	Numerik
<i>Organ Involved</i>	Numerik	Medium right	Numerik
<i>Suspicious Patient Nodes</i>	Numerik	Medium left	Numerik
<i>Abnormal Patient Bones</i>	Numerik	Base right	Numerik
<i>PIRADS</i>	Numerik	Base left	Numerik
<i>DECISION</i>	Aksara	Seminal vesicles	Numerik
		External sphincter	Numerik
		Rectum	Numerik
		Bladder	Numerik

Levator mucleus	Numerik
Pelvic wall	Numerik

c) Tranformasi Data

Dalam proses ini, transformasi data yang berlaku ialah pendiskretan data yang bertujuan untuk menukarkan data bentuk numerik ke bentuk nominal mudah. Kaedah ini adalah untuk memudahkan permodelan data dan pembangunan sistem dengan menggunakan perwakilan nilai atribut yang senang difahami dan digunakan seperti 0,1,2,3,4 di samping memudahkan proses menganalisis data. Atribut yang mengandungi julat angka selanjur perlu ditukar ke bentuk nominal melalui pembahagian skala tertentu. Dalam kajian ini, atribut yang melalui proses pendiskretan data adalah seperti Rajah 5.1 di bawah: -

Tumour	
Image Description	wt
Lesion Location	1a
Mutiple Lesion Location	1b
No Lesion Location	1c
Extracapsular Extension	2a
No Extracapsular Extension	2b
Organ Involved	4a
Mutiple Organ Involved	4b
No Organ Involved	4c
Suspicious Patient Nodes	5a
Non Suspicious Patient Nodes	5b
Abnormal Patient Bones	6a
Normal Patient Bones	6b
For Volume	
CC Description	Wt
0<=30 (Small)	1
31-80 (Large)	2
>=81 (Very Large)	3
Desicion	
Step Description	Wt
Follow - up	1
Biopsy	2
Treatment Oncology	3

Rajah 5.1 Pendiskretan data

5.4 KONFIGURASI EKPERIMEN

Beberapa jenis teknik pembelajaran mesin telah dilakukan diuji untuk mengkaji keupayaan bagi mendapatkan keputusan pengelasan yang baik bagi set data diagnosis kanser prostat. Antara teknik pembelajaran mesin yang dikaji bagi kaedah pengelasan ialah J48, hutan rawak dan pokok rawak. Bagi setiap teknik yang digunakan akan dilatih dengan beberapa parameter yang telah ditetapkan untuk mendapatkan keputusan kajian yang lebih baik dan berkualiti. Model yang paling baik untuk setiap teknik pembelajaran mesin akan diteruskan dengan set pengujian data. Kemudiaannya, setelah menjalani pengujian data, pemilihan teknik yang mendapat ketepatan ramalan yang tinggi akan dijalankan.

5.4.1 Eksperimen Pokok Rawak

Teknik Pokok Rawak merupakan gabungan daripada dua algoritma pembelajaran mesin yang sedia ada iaitu model pokok tunggal digabungkan dengan hutan rawak. Konsep pokok rawak adalah menghasilkan seberapa banyak pokok daripada atribut yang dipilih secara rawak. Pokok yang dianggarkan seimbang akan berpecah pada median untuk beberapa atribut. Random tree menggunakan hasil ini untuk memilih pecahan bagi atribut dan menghasilkan pokok yang seimbang di mana penetapan global bagi setiap nod daun ditentukan bagi memudahkan prosedur pengoptimuman (Pfahringner 2010).

Bagi penilaian ketepatan terhadap pokok rawak, kaedah peratus pecahan dan kaedah keesahan silang digunakan. Peratus pecahan tersebut terdiri daripada 50%, 55%, 60%, 65%, 70% dan 75% Manakala keesahan silang terdiri daripada 10 lipatan dan 25 lipatan.

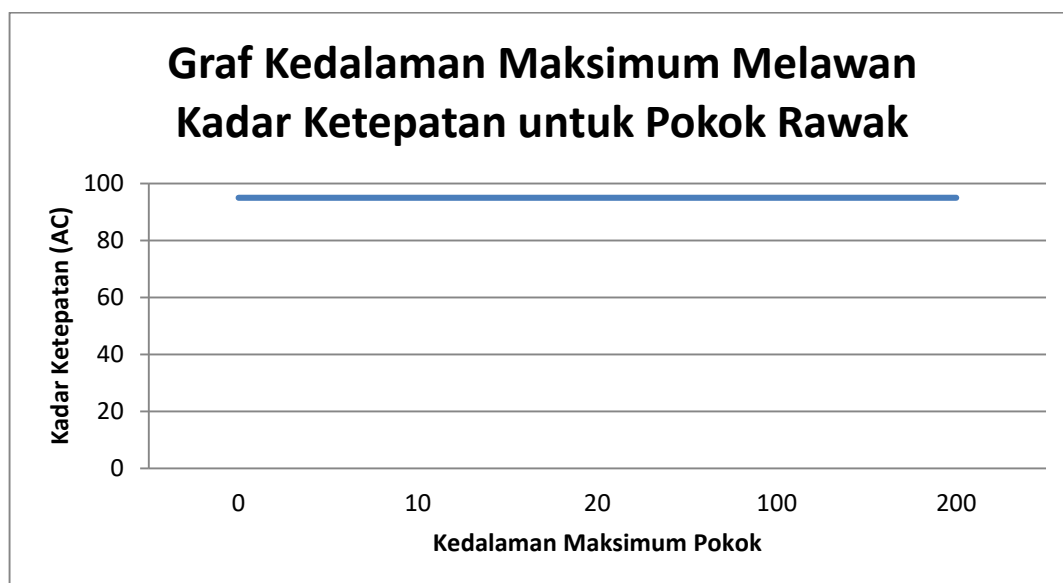
a. Pemilihan Kedalaman Maksimum Pokok Rawak

Dalam kajian ini, pokok rawak digunakan untuk mendapatkan kedalaman maksimum pokok yang sesuai menggunakan set data prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 5.4 dan Rajah 5.2, prestasi terbaik yang ditunjukkan Pokok Rawak ditunjukkan pada kedalaman maksimum pokok adalah pada nilai 0 default. Kadar

ketepatan atau *Accuracy Rate* (AC) yang dicapai adalah sebanyak 95% dengan ralat min kuasa dua atau *Mean Sum-squared Error* (MSE) sebanyak 0.1826. Oleh itu, kedalaman maksimum pokok iaitu 0 ditetapkan pada semua eksperimen. Dapat dilihat bahawa purata peratusan ketepatan yang sama adalah apabila kedalaman maksimum ialah 0 sehingga 200. Ralat min kuasa dua bagi setiap kedalaman maksimum juga adalah sama.

Jadual 5.4 Pemilihan kedalaman maksimum pokok untuk latihan pokok rawak

Kedalaman Maksimum Pokok	0 (Tidak Terhad)	10	20	100	200
AC(%)	95	95	95	95	95
MSE	0.1826	0.1826	0.1826	0.1826	0.1826



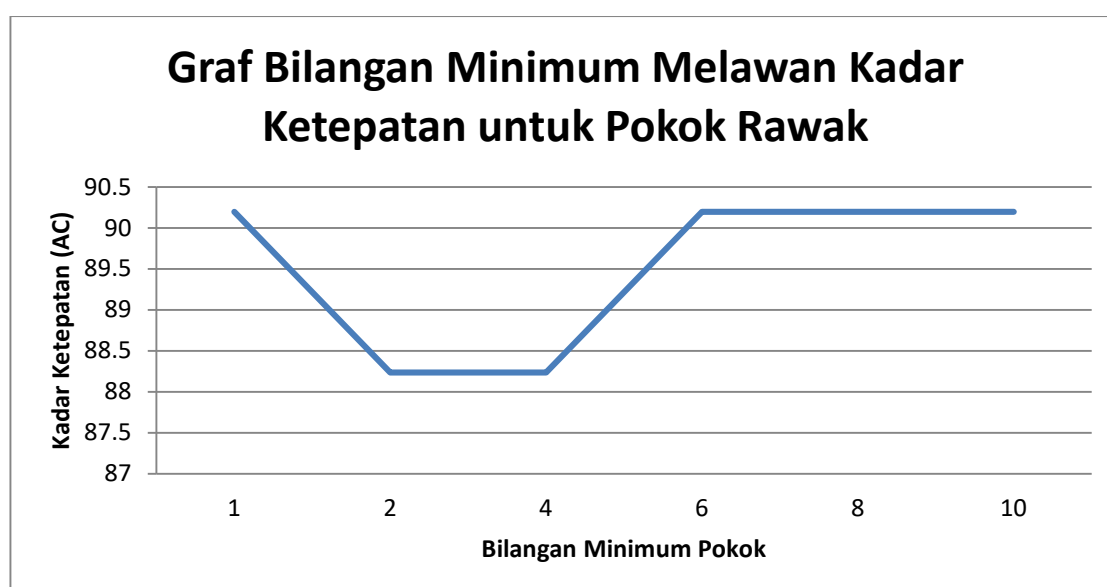
Rajah 5.2 Bilangan slot perlakuan melawan kadar ketepatan untuk pokok rawak

b. Pemilihan Bilangan Minimum Pokok Rawak

Dalam kajian ini, pokok rawak digunakan untuk mendapatkan jumlah berat minimum yang sesuai menggunakan set data prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 5.5 dan Rajah 5.3, prestasi terbaik yang ditunjukkan Pokok Rawak ditunjukkan pada bilangan minimum pokok adalah pada nilai 1.0 default. Kadar ketepatan atau *Accuracy Rate* (AC) yang dicapai adalah sebanyak 95% dengan ralat min kuasa dua atau *Mean Sum-squared Error* (MSE) sebanyak 0.1826. Oleh itu, kedalaman maksimum pokok iaitu 1.0 ditetapkan pada semua eksperimen. Dapat dilihat bahawa purata peratusan ketepatan yang sama adalah apabila bilangan minimum 0.0 sehingga 10.0 dan peratusan ralat min kuasa dua ramalan semakin rendah apabila nilai bilangan minimum meningkat.

Jadual 5.5 Pemilihan bilangan minimum untuk latihan pokok rawak

Bilangan Minimum Pokok	0.0 (Tidak Terhad)	1.0	5.0	10.0	50.0	100.0
AC	95	95	95	95	30	30
MSE	0.1826	0.1826	0.1925	0.1925	0.4724	0.4724



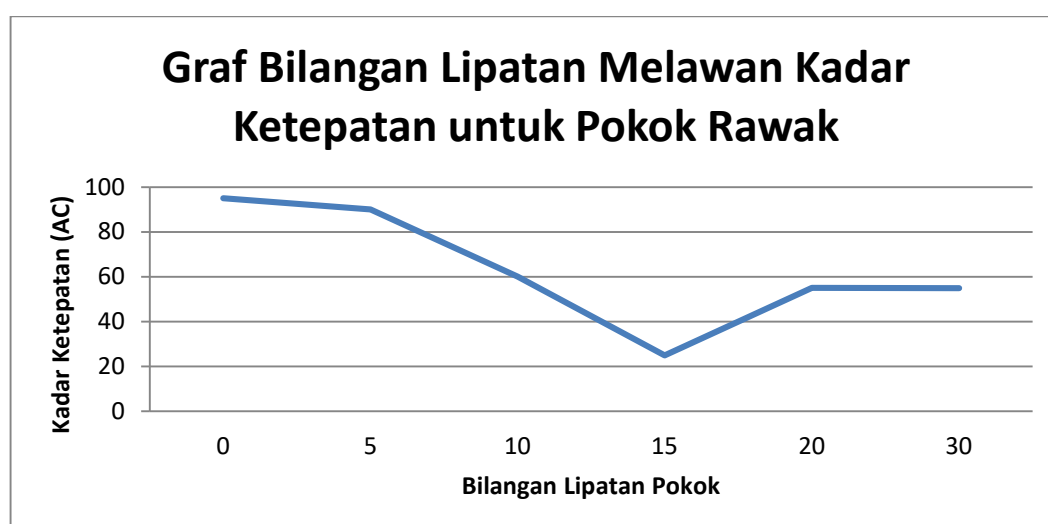
Rajah 5.3 Bilangan minimum pokok melawan kadar ketepatan untuk pokok rawak

c. Pemilihan Bilangan Lipatan Pokok Rawak

Dalam kajian ini, pokok rawak digunakan untuk mendapatkan bilangan lipatan yang sesuai dengan menggunakan set data imej prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk menentukan jumlah data digunakan untuk penyesuaian semula. Satu lipatan digunakan untuk penyesuaian semula, yang lain untuk menanam pokok dan pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 5.6 dan Rajah 5.4, prestasi terbaik ditunjukkan pokok rawak adalah sebanyak 95% dengan ralat min kuasa dua (MSE) sebanyak 0.1826 pada lipatan adalah 0 default. Oleh itu, lipatan adalah 0 ditetapkan pada semua eksperimen. Dapat dilihat purata peratusan ketepatan ramalan semakin rendah apabila nilai bilangan lipatan semakin meningkat. Purata peratusan ramalan paling tinggi apabila bilangan lipatan adalah 0 dengan ralat min kuasa dua yang paling rendah apabila parameter diubah.

Jadual 5.6 Pemilihan bilangan lipatan pokok untuk latihan pokok rawak

Bilangan Lipatan (Tidak Terhad)	0	5	10	15	20	30
AC	95	90	60	25	55	35
MSE	0.1826	0.2357	0.2833	0.05	0.5477	0.6583



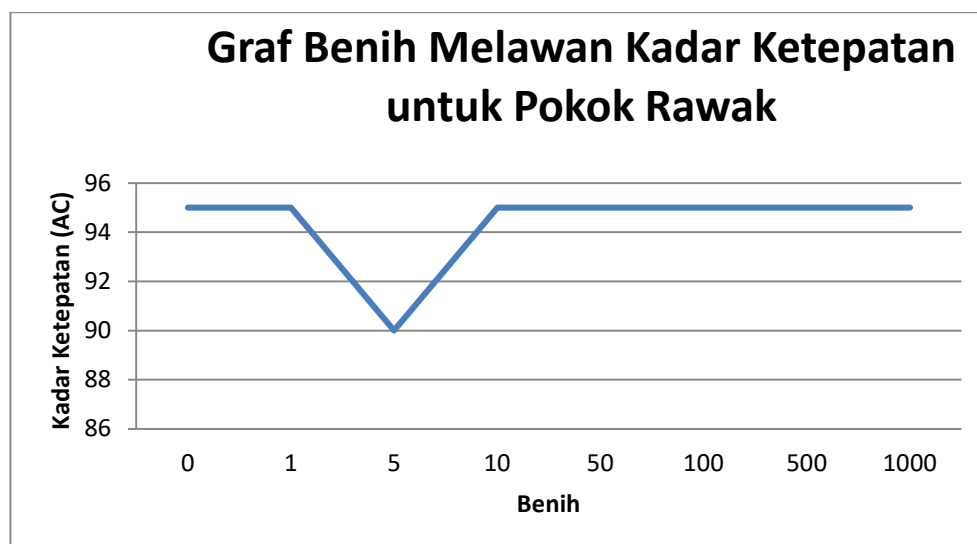
Rajah 5.4 Bilangan lipatan pokok melawan kadar ketepatan untuk pokok rawak

d. Pemilihan Benih Pokok Rawak

Dalam kajian ini, pokok rawak digunakan untuk mendapatkan benih yang sesuai dengan menggunakan set data diagnosis kanser prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 5.7 dan Rajah 5.5, prestasi terbaik ditunjukkan Pokok Rawak adalah sebanyak 95% dengan ralat min kuasa dua (MSE) sebanyak 0.1826 pada benih adalah 1 default. Oleh itu, benih adalah 1 ditetapkan pada semua eksperimen. Dapat dilihat purata peratusan ketepatan ramalan sama apabila nilai benih 0 sehingga 1000. Purata peratusan ralat min kuasa dua tinggi apabila benih adalah 5 apabila parameter diubah.

Jadual 5.7 Pemilihan benih untuk latihan pokok rawak

Benih	0	1	5	10	50	100	500	1000
(AC)	95	95	90	95	95	95	95	95
(MSE)	0.1826	0.1826	0.2258	0.1826	0.1826	0.1826	0.1826	0.1826



Rajah 5.5 Benih melawan kadar ketepatan untuk pokok rawak

e. Keputusan dan Analisis Pokok Rawak

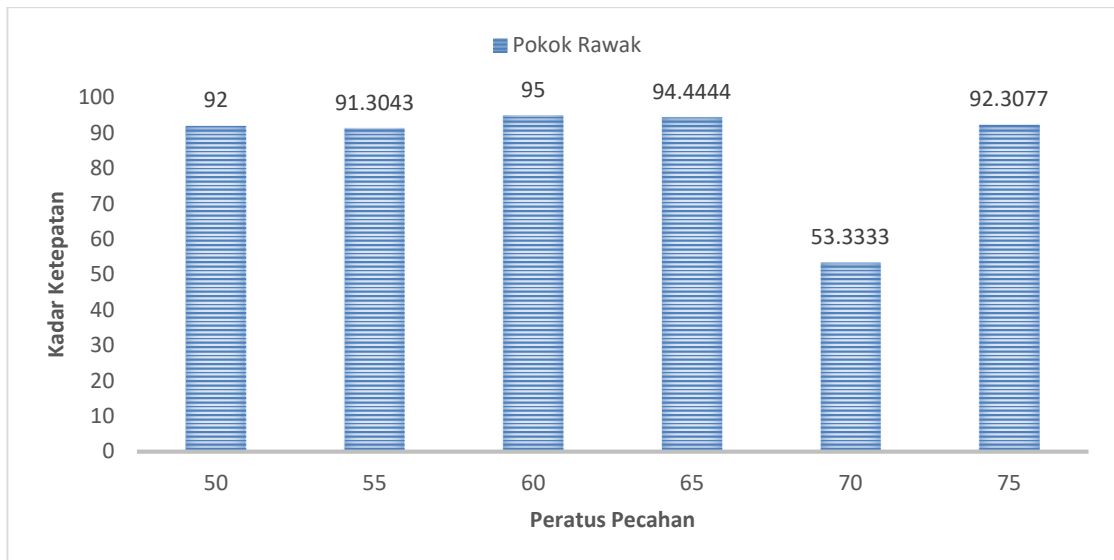
Ujikaji dilakukan ke atas pokok rawak di mana parameter bagi kedalaman maksimum, bilangan minimum, bilangan lipatan dan benih akan diubah- ubah untuk mendapatkan model yang terbaik. Parameter maxDepth bermaksud kedalaman maksimum pokok. Parameter minNum bermaksud variasi minimum untuk setiap pecahan yang terhasil daripada pokok. Manakala Parameter numFold pula ialah bilangan lipatan untuk backfitting dan akhir sekali ialah parameter seed, digunakan untuk menjana nombor secara rawak.

Bahagian ini menerangkan penilaian awal tentang pengelas pokok rawak. Nilai yang terbaik dipilih bagi kedalaman maksimum pokok (maxDepth) adalah 0, bilangan minimum (minNum) adalah 1.0, bilangan lipatan (numFolds) adalah 0 dan benih (Seed) adalah 1. Peratus pecahan yang digunakan dalam eksperimen ini adalah 60% untuk latihan dan 40% untuk ujian.

Rajah 5.6 dan Jadual 5.8 smenunjukkan prestasi pokok rawak mengikut peratus pecahan yang berbeza. Peratus pecahan yang digunakan adalah 50%, 55%, 60%, 65%, 70% dan 75%. Pokok rawak mencapai ketepatan tertinggi iaitu 95% pada peratus pecahan adalah 60% berbanding peratus pecahan yang lain.

Jadual 5.8 Prestasi antara pokok rawak berdasarkan peratus pecahan yang berbeza

Pengelas	Peratus Pecahan	Kadar Ketepatan (%)	Ralat Min Kuasa Dua
Pokok Rawak	50	92.00	0.2309
	55	91.30	0.2408
	60	95.00	0.1826
	65	94.44	0.1925
	70	53.33	0.4717
	75	92.30	0.2265



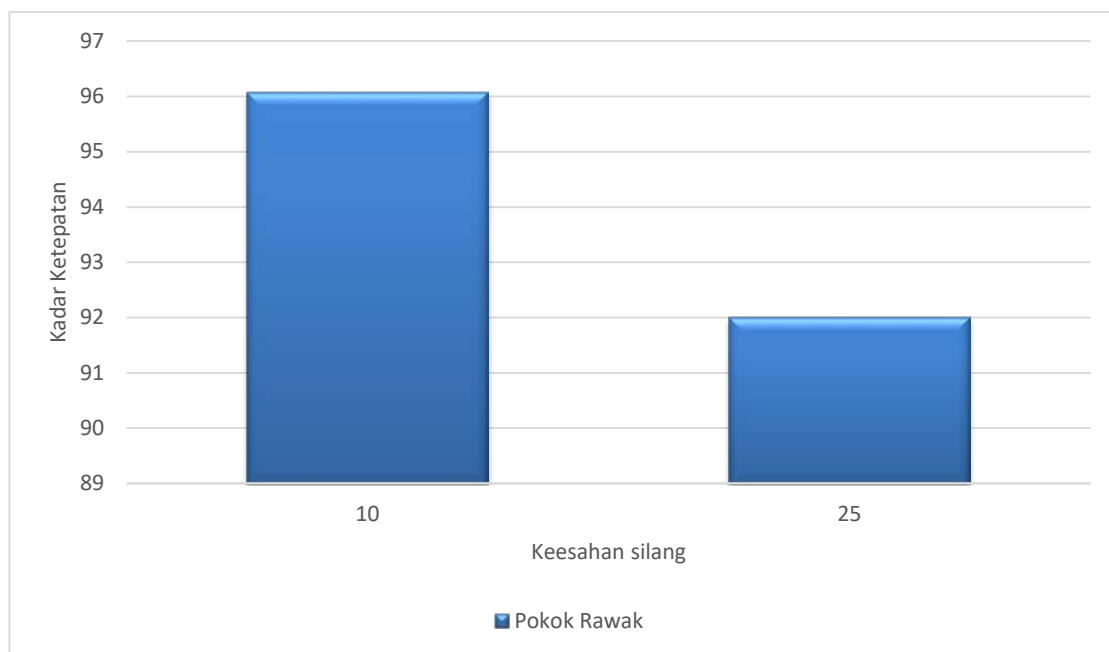
Rajah 5.6 Prestasi antara pokok rawak berdasarkan peratus pecahan yang berbeza

Rajah 5.7 dan Jadual 5.9 menunjukkan prestasi pokok rawak dan mengikut keesahan silang yang berbeza. Keesahan silang yang digunakan adalah 10 lipatan dan 25 lipatan. Pokok rawak memperoleh ketepatan tertinggi iaitu 96% pada keputusan keesahan silang adalah 25 lipatan.

Jadual 5.9 Prestasi pokok rawak berdasarkan keesahan silang yang berbeza

Pengelas	Kaedah Keesahan Silang	Kadar Ketepatan	Ralat Min Kuasa Dua
Pokok Rawak	10	90.19	0.1617
	25	96.07	0.2393

Secara keseluruhannya, berdasarkan daripada keputusan yang diperoleh, Rajah 5.7 mendapat peratusan ketepatan yang paling tinggi iaitu 96.0784% apabila nilai parameter *maxDepth* ialah 0, *minNum* ialah 1.0, *numFold* ialah 0 dan *seed* ialah 1. Dapat disimpulkan bahawa Pokok Rawak menunjukkan prestasi yang baik berbanding pokok keputusan yang lain.



Rajah 5.7 Prestasi antara Pokok Rawak berdasarkan keesahan silang yang berbeza.

5.4.2 Pengujian Anova

Statistik deskriptif digunakan untuk mencirikan semua pengelasan. Sebelum analisis statistik, data berangka log-berubah untuk normalisasi. Ujian trafik dua arah (ANOVA) digunakan untuk menganalisis perbezaan dalam data (Hutan Rawak, Pokok Rawak, dan J48) di antara pengelasan yang berlainan. Berikut merupakan keputusan yang diperolehi dalam pengujian ANOVA.

Jadual 5.10 Keputusan pengujian ANOVA pokok rawak berdasarkan peratus pecahan yang berbeza

RINGKASAN	Kiraan	Jumlah	Purata	Varians
50	3	276.00	92.00	0.0000
55	3	273.91	91.30	0.0000
60	3	285.00	95.00	0.0000
65	3	283.33	94.44	3.0333
70	3	239.99	79.99	533.3333
75	3	276.92	92.30	0.0000
Pokok Rawak	6	518.38	86.398	264.4784
J48	6	558.38	93.064	2.1053
Hutan Rawak	6	558.38	93.064	2.1053

Jadual 5.11 Keputusan pengujian ANOVA pokok rawak berdasarkan keesahan silang yang berbeza

ANOVA						
Sumber Variasi	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Baris	454.55	5	90.911	1.022	0.45	3.32
Lajur	177.77	2	88.88	1	0.40	4.10
Ralat	888.88	10	88.88			
Jumlah	1521.22	17				

Hasil daripada keputusan mendapati *h null* tidak terdapat perbezaan antara setiap kaedah klasifikasi dan nilai *p* 0.401878 adalah lebih tinggi daripada 0.05 maka kami menerima *h null* tidak ada beza dengan setiap kaedah pengelasan.

5.5 PENGUJIAN MODEL

Pengujian model dijalankan untuk menilai pencapaian model yang telah dihasilkan semasa proses latihan data dan juga untuk mendapatkan ketepatan pengelasan. Pengujian data bertujuan untuk menguji sama ada model yang dihasilkan berpotensi untuk menganggar ketepatan pengelasan data. Sebanyak 51 set data yang diambil dari laporan pesakit melalui pengetahuan pakar diuji untuk mendapatkan model yang terbaik daripada ketiga-tiga teknik yang dipilih iaitu pokok keputusan J48, pokok rawak dan hutan rawak.

Bagi teknik pokok keputusan J48, terdapat beberapa parameter yang akan berubah bagi menyiasat keberkesanannya ke atas data diagnosis kanser prostat iaitu bilangan minimum objek, bilangan lipatan dan benih pokok digunakan untuk merawak data. Keputusan yang diperoleh daripada jadual divisualkan ke dalam bentuk graf untuk melihat hubungan antara parameter yang digunakan terhadap perubahan peratusan ketepatan bagi setiap parameter yang berubah.

Struktur pengelas Hutan Rawak mengandungi komponen seperti nod akar, nod terminal dan daun. Ketepatan model kebanyakannya sangat bergantung pada pengawalan model seperti yang digunakan oleh parameter kedalaman maksimum pokok, bilangan slot perlakuan, bilangan ciri, bilangan pokok dan benih bagi merawak

data. Oleh itu, pemilihan parameter yang optimum adalah sangat penting untuk meningkatkan kadar ketepatan.

Bagi pokok rawak, terdapat beberapa parameter yang akan berubah bagi menyiasat keberkesanannya ke atas data diagnosis kanser prostat iaitu kedalaman maksimum, bilangan minimum, bilangan lipatan dan benih bagi merawak data. Hasil keputusan yang diperoleh daripada jadual digambarkan ke dalam bentuk graf untuk melihat hubungan antara parameter-parameter yang digunakan terhadap perubahan peratusan ketepatan model bagi pokok rawak.

Jadual 5.12 menunjukkan keputusan ujikaji set pengujian data bagi model-model yang telah dipilih semasa proses latihan data bagi ketiga-tiga teknik pembelajaran mesin.

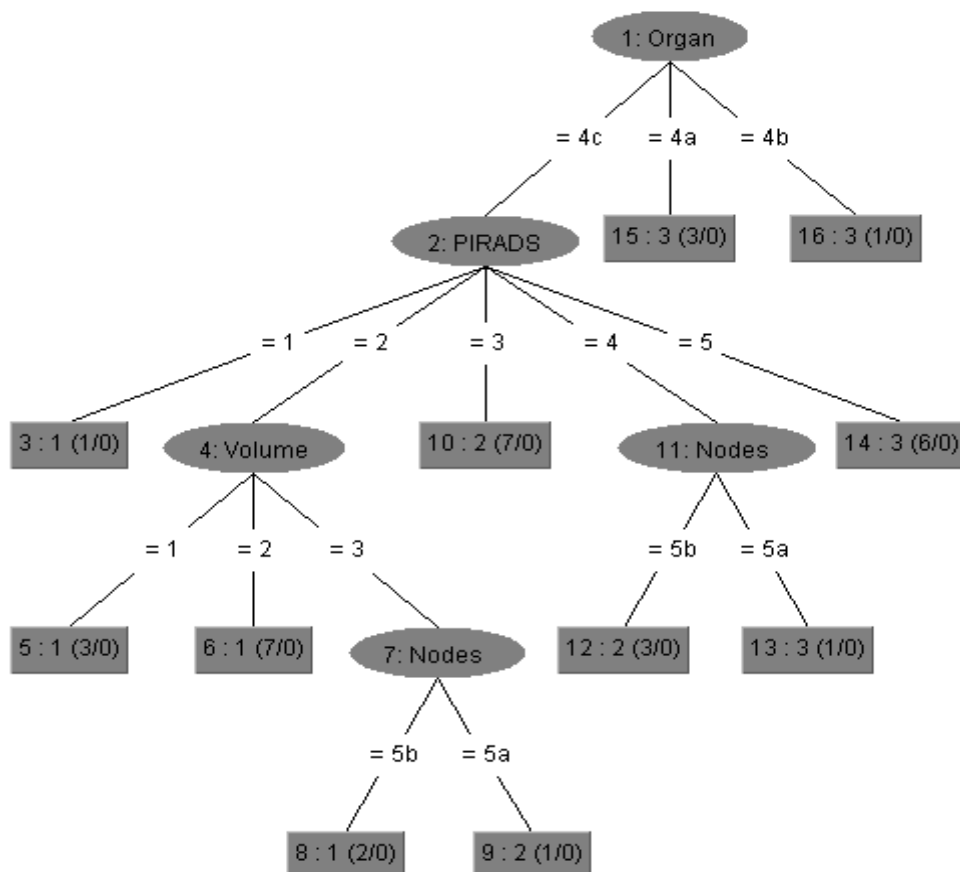
Jadual 5.12 Keputusan pengujian set data ketepatan untuk teknik pembelajaran mesin pokok keputusan J48, pokok rawak dan hutan rawak

Teknik Pembelajaran Mesin	Ketepatan (%)	Ralat Min Kuasa Dua
J48	95.00	0.1826
Pokok Rawak	96.07	0.1642
Hutan Rawak	95.00	0.1897

Hasil keputusan mendapati pokok keputusan pokok rawak mendapat peratusan ketepatan ramalan yang paling tinggi iaitu 96%. Jika dilihat nilai ralat min kuasa dua bagi teknik pokok rawak juga lebih rendah jika dibandingkan dengan teknik J48 dan hutan rawak rujuk Lampiran B dan Lampiran C. Secara keseluruhannya, teknik pembelajaran mesin pokok rawak menghasilkan model yang lebih baik daripada pengelas J48 dan pokok rawak dengan kadar ketepatan sebanyak 96% dan jika dilihat pada nilai ralat min kuasa dua, pokok rawak mendapat nilai yang lebih rendah berbanding model yang lain selepas pengujian dijalankan.

5.6 HASIL KEPUTUSAN PETUA BAGI MODEL POKOK RAWAK

Teknik pembelajaran mesin pokok rawak merupakan teknik pengelasan yang baik jika dibandingkan dengan teknik pokok keputusan yang lain. Tambahan pula, teknik tersebut telah dilatih dengan menggunakan pelbagai parameter yang berbeza untuk melihat keupayaan teknik tersebut dalam meramal dan membuat pengelasan pada data. Berdasarkan pengesetan eksperimen, telah dinyatakan parameter-parameter yang memberikan kesan dalam pembinaan model dijalankan. Keputusan yang telah dijana juga boleh dihuraikan dalam bentuk pokok keputusan ataupun petua. Pokok rawak yang dihasilkan daripada pengesetan eksperimen boleh dilihat pada Rajah 5.8.



Rajah 5.8 Model pokok rawak

Berdasarkan Rajah 5.8 dan Rajah 5.1, petua-petua yang boleh dihasilkan daripada pokok keputusan tersebut ialah:

1. JIKA $organs_prostate = false$ DAN penilaian $PIRADS = 1$, MAKA *followup*
2. JIKA $organs_prostate = false$ DAN penilaian $PIRADS = 2$ DAN $Volume \leq 30$ Atau $organs_prostate = false$ DAN penilaian $PIRADS = 2$ DAN $Volume \geq 31$ DAN $Volume \leq 80$, MAKA *followup*
3. JIKA $organs_prostate = false$ DAN penilaian $PIRADS = 2$ DAN $Volume \geq 81$ DAN $nodes_prostate = false$, MAKA *followup*
4. JIKA $organs_prostate = false$ DAN penilaian $PIRADS = 2$ DAN $Volume \geq 81$ DAN $nodes_prostate = true$, Maka *biopsy*
5. JIKA $organs_prostate = false$ DAN $PIRADS = 3$, MAKA *biopsy*
6. JIKA $organs_prostate = false$ DAN penilaian $PIRADS = 4$ Dan $nodes_prostate = false$, MAKA *biopsy*
7. JIKA $organs_prostate = false$ DAN penilaian $PIRADS = 4$ Dan Nodes Terlibat $= true$, MAKA *treatment*
8. JIKA $organs_prostate = false$ DAN penilaian $PIRADS = 5$, MAKA *treatment*
9. Jika $organs_prostate = true$ ATAU $organs_prostate > multiple$, MAKA *treatment*

Petua RP.1	$: patient_id(x) \wedge \left(organs_prostate(x, z) \wedge pirads(x, a) \vee \right.$ $organs_prostate(x, z) \wedge pirads(x, a) \wedge volume(x, b) \vee$ $volume(x, d) \vee organs_prostate(x, z) \wedge pirads(x, a) \wedge$ $volume(x, d) \wedge nodes_prostate(x, z) \wedge$ $risktype(y) \left. \right) \exists x: x \in I, \forall y: \{Y = 1\}, \exists z: z \in$ $\{false\}, \exists a: a \in \{1, 2\}, \exists b: b \geq \{30, 80\}, \exists d: d \leq \{31, 81\}.$
Petua RP.2	$: patient_id(x) \wedge (organs_prostate(x, z) \wedge pirads(x, a) \vee$ $organs_prostate(x, z) \wedge pirads(x, c) \wedge volume(x, b) \wedge$ $nodes_prostate(x, z) \wedge organs_prostate(x, z) \wedge$

	$pirads(x, c) \wedge nodes_{prostate}(x, z) \wedge$ $result(y), \exists x: x \in \mathbb{I}, \forall y: \{Y = 2\}, \exists z: z \in$ $\{false, true\}, \exists c: c \{2,3,4\}, \exists b: b \geq \{81\}.$
Petua RP.3	$: patient_{id(x)} \wedge (organs_{prostate}(x, b) \vee organs_{prostate}(x, b) \wedge$ $pirads(x, c) \vee organs_{prostate}(x, b) \wedge pirads(x, c) \wedge$ $nodes_{prostate}(x, z)) \wedge result(y) \exists x: x \in \mathbb{I}, \forall y: \{Y =$ $3\} \exists z: z \in \{true, false, mutiple\}, \exists c: c \{4,5\}.$
Petua RPS.1	$: decision(x, y, z) \leftarrow patientrisk(x, y) \wedge treatmenttype(z)$ $treatmenttype(z)$ $= \begin{cases} followup & \text{if } y = low, \\ biopsy & \text{elseif } y = moderate, \\ treatment & \text{elseif } y = high, \end{cases}$ <p style="text-align: center;">where x is patient id, $\exists x: x \in \mathbb{I}$, y is type of risk, z is treatment type</p>

Rajah 5.9 Perwakilan pengetahuan secara secara predikat logik tertib kedua

Antara kelebihan menggunakan pokok rawak dalam kajian ini ialah keputusan yang diperolehi mudah difahami dan bagaimana tahap kanser prostat dikelaskan boleh ditafsirkan berbanding teknik- teknik yang lain.

5.7 RUMUSAN

Secara ringkasnya, Bab V menerangkan proses yang terlibat dalam mengenalpasti kanser prosat dan hasil daripada proses pengkelasan set data. Rajah ditunjukkan untuk melihat keseluruhan hasil process yang terlibat. Pemilihan parameter yang sesuai juga menyumbang kepada penghasilan model yang baik. Faktor-faktor ini yang akan menentukan keupayaan untuk pengelasan data bagi teknik pembelajaran mesin yang digunakan Di samping itu, melalui kesemua eksperimen pokok keputusan dapat disimpulkan bahawa pokok rawak mencapai prestasi yang baik pada peratus pecahan adalah 96% dan keesahan silang adalah 25 lipatan. Hasil ujian menunjukkan pengguna berpuas hati dengan petua yang dihasilkan dan maklum balas turut disumbang untuk penambahbaikan sistem pada masa hadapan.

BAB VI

RUMUSAN DAN CADANGAN

6.1 RINGKASAN KAJIAN

Di akhir tempoh pembangunan projek ini, semua objektif kajian yang telah dinyatakan dalam Bab I akan berjaya dicapai. Secara ringkasnya:

Objektif utama dalam eksperimen ini adalah untuk membuat penilaian bagi perbandingan petua pengeluaran dan pokok rawak. Perbandingan petua pengeluaran dan pokok rawak ini dilakukan dengan mengambil nilai ketepatan di mana petua pokok rawak mendapat peratusan yang tinggi sebanyak 96% berbanding petua pengeluaran sebanyak 76.4%. Dalam kajian ini, teknik pokok keputusan dipilih dalam mengkaji tahap kanser prostat pesakit menggunakan petua yang dikumpul kerana bersesuaian dengan set data sedia ada. Hal ini kerana data sedia ada mempunyai banyak data yang hilang dan atribut set data bertaip nominal dan numerik.

a. Sistem Pakar Dapat Membantu Dalam Membuat Cadangan Pandangan Kedua

Perwakilan pengetahuan dalam bentuk petua akan dilaksanakan. Seterusnya, membangunkan satu sistem pakar yang berupaya mencadangkan rawatan susulan kepada pesakit dengan menembak petua yang tepat. Selain itu, sistem pakar ini juga mampu merekod komen dari pengamal perubatan sekiranya cadangan rawatan diubah oleh pakar.

b. Sistem Pakar Dapat Menyediakan Informasi Yang Sangat Terperinci Mengenai Simpton Dapat Membantu Mendiagnosis Dan Menguruskan Penyakit Dengan Lebih Pantas

Dengan sistem pakar diagnostik kanser prostat ini akan dapat membantu para doktor dalam mengesan kanser berdasarkan tahap peringkat dengan cepat dan tepat. Tetapi, bukanlah sebagai pengganti kepada pakar perubatan tetapi ia dapat memudahkan pekerjaan mereka.

6.2 RUMUSAN KAJIAN

Seperti yang dinyatakan dalam bab pertama, objektif kajian ini adalah untuk membangunkan sebuah sistem pakar erasaskan petua pengeluaran untuk mendiagnosis pesakit kanser prostat di Malaysia di samping menjana sebuah model pokok keputusan yang boleh ditafsirkan. Justeru, sistem ini dapat memberikan simpton awal kanser prostat yang dapat membantu memberikan maklumat awal sebagai panduan dan langkah berjaga – jaga untuk mendapatkan rawatan.

Selepas pengujian dan penilaian sistem dilakukan, terdapat dua baris data menghasilkan keputusan yang sebaliknya apabila dibandingkan dengan data yang sebenar. Ini disebabkan oleh tiada petua yang dihasilkan untuk dipadankan dengan kes tersebut. Hal ini akan dapat diatasi sekiranya terdapat data yang mencukupi yang boleh mewakili kesemua kes. Sekiranya data – data yang didapati mencukupi, maka petua yang dijana juga kan pelbagai dan lebih banyak kes peramalan dapat diselesaikan.

Kelebihan kajian dalam meramal tahap penyakit kanser prostat ini ialah teknik pembelajaran mesin pokok rawak adalah teknik pengelasan yang terbaik jika dibandingkan dengan teknik yang lain. Tambahan pula, teknik tersebut telah dilatih dengan menggunakan pelbagai parameter yang berbeza untuk melihat keupayaan teknik tersebut dalam meramal dan membuat pengelasan pada data.

6.3 KELEBIHAN DAN KEKURANGAN KAJIAN

Terdapat beberapa kelebihan dan kekurangan yang dikenal pasti hasil proses sistem pakar. Berikut merupakan kelebihan dan kekurangan dalam sistem ini.

6.3.1 Kelebihan

- 1) Dapat membantu pakar mendiagnosis pesakit dengan pantas.
Hasil daripada pembangunan sistem ini pakar dapat mengetahui masalah pesakit prostat dengan lebih pantas.
- 2) Sistem yang ringkas
Sistem ini menggunakan keupayaan yang minimum daripada pakar dan dapat menghasilkan output yang mempunyai ketepatan sehingga 96%.

6.3.2 Kekurangan

- 1) Kekurangan maklumat untuk pembelajaran mesin.
Maklumat dan petua yang dihasilkan adalah tidak mencukupi untuk mendapatkan keputusan yang tepat.

6.4 PERSAMAAN SISTEM

Secara keseluruhan, sistem prostat ini dapat membantu pakar dalam memperoleh pandangan kedua. Secara tidak langsung mempercepatkan proses diagnosis kanser prostat. Tidak dinafikan sistem yang dibangunkan ini telah wujud dan diguna pakai, tapi sistem ini telah dinaiktaraf seiring dengan perkembangan teknologi maklumat masa kini dan pada masa hadapan. Antara teknik sistem yang dipakai dan diperbaharui adalah pembelajaran mesin atau kajian pengecaman pola yang digunakan untuk belajar dan membuat petua daripada data yang disediakan. Sistem diagnosis kanser prostat boleh memanfaatkan teknik-teknik ini dalam enjin tadbir sistem. Salah satu contoh yang mengaplikasikan konsep ini ialah Watson for Oncology yang mencadangkan rawatan berdasarkan bukti dengan memanfaatkan pangkalan rekod perubatannya yang luas (Watson 2015).

6.5 SUMBANGAN

Sumbangan yang diberikan hasil daripada kajian ini ialah sebuah sistem pakar berasaskan petua pengeluaran dan pokok keputusan yang dijana bagi mendiagnosis kanser prostat dengan lebih pantas. Antara sumbangan lain bagi kajian ini adalah :

- I. Sistem pakar dalam kajian ini dapat meramal keputusan tentang tahap risiko pada pesakit tertentu dengan secara ringkas, jelas dan tepat.
- II. Pembangunan sistem pakar dalam kajian ini dengan menggunakan kaedah pengelasan pokok keputusan telah menjanakan petua dan dianalisis kepada pengetahuan yang boleh ditafsir dan difahami oleh manusia.

6.6 CADANGAN MASA HADAPAN

Cadangan penambahbaikan ini adalah untuk digunakan pada masa akan datang bagi mengatasi kekurangan yang terdapat pada sistem yang dibangunkan dan mempertingkatkan lagi fungsi yang sedia ada. Antara cadangan yang dikemukakan ialah seperti berikut:

- 1) Memperbanyakkan petua dalam pembangunan sistem.
Petua yang lebih besar membantu untuk sesuatu sistem pakar itu berfungsi dengan lebih tepat dan baik. Justeru, jumlah petua boleh ditingkatkan untuk penghasilan sistem yang lebih dinamik.
- 2) Menukarkan sistem pakar ini kepada aplikasi mudah alih.
Sistem yang telah dibangunkan hanya terhad digunakan didalam komputer. Penambahbaikan sistem ini ialah dengan menukarkan sistem pakar kepada aplikasi mudah alih yang dapat digunakan oleh pakar terutamanya pelajar Pusat Perubatan Universiti Kebangsaan Malaysia (PPUKM).

6.7 RUMUSAN

Secara keseluruhannya, kelebihan dan kekurangan yang terdapat pada sistem yang dibangunkan, dijadikan sebagai panduan dalam cadangan penambahbaikan pada sistem tersebut supaya menjadikan sistem ini lebih efisien. Penyelenggaraan juga perlu dilakukan dari semasa ke semasa bagi mengatasi ralat pada sistem ini. Oleh itu, jelaslah bahawa teknik pokok keputusan yang diaplikasikan mampu untuk menghasilkan sebuah model yang terbaik berbanding petua pengeluaran dan seterusnya membangunkan sebuah sistem pakar yang berguna. Secara keseluruhannya, kajian ini telah berjaya mencapai objektif yang dinyatakan mengikut skop kajian yang ditetapkan. Justeru, daripada hasil pengujian hasil yang dijalankan, teknik pokok keputusan pokok rawak merupakan teknik yang paling baik dalam mendiagnosis kanser prostat dengan kadar ketepatan keputusan sebanyak 96% berbanding 76.4% menggunakan petua pengeluaran secara purata.

RUJUKAN

- Adam, A., Bulpitt, A. and Treanor, D. (2012). Grading Dysplasia in Barrett's Oesophagus Virtual Pathology Slides with Cluster Co-occurrence Matrices, In Proc. of Histopathology Image Analysis: Image Computing in Digital Pathology in conjunction with The 15th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI).
- AXA 200 (2016). Prostate Cancer: Fourth Most Common Cancer Among Malaysian Men, CancerCare 2016.
- Abdullah, A., Veltkamp, R. C. & Wiering, M. A. 2010. Fixed partitioning and salient points with MPEG-7 cluster correlograms for image categorization. Pattern Recognition 43(3): 650-662.
- Adam, A. 2015. Computer Aided Dysplasia Grading for Barrett's Oesophagus Virtual Slides. Tesis University of Leeds.
- Albashish, D., Sahran, S., Abdullah, A., Adam, A., Shukor, N. A. & Pauzi, S. H. M. 2015. Multi-scoring feature selection method based on SVM-RFE for prostate cancer diagnosis. Electrical Engineering and Informatics (ICEEI), 2015 International Conference on, hlm. 682-686.
- Alomari, A. K., Cohen, J., Vortmeyer, A. O., Chiang, A., Gettinger, S., Goldberg, S., Kluger, H. M. & Chiang, V. L. 2016. Possible interaction of anti-PD-1 therapy with the effects of radiosurgery on brain metastases. Cancer immunology research 4(6): 481-487.
- Alomoush, W., Abdullah, S. N. H. S., Sahran, S. & Hussain, R. I. 2014. MRI Brain Segmentation Via Hybrid Firefly Search Algorithm. Journal of Theoretical & Applied Information Technology 61(1).
- Alomoush, W. K., Abdullah, S. N. H. S., Sahran, S. & Hussain, R. I. 2014. Segmentation of MRI Brain Images Using FCM Improved by Firefly Algorithms. Journal of Applied Sciences 14: 66-71.
- Amit, Y. & Geman, D. 1997. Shape Quantization and Recognition with Randomized Trees. Neural computation 9(7): 1545-1588.
- Arimura, H., Magome, T., Yamashita, Y. & Yamamoto, D. 2009. Computer-Aided Diagnosis Systems for Brain Diseases in Magnetic Resonance Images. Algorithms 2(3): 925-952.
- Baharuddin, W. N. A., Abdullah, S. N. H. S., Sahran, S., Qasem, A., bin Abdullah, A., Iqbal, R. & Ismail, F. 2016. Type 2 Fuzzy Logic for mammogram breast tissue classification. Industrial Informatics and Computer Systems (CIICS), 2016 International Conference on, hlm. 1-6.

- Bernama, "Grow moustache for prostate and testicular cancer awareness," 19 August 2015. [Online]. Available: <http://makna.org.my/2015/08/grow-moustache-prostate-testicular-cancer-awareness/>.
- Bauer, S., Fejes, T., Slotboom, J., Wiest, R., Nolte, L.-P. & Reyes, M. 2012. Segmentation of brain tumor images based on integrated hierarchical classification and regularization. MICCAI BraTS Workshop. Nice: Miccai Society, hlm.
- Bezdek, J. C., Keller, J., Krisnapuram, R. & Pal, N. 2006. Fuzzy models and algorithms for pattern recognition and image processing. Springer Science & Business Media.
- Breiman, L. 2001. Random forests. *Mach Learn*45: 5-32.
- Castanho, M., Hernandez, F., De Ré, A. M., Rautenberg, S. & Billis, A. 2013. Fuzzy Expert System for Predicting Pathological Stage of Prostate Cancer. *Expert Systems with Applications* 40(2): 466-470.
- Chen, A. H. & Lin, C.-H. 2011. A Novel Support Vector Sampling Technique to Improve Classification Accuracy and to Identify Key Genes of Leukaemia and Prostate Cancers. *Expert Systems with Applications* 38(4): 3209-3219.
- Campbell, D. (2014). Almost half of cancer patients diagnosed too late, health correspondent, *TheGuardian*.
- Cancer.Net. 2013. Types of Oncologists. <http://www.cancer.net/navigating-cancer-care/cancer-basics/cancer-care-team/types-oncologists> [13 April 2015].
- Çınar, M., Engin, M., Engin, E. Z. & Ateşçi, Y. Z. 2009. Early prostate cancer diagnosis by using artificial neural networks and support vector machines. *Expert Systems with Applications* 36(3): 6357-6361.
- Canceraustralia.gov.au. 2015. Calculate your risk. <https://prostate-cancer.canceraustralia.gov.au/risk-factors> [18 Desember 2015].
- Comaniciu, D. 2010. Robust ultrasound image analysis using learning. *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, hlm. 280-280.
- Creswell, J. W. & Designm, R. 2003. Qualitative, quantitative, and mixed method approaches. Aufl. Thousand Oaks:
- Delahunt, B., Egevad, L., Grignon, D. J., Srigley, J. R. & Samaratunga, H. 2016. Prostate cancer grading: recent developments and future directions. *BJU international* 117: 7-8.
- DeMarco, T., Hatley, D., Pirbhai, I. & Petri, C. 18] Yourdon, E., and Edward, C. Structured Design. Prentice Hall, Eaglewood.

- DiFranco, M. D., O'Hurley, G., Kay, E. W., Watson, R. W. G. & Cunningham, P. 2011. Ensemble Based System for Whole-Slide Prostate Cancer Probability Mapping Using Color Texture Features. *Computerized medical imaging and graphics* 35(7): 629-645.
- Durkin, J. & Durkin, J. 1998. *Expert systems: design and development*. Prentice Hall PTR.
- Eble, J. 2004. *Classification of Tumours: Pathology and Genetics of Tumours of The Urinary System And Male Genital Organs*. World Health Organization Classification of Tumours: 255-257.
- Ellsworth, P. 2012. *100 Questions and Answers about Prostate Cancer*. Jones & Bartlett Publishers.
- Ellsworth, P. 2013. *The Evolving Role of Current and Emerging Therapies in the Management of Castrate Resistant Prostate Cancer*. *Educating Medical Directors of Employers, Health Plans and Provider Systems*: 41.
- Eng, J. 2003. Sample Size Estimation: How Many Individuals Should Be Studied? *Radiology* 227(2): 309-313.
- Fausto, A., Rizzatto, G., Preziosa, A., Gaburro, L., Washburn, M. J., Rubello, D. & Volterrani, L. 2012. A New Method to Combine Contrast-Enhanced Magnetic Resonance Imaging During Live Ultrasound of The Breast Using Volume Navigation Technique: A Study For Evaluating Feasibility, Accuracy And Reproducibility In Healthy Volunteers. *European Journal of Radiology* 81(3): e332-e337.
- Fairul Asmaini Mohd Pilus (2015) *Awas Kanser Prostat, Harian Metro*.
- Fang, R., Pouyanfar, S., Yang, Y., Chen, S.-C. & Iyengar, S. 2016. Computational health informatics in the big data age: A survey. *ACM Computing Surveys (CSUR)* 49(1): 12.
- Fayyad, U. M. 1997. Editorial. *Data mining and Knowledge discovery* 1(1): 5-10.
- Froelich, W., Papageorgiou, E. I., Samarinas, M. & Skriapas, K. 2012. Application of Evolutionary Fuzzy Cognitive Maps to The Long-Term Prediction of Prostate Cancer. *Applied soft computing* 12(12): 3810-3817.
- Gallo, L., Minutolo, A. & De Pietro, G. 2010. A User Interface for VR-Ready 3D Medical Imaging by Off-The-Shelf Input Devices. *Computers in Biology and Medicine* 40(3): 350-358.
- Garcia-Urbe, A., Kehtarnavaz, N., Marquez, G., Prieto, V., Duvic, M. & Wang, L. V. 2004. Skin Cancer Detection by Spectroscopic Oblique-Incidence Reflectometry: Classification and Physiological Origins. *Applied Optics* 43(13): 2643-2650.
- Gane, C. P. & Sarson, T. 1979. *Structured Systems Analysis: Tools and Techniques*.

- Gómez-Ferrer, Á. & Arlandis, S. 2011. Computer-Aided Analysis of Transrectal Ultrasound Images of The Prostate. *Actas Urológicas Españolas (English Edition)* 35(7): 404-413.
- Gutierrez-Osuna, R. 2002. Pattern Analysis for Machine Olfaction: A Review. *IEEE Sensors journal* 2(3): 189-202.
- Haider, S., Prabuwno, A. S. & Siti, N. H. S. A. 2012. Metal Parts Visual Inspection Based On Production Rules. *Applied Mechanics and Materials*, hlm. 4091-4095.
- Han, J. & Kamber, M. 2000. *Data Mining: Concepts and Techniques*. Series Editor Morgan Kaufmann Publishers, The Morgan Kaufmann Series in Data Management Systems, Jim Gray.
- Han, J., Pei, J. & Kamber, M. 2011. *Data mining: concepts and techniques*. Elsevier.
- Hand, D. J., Mannila, H. & Smyth, P. 2001. *Principles of data mining*. MIT press.
- Haykin, S. & Network, N. 2004. *A Comprehensive Foundation. Neural Networks* 2(2004): 41.
- Ho, T. K. 1995. Random Decision Forests. *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, hlm. 278-282.
- Hoffelt, S. C., Garzotto, M., Hung, A., Holland, J. & Beer, T. 2003. A Comparison of CT Scan to Transrectal Ultrasound Measured Prostate Volume in Untreated Prostate Cancer: P013. *The Cancer Journal*, hlm. 504.
- Hoffelt, S. C., Marshall, L. M., Garzotto, M., Hung, A., Holland, J. & Beer, T. M. 2003. A Comparison of CT Scan to Transrectal Ultrasound-Measured Prostate Volume in Untreated Prostate Cancer. *International Journal of Radiation Oncology* Biology* Physics* 57(1): 29-32.
- Hor, C.-L., Crossley, P. A. & Watson, S. J. 2007. Building Knowledge for Substation-Based Decision Support Using Rough Sets. *IEEE Transactions on Power Delivery* 22(3): 1372-1379.
- Horning, N. 2010. *Random Forests: An Algorithm for Image Classification and Generation of Continuous Fields Data Sets*. New York:
- Jordan, M. I. & Jacobs, R. A. 1994. Hierarchical mixtures of experts and the EM algorithm. *Neural computation* 6(2): 181-214.
- Kaidar, S. M., Hussain, R. I., Bohani, F. A., Sahran, S., binti Zainuddin, N., Ismail, F., Thanabalan, J., Kalimuthu, G. & Abdullah, S. N. H. S. 2013. Brain tumor treatment advisory system. *Dlm. (pnyt.)*. *Soft Computing Applications and Intelligent Systems*, hlm. 78-88. Springer.
- Karaboga, D. & Ozturk, C. 2011. A Novel Clustering Approach: Artificial Bee Colony (ABC) algorithm. *Applied soft computing* 11(1): 652-657.

- Kim, S. B., Temiyasathit, C., Bensalah, K., Tuncel, A., Cadeddu, J., Kabbani, W., Mathker, A. V. & Liu, H. 2010. An Effective Classification Procedure for Diagnosis of Prostate Cancer in Near Infrared Spectra. *Expert Systems with Applications* 37(5): 3863-3869.
- Kim, S. J., Kim, S. I., Ahn, H. S., Choi, J. B., Kim, Y. S. & Kim, S. J. 2010. Risk Factors for Acute Prostatitis After Transrectal Biopsy of The Prostate. *Korean Journal of Urology* 51(6): 426-430.
- Lanzi, P. L. & Loiacono, D. 2007. Classifier Systems That Compute Action Mappings. *Proceedings of The 9th Annual Conference On Genetic and Evolutionary Computation*:1822-1829.
- Liaw, A. 2013. Package randomForest: Breiman and Cutler's Random Forest for Classification and Regression. CRAN Repository. <http://cran.r-project.org/web/packages/randomForest/randomForest.pdf>:
- Lippmann, R. P. 1989. Pattern Classification Using Neural Networks. *IEEE Communications Magazine* 27(11): 47-50.
- Ma, Z., Tavares, J. M. R., Jorge, R. N. & Mascarenhas, T. 2010. A Review of Algorithms for Medical Image Segmentation and Their Applications to The Female Pelvic Cavity. *Computer Methods in Biomechanics and Biomedical Engineering* 13(2): 235-246.
- Maloof, M. A. & Michalski, R. S. 2000. Selecting Examples for Partial Memory Learning. *Machine Learning* 41(1): 27-52.
- Mannila, H. 1996. Data Mining: Machine Learning, Statistics, and Databases. *Scientific and Statistical Database Systems, 1996. Proceedings., Eighth International Conference on*: 2-9.
- Uzair, Abdullah, S. N. H. S., Khairuddin, Afzan, Syazarina 2015. Prostate Cancer Classification Using Logistic Regression and Random Forest On Pelvis Ct Images, 5th International Technical Conference 2016 (ITC2016) Proceeding.
- Nayef, B. H., Sahran, S., Hussain, R. I. & Abdullah, S. N. H. S. 2013. Brain Imaging Classification Based On Learning Vector Quantization. *Communications, Signal Processing, and Their Applications (ICCSPA), 2013 1st International Conference on*:1-6.
- Nguyen, K., Jain, A. K. & Sabata, B. 2011. Prostate Cancer Detection: Fusion of Cytological and Textural Features. *Journal of Pathology Informatics* 2:
- Omar, Z. A. & Tamin, N. S. I. 2011. National Cancer Registry Report 2007. Malaysia Cancer Statistics-Data and Figure, Kuala Lumpur: National Cancer Registry, Ministry of Health Malaysia 89:
- Ou, Y., Shen, D., Zeng, J., Sun, L., Moul, J. & Davatzikos, C. 2009. Sampling The Spatial Patterns of Cancer: Optimized Biopsy Procedures for Estimating

- Prostate Cancer Volume and Gleason Score. *Medical Image Analysis* 13(4): 609-620.
- Palouš, J. 2002. *Machine Learning and Data Mining*. Prague: Gerstner Laboratory for Intelligent Decision Making and Control Czech Technical University. URL: [Http://Citeseer. Nj. Nec. Com/506615. Html](http://Citeseer.Nj.Nec.Com/506615.Html) (accessed on 26/02/2009):
- Parfait, S., Walker, P. M., Créhange, G., Tizon, X. & Miteran, J. 2012. Classification of Prostate Magnetic Resonance Spectra Using Support Vector Machine. *Biomedical Signal Processing and Control* 7(5): 499-508.
- Park, Y. 1994. A comparison of neural net classifiers and linear tree classifiers: Their similarities and differences. *Pattern Recognition* 27(11): 1493-1503.
- Partin, A. W., Mangold, L. A., Lamm, D. M., Walsh, P. C., Epstein, J. I. & Pearson, J. D. 2001. Contemporary Update of Prostate Cancer Staging Nomograms (Partin Tables) For The New Millennium. *Urology* 58(6): 843-848.
- Penzkofer, T. & Tempany - Afdhal, C. M. 2014. Prostate cancer detection and diagnosis: the role of MR and its comparison with other diagnostic modalities—a radiologist's perspective. *NMR in Biomedicine* 27(1): 3-15.
- Pfahring, B. 2010. *Random Model Trees: An Effective and Scalable Regression Method*.
- Pillay, B., Wootten, A. C., Crowe, H., Corcoran, N., Tran, B., Bowden, P., Crowe, J. & Costello, A. J. 2016. The impact of multidisciplinary team meetings on patient assessment, management and outcomes in oncology settings: a systematic review of the literature. *Cancer treatment reviews* 42: 56-72.
- Provost, F. & Kohavi, R. 1998. Guest Editors' Introduction: On Applied Research in Machine Learning. *Machine Learning* 30(2): 127-132.
- Quinlan, J. R. 1993. *Constructing Decision Tree*. C4 5: 17-26.
- Razi, M. A. & Athappilly, K. 2005. A Comparative Predictive Analysis of Neural Networks (Nns), Nonlinear Regression and Classification and Regression Tree (CART) Models. *Expert Systems with Applications* 29(1): 65-74.
- Samuel, A. L. 2000. *Some Studies in Machine Learning Using the Game Of Checkers*. *IBM Journal of Research and Development* 44(1.2): 206-226.
- Saritas, I., Ozkan, I. A. & Sert, I. U. 2010. Prognosis of Prostate Cancer by Artificial Neural Networks. *Expert Systems with Applications* 37(9): 6646-6650.
- Schuermann, J. & Doster, W. 1984. A decision theoretic approach to hierarchical classifier design. *Pattern Recognition* 17(3): 359-369.
- Sheikh Abdullah, S. N. H., Bohani, F. A., Nayef, B. H., Sahran, S., Al Akash, O., Iqbal Hussain, R. & Ismail, F. 2016. *Round Randomized Learning Vector*

Quantization for Brain Tumor Imaging. Computational and mathematical methods in medicine 2016:

- Smith, W. P., Doctor, J., Meyer, J., Kalet, I. J. & Phillips, M. H. 2009. A Decision Aid for Intensity-Modulated Radiation-Therapy Plan Selection in Prostate Cancer Based On A Prognostic Bayesian Network And A Markov Model. *Artificial Intelligence in Medicine* 46(2): 119-130.
- Snow, P. B., Smith, D. S. & Catalona, W. J. 1994. Artificial Neural Networks in The Diagnosis and Prognosis of Prostate Cancer: A Pilot Study. *The Journal of Urology* 152(5 Pt 2): 1923-1926.
- Srimani, P. & Koti, M. S. 2011. A Comparison of Different Learning Models Used in Data Mining for Medical Data. *AIP Conference Proceedings*: 51-55.
- Ubaidillah, S., Sallehuddin, R. & Ali, N. 2013. Cancer Detection Using Artificial Neural Network and Support Vector Machine: A Comparative Study. *Jurnal Teknologi* 65(1): 73-81.
- Watson, P. A., Arora, V. K. & Sawyers, C. L. 2015. Emerging Mechanisms of Resistance to Androgen Receptor Inhibitors in Prostate Cancer. *Nature Reviews Cancer* 15(12): 701-711.
- Watson, I. 2015. Watson for Oncology. <http://www.ibm.com/smarterplanet/us/en/ibmwatson/watson-oncology.html> [15 April 2015].
- Ye, Y., Wu, Q., Huang, J. Z., Ng, M. K. & Li, X. 2013. Stratified Sampling for Feature Subspace Selection in Random Forests for High Dimensional Data. *Pattern Recognition* 46(3): 769-787.
- Zait, M. & Messatfa, H. 1997. A Comparative Study of Clustering Methods. *Future Generation Computer Systems* 13(2-3): 149-159.
- Zakana, M., Nasrudm, M., Omar, K., Sahran, S. & Nordin, M. 2011. The Development of Ball Control Techniques for Robot Soccer Based On Predefined Scenarios. *Journal of Applied Sciences* 11(1): 111-117.
- Zhang, G. P. 2000. Neural Networks for Classification: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 30(4): 451-462.

LAMPIRAN A
PEROLEHAN PENGETAHUAN

Perwakilan pengetahuan merangkumi penemuan daripada psikologi tentang bagaimana manusia menyelesaikan masalah dan mewakili pengetahuan untuk merancang *formalisms* yang akan menjadikan sistem yang kompleks lebih mudah untuk mereka bentuk dan dibina. Perwakilan pengetahuan dan pertimbangan juga merangkumi penemuan daripada logik untuk mengautomatiskan pelbagai jenis penaakulan, seperti penggunaan kaedah-kaedah atau hubungan yang set dan subsets.

Setiap jenis kepakaran mempunyai 1 jenis sub modul yang digunakan untuk penilaian. Penjelasan mengenai jenis-jenis sub modul petua dan logik perwakilan yang dibentangkan dalam Rajah 6.1(a)(b), Rajah 6.2(a)(b) dan Rajah 6.3(a)(b).

Petua U1	: JIKA <i>bladder_empty</i> = true, MAKA <i>low_risk</i> = <i>routinefollowup</i>
Petua U2	: JIKA <i>bladder_less</i> = true, MAKA <i>low_risk</i> = <i>routinefollowup</i>
Petua U3	: JIKA <i>bladder_delay</i> = true, MAKA <i>low_risk</i> = <i>routinefollowup</i>
Petua U4	: JIKA <i>bladder_weak</i> = true, MAKA <i>low_risk</i> = <i>routinefollowup</i>
Petua U5	: JIKA <i>bladder_hard</i> = true, MAKA <i>low_risk</i> = <i>routinefollowup</i>
Petua U6	: JIKA <i>bladder_stop</i> = true, MAKA <i>low_risk</i> = <i>routinefollowup</i>
Petua U7	: JIKA <i>getbladder_month</i> = true, MAKA <i>low_risk</i> = <i>routinefollowup</i>
Petua U8	: JIKA <i>lost_weight</i> = true, MAKA <i>moderate_risk</i> = <i>followup</i>
Petua U9	: JIKA <i>bladder_blood</i> = true, MAKA <i>moderate_risk</i> = <i>followup</i>
Petua U10	: JIKA <i>age_cancer</i> >=50, MAKA <i>high_risk</i> = <i>biopsy</i>
Petua U11	: JIKA <i>family_history_with_prostate</i> = true, MAKA <i>high_risk</i> = <i>biopsy</i>
Petua U12	: JIKA <i>psa</i> > 4, MAKA <i>high_risk</i> = <i>biopsy</i>
Petua U13	: JIKA <i>Dre</i> = true, MAKA <i>high_risk</i> = <i>biopsy</i>

(a)

Petua U.1	: $patientrisk(x, y) \leftarrow patient_{id}(x) \wedge (bladder_{weak}(x, z) \vee bladder_{less}(x, z) \vee bladder_{stop}(x, z) \vee bladder_{delay}(x, z) \vee bladder_{weak}(x, z) \vee bladder_{hard}(x, z) \vee bladder_{month}(x, z)) \wedge$
-----------	---

	$risktype(y), \exists x: x \in I, \forall y: \{Y = low\}, \exists z: z \in \{no, less, intermediate, frequent, often\}.$
Petua U.2	$: patientrisk(x, y) \leftarrow patient_{id}(x) \wedge (lost_{weight}(x, a) \vee bladder_{blood}(x, a) \wedge risktype(y), \exists x: x \in I, \forall y: \{Y = moderate\}, \exists z: a \in \{True, False\}).$
Petua U.3	$: patientrisk(x, y) \leftarrow patient_{id}(x) \wedge (age(x, a) \vee family_{history}(x, b) \vee psa(x, c) \vee dre(x, b)) \wedge risktype(y) \exists x: x \in I, \forall y: \{Y = high\} \exists a: a \geq 50, \exists b: b \in \{true, false\}, \exists c: c \leq 4.$
Petua US.1	$: decision(x, y, z) \leftarrow patientrisk(x, y) \wedge treatmenttype(z)$ $treatmenttype(z)$ $= \begin{cases} routinefollowup & \text{if } y = low, \\ followup & \text{elseif } y = moderate, \\ biopsy & \text{elseif } y = high, \end{cases}$ where x is patient id, $\exists x: x \in I$, y is type of risk, z is treatment type

(b)

Rajah 6.1 (a) Petua-petua bagi Modul Urologi yang apabila dibincang dengan pakar Urologi, dan (b) perwakilan pengetahuan secara secara predikat logik tertib kedua

Petua R1	: JIKA $extension_prostate = false$, MAKA $low_risk = followup$
Petua R2	: JIKA $bones_prostate = false$, MAKA $low_risk = followup$
Petua R3	: JIKA $nodes_prostate = false$, MAKA $low_risk = followup$
Petua R4	: JIKA $organs_prostate = false$, MAKA $low_risk = followup$
Petua R5	: JIKA penilaian $PIRADS = 1$, MAKA $low_risk = followup$
Petua R6	: JIKA penilaian $PIRADS = 2$, MAKA $low_risk = followup$
Petua R7	: JIKA penilaian $PIRADS = 3$, MAKA $moderate_risk = biopsy$
Petua R8	: JIKA penilaian $PIRADS = 4$, MAKA $moderate_risk = biopsy$
Petua R9	: JIKA $extension_prostate = true$, MAKA $moderate_risk = biopsy$
Petua R10	: JIKA $bones_prostate = true$, MAKA $moderate_risk = biopsy$
Petua R11	: JIKA $nodes_prostate = true$, MAKA $moderate_risk = biopsy$
Petua R12	: JIKA $organs_prostate = true$ MAKA $high_risk = treatment$
Petua R13	: JIKA penilaian $PIRADS = 5$, MAKA $high_risk = treatment$

(a)

Petua R.1	: $radiologyrisk(x, y) \leftarrow patient_{id(x)} \wedge$ $(extension_{prostate}(x, z) \vee bones_{prostate}(x, z) \vee$ $nodes_{prostate}(x, z) \vee organs_{prostate}(x, z) \vee pirads(x, a) \wedge$ $risktype(y)) \exists x: x \in I, \forall y: \{Y = low\}, \exists z: z \in$ $\{false\}, \exists a: a \{1,2\}.$
Petua R.2	: $radiologyrisk(x, y) \leftarrow patient_{id(x)} \wedge (bones_{prostate}(x, z) \vee$ $nodes_{prostate}(x, z) \vee organs_{prostate}(x, z) \vee pirads(x, a) \wedge$ $risktype(y), \exists x: x \in I, \forall y: \{Y = moderate\}, \exists z: z \in$ $\{True\}, \exists a: a \{3,4\}.$
Petua R.3	: $radiologyrisk(x, y) \leftarrow patient_{id(x)} \wedge (organs_{prostate}(x, b) \vee$ $pirads(x, c)) \wedge risktype(y) \exists x: x \in I, \forall y: \{Y =$ $high\} \exists b: b \in \{true, false\}, \exists c: c \{5\}.$
Petua RS.1	: $decision(x, y, z) \leftarrow patientrisk(x, y) \wedge treatmenttype(z)$ $treatmenttype(z)$ $= \begin{cases} followup & \text{if } y = low, \\ biopsy & \text{elseif } y = moderate, \\ treatment & \text{elseif } y = high, \end{cases}$ <p style="text-align: center;">where x is patient id, $\exists x: x \in I, y$ is type of risk, z is treatment type</p>

(b)

Rajah 6.2 (a) Petua-petua bagi Modul Radiologi yang apabila dibincang dengan pakar radiologi, dan (b) perwakilan pengetahuan secara secara predikat logik tertib kedua.

Petua H1	: JIKA $gleason_overall \leq 6$, MAKA $low_progression = surveillance,$ $or radiotherapy or surgery$
Petua H2	: JIKA $gleason_overall = 7$, MAKA $moderate_progression =$ $radiotherapy or surgery$
Petua H3	: JIKA $gleason_overall \geq 8$, MAKA $high_progression = radiotherapy$ $or surgery or hormonotherapy$

(a)

Petua H.1	: $histopathologyprogression(x, y) \leftarrow patient_{id}(x) \wedge (gleason_{overall}(x, a) \wedge risktype(y), \exists x: x \in I, \forall y: \{Y = low\}, \exists z: z \leq 6.$
Petua H.2	: $histopathologyprogression(x, y) \leftarrow patient_{id}(x) \wedge (gleason_{overall}(x, b) \wedge risktype(y), \exists x: x \in \mathbb{I}, \forall y: \{Y = moderate\}, \exists z: b == 7.$
Petua H.3	: $histopathologyprogression(x, y) \leftarrow patient_{id}(x) \wedge (gleason_{overall}(x, c) \wedge risktype(y), \exists x: x \in \mathbb{I}, \forall y: \{Y = high\}, \exists z: c == 7.$
Petua HS.1	: $decision(x, y, z) \leftarrow patientrisk(x, y) \wedge treatmenttype(z)$ $treatmenttype(z)$ $= \begin{cases} surveillance, or radiotherapy or surgery & if & y = low, \\ radiotherapy or surgery & elseif & y = moderate, \\ radiotherapy or surgery or hormone therapy & elseif & y = high, \end{cases}$ where x is patient id, $\exists x: x \in \mathbb{I}$, y is type of risk, z is treatment type

Rajah 6.3 (a) Petua-petua bagi Modul Histopatologi yang apabila dibincang dengan pakar histopatologi, dan (b) perwakilan pengetahuan secara secara predikat logik tertib kedua.

a. Model Petua Pengeluaran Bagi Jabatan Pembedahan (Urologi), Radiologi Dan Hispatologi

Seterusnya merekabentuk bahagian belakang dan bahagian depan sistem diagnosis kanser prostat, logik petua diagnosis kanser prostat harus direka untuk mencapai hajat pengguna semasa menggunakan sistem. Pseudokod telah digunakan sebagai rekabentuk algoritma sistem diagnosis kanser prostat kerana lebih senang untuk diterjemahkan dalam bahasa pengaturcaraan dan menjimat masa. Seksyen ini akan membincangkan algoritma yang digunakan sebagai enjin tadbir dalam proses diagnosis mengikut bidang masing-masing iaitu pembedahan (urologi), radiologi dan histopatologi dalam sistem diagnosis kanser prostat.

Reka bentuk algoritma ini terdiri daripada proses-proses yang diperlukan setelah kajian ini selesai iaitu hasil daripada petua yang diperolehi. Pengguna perlu memasukkan input set data yang telah ditetapkan bagi mendapatkan hasil keputusan bagi setiap jabatan. Jika pesakit tersebut adalah menghidapai kanser prostat, maka perjalanan proses kembali kepada algoritma diagnosis di urologi untuk diproses dan dianalisis semula bagi mendapatkan keputusan yang akhir.

Rajah 6.4 merupakan pseudokod yang dihasilkan berdasarkan petua pakar untuk mendiagnosis pesakit pada peringkat pembedahan (urologi). Untuk menghasilkan keputusan pada peringkat ini, doktor pembedahan (urologi) harus membekalkan maklumat pesakit dan rekod pembedahan (urologi) seperti yang diperlukan oleh sistem. Algoritma ini akan menghasilkan keputusan berdasarkan peraturan yang ditetapkan sama ada keluar wad, membuat radiologi ataupun serta merta perlu mendapatkan rawatan.

```

Algoritma diagnosis urologi
// Input: Laporan pesakit urologi
// Output: Keputusan selepas

JIKA (bladder_empty >0 ATAU bladder_less >0 ATAU bladder_delay >0
      ATAU bladder_weak ATAU bladder_hard >0 ATAU bladder_stop >0
      ATAU bladder_month >0)
      MAKA decision ← routine_follow_up

JIKA (lost_weight >0 ATAU bladder_blood >0)
      MAKA decision ← follow_up

JIKA (age_cancer >0 ATAU family_history_with_prostate > 0 ATAU psa >4
      ATAU dre >0)
      MAKA decision ← biopsy

JIKA (decision == routine_follow_up)
      MAKA risk ← low_risk
JIKA (decision == follow_up)
      MAKA risk ← moderate_risk
JIKA (decision == biopsy)
      MAKA risk ← high_risk

```

Rajah 6.4 Pseudokod diagnosis urologi

Pseudokod untuk menghasilkan keputusan diagnosis daripada laporan radiologi adalah seperti yang ditunjukkan di Rajah 6.5. Algoritma ini telah dibina dengan merujuk kepada petua keputusan yang diperolehi daripada pakar.


```

Algoritma diagnosis radiologi
// Input: Laporan pesakit radiologi
// Output: Keputusan selepas

JIKA (extension_prostate == 0 ATAU bones_prostate ==0 ATAU
       organs_prostate==0 ATAU PIRADS == 1 ATAU PIRADS == 2)
       MAKA decision ← follow_up

JIKA (PIRADS == 3 ATAU PIRADS == 4 ATAU nodes_prostate > 0
       ATAU extension_prostate > 0 ATAU bones_prostate > 0)
       MAKA decision ← biopsy

JIKA (organs_prostate > 0 ATAU PIRADS == 5)
       MAKA decision ← treatment

JIKA (decision == follow_up)
       MAKA risk ← low_risk
JIKA (decision == biopsy)
       MAKA risk ← moderate_risk
JIKA (decision == treatment)
       MAKA risk ← high_risk

```

Rajah 6.5 Pseudokod diagnosis radiologi

Rajah 6.6 menunjukkan algoritma yang digunakan untuk menentukan keputusan diagnosis berdasarkan pentaksiran laporan histopatologi. Sekiranya pesakit didapati mempunyai nilai tumor besar dari [≥ 0] maka adalah positif, dia-disarankan untuk menerima rawatan daripada pakar onkologi. Pesakit yang mempunyai tumor benign pula disarankan untuk melakukan pemeriksaan radiologi setiap satu atau dua tahun untuk menjejak saiz tumor.

```

Algorithm diagnosis histopatologi
// Input: Jumlah gleason score pesakit hadapi
// Output: Rawatan yang dicadangkan

JIKA (gelason_overall <= 6)
    MAKA treatment ← surveillance, radiotherapy, surgery
JIKA (gelason_overall == 7) // solid
    MAKA treatment ← radiotherapy, surgery
JIKA (gelason_overall >=8)
    MAKA treatment ← radiotherapy, surgery, hormontherphy

JIKA (decision <= 6)
    MAKA risk ← low_progression
JIKA (decision == 7)
    MAKA risk ← moderate_progression
JIKA (decision >= 8)
    MAKA risk ← high_progression
JIKA
    MAKA risk ← no_progression

```

Rajah 6.6 Pseudokod diagnosis histopatologi

b. Penemberengan

Bahagian ini mengandungi prapemprosesan imej, penemberengan imej, post-pemprosesan imej dan pengekstrakan fitur tekstur imej. Maklumat yang telah dikumpulkan perlu membentuk perkaitan atau hubungan diantara maklumat dalam bidang perubatan yang sama atau berlainan. Disinilah proses membina peraturan – peraturan dan proses pengesahan peraturan dilakukan. Perincian pemeriksaan imbasan CT dengan mengenalpasti parameter imej CT yang dihadapi oleh pesakit yang menhadapi ketubuhan pada PCA. Selain itu, parameter umum dan 15 parameter khusus telah dikenalpasti bagi mengesan kewujudan ketumbuhan pada PCA pesakit. Sekiranya pesakit tidak mempunyai salah satu parameter imej CT, maka pesakit diletakan di bawah pengawalan di bawah pakar perubatan neuro dan radiologi atau pesakit diminta untuk dating berjumpa dengan doctor bagi tujuan pemeriksaan berkala.

Seterusnya, pesakit yang mempunyai salah satu parameter imej CT, pesakit dikendaki untuk mengambil histopatologi, pesakit diminta melakukan biopsi untuk mendapatkan keratin sel-sel dalam bahagian yang disyaki. Jadual 6.1 menunjukkan

beberapa jenis parameter yang digunakan dalam mengdiagnosis kanser prostat. Maka, bagi projek ini, Dr. Amir Radhi Othman, Pakar Radiologi yang telah dipertanggungjawabkan dalam mendapatkan maklumat perwakilan pengetahuan mengenai kanser prosatat telah mengariskan 5 parameter utama dan 15 sub parameter imej CT (CT Features) bagi mendiagnosis kanser prostat.

Jadual 6.1 Pengumpulan maklumat penemberengan CT

CT Features

01. Prostate imaging (Radiologist)

P001. Size

P002. Location

P003. Texture

P004. Colour intensity

02. Type of Cut

P005. Coronal

P006. Sagittal

P007. Axial

03. Prostate cancer staging (TNM staging)

P008. Evaluation of primary tumor ('T')

P009. Evaluation of the regional lymph nodes ('N')

P010. Evaluation of distant metastasis ('M')

P011. Evaluation of the histologic grade ('G')

04. Prostate Biopsy (Gleason score)

P012. Primary grade

P013. Secondary grade

05. Prostate cancer screening

P014. Digital rectal examination (DRE)

P015. Prostate-specific antigen (PSA) level

LAMPIRAN B
EKSPERIMEN J48

Parameter J48 bilangan minimum objek ditetapkan untuk menentukan nilai treshold. Apabila pokok keputusan teraruh, setiap pecahan akan diperiksa jumlah minimum objek pada nod daun. Jika nod daun adalah lebih besar daripada bilangan minimum objek, pemangkasan akan dilakukan. Manakala parameter bilangan lipatan digunakan bagi menentukan jumlah data yang digunakan untuk mengurangkan kesilapan ketika pemangkasan. Parameter benih pula digunakan untuk merawak data.

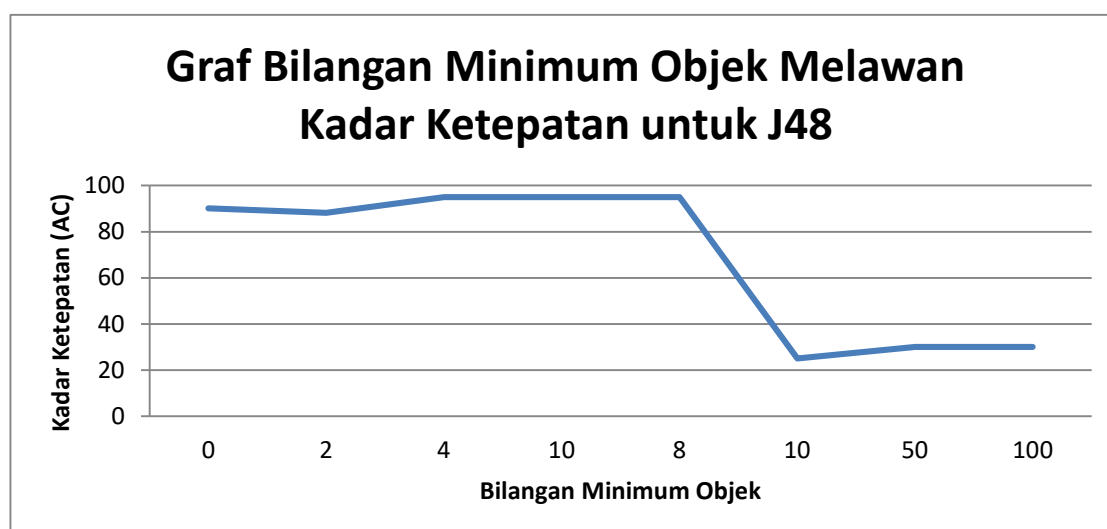
LAMPIRAN B.1

PEMILIHAN BILANGAN MINIMUM OBJEK J48

Parameter bilangan minimum objek ditetapkan untuk menentukan nilai treshold. Apabila pokok keputusan teraruh, setiap pecahan akan diperiksa bilangan minimum objek pada nod daun. Jika nod daun adalah lebih besar daripada bilangan minimum objek, pemangkasan akan dilakukan. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 6.2 dan Rajah 6.7, Kadar ketepatan atau *Accuracy Rate* (AC) yang dicapai adalah sebanyak 95% dengan ralat min kuasa dua atau *Mean Sum-squared Error* (MSE) sebanyak 0.1826 pada bilangan minimum 0. Oleh itu, bilangan minimum objek iaitu 0 ditetapkan pada semua eksperimen.

Jadual 6.2 Pemilihan bilangan minimum objek untuk latihan J48

Bilangan Minimum Objek	0 (Tidak Terhad)	2	4	10	50	100
AC	95	95	95	25	30	30
MSE	0.1826	0.1925	0.1925	0.4863	0.4724	0.4724



Rajah 6.7 Bilangan minimum objek melawan kadar ketepatan untuk J48

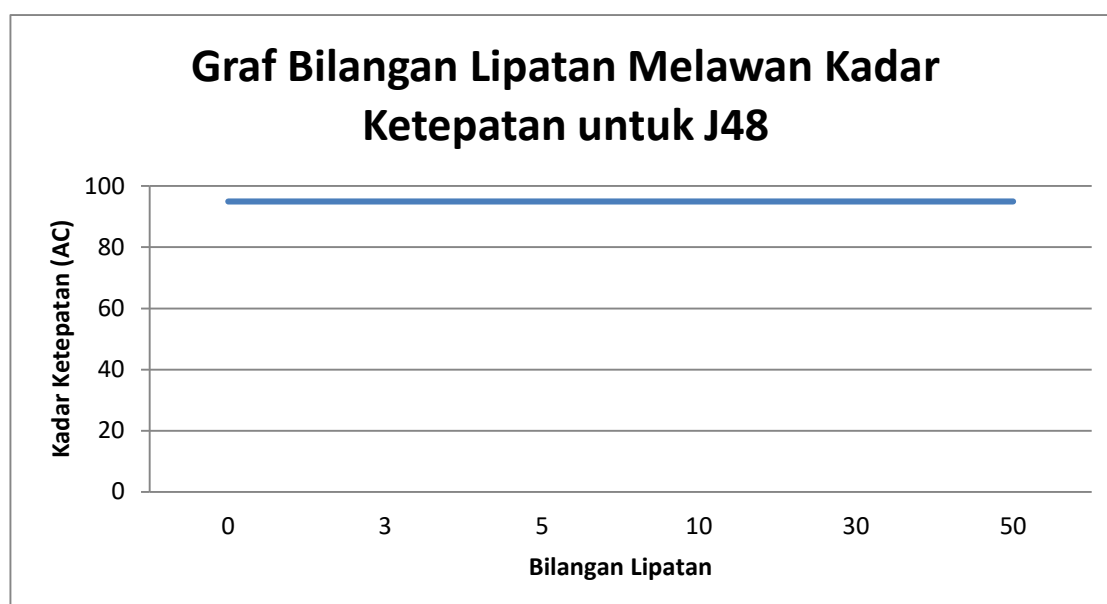
LAMPIRAN B.2

PEMILIHAN BILANGAN LIPATAN J48

Dalam kajian ini, pokok rawak digunakan untuk mendapatkan bilangan lipatan yang sesuai dengan menggunakan set data imej prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk menentukan jumlah data digunakan untuk penyesuaian semula. Satu lipatan digunakan untuk penyesuaian semula, yang lain untuk menanam pokok dan pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 6.3 dan Rajah 6.8, prestasi terbaik ditunjukkan pokok rawak adalah sebanyak 95% dengan ralat min kuasa dua (MSE) sebanyak 0.1925 pada lipatan adalah 3 default. Oleh itu, lipatan adalah 3 ditetapkan pada semua eksperimen.

Jadual 6.3 Pemilihan bilangan lipatan pokok untuk latihan J48

Bilangan Lipatan (Tidak Terhadap)	0	3	5	10	30	50
AC	95	95	95	95	95	95
MSE	0.1925	0.1925	0.1925	0.1925	0.1925	0.1925



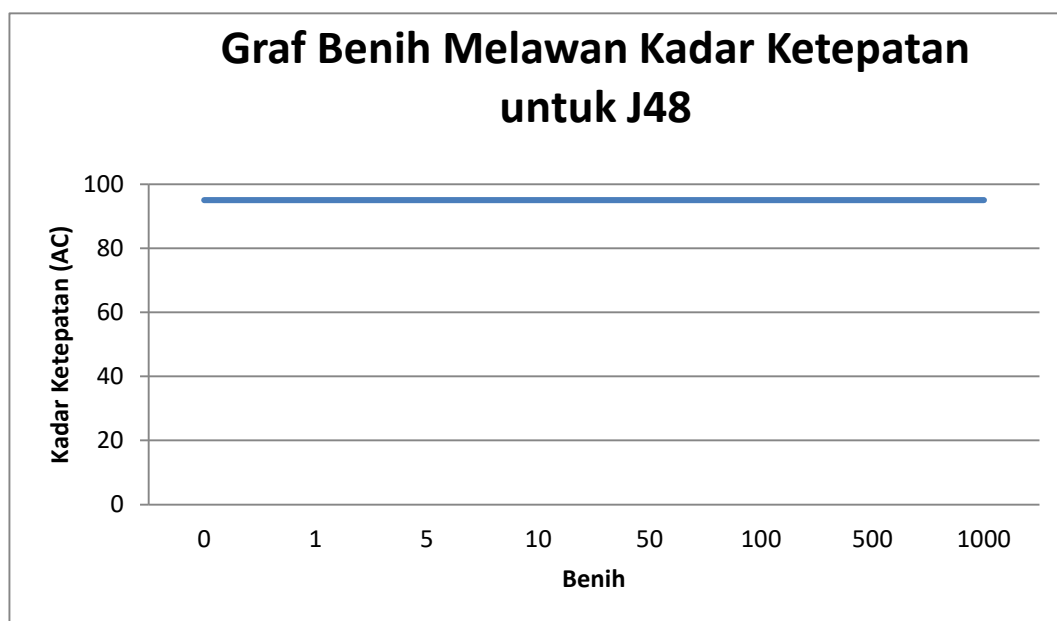
Rajah 6.8 Bilangan lipatan pokok melawan kadar ketepatan untuk J48

LAMPIRAN B.3
PEMILIHAN BENIH J48

Dalam kajian ini, pokok rawak digunakan untuk mendapatkan benih yang sesuai dengan menggunakan set data diagnosis kanser prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 6.4 dan Rajah 6.9, prestasi terbaik ditunjukkan Hutan Rawak adalah sebanyak 95% dengan ralat min kuasa dua (MSE) sebanyak 0.1925 pada benih adalah 1 default. Oleh itu, benih adalah 1 ditetapkan pada semua eksperimen.

Jadual 6.4 Pemilihan benih untuk latihan J48

Benih	0	1	5	10	50	100	500	1000
(AC)	95	95	95	95	95	95	95	95
(MSE)	0.1925	0.1925	0.1925	0.1925	0.1925	0.1925	0.1925	0.1925



Rajah 6.9 Benih melawan kadar ketepatan untuk J48

LAMPIRAN B.4

KEPUTUSAN DAN ANALISA J48

Bahagian ini menerangkan penilaian awal tentang pengelas J48. Nilai yang terbaik dipilih bagi bilangan minimum objek (minNumObj) adalah 0, bilangan lipatan (numFolds) adalah 3 dan benih (Seed) adalah 1. Peratus pecahan yang digunakan dalam eksperimen ini adalah 60% untuk latihan dan 40% untuk ujian.

Berdasarkan daripada ujikaji yang dijalankan terhadap set data latihan bagi teknik pokok keputusan J48 model yang terbaik dan mempunyai ketepatan ramalan yang paling tinggi jika dibandingkan dengan model- model yang lain. Jadual 6.5 menunjukkan prestasi pokok rawak mengikut peratus pecahan yang berbeza. Peratus pecahan yang digunakan adalah 50%, 55%, 60%, 65%, 70% dan 75%. J48 mencapai ketepatan tertinggi iaitu 95% pada peratus pecahan adalah 60% berbanding peratus pecahan yang lain.

Jadual 6.5 Prestasi antara J48 berdasarkan peratus pecahan yang berbeza

Pengelas	Peratus Pecahan	Kadar Ketepatan (%)	Ralat Min Kuasa Dua
J48	50	92.00	0.2309
	55	91.30	0.2408
	60	95.00	0.1826
	65	94.44	0.1925
	70	93.33	0.2108
	75	92.30	0.2265

Jadual 6.6 menunjukkan prestasi J48 dan mengikut keesahan silang yang berbeza. Keesahan silang yang digunakan adalah 10 lipatan dan 25 lipatan. J48 memperoleh ketepatan yang sama iaitu 94% pada keesahan silang adalah 10 lipatan dan 25 lipatan.

Jadual 6.6 Prestasi J48 berdasarkan keesahan silang yang berbeza

Pengelas	Kaedah Keesahan Silang	Kadar Ketepatan	Ralat Min Kuasa Dua
J48	10	94.11	0.1985
	25	94.11	0.198

LAMPIRAN C

EKSPERIMEN HUTAN RAWAK

Objektif eksperimen ini adalah untuk menilai prestasi pengelas Hutan Rawak. Hutan Rawak ini adalah pembesaran pokok pengelasan yang banyak. Vektor baru yang dikelaskan adalah dari setiap pokok yang ada pada Hutan Rawak. Setiap pokok yang ada pada Hutan Rawak ini mengelaskan objek yang baru dan undian (*votes*) dilakukan ke atas kelas tersebut. Sekiranya, pokok pengelasan memperoleh undian yang banyak, maka kelas tersebut dipilih.

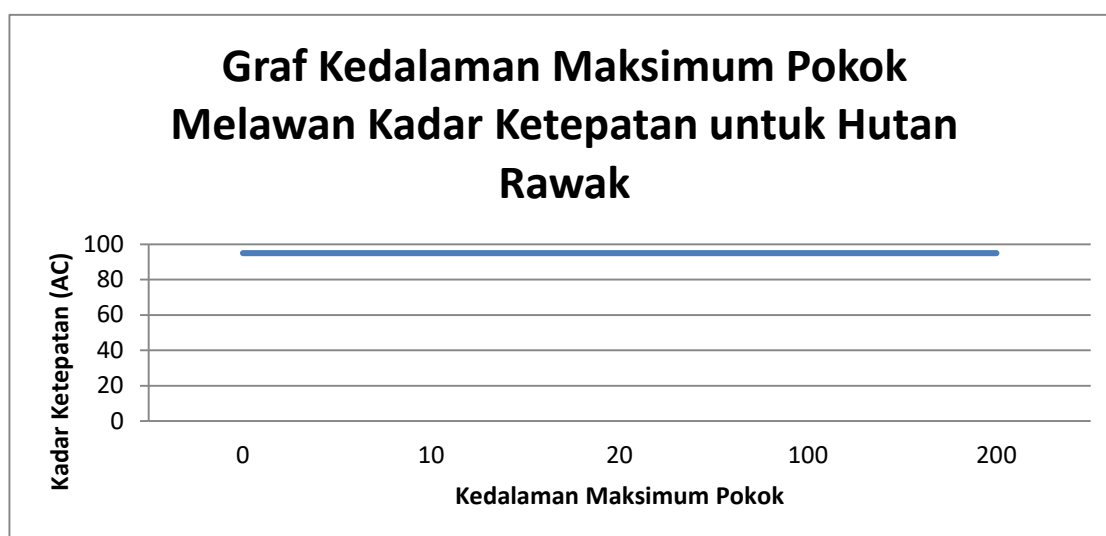
LAMPIRAN C.1

PEMILIHAN KEDALAMAN MAKSIMUM HUTAN RAWAK

Dalam kajian ini, Hutan Rawak digunakan untuk mendapatkan kedalaman maksimum yang sesuai menggunakan set laporan pesakit kanser prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 6.7 dan Rajah 6.10, prestasi terbaik yang ditunjukkan hutan rawak ditunjukkan pada kedalaman maksimum pokok adalah pada nilai 0 default. Kadar ketepatan atau *Accuracy Rate* (AC) yang dicapai adalah sebanyak 95% dengan ralat min kuasa dua atau *Mean Sum-squared Error* (MSE) sebanyak 0.1976. Oleh itu, kedalaman maksimum pokok iaitu 0 ditetapkan pada semua eksperimen.

Jadual 6.7 Pemilihan kedalaman maksimum pokok untuk latihan hutan rawak

Kedalaman Maksimum Pokok	0				
	(Tidak Terhad)	10	20	100	200
AC	95	95	95	95	95
MSE	0.1976	0.1976	0.1976	0.1976	0.1976



Rajah 6.10 Kedalaman maksimum pokok melawan kadar ketepatan untuk hutan rawak

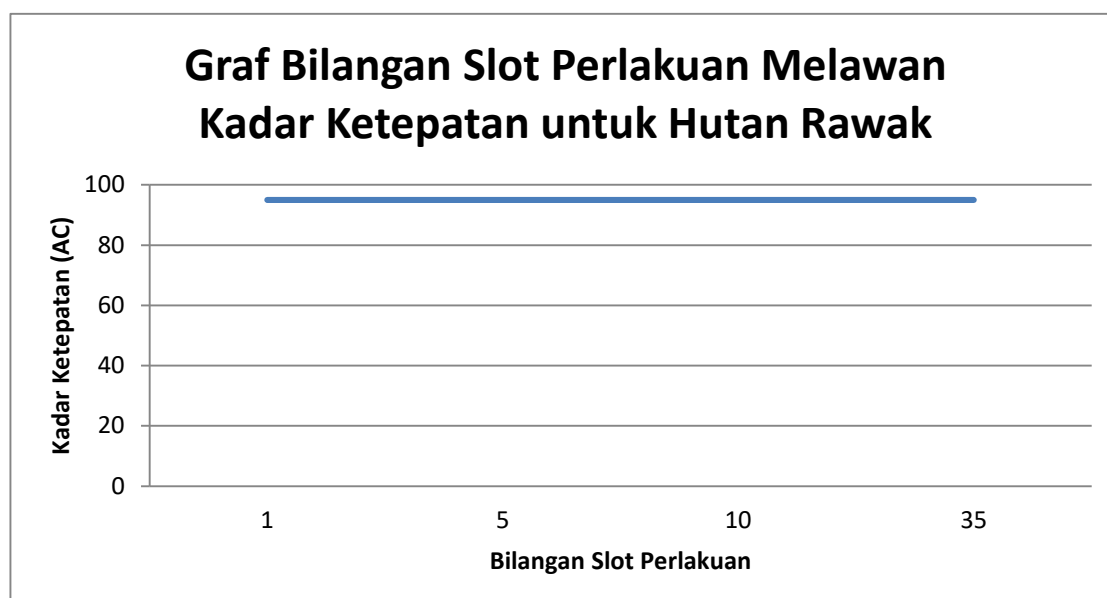
LAMPIRAN C.2

PEMILIHAN BILANGAN SLOT PERLAKUAN HUTAN RAWAK

Dalam kajian ini, Hutan Rawak digunakan untuk mendapatkan bilangan slot perlakuan yang sesuai menggunakan set data imej prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 6.8 dan Rajah 6.11, prestasi menunjukkan hasil yang sama iaitu sebanyak 95% dengan ralat min kuasa dua (MSE) sebanyak 0.3414 pada bilangan slot perlakuan adalah dari 0 hingga 35. Oleh itu, bilangan slot perlakuan adalah 1 default ditetapkan pada semua eksperimen.

Jadual 6.8 Pemilihan bilangan slot perlakuan untuk latihan hutan rawak

Bilangan Slot Perlakuan	0	1	5	10	35
AC	95	95	95	95	95
MSE	0.1976	0.1976	0.1976	0.1976	0.1976



Rajah 6.11 Bilangan slot perlakuan melawan kadar ketepatan untuk hutan rawak

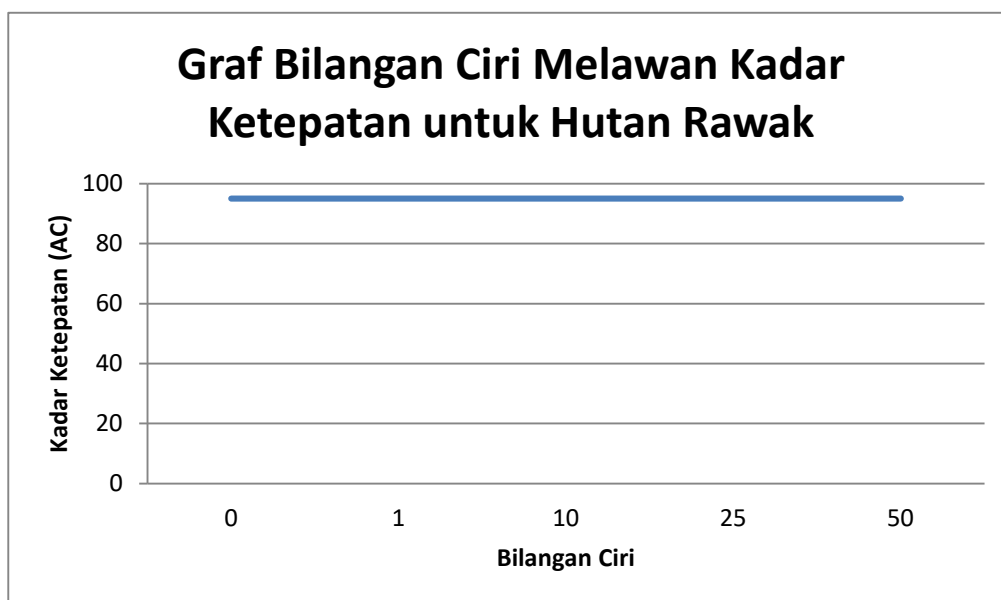
LAMPIRAN C.3

PEMILIHAN BILANGAN CIRI HUTAN RAWAK

Dalam kajian ini, Hutan Rawak digunakan untuk mendapatkan bilangan ciri yang sesuai dengan menggunakan set data imej prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 6.9 dan Rajah 6.12, prestasi terbaik ditunjukkan Hutan Rawak adalah sebanyak 95% dengan ralat min kuasa dua (MSE) sebanyak 0.1976 pada bilangan ciri adalah 0 default. Oleh itu, bilangan ciri adalah 0 ditetapkan pada semua eksperimen.

Jadual 6.9 Pemilihan bilangan ciri untuk latihan hutan rawak

Bilangan Ciri	0	1	10	25	50
(AC)	95	90	95	95	95
(MSE)	0.1976	0.2493	0.1987	0.1987	0.1987



Rajah 6.12 Bilangan ciri melawan kadar ketepatan untuk hutan rawak

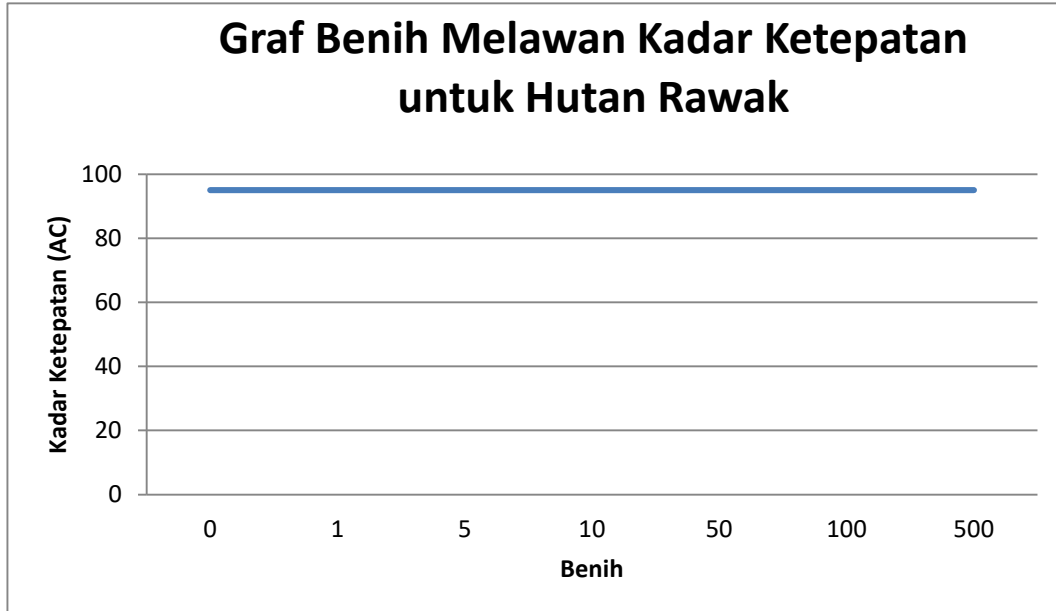
LAMPIRAN C.4

PEMILIHAN BENIH HUTAN RAWAK

Dalam kajian ini, Hutan Rawak digunakan untuk mendapatkan benih yang sesuai dengan menggunakan set data imej prostat. Kaedah penilaian berdasarkan peratus pecahan digunakan untuk pecahkan set data kepada latihan (60%) dan ujian (40%). Berdasarkan Jadual 6.10 dan Rajah 6.13, prestasi terbaik ditunjukkan Hutan Rawak adalah sebanyak 95% dengan ralat min kuasa dua (MSE) sebanyak 0.1897 pada benih adalah 5. Oleh itu, benih adalah 5 ditetapkan pada semua eksperimen.

Jadual 6.10 Pemilihan benih untuk latihan hutan rawak

Benih	0	1	5	10	50	100	500
(AC)	95	95	95	95	95	95	95
(MSE)	0.1912	0.1976	0.1897	0.1944	0.2006	0.1904	0.1989



Rajah 6.13 Benih melawan kadar ketepatan untuk hutan rawak

LAMPIRAN C.5

KEPUTUSAN DAN ANALISA HUTAN RAWAK

Bahagian ini menerangkan penilaian awal tentang pengelas Hutan Rawak. Nilai yang terbaik dipilih bagi kedalaman maksimum pokok (maxDepth) adalah 0, bilangan slot perlakuan (numExecutionSlots) adalah 1, bilangan ciri (numFeature) adalah 0 dan benih (Seed) adalah 5. Peratus pecahan yang digunakan dalam eksperimen ini adalah 60% untuk latihan dan 40% untuk ujian.

Berdasarkan Jadual 6.11, Hutan Rawak menghasilkan peratus ketepatan yang terendah berbanding pokok keputusan yang lain pada setiap peratus pecahan. Peratus pecahan yang digunakan adalah 50%, 55%, 60%, 65%, 70% dan 75%. Hutan rawak mencapai ketepatan tertinggi iaitu 95% pada peratus pecahan adalah 60% berbanding peratus pecahan yang lain.

Jadual 6.11 Prestasi antara hutan rawak berdasarkan peratus pecahan yang berbeza

Pengelas	Peratus Pecahan	Kadar Ketepatan (%)	Ralat Min Kuasa Dua
Hutan Rawak	50	92.00	0.2297
	55	91.30	0.2377
	60	95.00	0.1897
	65	94.44	0.1987
	70	93.33	0.2203
	75	92.30	0.2223

Jadual 6.12 menunjukkan prestasi hutan rawak dan mengikut keesahan silang yang berbeza. Keesahan silang yang digunakan adalah 10 lipatan dan 25 lipatan. Hutan rawak memperoleh ketepatan yang sama iaitu 94% pada keesahan silang adalah 10 lipatan dan 25 lipatan.

Jadual 6.12 Prestasi pokok rawak berdasarkan keesahan silang yang berbeza

Pengelas	Kaedah Keesahan Silang	Kadar Ketepatan	Ralat Min Kuasa Dua
Hutan Rawak	10	94.11	0.1853
	25	94.11	0.1801

LAMPIRAN D

SURAT LATIHAN SANGKUTAN



Fakulti Teknologi dan Sains Maklumat

Faculty Of Information Science and Technology

Pusat Teknologi Kecerdasan Buatan (CAIT)

UKM 1.22.7/237/3
17 April 2015

PROF. MADYA DR. HAMZAINI ABD. HAMID

Ketua
Jabatan Radiologi
Fakulti Perubatan Universiti Kebangsaan Malaysia
Jalan Ya'acob Latif, Bandar Tun Razak,
Cheras, 56000,
Kuala Lumpur.

Melalui

PROF. DATO' DR. RAYMOND AZMAN ALI

Dekan
Fakulti Perubatan Dan Pengarah Pusat Universiti Kebangsaan Malaysia,
Jalan Ya'acob Latif, Bandar Tun Razak,
Cheras, 56000,
Kuala Lumpur.

PERMOHONAN PENEMPATAN PELAJAR UNTUK LATIHAN SANGKUTAN DI PPUKM

Merujuk kepada perkara di atas,

2. Dimaklumkan bahawa penama di dalam jadual berikut adalah pelajar daripada Fakulti Teknologi Dan Sains Maklumat, Universiti Kebangsaan Malaysia (UKM), Bangi. Pelajar ini terlibat dengan penyelidikan di bawah Pusat Teknologi Kecerdasan Buatan (CAIT) berkaitan Projek Arus Perdana (AP-2012-019) "AUTOMATED MEDICAL IMAGING DIAGNOSTIC BASED ON FOUR CRITICAL DISEASES BRAIN, BREAST, PROSTATE AND LUNG CANCER" serta sedang maju menjalankan projek tahun akhir bertajuk "Prostate Cancer Detection System" Pelajar ini memerlukan khidmat nasihat daripada doktor-doktor pakar di Jabatan Radiologi. Oleh itu, pelajar ini dikehendaki menjalani latihan sangkutan dalam tempoh masa yang dipersetujui oleh kedua-dua pihak iaitu CAIT dan PPUKM mengikut keperluan penyelidikan di bawah seliaan Dr. Aini Ab. Aziz, dan Dr Ruziana Iqbal Hussain dari Jabatan Radiologi bermula daripada 30 April 2015 hingga 30 Julai 2015.

BIL	NAMA PELAJAR	NO.MATRIK	DATA	BAHAGIAN
1	Muhammad Uzair Bin Sahrin	A142770	Prostate cancer	Radiologi

Pusat Teknologi Kecerdasan Buatan , Fakulti Teknologi dan Sains Maklumat,
Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor.
Telefon: +603-8921 6089/6086 Faksimili: +603-8925 6732 E-mel: k.cait@ftsm.ukm.my
Laman Web: <http://www.ftsm.ukm.my>

2. Untuk pengetahuan Prof. Madya Dr. projek Arus Perdana (AP-2012-019) ini sememangnya melibatkan beberapa ahli daripada PPUKM. Berikut adalah nama ahli yang terlibat:-

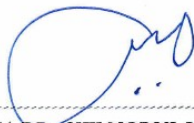
PROF. DR. ZULKIFLI MD ZAINUDDIN	- JABATAN SURGERI
DR. REENA RAHAYU MD ZIN	- JABATAN PATOLOGI
PROF. MADYA DATO' DR. FUAD ISMAIL	- JABATAN ONKOLOGI
DR. RUZIANA IQBAL HUSSAIN	- JABATAN RADIOLOGI

3. Masa yang dicadangkan adalah dalam tempoh 30 April 2015 hingga 30 Julai 2015. Secara umumnya, skop latihan sangkutan ini adalah meliputi lawatan ke hospital (jabatan yang terlibat sahaja), pengumpulan data, atau sebarang aktiviti yang difikirkan bersesuaian untuk projek ini.

4. Sebarang pertanyaan boleh diajukan kepada Prof. Madya Dr. Siti Norul Huda Sheikh Abdullah di talian 019-2377984. Kerjasama dan perhatian yang diberikan oleh pihak Prof. Madya Dr. sangat kami hargai dan didahului dengan ucapan terima kasih.

Sekian untuk makluman dan tindakan selanjutnya.

Yang benar,




PROF. MADYA DR. SITI NORUL HUDA SHEIKH ABDULLAH
PUSAT TEKNOLOGI KECERDASAN BUATAN
FAKULTI TEKNOLOGI SAINS MAKLUMAT
UNIVERSITI KEBANGSAAN MALAYSIA
43600 BANGI, SELANGOR.

PROF. MADYA DR. SITI NORUL HUDA SHEIKH ABDULLAH
Penyelidik dan Penyelia,
Pusat Teknologi Kecerdasan Buatan (CAIT),
Fakulti Teknologi dan Sains Maklumat,

sk: Dr. Rizuana Iqbal Hussain
Jabatan Radiologi PPUKM
Pusat Perubatan Universiti Kebangsaan Malaysia (PPUKM)
Jalan Ya'acob Latif, Bandar Tun Razak,
Cheras, 56000, Kuala Lumpur.

LAMPIRAN E
SURAT KELULUSAN ETIKA


UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

Pusat Pengurusan Penyelidikan dan Instrumentasi Centre for Research and Instrumentation Management

Rujukan : UKM PPI/111/8/JEP-2016-225
Tarikh : 22 Ogos 2016

Profesor Madya Dr. Siti Norul Huda Sheikh Abdullah
Pusat Teknologi Kecerdasan Buatan
Fakulti Teknologi & Sains Maklumat
Universiti Kebangsaan Malaysia

Y.bhg. Profesor/Datuk/Dato'/Datin/Tuan/Puan,

Kelulusan Etika Menjalankan Penyelidikan di UKM

Tajuk : *Prostate Cancer Detection And Recognition System (PCDRS)*

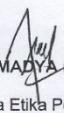
Perkara yang tersebut di atas adalah dirujuk.

Sukacita dimaklumkan, Jawatankuasa Etika Penyelidikan UKM meluluskan permohonan penyelidikan Y. Bhg. Profesor/Datuk/Dato'/Datin/Tuan/Puan bagi tajuk diatas. Tempoh kelulusan penyelidikan adalah daripada **17 Ogos 2016 - 16 Ogos 2019**. Sila kemukakan sebarang Laporan Kesan Sampingan, Laporan Kemajuan Setiap 6 Bulan dan Laporan Akhir sebaik sahaja penyelidikan tamat kepada Jawatankuasa Etika Penyelidikan UKM.

Sukacita diingatkan projek penyelidikan ini hanya boleh dijalankan setelah mendapat **surat kelulusan menjalankan penyelidikan** dari Timbalan Dekan Penyelidikan Fakulti.

Sekian, terima kasih.

Yang benar,


PROFESOR MADYA (K) DATO' DR. FUAD ISMAIL
Pengerusi
Jawatankuasa Etika Penyelidikan
Universiti Kebangsaan Malaysia

s.k. - **Pengarah**
Pusat Pengurusan Penyelidikan dan Instrumentasi (CRIM)
Universiti Kebangsaan Malaysia

- **Timbalan Dekan (Penyelidikan, Inovasi & Antarabangsa)**
Fakulti Teknologi & Sains Maklumat
Universiti Kebangsaan Malaysia

- **Pengerusi Pusat Teknologi Kecerdasan Buatan**
Fakulti Teknologi & Sains Maklumat
Universiti Kebangsaan Malaysia

Sekretariat Etika Penyelidikan Universiti Kebangsaan Malaysia, Tingkat 1, Blok Klinikal, Hospital Canselor Tuanku Muhriz, Pusat Perubatan UKM
Jalan Yaacob Latif, Bandar Tun Razak, 56000 Cheras Kuala Lumpur, MALAYSIA. Telefon: +603-9145 5046 / 5048

Mengilham Harapan, Mencipta Masa Depan • *Inspiring Futures, Nurturing Possibilities* www.ukm.my

LAMPIRAN F
SURAT MINIT MESYUARAT ETIKA



UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

**MINUTES OF MEETING
RESEARCH ETHICS COMMITTEE
THE NATIONAL UNIVERSITY OF MALAYSIA
NO. 10/2016**

Date : 16th June 2016 (Thursday)
Venue : Seminar Room 2,
Department of Radiotherapy & Oncology
1st Floor, Block Clinical
UKM Medical Centre, Cheras
Time : 2:15 p.m

3. New Application

3.8

Researcher	: Associate Professor Dr. Siti Norul Huda Sheikh Abdullah (FTSM), Associate Professor Dato' Dr. Fuad Ismail, Associate Professor Dr. Shahnorbanun Sahran, Dr. Syazarina Sharis Osman, Dr. Suria Hayati Md Pauzi, Dr. Praveen Singam, Dr. Afzan Adam, Muhammad Uzair Sahrin (P83911) , Mohd Fazlan Abd Azis (A146388) , Muhd Ridzuan Muhd Yunus (P75864) , Yasmin Mumtaz Ahmad (P82814) , Nursyuhadah Abdul Razak (A146368)
Title	: Prostate Cancer Detection And Recognition System (PCDRS)
Duration	: 36 Month
No. of subject	: 30
Result	: Approved with corrections
Correction	: 1. Terminology normal prostate – non-malignant prostate biopsy. 2. Consent not necessary provided patient anonymity is preserved. 3. Need director permission for data collection.
Declaration	: Associate Professor Dato' Dr. Fuad Ismail is a co-investigator in this study and also the chairman of Research Ethics Committee UKM. He is not involved in discussion of approval of this study. The presentation chaired by Professor Dr. Mohd Shahrir Mohamed Said.

LAMPIRAN G

LAPORAN PESAKIT

59yr old, u/l DM/HPT, persistently elevated PSA btwn 3-4 since 2008. Previous TRUS biopsy in 2008 & 2014 normal. Latest reading 4.34 on 5 A-reductase inhibitor. Increasing PSA level 5.12 (1.8.2017). Not keen on TRUS biopsy at the moment in view of morbidity and on aspirin. .

Sequences:

Ax: T2WI, T1WI, DWI/ ADC, DCE ; Cor: T2WI ; Sag: T2WI

Findings:

Prostate gland is heterogeneously enlarged measuring about 3.1 x 4.5 x 4.3cm (AP x W x CC), with calculated volume is 30mls.

Peripheral zone:

The peripheral zone is compressed by the enlarged transitional zone.

Normal signal intensity in the rest of the peripheral zone. No definite nodules showing fluid restriction on DWI/ADC seen. Only vague hypointensity on ADC is noted in the left peripheral zone at about 4-5 o'clock in the apex (PIRADS 2)

Central/transitional zone:

There are foci of T2 hyperintense and T1 hypointense signal seen in the transitional zon at 4-5 o'clock. These could represent focal hemorrhage or proteinaceous content prostate cyst. No significant enhancement post contrast. There is hypertrophy of the prostate gland involving the transition and central zone/central gland with diffuse enhancement and multiple encapsulated nodules within with overall **PIRADS 2 score**.

Intact prostatic capsule and anterior fibromuscular stroma.

The urinary bladder is distended. Clear plane is demonstrated between the prostate gland with the urinary bladder anteriorly and rectum posteriorly.

Bilateral neurovascular bundles are preserved.

Bilateral seminal vesicles are dilated with normal high T2 signal intensity within.

No intraabdominal free fluid in the pelvis. Levator ani muscles are normal.

The visualized anal canal is unremarkable. No enlarged pelvic or iliac nodes.

No suspicious abnormal signal within the visualized bones.

Impression:

Prostatomegaly with transitional zone hypertrophy containing PIRADS 2 nodules consistent with benign prostatic hyperplasia.

No suspicious malignant nodule within the compressed peripheral zone.

Dr Janice/ Nurdillah (MO)

Dr Syazarina (Radiologist)

Prostate volume: 30mls

Lesion location: R peripheral zone, apex
PIRADS Version 2 lesion score: PIRADS 2

Extracapsular extension : No

Involvement of adjacent local organ: No

Organ involved:-

Lymph nodes involvement : non suspicious

Visualized bones: Normal/benign

DESCISION : Follow up

LAMPIRAN H
MANUAL PENGGUNA



**USER MANUAL
PROSTATE CANCER
EXPERT SYSTEM**



User Manual Prostate Cancer Expert System 2018

Table of Contents

1.0	Introduction.....	2
1.1	User Expert System	2
2.0	Logon	2
3.0	General ward.....	3
4.0	Radiology Ward.....	7
5.0	Histopathology ward	11

User Manual Prostate Cancer Expert System 2018

1.0 Introduction

This user manual is available to provide guidance to Users of prostate cancer expert system about how to use the functions available in the expert system

1.1 User Expert System

Listed below are definitions and category "user expert system". An "User expert system " can granted access to relevant departments only.

a. User

"User" is a user an expert who has been registered as a user.




1.2 Overview User Manual

This document explains on how to use an expert system. User roles in this system is divided into three parts namely Urology, Radiology and Histology.

2.0 Logon

This section describes how to log on to an expert system.

1. Double click on bcds.jar to start the application.

Name	Date modified	Type
 pces.accdb	14/7/2015 5:30 AM	Microsoft Access ...
 pces.jar	14/7/2015 5:33 AM	Executable Jar File
 readme.txt	1/5/2017 11:11 PM	Text Document

Note: This application requires Java Runtime Environment (JRE) to be installed in the machine in order to run properly. If JRE is not installed, it can be downloaded at the following link:

<http://www.oracle.com/technetwork/java/javase/downloads/jre7-downloads-1880261.html>

User Manual Prostate Cancer Expert System 2018

3.0 General ward

- 1 When the application starts executing, the login screen will appear as follows:

Please enter the following details to log into the application.

Username: general

Password: general

- 2 The main menu will appear as follows:

To start diagnosis a patient, user will have to first create a new patient in the system database. Click on the Add **new patient button** to proceed.

- 3 The Patient Information page will appear as follows:

User Manual Prostate Cancer Expert System 2018

- (i) Enter the required patient information.
- (ii) Click **Save** button to save the record.
- (iii) **Clear** button can be used to clear all the fields to be empty.
- (iv) Click on the **Home** button to get back into the main menu after finish adding required patients into system.

Note: The **Date of Birth** field requires date to be entered in dd/MM/yyyy format. For e.g. 12/3/1978 for March 12, 1978.

- 4 To add a new medical record for a patient, click on the New medical record button from main menu and the following page will appear:**

The screenshot shows a web-based form titled "GENERAL MEDICAL (Adding)". It has a "Home" button in the top right. The form is organized into three main sections, each with a yellow header bar:

- PATIENT INFORMATION:** Includes a "Select patient" dropdown menu (currently showing "Et: 24, Name: Sulaiman"), and input fields for "Patient ID: 24", "Age: 28", "Height: 178.0", "Patient Name: Sulaiman", "Weight: 67.0", and "BMI: 21.15". An "Update" button is located at the bottom right of this section.
- IPSS SCORE:** Contains five questions with dropdown menus for answers:
 - Question 1: "Over the past month, how often have you had a sensation of not emptying your bladder completely after you finished urinating?" (Options: About half the time, Less than half the time)
 - Question 2: "Over the past month, how often have you had to urinate again less than two hours after you finished urinating?" (Options: Less than half the time, More than half the time)
 - Question 3: "Over the past month, how often have you had a weak urinary stream?" (Options: Less than half the time, More than half the time)
 - Question 4: "Over the past month, how many times did you most typically get up to urinate, from the time you went to bed at night until the time you got up in the morning?" (Options: None, 1-2, 3-4, 5 or more)
 - Question 5: "Over the past month, how often have you found you have stopped and started again several times when you urinated?" (Options: Less than half the time, More than half the time)
 - Question 6: "Over the past month, how often have you found it difficult to postpone urinating?" (Options: Less than half the time, More than half the time)
 - Question 7: "Over the past month, how often have you had to push or strain to begin urination?" (Options: Less than half the time, More than half the time)
- FACTOR:** Contains two Yes/No questions:
 - "Patients aged 50 years and older?" (Options: Yes, No)
 - "Patients lose weight?" (Options: Yes, No)
 - "Patients have a close relative with cancer of the prostate?" (Options: Yes, No)
 - "Urine comes with blood?" (Options: Yes, No)

At the bottom of the form, there is a "Doctor's comment" text area, a "Diagnose" button, a "Decision" dropdown menu (set to "Follow-up"), and a "Save" button.

- (i) Select the patient before update medical record.
- (ii) Update the information (when necessary). The BMI will be calculated automatically from weight and height and age will be calculated automatically from date of birth.
- (iii) Update the IPSS SCORE section by selecting **Not at all**, **Less than 1 time in 5**, **Less than half the time**, **About half the time**, **More than half the time** or **Almost always** from the dropdown. By default, all values are **Not at all**.
- (iv) Update the FACTOR section by selecting **Yes** or **No** from the dropdown. By default, all values are **No**.
- (v) User may add in addition comments on the diagnosed result.
- (vi) Click on **Diagnose** button to start diagnosing. A pop up window showing the result will be shown as following:

User Manual Prostate Cancer Expert System 2018

Results Of General Medical Record

The risk level to have prostate cancer is : **HIGH**

Factors contributing to the result :

- Patients lose weight.
- Patients have a close relative with cancer of the prostate.
- Over the past month, how often have you had to push or strain to begin urination.
- Over the past month, how often have you found you have stopped and started again several times when you urinated.
- Over the past month, how many times did you most typically get up to urinate, from the time you went to bed at night until the time you got up in the morning.
- Over the past month, how often have you had a sensation of not emptying your bladder completely after you finished urinating?
- Over the past month, how often have you had to urinate again less than two hours after you finished urinating.
- Over the past month, how often have you found it difficult to postpone urinating.
- Over the past month, how often have you had a weak urinary stream.

Suggestion : Patients are recommended to perform the test PCA / DRE for further diagnosis.

OK

(vii) Update the final decision from the dropdown.

(viii) Click on the **Save** button to save the record.

(ix) Click on **Home** button to get back to the main menu when done.

5 To update a medical record for a patient, click on the Update medical record button from main menu and the following page will appear:

General Medical Record (Update)

Home (x)

PATIENT INFORMATION (i)

Select patient: [IC 24, Name: Sulaiman] Select report: [Report ID: 52, Date: 2010-01-01]

Patient ID: 24 Patient Name: Sulaiman

Age: 20 DM: 21.15

IPSS SCORE (ii)

Over the past month, how often have you had a sensation of not emptying your bladder completely after you finished urinating? Less than half the time Over the past month, how often have you found you have stopped and started again several times when you urinated? Less than half the time

Over the past month, how often have you had to urinate again less than two hours after you finished urinating? About half the time Over the past month, how often have you found it difficult to postpone urinating? Less than half the time

Over the past month, how often have you had a weak urinary stream? Less than half the time Over the past month, how often have you had to push or strain to begin urination? About half the time

Over the past month, how many times did you most typically get up to urinate, from the time you went to bed at night until the time you got up in the morning? 4 times

FACTOR (iii)

Patients aged 50 years and older? Yes No

Patients lose weight? Yes No

Score: 18 / 35

PSAORE (iv)

PSA Score Test: 0 DRE Examination: No

Doctor's comment: (v)

Diagnose (vi)

Decision: [Biopsy] (vii)

(viii) Update (ix) Delete

(i) Select the patient before adding medical record.

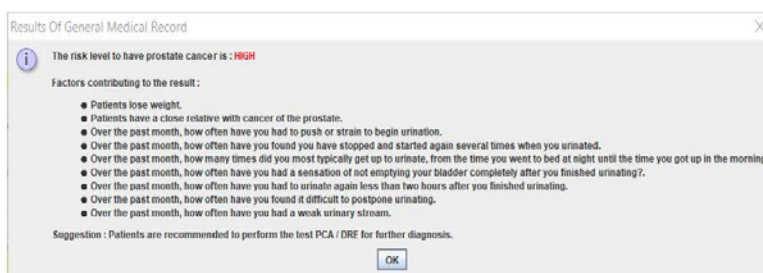
(ii) Update the IPSS SCORE section by selecting **Not at all**, **Less than 1 time in 5**, **Less than half the time**, **About half the time**, **More than half the time** or **Almost always** from the dropdown. By default, all values are **Not at all**.

(iii) Update the FACTOR section by selecting **Yes** or **No** from the dropdown. By default, all values are **No**.

5 | 36

User Manual Prostate Cancer Expert System 2018

- (iv) Update the PSA/DRE section by insert **Value** PSA and DRE **Yes** or **No** from the dropdown. By default, all values are **No**.
- (v) User may add in addition comments on the diagnosed result.
- (vi) Click on **Diagnose** button to start diagnosing. A pop up window showing the result will be shown as following:



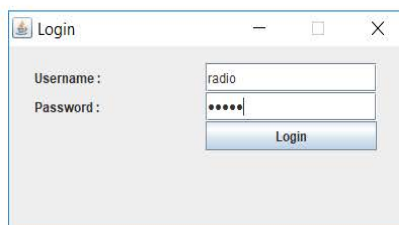
- (vii) Update the final decision from the dropdown.
- (viii) Click on the **Save** button to save the record.
- (ix) Click on the **Delete** button to remove patient record from the system.
- (x) Click on **Home** button to get back to the main menu when done.

6 To view a medical record for a patient, click on the **View medical record** button from main menu and the following page will appear:

- (i) Select the patient before view medical record.
- (ii) Click on the **Save pdf** button to preview report in document (pdf) from the system.
- (iii) Click on **Home** button to get back to the main menu when done.

4.0 Radiology Ward

1. When the application starts executing, the login screen will appear as follows:

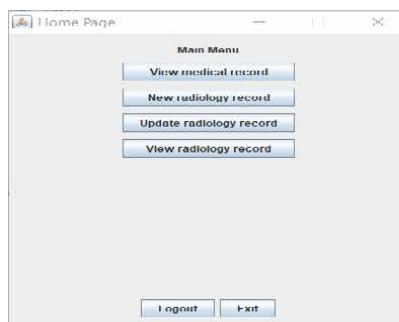


Please enter the following details to log into the application.

Username: radio

Password: radio

2. The main menu will appear as follows:

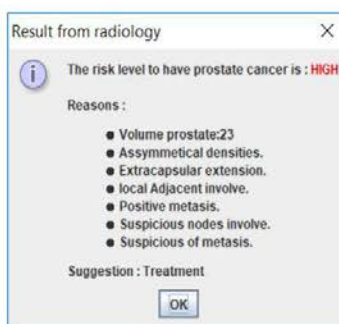


To start diagnosis a patient, user will have to first see record from General ward to know patient report. Click on the Add **view medical record** button to proceed.

3. To add a new radiology record for a patient, click on the **New radiology record** button from main menu and the following page will appear:

User Manual Prostate Cancer Expert System 2018

- (i) Select the patient before adding radiology record.
- (ii) Browse the IMAGE INFORMATION section by selecting the **Source Image**.
- (iii) Update the RADIOLOGY ASSESSMENT section by selecting necessary value according to Image Information.
- (iv) User may add in addition comments on the diagnosed result.
- (v) Click on **Diagnose** button to start diagnosing. A pop up window showing the result will be shown as following:



- (vi) Update the final decision from the dropdown.
- (vii) Click on the **Save** button to save the record.
- (viii) Click on the **Clear** button to clear the record.
- (ix) Click on **Home** button to get back to the main menu when done.

User Manual Prostate Cancer Expert System 2018

4. To update a radiology record for a patient, click on the Update radiology record button from main menu and the following page will appear:

- (i) Select the patient before update radiology record.
- (ii) Select the report before update radiology record.
- (iii) Browse the IMAGE INFORMATION section by selecting the **Source Image**.
- (iv) Update the RADIOLOGY ASSESSMENT section by selecting necessary value according to Image Information.
- (v) User may add in addition comments on the diagnosed result.
- (vi) Click on **Diagnose** button to start diagnosing. A pop up window showing the result will be shown as following:

- (vii) Update the final decision from the dropdown.
- (viii) Click on the **Update** button to update the record.
- (ix) Click on the **Delete** button to remove patient record from the system.
- (x) Click on **Home** button to get back to the main menu when done.

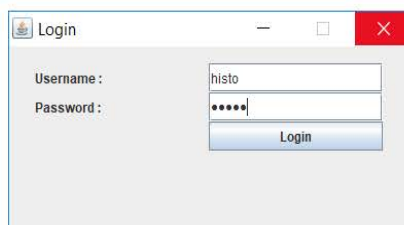
User Manual Prostate Cancer Expert System 2018

5. To view a radiology record for a patient, click on the View radiology record button from main menu and the following page will appear:

- (i) Select the patient before view radiology record.
- (ii) Select the report before view radiology record.
- (iii) Click on the **Save pdf** button to preview report in document (pdf) from the system.
- (iv) Click on **Home** button to get back to the main menu when done.

5.0 Histopathology ward

1. When the application starts executing, the login screen will appear as follows:

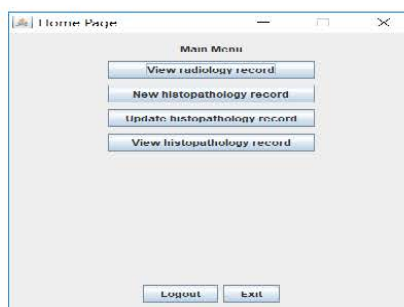


Please enter the following details to log into the application.

Username: histo

Password: histo

2. The main menu will appear as follows:



To start diagnosis a patient, user will have to first see record from General ward and Radiology Ward to know patient report. Click on the Add **view medical record** button and **view radiology record** button to proceed.

3. To add a new histopathology record for a patient, click on the **New histopathology record** button from main menu and the following page will appear:

User Manual Prostate Cancer Expert System 2018

- (i) Select the patient before adding histopathology record.
- (ii) Browse the IMAGE INFORMATION section by selecting the **Source Image**.
- (iii) Update the HISTOPATHOLOGY ASSESSMENT section by selecting necessary value according to Image Information.
- (iv) User may add in addition comments on the diagnosed result.
- (v) Click on **Diagnose** button to start diagnosing. A pop up window showing the result will be shown as following:

- (vi) Update the final decision from the dropdown.
- (vii) Click on the **Save** button to save the record.
- (viii) Click on **Home** button to get back to the main menu when done.

4. To update a histopathology record for a patient, click on the Update histopathology record button from main menu and the following page will appear:

User Manual Prostate Cancer Expert System 2018

- (i) Select the patient before adding histopathology record.
- (ii) Select the report before update histopathology record.
- (iii) Browse the IMAGE INFORMATION section by selecting the Source Image.
- (iv) Update the HISTOPATHOLOGY ASSESSMENT section by selecting necessary value according to Image Information.
- (v) User may add in addition comments on the diagnosed result.
- (vi) Click on **Diagnose** button to start diagnosing. A pop up window showing the result will be shown as following:

- (vii) Update the final decision from the dropdown.
- (viii) Click on the **Save** button to save the record.
- (ix) Click on the **Delete** button to remove patient record from the system.
- (x) Click on **Home** button to get back to the main menu when done.

5. button from main menu and the following page will appear:

Expert System UKM

User Manual Prostate Cancer Expert System 2018

- (i) Select the patient before view medical record.
- (ii) Select the report before view medical record.
- (iii) Click on the **Save pdf** button to preview report in document (pdf) from the system.
- (iv) Click on **Home** button to get back to the main menu when done.

LAMPIRAN I

CONTOH BORANG SOAL SELIDIK

Evaluation Form of Automation Integrated Diagnosis Of Prostate Cancer

Purpose

The purpose of the evaluation form is to summarize the results of a study of the Automation Integrated Diagnosis Of Prostate Cancer expertise acceptance criteria. The following table contains a question evaluation used in experiments.

PART A: PERSONAL INFORMATION

Instruction: Mark ✓ in the box provided

a) Level of study:

- 1) Undergraduate
- 2) Postgraduate
- 3) Pre- University
- 4) Teacher/Lecturer

b) Age:

- 5) 20 ~ 35
- 6) 35 ~ 45
- 7) 45 and Above

c) Position:

- 1) Professor
- 2) Associate Professor
- 3) Doctor
- 4) Others

d) Experience:

- 1) 1 ~ 3 Years
- 2) 3 ~ 6 Years
- 3) 6 ~ 9 Years
- 4) 10 Years & Above

c) Fields of study:

- 1) Science
- 2) Arts
- 3) Account
- 4) Others

d) Gender:

- 1) Male
- 2) Female

e) Religion:

- 1) Islam
- 2) Christian
- 3) Buddha
- 4) Hindu
- 5) Other

Herewith, IC No:, agrees / does not agree to participate in this Automation Integrated Diagnosis Of Prostate Cancer research conducted by Mr Muhammad Uzair Sahrin who is a Master candidate from Center for Artificial Intelligence Technology, Universiti Kebangsaan Malaysia.

Yours sincerely,

.....

()

Stamp:

Date:

Witnessed by,

.....

()

Stamp:

Date:

User Response for Typical Usability Design Factors

Project Identification

Project Name :

Date Created :

Researcher :

User Response for Typical Usability

Please rate from 1.0 to 5.0 marks.

Measurement	1	2	3	4	5	Comment
Ease of Use (1=not easy , 5 = very Easy)						
Starting the System						
Obtaining Explanation						
Help Facilities						
Interface Technique						
Nature of Question (1= not helpful , 5= extremely Helpful)						
Clarity of Term						
Clarity of Question						
Nature of Explanations (1=not useful , 5 =extremely useful)						
WHY Explanation						
Presentation of the Result (1=not easy , 5 = very Easy), (1=not confidence , 5 = extremely confidence)						
Ease to Follow						
Complete						
Helpful Visualization						
General Consideration(1=low , 5= high), (1=not useful , 5 =extremely useful)						
Speed of the system						
System is Useful						

Feedback from reviewers to validate system

Review Session	Score					Comment
	1= below Expectation, 2= slightly below expectation, 3= Acceptable, 4= slightly more than Expected, 5=above Expectation					
	1	2	3	4	5	
Knowledge Acquired						
Explanation facilities						
Help facilities						
System Result						
Recommended to Other						

General Comments (Wish List)
