

PERSEPSI PELAJAR TERHADAP PEMBELAJARAN DALAM TALIAN DAN PEMBELAJARAN LUAR TALIAN MENGGUNAKAN ASPEK ANALISIS SENTIMEN

ZULAIKHA DAWAMI
LAILATUL QADRI ZAKARIA

Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

Pandemik Koronavirus 2019 (Covid-19) menjadi masalah utama yang telah merebak secara global pada awal tahun 2020, di mana ia telah memberi kesan teruk kepada proses pembelajaran di pelbagai universiti dan institusi seluruh negara. Hal ini menyebabkan, pembelajaran dalam talian dijadikan sebagai medium pembelajaran bagi mengelakkan sentuhan fizikal semasa proses pembelajaran dan juga mengelakkan virus daripada terus merebak. Ramai pelajar cenderung menggunakan media sosial seperti aplikasi Twitter sebagai platform untuk meluahkan masalah dan perasaan mereka tentang perspektif pengalaman pembelajaran dalam talian dan luar talian. Projek ini bertujuan untuk menjalankan analisis persepsi pelajar sama ada mereka lebih cenderung ke arah pembelajaran dalam talian atau pembelajaran luar talian dengan mengumpul data berdasarkan ciapan yang ditulis oleh pengguna di Twitter. Justeru, API Twitter berperanan untuk mengumpul ciapan pengguna daripada Twitter sebagai set data. Teknik Pemprosesan Bahasa Tabii (NLP), Pembelajaran Mesin (ML) dan klasifikasi akan digunakan dalam projek ini. Selain itu, teknik pembersihan data seperti normalisasi teks akan menyingkirkan emoji, nombor, tanda baca, kata henti, ruang putih dan karakter tidak bermakna akan dilaksanakan. Kemudian, kaedah pengekstrakan ciri seperti TF-IDF (Frekuensi Istilah – Frekuensi Dokumen Terbalik akan dilakukan. Seterusnya, kaedah analisis sentimen memberikan hasil yang melibatkan emosi dan perasaan pelajar sama ada positif, neutral atau negatif. Teknik pembelajaran mesin seperti Naïve Bayes akan digunakan untuk mengklasifikasikan kategori berkenaan topik yang dibincangkan oleh pengguna. Sentimen pengguna terhadap perspektif pembelajaran dalam talian dan luar talian akan dianalisis dalam projek ini dan boleh diguna pakai oleh pakar pendidikan sebagai rujukan pada masa hadapan bagi menyelesaikan masalah terhadap proses pembelajaran terbaik yang perlu digunakan dalam kalangan pelajar ketika ini. Harapan untuk projek ini adalah, agar dapat mengetahui tahap keberkesanan dan keinginan pelajar terhadap proses pembelajaran terbaik yang perlu dilaksanakan sepanjang wabak Covid-19 ini yang tidak pasti bila penghujungnya.

1 PENGENALAN

Pembelajaran secara atas talian (*online learning*) atau istilah lainnya e-pembelajaran (*elearning*) merupakan kaedah pembelajaran menggunakan rangkaian teknologi internet yang semakin berkembang pesat dari semasa ke semasa sejajar dengan kemajuan teknologi di negara Malaysia. Berikutan daripada penularan Covid-19 yang dikesan oleh Pertubuhan Kesihatan Sedunia (WHO) sekitar bulan November 2019 di Bandar Wuhan, China dan ia telah merebak ke serata negara sehingga ke hari ini. Susulan daripada penularan Covid-19 ini, kerajaan Malaysia telah memperkenalkan Perintah Kawalan Pergerakan (PKP) pada 18 Mac 2020 yang lalu, yang menuntut seluruh masyarakat untuk kekal di kediaman masing-masing bagi mengawal penularan wabak Covid-19 daripada terus merebak. Disebabkan hal itu, sektor

pendidikan juga turut menerima kesan daripada pelaksanaan PKP ini. Justeru, memberi impak yang besar kepada semua institusi dan universiti di seluruh negara. Akibatnya, proses pengajaran dan pembelajaran secara bersemuka yang telah dilaksanakan sejak dahulu lagi terpaksa ditangguhkan bagi mengengkang penularan wabak ini daripada tersebar luas akibat sentuhan fizikal dan interaksi sekeliling semasa pembelajaran.

Kementerian Pendidikan Malaysia (KPM) telah memutuskan supaya semua pembelajaran dilakukan secara dalam talian bagi seluruh negara termasuklah peringkat rendah, peringkat menengah mahupun institusi pengajian tinggi. Ia sebagai satu usaha membolehkan pelajar tidak ketinggalan dalam mengikuti sesi pembelajaran mereka. Dengan ini, bilangan pelajar yang menggunakan kaedah pembelajaran secara atas talian menunjukkan peningkatan secara drastik. Oleh itu, ramai pelajar dan tenaga pengajar beralih ke platform digital seperti *Microsoft Teams*, *Google Meet*, dan juga *Zoom* sepanjang sesi pembelajaran atas talian. Platform ini memang menjadi salah sebuah medium yang popular dalam kalangan pelajar kerana penggunaannya yang mudah dicapai, akses percuma dan juga sesuai digunakan oleh semua golongan usia. Selari dengan pembelajaran abad ke-21 (PAK21), tuntutan pembelajaran dalam talian yang dituntut dalam norma baharu ini bukanlah sesuatu perkara yang janggal dalam pendidikan. Objektif PAK21 ialah melatih pelajar menjadi lebih komunikatif, kolaboratif, kreatif dan kritis berfikir, yang sudah pasti pelajar dan tenaga pengajar perlu lebih maju ke hadapan dalam bidang teknologi agar dapat menimba dan berkongsi ilmu dengan lebih efektif.

Bagi pembelajaran secara luar talian (*offline learning*) atau lebih dikenali dengan belajar secara bersemuka dan merupakan proses pembelajaran secara langsung dan terus antara tenaga pengajar, pensyarah dengan pelajar. Pembelajaran ini juga berlaku sama ada sehala mahupun dua hala iaitu (*two-way communication*) di mana tenaga pengajar menjalankan pengajaran secara bersemuka di dalam kelas atau ruang pembelajaran. Melalui kaedah ini, pelajar dapat berinteraksi dengan pensyarah secara langsung dan secara terus jika mempunyai sebarang keraguan terhadap perkara yang diajar. Seperti sedia maklum, pembelajaran ini telah menjadi kebiasaan dan dilaksanakan sejak dari dulu lagi oleh semua pelajar di seluruh negara sehinggalah Malaysia dilanda wabak Covid-19 dan kaedah ini perlu ditangguh buat sementara waktu. Namun begitu, pada permulaan bulan Oktober 2021, kes jangkitan Covid-19 di negara Malaysia telah menyaksikan penurunan dan pada bulan ini juga, banyak sektor seperti sektor perniagaan, sektor hiburan dan juga sektor pendidikan dibuka secara berperingkat berikutan dengan 90% kadar vaksinasi penuh oleh golongan dewasa di Malaysia yang terus meningkat sehingga kini. Dengan itu, Kementerian Kesihatan Malaysia (KKM), menyasarkan negara ini

bakal memasuki fasa transisi 'hidup bersama virus COVID-19' sebelum beralih ke fasa endemik, dengan bilangan kes berada dalam kadar yang dijangkakan dan tidak membebani kesihatan awam, bergantung pada kekuatan dan tempoh masa perlindungan ketahanan daripada vaksinasi dan infeksi semula jadi.

Dalam peralihan fasa pandemik kepada fasa endemik di Malaysia, beberapa perkara dapat dilihat berkaitan sistem pembelajaran di Malaysia di mana terdapat beberapa pengecualian yang dibuat kepada belajar untuk belajar secara besemuka. Perkara ini lebih tertumpu kepada pelajar yang memerlukan pengajaran dan pembelajaran (PdP) secara hibrid di kampus, memerlukan makmal bagi tujuan penyelidikan, kerja-kerja amali, klinikal, studio, bengkel dan peralatan khas. Kes ini telah diteliti dan disemak semula oleh pihak KPM bagi memberi kelonggaran kepada kaedah pembelajaran bagi kepentingan pelajar dalam menyambung pengajian dengan suasana yang kondusif dan sesuai dalam menyiapkan tugas. Selain daripada golongan pengecualian ini, pelajar yang lain masih perlu belajar secara maya walaupun terdapat sesetengah pelajar institusi pengajian tinggi sudah pulang ke kampus masing-masing.

Seterusnya, bagi pengendalian kelas secara atas talian, kebolehcapaian internet menjadi salah satu masalah atau halangan terbesar kerana terdapat pelajar yang gagal untuk mendapat capaian internet dengan lancar akibat kawasan penempatan atau rumah pelajar yang kurang mendapat sambungan internet. Pengagihan tugas juga sering diberikan secara atas talian melalui aplikasi *Whatsapp*, *Telegram*, dan juga dari laman web atau portal institusi masing-masing. Masalah kesihatan juga berlaku akibat tekanan yang dialami oleh pelajar sewaktu sesi pembelajaran. Sebaliknya, terdapat juga pelajar yang memberi respons baik terhadap kaedah pembelajaran secara atas talian ini kerana mereka boleh berdekatan dengan keluarga masing-masing serta menjimatkan waktu, duit dan tenaga. Selain itu, mereka yang berada di rumah juga tidak perlu memikirkan kos makanan harian, tempat tinggal dan juga pengangkutan. Kemudian, bagi pengendalian kelas secara luar talian, sesetengah pelajar beranggapan ini adalah kaedah yang baik kerana mereka mempunyai teman dan juga pensyarah yang dapat membantu mereka secara langsung jika mereka terdapat masalah dalam pembelajaran. Suasana belajar yang kondusif dan selesa juga merupakan faktor penting pelajar lebih berminat untuk belajar secara luar talian. Walau bagaimanapun, kaedah pembelajaran secara luar talian ini, memerlukan tenaga yang lebih banyak untuk bergerak dari satu kelas ke kelas yang lain dan juga kos sara hidup yang tinggi ketika melanjutkan sesi pembelajaran di kampus atau institusi masing-masing.

Oleh hal yang demikian, isu pembelajaran dalam talian dan luar talian banyak dibicarakan dan dibahas oleh masyarakat di Malaysia terutama sekali pelajar di Malaysia terhadap cara yang sepatutnya dilakukan dalam meneruskan pembelajaran. Setiap pelajar kini, lebih menyesuaikan diri dengan keadaan semasa dan cuba untuk menghadapi ketidakpastian proses pembelajaran pada masa kini berikutan kesan penularan Covid-19 yang tidak tahu bila penghujungnya. Pada zaman ini, kebanyakan anak muda menggunakan media sosial sebagai platform untuk meluahkan pendapat, perasaan, pandangan dan pengkritikan mengenai pembelajaran dalam talian dan juga pembelajaran secara bersemuka di media sosial seperti Twitter. Oleh itu, kajian ini dapat mempertimbangkan pandangan pelajar bagi lebih memahami luahan dan pendapat mereka supaya dapat digunakan untuk menganalisis sentimen tentang pembelajaran dalam talian dan luar talian. Proses analisis sentimen ini dapat mengenal pasti bentuk data supaya lebih mudah dikelaskan dan difahami. Secara umumnya, proses kajian ini dapat mengetahui sentimen sesuatu teks dan dapat mengenal pasti corak positif, negatif dan neutral tentang kandungan yang dikongsi dalam media sosial tersebut. Data dan maklumat yang diperoleh dapat membantu pakar pendidikan seperti Kementerian Pengajian Tinggi (KPT) dan juga KPM dalam menentukan cara yang terbaik dalam menangani masalah dan krisis yang dihadapi oleh pelajar serta berharap membawa perubahan kepada situasi yang lebih positif dan berkesan.

2 PENYATAAN MASALAH

Dalam kajian ini, topik ini dilihat sebagai satu cabaran besar kepada pelajar untuk menjalankan pembelajaran secara atas talian. Oleh itu, pengumpulan data dan juga maklumat akan diambil dari aplikasi Twitter iaitu salah satu media sosial yang berkaitan rapat dengan masyarakat di Malaysia dan juga ramai golongan muda seperti pelajar menggunakan aplikasi ini. Twitter digunakan sebagai medium pengumpulan data kerana kebanyakan pelajar menggunakannya untuk meluahkan perasaan mereka terhadap belajar secara atas talian dan juga ramai pelajar yang menyatakan keinginan mereka dalam ciapan supaya melaksanakan semula pembelajaran secara bersemuka. Selain itu, data-data yang terdapat di dalam ciapan dapat membantu untuk mengklasifikasikannya kepada positif, negatif dan juga neutral sentimen. Dengan had 280 aksara, sekatan saiz mesej ini menggalakkan penggunaan bahasa yang bijak, menjadikan *tweets* sangat muah diimbas, dan juga sangat mencabar bagi pengguna Twitter untuk menulis dengan baik. Namun begitu, kajian ada membuat penilaian bahawa pemilihan data dari Twitter mampu memberi kejayaan dalam menentukan strategi dan juga penyelesaian masalah kerana

kebanyakan ciapan yang digunakan adalah tertumpu kepada teks. Terdapat juga beberapa masalah yang boleh dikenalpasti dalam proses pengumpulan data di mana percampuran penggunaan bahasa yang digunakan oleh pelajar Malaysia dalam membuat ulasan terdiri dalam dwibahasa iaitu Bahasa Melayu dan juga Bahasa Inggeris. Penggunaan kata singkatan, tanda pagar (*hashtag*), tanda baca, aksara khas dan juga emoji mesej yang dikongsikan oleh pengguna sukar untuk difahami.

Selain itu juga, terdapat beberapa kajian yang dilakukan terhadap pembelajaran secara atas talian dan pembelajaran luar talian tetapi tiada dataset yang sedia ada di mana kebanyakan dataset hanya tertumpu kepada pencarian pembelajaran secara atas talian sahaja dan kurang data sedia ada tentang pembelajaran secara luar talian. Masalah dalam mencari latar belakang kajian serta pengelasan aspek terhadap pembelajaran dalam dan luar talian juga sukar dijumpai.

Seterusnya, terdapat kajian yang hampir sama dijalankan oleh Anieka Horo, Nupur Biswas dan Jagruti Das, 2020 tentang persepsi pelajar terhadap pembelajaran secara atas talian dan luar talian semasa sekatan Covid-19. Namun, dalam kajian tersebut pengkaji hanya menggunakan teknik yang biasa dengan mengumpulkan maklumat melalui tinjauan dan mendapatkan hasil akhir dari peratusan carta pai. Kaedah sentimen analisis tidak digunakan. Dalam kajian ini, sentimen analisis yang digunakan mampu untuk memberikan keputusan sentimen yang lebih baik dan lebih tepat.

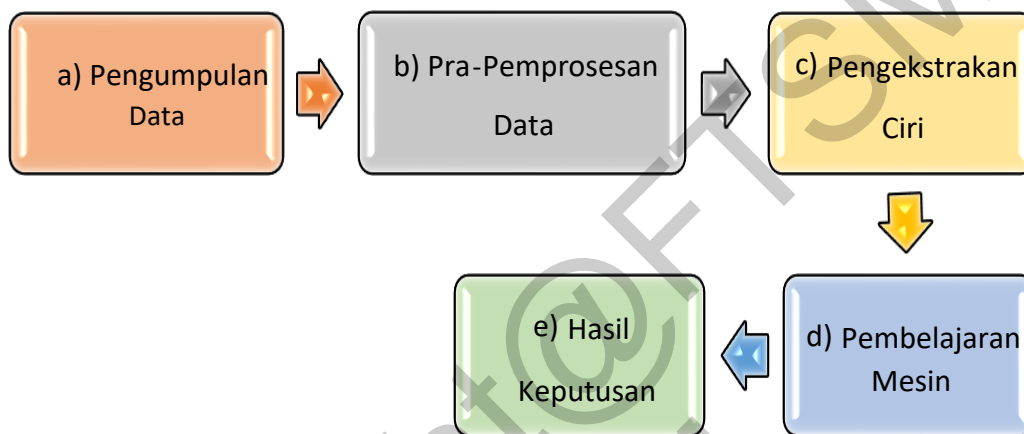
3 OBJEKTIF KAJIAN

Kajian ini ini memfokuskan pembangunan model dalam mengenal pasti persepsi pelajar terhadap pembelajaran dalam talian dan pembelajaran luar talian berikutan wabak virus Covid-19 yang merebak di Malaysia. Terdapat beberapa objektif bagi kajian ini iaitu:

1. Membangunkan dataset yang terdiri daripada perspektif pelajar terhadap pembelajaran secara dalam talian atau luar talian.
2. Mengelaskan aspek berkaitan pembelajaran dalam talian dan pembelajaran luar talian menggunakan teknik pembelajaran mesin.
3. Mengenal pasti sentimen pelajar terhadap pembelajaran secara dalam talian atau luar talian menggunakan analisis sentimen.

4 METOD KAJIAN

Model metodologi akan digunakan untuk melaksanakan algoritma dan model ini merupakan Model Klasifikasi Aspek Sentimen (*Sentiment Aspect Classification Model*). Model ini mudah untuk difahami dan menjadi panduan dalam proses analisis sentimen. Fasa pembangunan algoritma dapat dilihat dan dipantau dengan baik serta risiko untuk kegagalan bagi membangunkan algoritma ini adalah rendah.



Rajah 1 Model Klasifikasi Aspek Sentimen

4.1 Fasa Perancangan

Fasa ini melibatkan proses pengumpulan dan analisis keperluan kajian seperti set data yang berkaitan dengan objektif kajian dan permasalahan kajian serta menganalisis sorotan kajian susastera yang melibatkan pencarian jurnal serta artikel, pengumpulan maklumat, pembacaan jurnal dan kajian lepas sebagai rujukan kajian. Dengan itu, antara topik yang dikaji adalah berkaitan dengan sentimen terhadap pelajar dalam menjalankan pembelajaran secara dalam dan luar talian daripada media sosial menggunakan pembelajaran mesin sedia ada. Data yang berkaitan dengan kajian daripada Twitter dikumpulkan. Tambahan, tempoh waktu dalam mengendalikan projek kajian ini juga dilakukan. Dengan berpandukan pelan perancangan carta Gantt yang dibina, setiap aktiviti dan proses yang dilaksanakan dapat memastikan pembangunan projek mengikut masa yang ditetapkan.

4.2 Fasa Analisis

Fasa ini memfokuskan analisis terhadap data yang dikumpulkan. Hal ini bertujuan bagi memastikan data yang dikumpulkan adalah relevan dan sesuai dengan topik kajian serta diselaraskan kepada penggunaan Bahasa Inggeris sahaja. Seterusnya, data yang telah ditapis akan dibersihkan supaya menjadi dataset yang lebih bermakna melalui pra-pemprosesan data dan pengekstrakan ciri. Sebanyak 2,037 data iaitu ciapan daripada Twitter berjaya dikumpul dan dilabelkan secara manual kepada enam kategori yang menjadi skop kajian ini iaitu aspek infrastruktur, kewangan, kesihatan, persekitaran, keberkesanan pembelajaran dan kesosialan. Di samping itu, analisis tentang perkakasan dan perisian juga dijalankan bagi memastikan setiap perkakasan dan perisian yang sedia ada seperti Jadual 1 dan Jadual 2 adalah bersesuaian dalam pembangunan projek kajian ini.

Jadual 1 Spesifikasi Keperluan Perkakasan

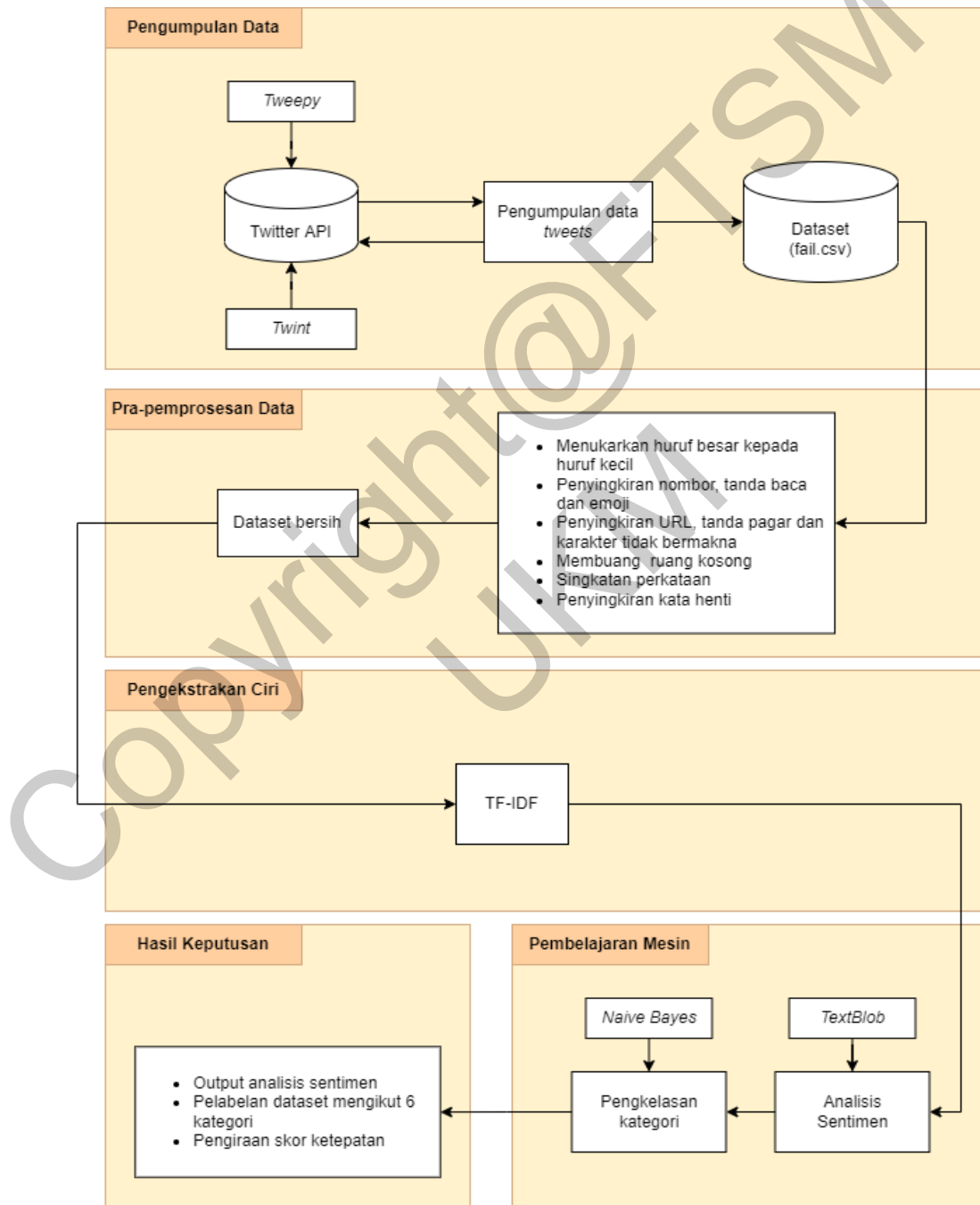
Perkakasan	Perincian
Model	Acer Aspire E 14
Sistem Pengoperasian	Windows 10 (64-bit)
Pemproses	Intel(R) Core(TM) i5-6200U CPU 2.30GHz
Saiz Memori (RAM)	4.00 GB
Saiz SSD	256 GB
Keluasan Skrin	14 inches

Jadual 2 Spesifikasi Keperluan Perisian

Perisian	Justifikasi
Google Chrome	Mencari maklumat dan info
Python	Bahasa pengaturcaraan utama
Google Colaboratory	Membuat dan melaksanakan kod program
Windows 10	Perisian sistem pengendalian komputer
Microsoft Excel Office 2016	Menyimpan data dan maklumat dari Twitter dalam format fail .csv

4.3 Fasa Reka Bentuk

Fasa ini merupakan reka bentuk seni bina yang secara umumnya memaparkan langkah demi langkah bertujuan memberikan gambaran awal secara menyeluruh bagi pembangunan model kajian ini. Seperti yang dipaparkan dalam Rajah 2 di bawah, reka bentuk kajian ini terdiri daripada lima fasa utama iaitu fasa pengumpulan data, fasa pra-pemrosesan data, fasa pengekstrakan ciri, fasa pembelajaran mesin dan juga fasa hasil keputusan atau output.



Rajah 2 Reka Bentuk Seni Bina Kajian

i) Pengumpulan data

Fasa ini telah menggunakan kaedah Tweepy dan juga Twint dengan menggunakan kata kunci berkaitan persepsi pelajar terhadap pembelajaran dalam talian dan juga luar talian serta kata kunci yang berkaitan dengan enam aspek yang telah dibincangkan. Antara contoh kata kunci yang digunakan adalah seperti *'laptop'*, *'internet'* yang berkaitan dengan aspek infrastruktur, kata kunci *'money'*, *'fee'* untuk aspek kewangan, kata kunci *'lecture'*, *'room'* bagi aspek persekitaran, kata kunci seperti *'stress'*, *'tired'*, *'depress'* untuk aspek kesihatan, manakala kata kunci *'focus'*, *'sleep'* digunakan untuk aspek keberkesanan pembelajaran, kata kunci seperti *'friend'*, *'teacher'* untuk aspek kesosialan dan akhir sekali kata kunci *'online learning'*, *'offline learning'* serta seumpamanya yang meliputi topik projek ini turut digunakan. Kata kunci ini dipilih kerana, Twitter dapat mencari dan mengeluarkan ciapan berkaitan kata kunci dan topik mengenai situasi pembelajaran yang dihadapi oleh pelajar. Dalam kajian ini, data akan diambil bermula bulan Januari 2021 kerana pada jangka masa ini terdapat sesetengah pelajar yang sedang melakukan pembelajaran di rumah, sekolah atau kampus. Data yang dikumpulkan ini akan digunakan bagi mengenal pasti sentimen pengguna dan pengelasan mengikut kategori aspek perbincangan pengguna Twitter.

ii) Pra-pemprosesan data

Fasa ini dilakukan untuk menyingkirkan data hingar (*noise*) dan melibatkan pra-pemprosesan data supaya data yang bersih dapat diperoleh dan digunakan untuk proses selanjutnya. Antara aktiviti yang akan dilakukan dalam pra-pemprosesan ialah normalisasi teks iaitu penyingkiran tanda pagar (*hashtag*), URL, emotikon, penukaran huruf besar kepada huruf kecil, singkatan perkataan, penyingkiran nombor, karakter tidak bermakna, tanda baca, dan juga membuang ruang kosong atau tab dengan menggunakan fungsi Regular Expression (*Regex*). Selanjutnya, pembuangan kata henti (*stopword*) akan dilakukan agar senang untuk dibaca dan supaya hasil analisis kajian menjadi lebih tepat.

iii) Pengekstrakan ciri

Vektorisasi Frekuensi Istilah-Frekuensi Dokumen Terbalik (TF-IDF) digunakan sebagai pengekstrakan ciri utama kajian. TF-IDF mencerminkan betapa pentingnya perkataan dalam sesuatu dokumen atau korpus. Ianya juga merupakan ukuran statistik bagi menilai kerelevanan

sesuatu perkataan dengan sesebuah dokumen dalam sekumpulan dokumen. TF-IDF dapat dilakukan dengan mendarab dari dua metrik iaitu berapa kali kekerapan perkataan yang muncul dalam dokumen dan kekerapan perkataan dokumen terbalik dalam sekumpulan dokumen. TF-boleh dikira menggunakan pakej *Sklearn* yang mempunyai fungsi yang boleh mendapatkannya secara langsung iaitu menggunakan *TfidfVectorizer()* dalam bahasa pengaturcaraan Python.

iv) Pembelajaran mesin

Dalam fasa ini, teks yang telah dibersihkan akan menjalankan proses analisis sentimen dengan menggunakan pakej *TextBlob* dan digunakan untuk proses pengkelasan kategori menggunakan model pengelas. Dalam kajian ini, pengelas yang dipilih dan diguna pakai ialah pengelas *Naïve Bayes*. Perpustakaan *Scikit-learn (Sklearn)* digunakan bagi menentukan ramalan bagi setiap kategori aspek seterusnya menilai ketepatan model *Naïve Bayes*.

v) Hasil keputusan

Fasa hasil keputusan akan memaparkan keputusan analisis sentimen sama ada ciapan yang dianalisis tergolong dalam sentimen positif, neutral atau negatif. Hasil seterusnya, menunjukkan pengkelasan kategori yang terdiri daripada enam aspek yang berkaitan pembelajaran atas talian dan luar talian iaitu dari aspek infrastruktur, kewangan, kesihatan, persekitaran, keberkesanan pembelajaran dan juga kesosialan. Seterusnya, pengiraan skor ketepatan model pengelas dipaparkan. Akhir sekali, hasil keputusan akan divisualisasikan dalam bentuk graf yang menarik.

4.4 Fasa Pengujian

Fasa ini dilaksanakan bagi memilih teknik analisis dan pengelasan yang sesuai bagi projek ini. Terdapat dua teknik yang digunakan dalam menganalisis sentimen dalam kajian ini iaitu menggunakan *TextBlob Sentiment Analysis* dan *VADER Sentiment Analysis*. Kedua-dua teknik ini dibandingkan bagi mengenal pasti teknik yang sesuai dan terbaik untuk diguna pakai bagi pembangunan model kajian ini. Hasil pengujian analisis sentimen mendapati *TextBlob Sentiment Analysis* yang mempunyai struktur yang lebih ringkas dan tepat adalah lebih sesuai digunakan bagi menganalisis sentimen persepsi pelajar dalam kajian ini. Pengujian seterusnya adalah pengklasifikasian ciapan kepada enam kategori iaitu infrastruktur, kewangan, kesihatan, persekitaran, keberkesanan pembelajaran dan kesosialan menggunakan vektor TF-IDF. Dalam

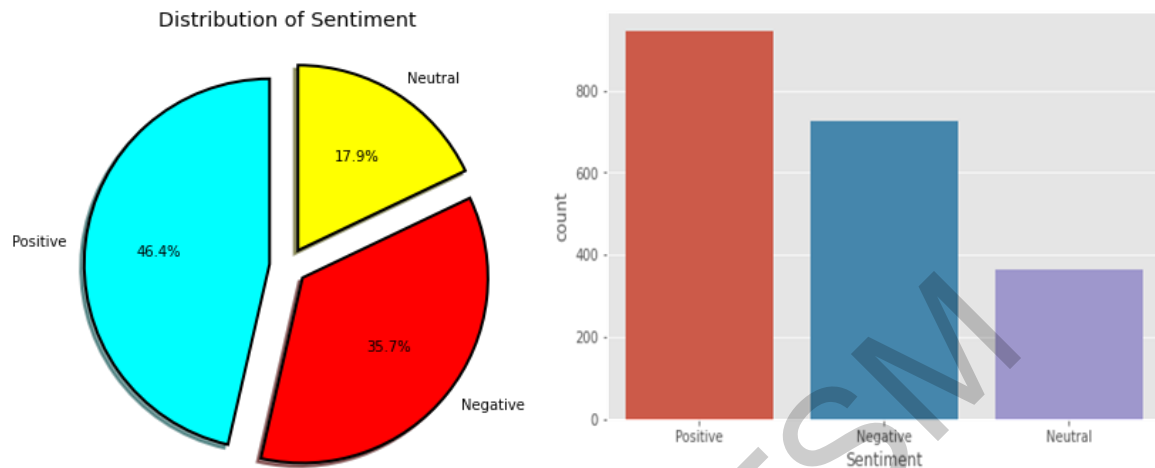
kajian ini, pengujian model pengelasan seperti *Random Forest*, *Logistic Regression*, *Support Vector Machine* dan *Multinomial Naive Bayes* digunakan bagi membandingkan ketepatan dengan pengelas utama kajian iaitu *Bernoulli Naive Bayes*. Namun begitu, keputusan ketepatan bagi setiap pengelas yang diuji didapati adalah lebih rendah berbanding pengelas *Bernoulli Naive Bayes*. Ini membuktikan bahawa pengelas *Naive Bayes* merupakan pengelas yang sesuai dan tepat dalam mengelaskan data mengikut setiap kategori. Jadual 3 menunjukkan perbandingan keputusan ketepatan pengelas-pengelas yang telah digunakan dan *Bernoulli Naive Bayes* mencatatkan nilai ketepatan yang paling tinggi antara pengelas yang lain.

Jadual 3 Nilai Ketepatan Bagi Setiap Pengelas

Jenis Pengelas	Ketepatan
Random Forest	0.7646
Logistic Regression	0.7721
Support Vector Machine	0.7745
Multinomial Naive Bayes	0.6862
Bernoulli Naive Bayes	0.8128

5 HASIL KAJIAN

Hasil keputusan divisualisasi dalam bentuk graf bar dan carta pai bagi memudahkan pengguna menilai sentimen pelajar iaitu positif, negatif mahupun neutral. Rajah 3 menunjukkan bahawa sentimen yang mencatat peratusan tertinggi adalah sentimen positif sebanyak 46.4% manakala sebanyak 35.7% telah dicapai oleh sentimen negatif dan selebihnya adalah 17.9% meliputi sentimen neutral. Keputusan ini membawa maksud pelajar memberi persepsi yang baik dalam melakukan pembelajaran dalam talian dan juga luar talian. Hal ini boleh ditakrifkan bahawa, pembelajaran secara hibrid iaitu percampuran pembelajaran dalam dan luar talian dalam waktu yang sama merupakan cara yang sesuai untuk dilaksanakan oleh para pelajar ketika fasa endemik di Malaysia.



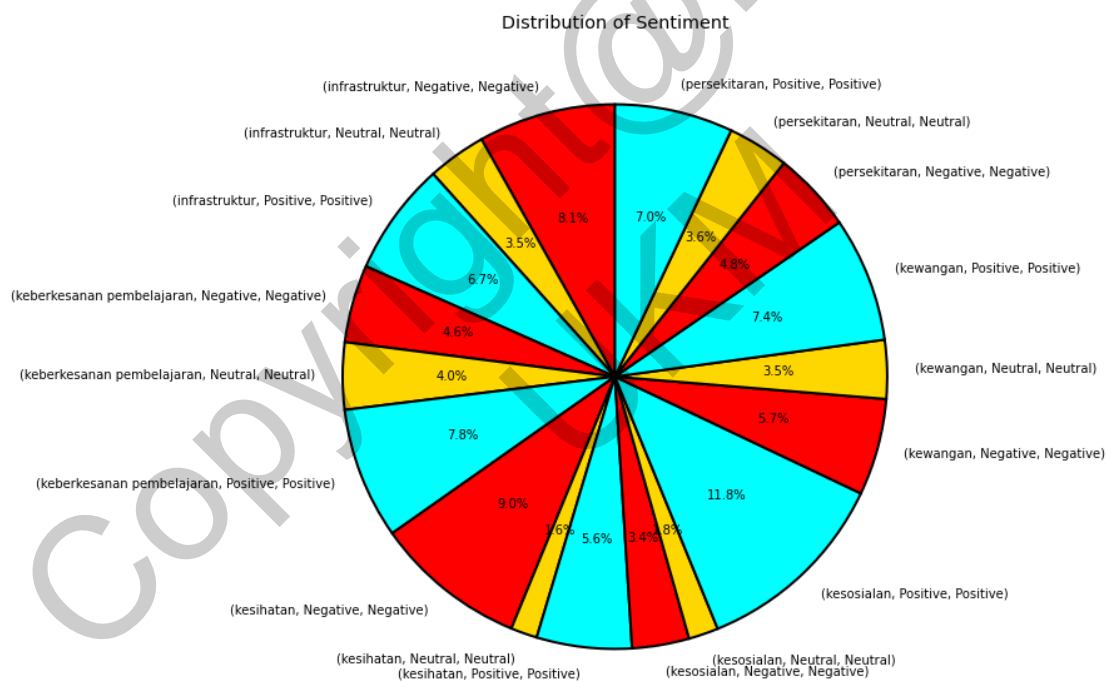
Rajah 3 Keputusan Analisis Sentimen Dalam Bentuk Carta Pai Dan Graf Bar

Dalam Rajah 4 menunjukkan antara contoh ciapan positif, negatif dan neutral. TextBlob akan memaparkan skor sentimen seberapa positif atau negatif teks tersebut dan skor subjektiviti pendapat sesuatu teks itu. Pengiraan ciri sentimen ciapan dibuat dalam bentuk polariti dan subjektiviti. Polariti merupakan nilai yang terletak dalam julat antara -1 hingga +1, di mana -1 mentakrifkan sentimen negatif dan +1 mentakrifkan sentimen positif sesuatu teks yang diberikan. Subjektiviti merupakan nilai yang terletak dalam julat antara 0 hingga +1 dimana nilai 0 adalah fakta dan nilai +1 adalah sangat banyak pendapat. Ciapan dilabel mengikut keputusan sentimen yang telah diperolehi.

Cleaned_Text	Polarity	Subjectivity	Sentiment
online class is so boring	-1	1	Negative
currently at home waiting for class but the weather is bad today cannot afford to be disconnected today	-0.35	0.53333333	Negative
thanks to the online class today i more confident of my own answers in the assignment	0.45	0.63333333	Positive
i love this kind of online class even play quiz the knowledge get into my mind	0.55	0.75	Positive
had to bring the monitor to uni lets go online class and fyp bismillah	0	0	Neutral
the government should develop on a government online learning platform rather depending on google meet and in	0	0	Neutral

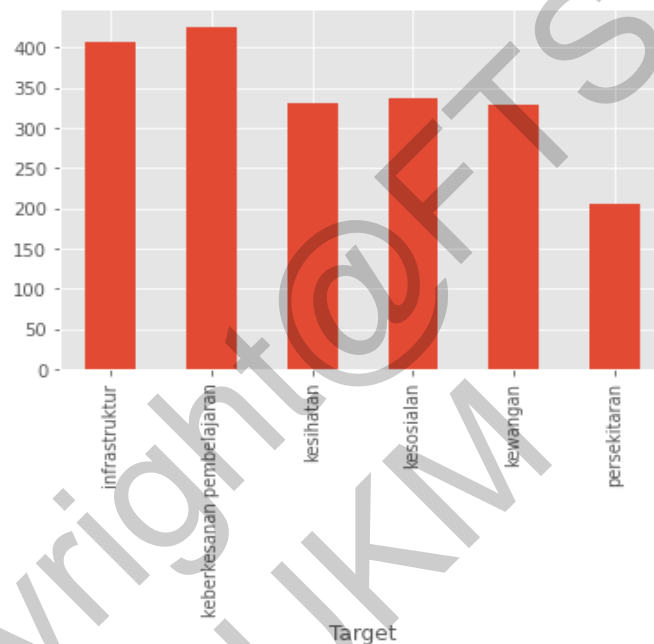
Rajah 4 Output Sentimen Ciapan Oleh TextBlob

Seterusnya, hasil keputusan sentimen analisis bagi setiap kategori juga turut divisualisasikan dalam bentuk carta pai seperti rajah di bawah. Ianya adalah bertujuan untuk mengetahui keputusan sentimen yang paling banyak digunakan oleh pelajar bagi setiap kategori. Seperti yang ditunjukkan dalam Rajah 5 aspek kesosialan mendapat peratusan paling tinggi iaitu 11.8% bagi keseluruhan sentimen positif. Hal ini menunjukkan pelajar yang berada di rumah mereka dapat bersama-sama keluarga dan tidak berjauhan antara satu sama lain, manakala bagi pelajar yang melakukan pembelajaran secara bersemuka mereka dapat berjumpa dan belajar bersama-sama rakan dengan bimbingan pendidik yang berada di samping mereka. Kemudian, aspek kesihatan mendapat peratusan tertinggi bagi sentimen negatif sebanyak 9.0% kerana kebanyakan pelajar merasai tekanan serta masalah kesihatan fizikal yang membuat pelajar berasa keletihan ketika melaksanakan pembelajaran secara dalam dan luar talian.



Rajah 5 Hasil Analisis Sentimen Bagi Keseluruhan Kategori

Hasil pengkelasan *Naive Bayes* dikelaskan mengikut setiap kategori supaya pembaca dapat melihat aspek-aspek yang dihadapi oleh pelajar-pelajar di Malaysia. Aspek berikut adalah infrastruktur, kewangan, kesihatan, persekitaran, keberkesanan pembelajaran dan kesosialan. Empat dari enam aspek mencatatkan sentimen yang positif daripada pelajar iaitu aspek kesosialan sebanyak 69.6%, keberkesanan pembelajaran sebanyak 47.6%, persekitaran sebanyak 45.5% dan kewangan sebanyak 44.5%. Manakala dua aspek mencatatkan nilai sentimen yang negatif daripada pelajar iaitu aspek kesihatan sebanyak 55.6% dan juga aspek infrastruktur sebanyak 44.4%.



Rajah 6 Hasil Keputusan Pengelasan *Naive Bayes*

Merujuk kepada Rajah 7 di bawah merupakan ciapan yang telah dikenalpasti sebagai kategori aspek infrastruktur. Beberapa perkataan unik dapat digambarkan dalam ciapan seperti 'internet', 'slow', 'connection', 'buffer' dan seumpamanya.

my internet came back everyday the video call still buffering i can hear them but the video cam is buffer connection slow fix it another problem on class how	infrastruktur
online class is difficult when you have slow internet connection	infrastruktur
not the internet crashing in the middle of my virtual class	infrastruktur

Rajah 7 Contoh Ciapan Kategori Infrastruktur

Merujuk kepada Rajah 8 merupakan ciapan yang telah dikenalpasti sebagai kategori aspek keberkesanan pembelajaran. Beberapa perkataan unik dapat digambarkan dalam ciapan seperti *'sleep'*, *'understand'*, *'lecture'*, *'zoning out'* dan seumpamanya.

felt tired already sleep times in this lecture	keberkesanan pembelajaran
i keep zoning out every single time when i read my lecture notes idk whats wrong with me	keberkesanan pembelajaran
cant believe i understand jurisprudence from youtube more than in lecture	keberkesanan pembelajaran

Rajah 8 Contoh Ciapan Kategori Keberkesanan Pembelajaran

Merujuk kepada Rajah 9 merupakan ciapan yang telah dikenalpasti sebagai kategori aspek kesihatan. Beberapa perkataan unik dapat digambarkan dalam ciapan seperti *'tired'*, *'stressed'*, *'stress'* dan seumpamanya.

class timing in an offline setting about to have everyone stressed	kesihatan
at home learning was a disease because whenever i got too stressed about assignments	kesihatan
im tired as shit all my stress is boiling over	kesihatan

Rajah 9 Contoh Ciapan Kategori Kesihatan

Merujuk kepada Rajah 10 merupakan ciapan yang telah dikenalpasti sebagai kategori aspek kesosialan. Beberapa perkataan unik dapat digambarkan dalam ciapan seperti *'friend'*, *'college'*, *'study'* dan seumpamanya.

i wanna go to campus but i scared that i wouldnt be able to make any friends since i been home for more than years now	kesosialan
got no friends at all and i like it	kesosialan
rightly treated we had a friend back in college who used to study all night	kesosialan

Rajah 10 Contoh Ciapan Kategori Kesosialan

Merujuk kepada Rajah 11 merupakan ciapan yang telah dikenalpasti sebagai kategori aspek kewangan. Beberapa perkataan unik dapat digambarkan dalam ciapan seperti *'fee'*, *'pandemic'*, *'income'*, *'rent'*, *'pay'* dan seumpamanya.

they forced us to pay a life fee during the pandemic when we literally weren allowed on campus	kewangan
so far pandemic isn that fair many suffer and lack daily income even among college students parents do	kewangan
im a student i want to pay my house rent	kewangan

Rajah 11 Contoh Ciapan Kategori Kewangan

Merujuk kepada Rajah 12 merupakan ciapan yang telah dikenalpasti sebagai kategori aspek persekitaran. Beberapa perkataan unik dapat digambarkan dalam ciapan seperti ‘*mask*’, ‘*home*’, ‘*classroom*’, ‘*learning*’, ‘*class*’ dan seumpamanya.

yeah the first post pandemic class day a friend of mine entered school with an ahegao hoodie and mask	persekitaran
for some people nothing beats the experience of learning in a classroom environment	persekitaran
home learning needs to end and go back to face to face stop pushing this garbage	persekitaran

Rajah 12 Contoh Ciapan Kategori Persekitaran

6 KESIMPULAN

Secara kesimpulannya, objektif kajian berjaya dicapai dalam membangunkan model yang dapat mengenal pasti sentimen pelajar. Selain itu, pengelasan enam kategori yang dilakukan dalam kajian ini dapat memaparkan aspek-aspek yang perlu ditangani seperti aspek infrastruktur iaitu masalah kebolehcapaian internet yang perlu ditangani segera untuk memastikan pelajar dapat melakukan sesi pembelajaran dengan sempurna sehinggalah ke aspek yang kurang untuk ditangani. Data dan juga hasil kajian amat berguna untuk dijadikan sebagai penanda aras serta rujukan kepada pakar pendidikan, Kementerian Pendidikan Malaysia dan Kementerian Pengajian Tinggi dalam mengatasi masalah pelajar serta memahami kehendak dan keinginan pelajar dengan lebih baik. Akhir sekali, diharap penggunaan kaedah analisis sentimen dan juga pembelajaran mesin ini dapat ditambah baik dan diguna pakai oleh ramai pihak seperti industri besar mahupun industri kecil dalam membuat penilaian atau mengesan sentimen sesuatu perkara.

7 RUJUKAN

- Abbasi, N. (2020). *What is TF-IDF?* Interactive Courses for Software Developers, Educative. <https://www.educative.io/edpresso/what-is-tf-idf>
- Akash. (2021). *Making Natural Language Processing easy with TextBlob*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2021/10/making-natural-language-processing-easy-with-textblob/>
- Ali, I., M. Ameen Chhajro, Kamlesh Kumar, & Usman Khan. (2019). *Student Feedback Sentiment Analysis Model Using Various Machine Learning Schemes A Review*. Indian Journal of Science and Technology 14(12):1-9 <https://doi.org/10.17485/ijst/2019/v12i14/143243>
- Althagafi, G. Althobaiti, H. Alhakami, & T. Alsubait. L. (2021). *Arabic Tweets Sentiment Analysis about Online Learning during COVID-19 in Saudi Arabia*. International Journal of Advanced Computer Science and Applications, vol. 12, no. 3. <https://doi.org/10.14569/ijacsa.2021.0120373>
- Aniketa Horo, Nupur Biswas & Jagruti Das. (2020). *Perception of Students towards the Offline and Online Modes of Learning during COVID-19 Lockdown*. Int.J.Curr.Microbiol.App.Sci. 9(09): 183-187. <https://doi.org/10.20546/ijemas.2020.909.023>
- Anon. (2014). *What are Word Clouds? The Value of Simple Visualizations*. Boost Labs. <https://boostlabs.com/blog/what-are-word-clouds-value-simple-visualizations/>
- Anon. (2015). *Sejarah Internet Malaysia*. MajalahSains | Sains & Teknologi Untuk Masyarakat. <https://www.majalahsains.com/sejarah-internet-malaysia/>
- Anon. (2017). *A Gentle Introduction to the Bag-of-Words Model*. Machine Learning Mastery. <https://machinelearningmastery.com/gentle-introduction-bag-words-model/>
- Anon. (2018). *Perbezaan Antara Keperluan Fungsian dan Bukan Fungsi*. Mldunbound.org. <https://ms.mldunbound.org/contrast/difference-between-functional-and-non-functional-requirements/>
- Anon. (2018). *Text Preprocessing in Python: Steps, Tools, and Examples*. KDnuggets. <https://www.kdnuggets.com/2018/11/text-preprocessing-python.html>
- Anon. (2018). *Using Student Feedback to Create Effective Online Learning Experiences*. Wiley. <https://edservices.wiley.com/how-student-feedback-creates-better-online-learning/>
- Anon. (2019). *Create Word Clouds for Free - Live & Interactive*. Mentimeter. <https://www.mentimeter.com/features/word-cloud>

- Anon. (2020). *1.7 juta pelajar hadapi “miskin digital.”* Berita Harian; Berita Harian. <https://www.bharian.com.my/berita/nasional/2020/11/757190/17-juta-pelajar-hadapi-miskin-digital>
- Anon. (2020). *Get data from Twitter API in Python - Step by Step Guide.* Data Science Parichay. <https://datascienceparichay.com/article/get-data-from-twitter-api-in-python-step-by-step-guide/>
- Anon. (2020). *Pembelajaran Online - Cabaran dan Cadangan Mengatasinya.* TVET & Vocational Training – Malaysian News. <https://jpkmalaysia.com/cabaran-pembelajaran-online-cadangan-mengatasinya/>
- Anon. (2021). *Naive Bayes Classifier in Machine Learning – Javatpoint.* www.javatpoint.com. <https://www.javatpoint.com/machine-learning-naive-bayes-classifier>
- Antaripasaha. (2020). *NLP - Expand contractions in Text Processing.* GeeksforGeeks. <https://www.geeksforgeeks.org/nlp-expand-contractions-in-text-processing/>
- Band, A. (2020). *Twitter Sentiment Analysis using Vader & Tweepy | Python in Plain English.* Medium; Python in Plain English. <https://python.plainenglish.io/twitter-sentiment-analysis-using-vader-tweepy-b2a62fba151e>
- Beri, A. (2020). *SENTIMENTAL ANALYSIS USING VADER - Towards Data Science.* Medium; Towards Data Science. <https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664>
- Dr. Rosninawati H. (2021). *Pembelajaran dalam talian sebagai norma baharu.* USIM UNIVERSITI SAINS ISLAM MALAYSIA. <https://www.usim.edu.my/news/in-our-words/pembelajaran-dalam-talian-sebagai-norma-baharu/>
- Duong Wu (2019). *Python Word Clouds Tutorial: How to Create a Word Cloud.* Datacamp.com. <https://www.datacamp.com/tutorial/wordcloud-python>
- Hari Suparwito, Agnes Maria Polina, & Budiraharjo, M. (2021). *Student Perceptions Analysis of Online Learning: A Machine Learning Approach.* Indonesian Journal of Information Systems, 4(1), 64–75. <https://doi.org/10.24002/ijis.v4i1.4594>
- Israr K. (2021). *Adakah kita sudah memasuki fasa endemik? Ini penjelasan KJ.* Astroawani.com. <https://www.astroawani.com/berita-malaysia/adakah-kita-sudah-memasuki-fasa-endemik-ini-penjelasan-kj-324703>
- Jawatankuasa Khas Jaminan Akses Bekalan Vaksin COVID-19 (JKJAV) (2021). *Statistik Perkembangan Program Imunisasi Covid-19 Kebangsaan.* Jawatankuasa Khas Jaminan Akses Bekalan Vaksin COVID-19 (JKJAV). <https://www.vaksin covid.gov.my/statistik/>
- Lum Choi K. (2019). *Analisis Sentimen Dalam Bitcoin Tweets.* Bangi: Universiti Kebangsaan Malaysia.

- M. Izzat Mailis, Z.Hanim, & N. Hafizaliyana H. (2020). *Persepsi Pelajar Kolej Universiti Islam Melaka Terhadap Pelaksanaan Pembelajaran Secara Atas Talian Dalam Era Pandemik Covid-19*. *Jurnal Kesidang*, 5(1), 88–99. <http://journal.kuim.edu.my/index.php/JK/article/view/768>
- Mohd Fadhil. (2021). *3.10 Penyediaan Spesifikasi Keperluan Sistem [F2.6]*. Mampu.gov.my. <https://sqa.mampu.gov.my/index.php/ms/3-10-penyediaan-spesifikasi-keperluan-sistem-f2-6>
- Nederbouchnak (2022). *ESG-Tweets-Analysis*. Kaggle.com; Kaggle. <https://www.kaggle.com/code/nederbouchnak/esg-tweets-analysis/notebook>
- Nik Najwa Binti Nik A. (2021). *Analisis Sentimen Kesan Pembelajaran Atas Talian Terhadap Pelajar Ipt Berikutan Pandemi Covid-19 Di Malaysia Dengan Menggunakan Kaedah Naïve Bayes*. Bangi: Universiti Kebangsaan Malaysia
- Noorazura Abdul R. (2021). *Pelajar UiTM Kedah meninggal dunia disyaki akibat tekanan belajar*. *Berita Harian*; *Berita Harian*. <https://www.bharian.com.my/berita/kes/2021/07/837942/pelajar-uitm-kedah-meninggal-dunia-disyaki-akibat-tekanan-belajar>
- Nor Hidayati M. (2020, April 7). *Pembelajaran Atas Talian Untuk Pembelajaran Sepanjang Hayat*. *UTM NewsHub*. *News.utm.my*. <https://news.utm.my/ms/2020/04/pembelajaran-atas-talian-untuk-pembelajaran-sepanjang-hayat/>
- Nor Musfirah M. (2021). *Cabaran Pedagogi Norma Baharu di Kolej Universiti Islam Perlis (Kuips) Ketika Pandemik Wabak Koronavirus Covid-19*. *Jurnal Pengajian Islam*, 243–254. <https://jpi.kuis.edu.my/index.php/jpi/article/view/71>
- Rafiza A. R. (2013). *Strategi Pembelajaran Aktif Secara Kolaboratif Atas Talian Dalam Analisis Novel Bahasa Melayu*. *JuKu: Jurnal Kurikulum & Pengajaran Asia Pasifik*, 1(3), 34–46. <http://ajba.um.edu.my/index.php/JUKU/article/view/7961>
- Raiham M. S. (2021). *Sistem bersemuka paling baik*. *Sinarharian*; *sinarharian*. <https://www.sinarharian.com.my/article/160749/BERITA/Nasional/Sistem-bersemuka-paling-baik>
- Rude, B. (2021). *Using Tweepy to analyze Twitter Data with Python*. *Github*. <https://brittarude.github.io/blog/2021/08/01/Using-Tweepy-to-analyze-twitter-data>
- Scott, W. (2019). *TF-IDF from scratch in python on a real-world dataset*. *Towards Data Science*. <https://towardsdatascience.com/tf-idf-for-document-ranking-from-scratch-in-python-on-real-world-dataset-796d339a4089>
- Shubham Singh. (2019). *What is Tokenization | Methods to Perform Tokenization*. *Analytics Vidhya*. <https://www.analyticsvidhya.com/blog/2019/07/how-get-started-nlp-6-unique-ways-perform-tokenization/>

- Sindhu, I., Muhammad Daudpota, S., Badar, K., Bakhtyar, M., Baber, J., & Nurunnabi, M. (2019). *Aspect-Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation*. IEEE Access, 7, 108729–108741. <https://doi.org/10.1109/access.2019.2928872>
- Stecanella B. (2019). *Understanding TF-IDF: A Simple Introduction*. MonkeyLearn Blog. <https://monkeylearn.com/blog/what-is-tf-idf/>
- Subramanian, D. (2019). *Sentiment analysis on Ellen's DeGeneres tweets using TextBlob*. Analytics Vidhya. <https://medium.com/analytics-vidhya/sentiment-analysis-on-ellens-degeneres-tweets-using-textblob-ff525ea7c30f>
- T. Zara A. R. (2012). *Kemahiran membaca*. Zdocs.tips. <https://zdocs.tips/doc/kemahiran-membaca-op0lqye9krpy>
- Tay Fui K. (2019). *Analisis Sentimen Twitter Mengenai Peristiwa-Persitiwa penting yang Berlaku di Sekitar UKM, Bangi*. Bangi: Universiti Kebangsaan Malaysia.
- Thanusha K. R. (2021). *Analisis Sentimen Filem Dan Drama Bersiri Di Netflix Menggunakan Pembelajaran Mesin*. Bangi: Universiti Kebangsaan Malaysia
- Yahaya, Mahathir, Wardatul & Hayat A. (2021). *Cabaran pelajar melalui kaedah pembelajaran atas talian: kajian institusi pengajian tinggi awam Malaysia*. Journal of Media and Information Warfare (JMIW), 14. pp. 11-20. Uitm.edu.my. <https://doi.org/https://ir.uitm.edu.my/id/eprint/46977/1/46977.pdf>
- Zenun Kastrati, Blend Arifaj, Blend Arifaj, & Engjëll Nishliu. (2020). *Aspect-Based Opinion Mining of Students' Reviews on Online Courses*. Acm.org. <https://doi.org/10.1145/3404555.3404633>

Zulaikha Dawami (A174805)
Lailatul Qadri Zakaria
Fakulti Teknologi & Sains Maklumat,
Universiti Kebangsaan Malaysia