

# PENGELASAN SOALAN PEPERIKSAAN BERLANDASKAN TAKSONOMI BLOOM MENGGUNAKAN PEMBELAJARAN MESIN

YEO YONG SHENG  
ASSOC PROF. DR. NAZLIA BINTI OMAR

*Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia*

## ABSTRAK

Tidak dapat dinafikan bahawa peperiksaan memainkan peranan yang sangat penting dalam kehidupan manusia. Peperiksaan perlu untuk menguji tahap kefahaman atas sesuatu perkara atau bidang tersebut melalui peperiksaan. Untuk memperoleh peperiksaan yang berkesan, penguji haruslah mempunyai garis panduan untuk membina soalan yang seimbang daripada tahap kognitif yang berlainan yang dapat menilai pelajar dengan efektif. Oleh itu, terdapatnya garis panduan Taksonomi Bloom yang digunapakai oleh pengajar pada hari ini. Walau bagaimanapun, pengklasifikasian secara manual berdasarkan Taksonomi Bloom adalah satu perkara yang amat mencabar dan memerlukan masa yang panjang. Justeru, kajian ini akan membina satu model pengklasifikasian soalan peperiksaan berlandaskan taksonomi bloom menggunakan teknik pembelajaran mesin. Kaedah pembelajaran mesin yang akan dijalankan dalam kajian ini ialah *Support Vector Machine(SVM)*, *Naïve Bayes*, *Random Forest*, and *K-Nearest Neighbour(KNN)*. Kaedah pembelajaran mesin yang dapat memberikan ketepatan yang paling tinggi akan dipilih. Untuk mendapatkan ketepatan yang lebih tinggi, set soalan data akan dijalankan pra-pemprosesan dan ciri pengekstrakan. Hasil daripada kajian ini akan digunakan untuk membina satu laman sesawang yang mesra pengguna kepada pengajar untuk mengklasifikasikan soalan peperiksaan mereka dengan mudah dan cepat. Satu laman sesawang yang mesra pengguna dibangunkan bagi memudahkan pengajar untuk mengklasifikasikan soalan peperiksaan mereka dengan mudah dan cepat dan melihat hasil analisis dari pengelas.

## 1 PENGENALAN

Peperiksaan perlu dirangka dengan berkesan supaya dapat menjadi satu kaedah yang dapat menggalakkan penyuburan dalam budaya ilmu, pembelajaran dan pemantapan budaya inovasi. Banyak kajian merumuskan bahawa peperiksaan merupakan penanda aras pencapaian semata-mata tidak dapat mendorong minat dan minda seseorang pelajar untuk bersifat kreatif dan inovatif. (Torgerson 1944).

Kertas peperiksaan yang berkualiti haruslah terdiri daripada pelbagai aspek dan tahap kesukaran untuk menguji dan menilai tahap pemahaman pelajar dalam sesuatu perkara yang diajar. Oleh itu, soalan peperiksaan haruslah seimbang daripada setiap aspek yang bersesuaian dengan isi kandungan yang diajar dan memenuhi objektif pembelajaran. Taksonomi berasal daripada dua perkataan bahasa Yunani iaitu *tassein* bermaksud mengklasifikasikan dan *nomia* yang bererti kaedah (Retno Utari). Taksonomi merupakan hirarki yang klasifikasi atas prinsip dasar atau aturan. Taksonomi dalam bidang pendidikan telah memberi kesedaran kepada para pengajar sejak lahirnya kurikulum bahawa tujuan pembelajaran harus dianalisis sebelum proses pembelajaran dan pengajaran berlangsung bagi mengenalpasti silibus yang penting dan apa yang tidak penting.

Benjamin Samuel Bloom iaitu seorang ahli psikologi dalam bidang pendidikan menggunakan istilah ini dan kemudian meneliti dan mengembangkan bidang berkaitan dengan kemampuan berfikir dalam proses pembelajaran. Beliau kemudian telah memperkenalkan satu hirarki kognitif dalam *Taxonomy Of Educational Objectives: Handbook I: Cognitive Domain* (New York: David McKay) yang dikenali sebagai taksonomi bloom pada hari ini. Untuk mempunyai peperiksaan yang lebih berkualiti, para pengajar pada hari ini telah dicadangkan untuk mengemukakan soalan peperiksaan berdasarkan taksonomi bloom.

Benjamin Bloom (1956) telah mengklasifikasikan aras tingkah laku keintelektualan dalam pembelajaran kepada tiga domain iaitu domain kognitif, domain afektif, dan domain psikomotor dalam tujuan pendidikan ini. Taksonomi Bloom lebih menitikberatkan dalam aspek kognitif atau pengetahuan seperti fakta, teori dan hukum yang dapat membantu guru dalam merancang pembelajaran dan pengajaran serta membuat soalan pelbagai aras dalam menguji tahap berfikir pelajar. Taksonomi Bloom tujuan untuk membina pemikiran, pemahaman dan kebolehan secara seimbang. Taksonomi Bloom mempunyai enam tahap kognitif yang mana setiap tahap mewakili serta menilai kemahiran yang berbeza.

## **2 PENYATAAN MASALAH**

Sehingga kini, Taksonomi Bloom masih lagi digunakan sebagai garis panduan yang agak penting dalam sektor pendidikan (Curzon & Tummons 2013). Walau bagaimanapun, terdapat juga segelintir soalan peperiksaan yang tidak dapat menggunakan garis panduan Taksonomi Bloom dengan tepat dan telah menyebabkan peperiksaan tersebut tidak dapat mencapai kesan yang terbaik kerana tidak seimbang daripada garis panduan Taksonomi Bloom. Hal ini kerana 6 peringkat kognitif berbeza dalam Taksonomi Bloom mempunyai kata kunci mereka yang tersendiri. Kata kunci dalam setiap peringkat meliputi skop yang agak besar dan ini telah menyebabkan kesusahan untuk manusia mengklasifikasikannya secara manual. Selain itu, alatan yang terdapat untuk mengklasifikasikan soalan memerlukan pakar-pakar komputer untuk mengendalikannya dan ia amat terhad dan tiada di pasaran.

## **3 OBJEKTIF KAJIAN**

Objektif utama projek ini adalah untuk membangunkan satu sistem pengelasan soalan peperiksaan Bahasa Inggeris kepada 6 kognitif yang terdapat dalam Taxonomi Bloom. Untuk mencapai objektif utama projek ini, beberapa langkah utama haruslah dijalankan seperti berikut:

1. Membangunkan model pembelajaran mesin dengan data yang dikumpul dalam pemrosesan Bahasa Tabii untuk mengklasifikasikan soalan-soalan peperiksaan berkualiti tinggi dan seimbang yang boleh menepati tahap kognitif yang berbeza.
2. Membina satu aplikasi web berintegrasi model pembelajaran mesin yang dipilih. Web aplikasi ini akan mengklasifikasikan dan menganalisis soalan peperiksaan yang disediakan.

## **4 METOD KAJIAN**

Fasa Permulaan Kajian iaitu perancangan projek pada awal kajian telah dijalankan untuk mengenalpasti isu-isu dan dapat memahami tajuk kajian dengan lebih lanjut. Kajian ini seterusnya terdiri daripada empat fasa iaitu Fasa Pengumpulan Data, Fasa Pra-Pemrosesan Data, Fasa Pengekstrakan Ciri, Fasa Pembinaan Model Pengelasan, Fasa Pengujian dan akhirnya Fasa Pembangunan Web Aplikasi untuk mencapai objektif kajian.

#### 4.1 Fasa Pengumpulan Data

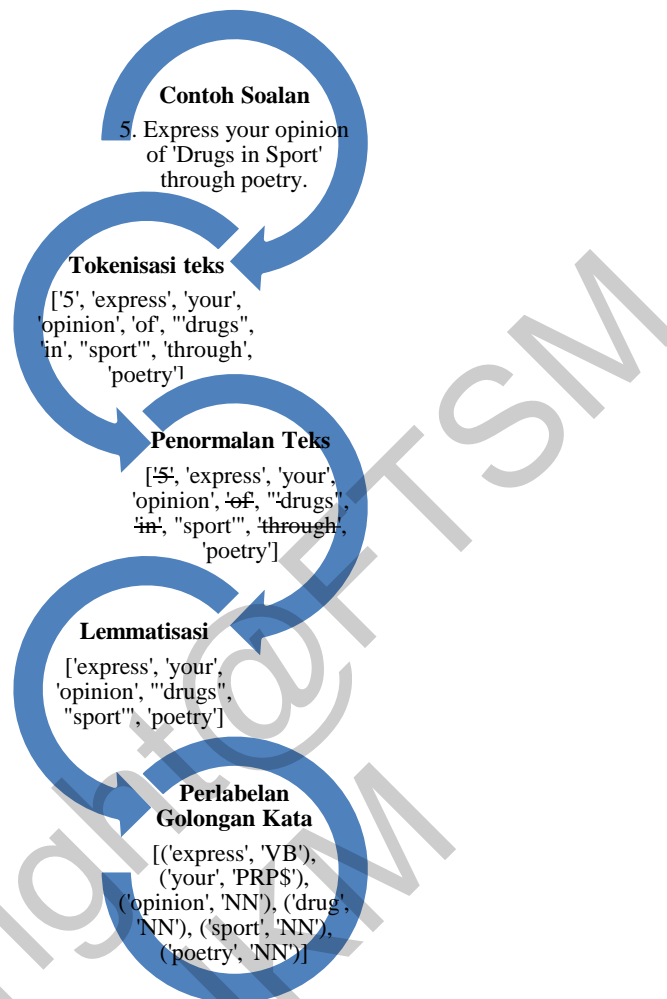
Kajian ini akan menggunakan dua domain set data terbuka untuk memastikan kestabilan dan keberkesanan teknik pengelasan dalam kajian ini. Set data yang pertama yang jumlahnya 141 soalan yang terdiri daripada tahap kognitif Taksonomi Bloom yang berbeza dikumpul dari laman web, buku dan juga kajian lepas (Haris & Omar). Set data yang kedua merangkumi jumlahnya 600 soalan (Yahya et al. 2012). Kedua-dua set data ini telah dahulu dilabelkan kepada enam tahap kognitif dalam Taksonomi Bloom iaitu: (1) *knowledge* (pengetahuan), (2) *comprehension* (pemahaman), (3) *application* (penerapan), (4) *analysis*(analisis), (5) *synthesis*(sintesis) dan (6) *evaluation*(penilaian).

#### 4.2 Fasa Pra-Pemprosesan Data

Pra-pemprosesan data untuk data yang tidak berstruktur amat penting sebelum soalan diproses menggunakan pemprosesan Bahasa Tabii bagi menyingkirkan data yang tidak penting, berganda, tidak relevan dan data yang bising ataupun mengemaskan data supaya dapat memudahkan pengendalian data untuk proses seterusnya.

Tokenisasi ialah satu proses digunakan untuk memecahkan data ayat menjadi data perkataan. Kajian ini telah mendapati bahawa pakej *Keras* boleh memperoleh hasil tokenisasi yang paling baik kerana boleh memperolehi hasil set data yang perkataan dalam set data telah ditukar menjadi huruf kecil dan pemansuhan tanda baca menjadikan hasil bilangan token yang paling kecil. Ini dapat memudahkan lagi proses pra-pemprosesan data seterusnya. Pemansuhan nombor dan kata hentiakan dilakukan dalam penormalan teks. Walau bagaimanapun, bukannya semua kata henti dalam pakej nltk akan digunapakai pada kajian ini kerana sesetengah kata henti dalam soalan memainkan peranan penting dalam klasifikasi tahap kognitif.

Seterusnya akan menjalankan proses lemmatisasi yang merupakan satu proses dalam menyingkirkan kata imbuhan awalan ataupun kata imbuhan akhiran dalam satu perkataan dan mendapatkan lemma perkataan. Langkah terakhir dalam pra-pemprosesan data ialah Perlabelan Golongan Kata(PGK) yang merupakan proses pemberian penanda label pada setiap perkataan di dalam corpus. Perlabelan Golongan Kata dapat membantu dalam proses pengestrakan ciri seterusnya. Rajah 1 akan menunjukkan set data selepas pra-pemprosesan data secara lengkap.



Rajah 1 Contoh Proses Algoritma Pra-pemprosesan data

### 4.3 Fasa Pengekstrakan Ciri

Dalam kaedah pembelajaran mesin, setiap set data pada mulanya hendaklah diubah menjadi perwakilan vektor dahulu untuk mengklasifikasikannya. Proses ini dipanggil pengekstrakan ciri. *Bag-of-Words* ialah salah satu cara yang umum untuk memproseskan bahasa semula jadi, di mana setiap ciri ialah satu perkataan dalam korpus, yang boleh bertukar menjadi nilai Boolean, menunjukkan sama ada perkataan tertentu wujud dalam soalan, dan yang juga boleh menjadi integer atau nombor nyata, serta menunjukkan frekuensi setiap perkataan tanpa mengira posisi perkataan tersebut. Setiap perkataan dalam ayat akan diekstrakan menjadi ciri dan dokumen tersebut akan bertukar menjadi vektor.

Teknik yang advance daripada *Bag-of-Words* banyak digunakan untuk mengekstrakan ciri data teks adalah TF-IDF telah digunakan dalam kajian ini yang merupakan ukuran statistik yang menggambarkan pentingnya suatu istilah terhadap sebuah dokumen dalam sebuah kumpulan atau korpus.

Kata kerja dalam setiap soalan merupakan kata kunci dalam klasifikasi kerana ia memainkan peranan yang sangat penting dalam klasifikasi soalan tersebut ke tahap kognitif Taksonomi Bloom. Kata kunci biasanya akan terdapat pada permulaan setiap soalan ataupun di tengah-tengah setiap soalan. Kadangkala satu soalan akan mempunyai lebih daripada satu kata kunci, Oleh itu, ciri taksonomi ialah mendapatkan dan menggunakan setiap kata kunci yang terdapat dalam soalan tersebut.

#### **4.4 Fasa Pembinaan Model Pengelasan Klasifikasi**

Proses yang dijalankan pada fasa ini adalah membangunkan model pengelasan yang sesuai bagi menyelesaikan masalah yang dikenalpasti dalam kajian ini. Untuk mengklasifikasikan soalan peperiksaan berlandaskan Taksonomi Bloom, kaedah pembelajaran mesin yang diselia yang paling umum telah digunakan iaitu SVM, KNN, Naïve Bayes dan Random Forest akan digunakan dalam fasa ini. Keempat-empat model ini akan dibandingkan dan akan menggunakan model yang boleh mendapatkan ketepatan yang paling tinggi. Fasa ini dibahagikan kepada dua langkah utama iaitu melatih model pada mulanya dan kemudian menguji model tersebut. Kedua-dua langkah ini akan menggunakan set data dengan mengagihkan set data yang sedia ada menjadikan 80% data latihan dan 20% data ujian.

#### **4.5 Fasa Pengujian**

Fasa pengujian ini akan menilai prestasi model pengelasan menggunakan parameter yang berlainan seperti kejituan (*precision*), dapatan (*recall*) dan ukuran-F. Beberapa istilah harus dikenali sebelum menggunakan parameter tersebut dalam kajian ini iaitu Positif Benar(TP), Positif Palsu(FP) dan Negatif Palsu(FN). Positif Benar(TP) ialah nombor soalan yang diklasifikasikan dengan betul oleh sistem berdasarkan tahap kognitif Taksonomi Bloom, Positif Palsu(FP) ialah nombor soalan yang diklasifikasikan dengan salah manakala Negatif Palsu(FN) ialah nombor soalan yang tidak diklasifikasikan. Nilai pengujian yang didapatkan dari parameter kejituan (*precision*), dapatan (*recall*) dan ukuran-F adalah di antara 0 dengan 1.

Semakin tinggi nilai pengujian yang semakin dekat dengan nilai 1 bermaksud keberkesanan sistem yang semakin baik dan sebaliknya semakin rendahnya nilai pengujian yang semakin dekat dengan nilai 0 bermaksud keberkesanan sistem yang tidak baik.

#### 4.6 Fasa Pembangunan Web Aplikasi

Selepas memperoleh model pembelajaran mesin yang terbaik. Langkah seterusnya adalah membina aplikasi web yang dapat menghubungkan model dan menghasilkan ramalan mengenai data baru dalam masa nyata. Oleh itu, kajian ini telah menggunakan “Gradio” untuk membangunkan web aplikasi. Gradio ialah *library* python sumber terbuka yang untuk mencipta komponen UI yang mudah digunakan, boleh disesuaikan untuk model ML, sebarang API, atau sebarang fungsi. Dengan adanya web aplikasi, objektif kajian ini dapat dicapai.

### 5 HASIL KAJIAN

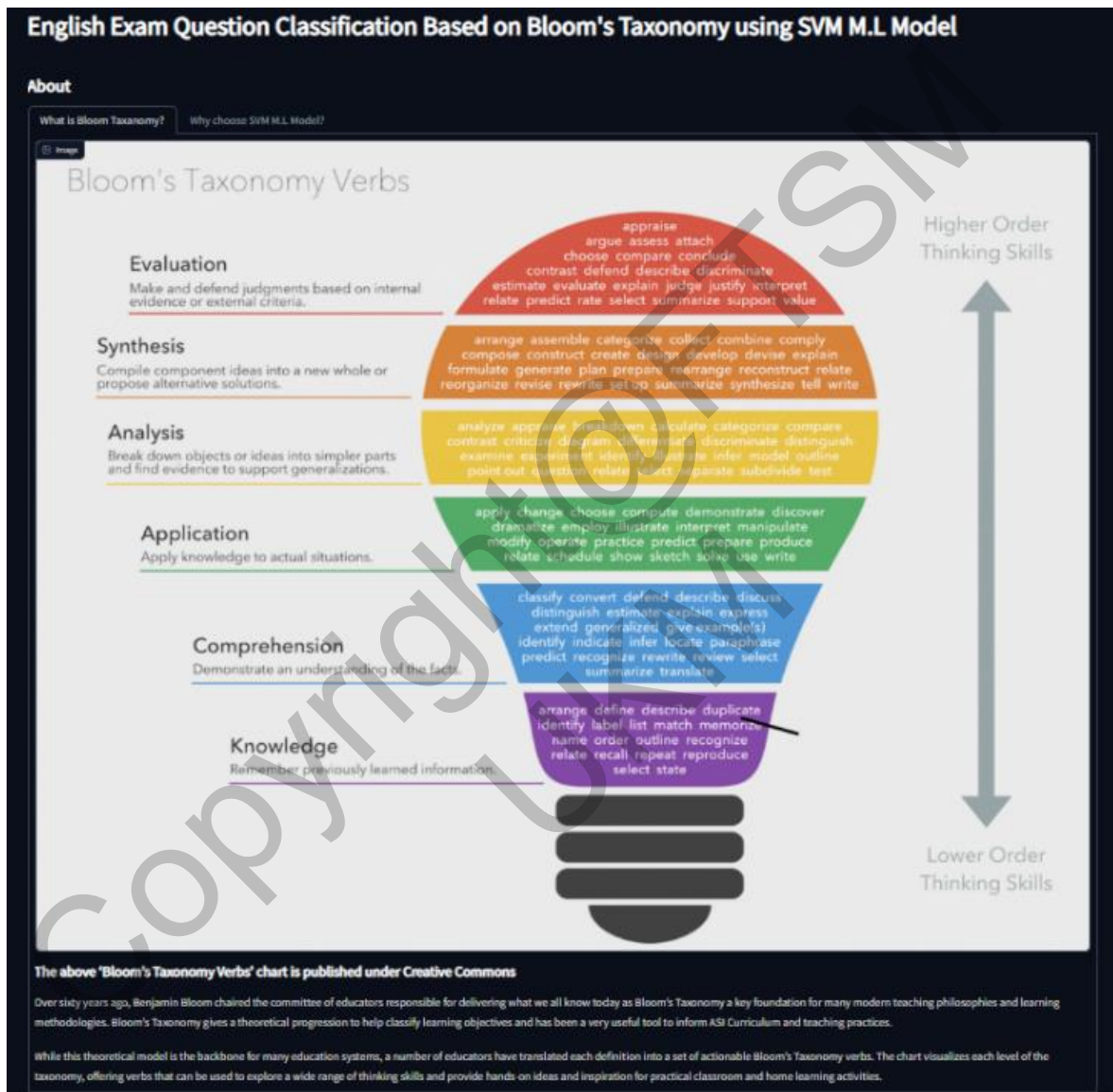
Setelah memperolehi prestasi model-model pengelasan seperti yang ditunjukkan di atas, hasil ketepatan akan dibandingkan di Jadual 1.

Model Pengelasan	SVM	KNN	Naïve Bayes	Random Forest
Ketepatan	77.18%	76.51%	70.47%	72.48%

Jadual 1 Ketepatan Model-model Pengelasan

Berdasarkan Jadual 1, model pengelasan membawa ketepatan yang paling tinggi di antara model pengelasan yang digunakan. Hasil ketepatan yang didapatkan dari model pengelasan SVM juga lebih tinggi berbanding dengan kajian lepas. Hal ini demikian kajian ini telah menjalankan ciri taksonomi sebagai pengekstrakan ciri dalam pemprosesan data yang dapat meningkatkan lagi keberkesanan pengelasan model. Oleh itu, sistem ini akan seterusnya dibina menggunakan model pengelasan SVM.

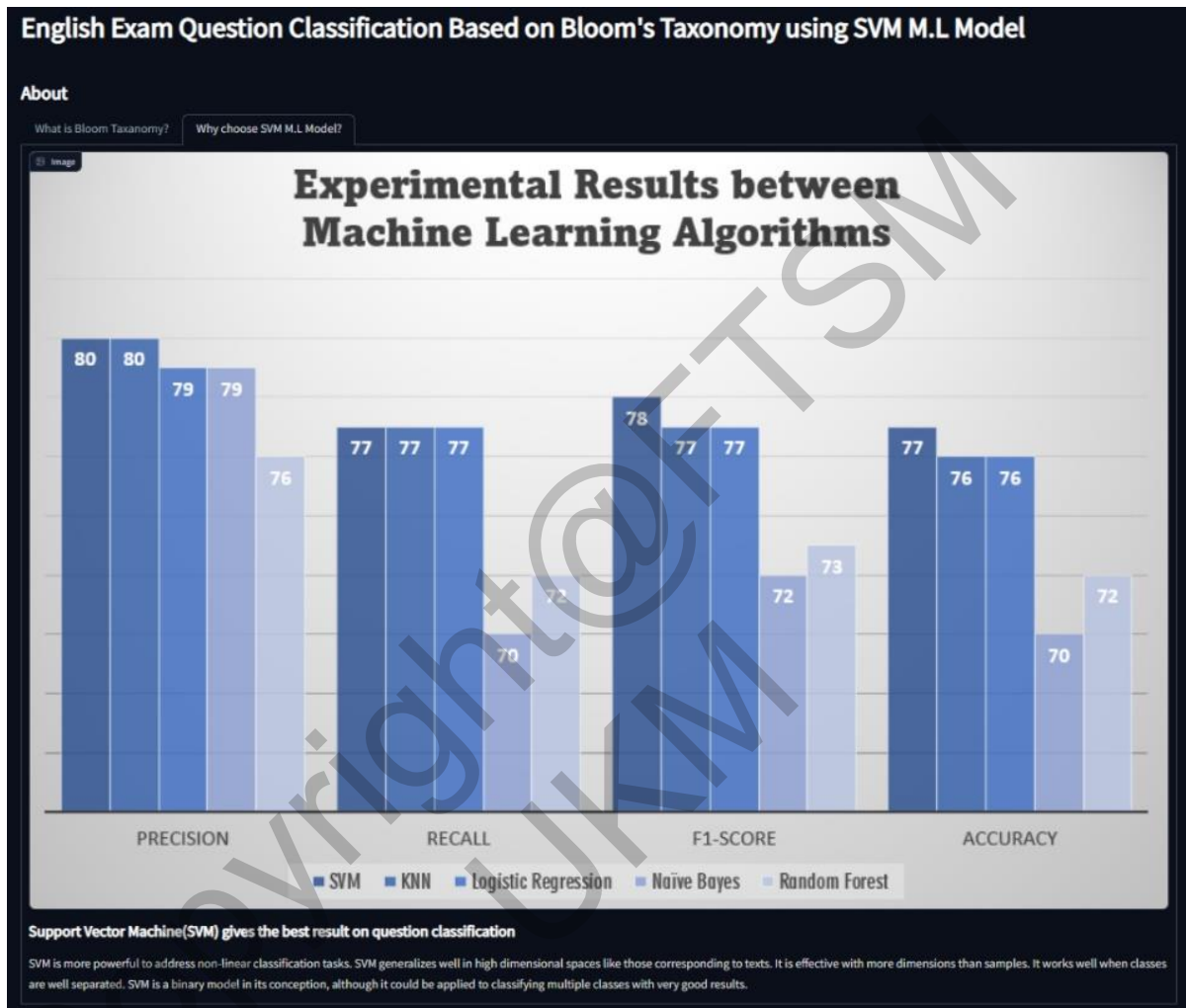
Web aplikasi yang dibina dibahagikan kepada dua bahagian iaitu bahagian informasi dan bahagian pengelasan soalan. Taksonomi Bloom telah diterangkan supaya pengguna dapat mengenali panduan untuk membina soalan yang dapat memenuhi tahap kognitif Taksonomi Bloom seperti yang ditunjukkan dalam Rajah 2.



Rajah 2 Pengenalan Taksonomi Bloom pada web aplikasi



Kaedah pembelajaran mesin yang digunakan untuk membina pengklasifikasian soalan juga diterangkan secara umum dan membandingkan kaedah pembelajaran mesin yang telah diuji sebelum membangunkan web aplikasi seperti yang ditunjukkan dalam Rajah 3.



Rajah 3 Pengenalan Model Pembelajaran Mesin pada web aplikasi

Untuk mengenalpastikan tahap kognif Taksonomi Bloom soalan peperiksaan, pengguna boleh mentaipkan soalan peperiksaan satu demi satu dahulu sebelum diletakkan dalam set soalan peperiksaan seperti yang ditunjukkan dalam Rajah 4.

Start classifying your exam question/s now~

Question Set Question(.txt file only)

Please type your question here:

Define compound interest.

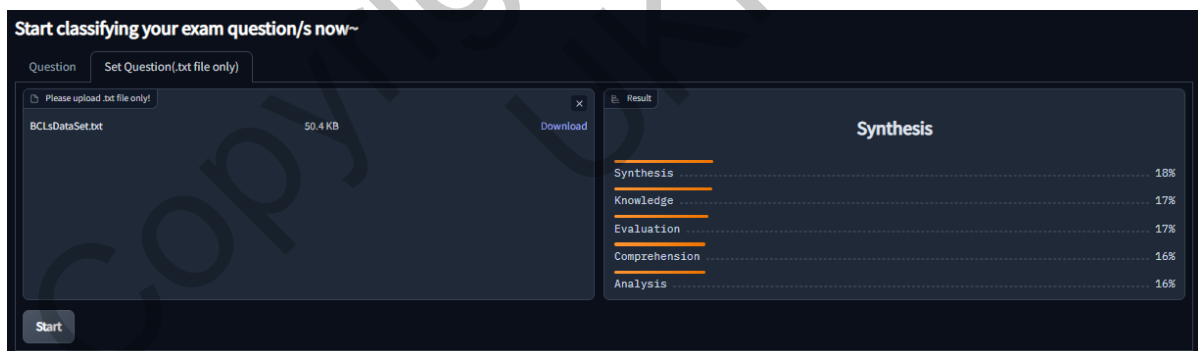
Result

Knowledge

Start

Rajah 4 Hasil pengklasifikasian satu soalan peperiksaan

Pengguna juga dapat mengklasifikasikan set soalan peperiksaan yang disediakan dalam format *.txt* dan dapat mengetahui pembahagian tahap kognitif untuk soalan peperiksaan yang disediakan seperti Rajah 5.



Rajah 5 Hasil pengklasifikasian satu set soalan peperiksaan

## 6 KESIMPULAN

Secara keseluruhannya, pembangunan web aplikasi berintergrasi model pembelajaran mesin telah disiapkan dalam tempoh masa yang telah ditetapkan. Pembangunan ini juga menepati objektif kajian yang telah disasarkan. Sistem ini dengan tujuan untuk memanfaatkan pengguna dapat mengklasifikasikan dan menganalisis soalan peperiksaan yang disediakan. Keperluan sistem dan reka bentuk sistem telah dijelaskan dengan nyata.

Algoritma mesin pembelajaran yang dapat memberikan ketepatan yang paling tinggi akan dipilih untuk membangunkan sistem aplikasi web. Pengekstrakan ciri seperti ciri taksonomi juga telah menambahkan lagi prestasi pengklasifikasian sistem. Oleh itu, prestasi sistem yang dibangunkan dalam projek ini adalah amat baik. Selain itu, web aplikasi ini sangat mudah digunakan oleh semua pengguna kerana pengguna boleh memilih untuk mengklasifikasikan set soalan peperiksaan ataupun satu demi satu soalan peperiksaan.

Walaupun, kekangan sistem ini ialah sistem ini hanya terhad untuk menganalisis soalan peperiksaan dalam Bahasa Inggeris dan juga set soalan hendaklah dalam bentuk .txt format sebelum dimuat naik ke sistem untuk dianalisis.

Untuk penambahbaikan masa hadapan sistem ini, sistem ini boleh ditambahbaikan lagi dengan menambah kepelbagaian bahasa soalan peperiksaan untuk dianalisis. Panduan untuk menganalisis soalan juga haruslah dipelbagaikan tetapi bukan sahaja terhad kepada Taksonomi Bloom.

## 7 RUJUKAN

- Zadina, J.N., Rita Smilkstein, Deborah Daiek & Anter 2014. *College Reading: The Science and Strategies of Expert Readers* Curzon dan Tummons (2013)
- Ghanem Nayef, E., Rosila, N., Yaacob, N. & Ismail, H. N. 2013. Taxonomies of Educational Objective Domain. *International Journal of Academic Research in Business and Social Sciences* 3(9): 2222–6990.
- Swart, A. J. 2010. Evaluation of Final Examination Papers in Engineering : A Case Study Using Bloom ' s Taxonomy. *IEEE Transactions on Education*.
- Krathwohl, D. R. 2002. A Revision of Bloom ' s Taxonomy: An Overview. *THEORY INTO PRACTICE* 41(4), pp 212–218
- Jayalakshmi, S. & Ananthi, S. 2015. Question Classification: A Review of State of the-Art Algorithms and Approaches. *Indian Journal of Science and Technology*, Vol 8(29)
- Shaalan, K. 2010. Rule-based Approach in Arabic Natural Language Processing. *International Journal on Information and Communication Technologies*, Vol. 3, No. 3.
- Marr, B. 2017. Supervised V Unsupervised Machine Learning -- What's The Difference? <https://www.forbes.com/sites/bernardmarr/2017/03/16/supervised-vunsupervised-machine-learning-whats-the-difference/#1fc28df485d8>, 27 Oct 2021
- Cortes, C. & Vapnik, V. N. 1995. *The Nature of Statistical Learning*
- Fix, E. dan Hodges, Jr. J.L., "Discriminatory Analysis: Non-parametric Discrimination: Small Sample Performance," Report No. 11, USAF School of Aviation Medicine, Randolph Field, Texas, 1952.
- Hassan, M. R., Hossain, M. M., Bailey, J., dan Ramamohanarao, K. "Improving k-nearest neighbor classification with distance functions based on receiver operating

characteristics,”

in W. Daelemans et al. (Eds.): ECML PKDD 2008, Part I, LNAI 5211, pp. 489-504. 2008.

K. Jayakodi, M. Bandara, I. Perera and D. Meedeniya, "WordNet and cosine similarity based classifier of exam questions using Bloom's taxonomy", *Int. J. Emerg. Technol. Learn.*, vol. 11, no. 4, pp. 142, Apr. 2016.

Sufi Haris Syahidah and Omar Nazlia, "A Rule-based Approach in Bloom's Taxonomy Question Classification Through Natural Language Processing", *IEEE 2012*, pp. 410-414.

Yusof Norazah and Jing Hui Chai, "Determination of Blooms Cognitive Level of Questions Item Using Artificial Neural Network", *International Conference on Intelligent System Design and Applicatons*, pp. 866-870, October 2010.

K. Osadi, M. Fernando and W. Welgama, "Ensemble classifier based approach for classification of examination questions into Bloom's taxonomy cognitive levels", *Int. J. Comput. Appl.*, vol. 162, no. 4, pp. 76-92, Mar. 2017.

Pincay, J. & Ochoa, X. 2015. Automatic Classification of Answers to Discussion Forums According to the Cognitive Domain of Bloom's Taxonomy using Text Mining and a Bayesian Classifier

Anderson, L. W., & Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing, Abridged Edition. Boston, MA: Allyn and Bacon

Yahya, A. A., Toukal, Z. & Osman, A. 2012. Bloom's Taxonomy – Based Classification for Item Bank Questions Using Support Vector Machines Bloom's Taxonomy – Based Classification for Item Bank. *Modern Advances in Intelligent Systems and Tools, SCI 431*, hlm. 135–140. Springer-Verlag Berlin Heidelberg 2012.

Haris, S. S. & Omar, N. 2015. Bloom's taxonomy question categorization using rules and N-gram approach. *Journal of Theoretical and Applied Information Technology* 76(3): 401–407.

Yeo Yong Sheng (A175749)  
Assoc Prof. Dr. Nazlia Binti Omar  
Fakulti Teknologi & Sains Maklumat,  
Universiti Kebangsaan Malaysia