

PORTAL GIG EKONOMI DAN IMPLEMENTASI ANALISIS SENTIMEN

Aiman Hasni Bin Zazani
Dr. Mohd Ridzwan Bin Yaakub

Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

Gig ekonomi merupakan segmen ekonomi yang menjadi sebuah fenomena dalam kalangan rakyat Malaysia kerana mempunyai gaya kerja yang fleksibel dan terdapat pelbagai pendapat dari kebanyakan masyarakat yang membuatkan domain ini sesuai untuk dianalisis. Dalam projek ini, sebuah portal akan dibangunkan sebagai sebuah platform untuk memperkasakan lagi sektor gig ekonomi dengan terdapat pelbagai fungsi seperti penawaran kerja daripada pengguna, pencarian kerja dan penyenaiaan kerja yang menumpukan konsep gig ekonomi. Projek ini mengimplimentasikan analisis sentimen sebagai suatu teknik untuk melombong pendapat masyarakat mengenai gig ekonomi daripada struktur perkataan seperti ciapan Twitter dan dari pendapat yang dilombong, terdapat nilai yang boleh diformulakan untuk memperoleh hasil analitik melalui perspektif masyarakat mengenai sektor ini. Analisis sentimen ini menggunakan pendekatan pembelajaran mesin yang diselia. Dalam portal ini, sebuah papan pemuka akan dipaparkan untuk memvisualisasikan data yang telah dianalisis menggunakan 'Natural Language Processing' iaitu analisis sentimen. Projek ini menggunakan data daripada pelbagai sumber dan platform sebagai contoh laman sesawang carian kerja dan daripada data terbuka kerajaan seperti MAMPU dan ILMIAH. Papan pemuka adalah merupakan paparan daripada data, sentimen dan maklum balas yang telah dianalisis hasil daripada pensekerapan data di laman sesawang carian kerja yang terkenal dan kerap digunakan seperti Glassdoor, Indeed dan Jobstreet berdasarkan pasaran pekerjaan pekerja bebas di Malaysia. Kesimpulannya, projek ini akan menggunakan teknik analisis sentimen dan kelebihan teknologi pembelajaran mesin untuk mengolah data bagi mendapatkan perolahan yang baik supaya pengguna dapat memahami dan mendalami potensi gig ekonomi di Malaysia.

1 PENGENALAN

Gig ekonomi merupakan sistem pasaran bebas yang melibatkan pelantikan pekerja secara sementara dalam sesebuah syarikat pada jangka masa yang singkat. Terma gig terhasil daripada perkataan yang membawa maksud pekerjaan yang berbentuk sementara. Sebagai contoh pekerja gig ialah mereka yang bekerja sebagai pekerja bebas, kontraktor persendirian, bekerja berdasarkan suatu projek dan pekerja yang dilantik sebagai pekerja sambilan pada tempoh yang singkat. Saban tahun, gig ekonomi menjadi suatu tren dalam kalangan rakyat Malaysia kerana mempunyai gaya bekerja yang fleksibel dan menjadi sebuah titik perubahan yang besar dalam ekonomi Malaysia. Menurut Jabatan Statistik Malaysia[1], 26 peratus

daripada jumlah 15.3 juta tenaga kerja adalah daripada sektor gig ekonomi mempunyai jumlah yang hampir bersamaan 4 juta pekerja bebas. Lanjutan itu, pekerjaan dalam sektor gig ekonomi telah meningkat sebanyak 31 peratus pada tahun 2017, melebihi jumlah peningkatan tenaga kerja konvensional. Melalui statistik yang diperoleh, gig ekonomi mempunyai kadar perkembangan yang memberangsangkan di mana ia memberi impak yang penting dalam pembangunan ekonomi negara.

Seiring dengan perkembangan platform digital, kebanyakan pekerja gig memerlukan akses digital untuk mendapatkan pekerjaan mereka kerana kebolehcapaian melalui digital amatlah memberangsangkan. Oleh itu, suatu platform digital yang khusus perlu diwujudkan untuk membantu para pekerja gig mendapatkan maklumat pekerjaan dengan lebih efektif. Pembinaan portal gig ekonomi akan menjadi sebuah platform yang sesuai sebagai medium untuk penawaran dan pencarian kerja secara gig. Bagi meningkatkan lagi kebolehcapaian terhadap pekerjaan, data yang sedia ada dianalisis menggunakan teknik analisis sentimen untuk mendapatkan perolehan data berkaitan isu yang melibatkan gig ekonomi, sebagai contoh kadar kebergantungan sesebuah syarikat terhadap pekerja gig, impak pekerja gig dalam produktiviti syarikat dan banyak lagi. Sejurus itu, pengguna boleh mendapatkan informasi yang berguna melalui perincian data yang dilakukan.

2 PENYATAAN MASALAH

Pekerja gig bergantung kepada akses digital untuk melakukan carian kerja. Sebagai contoh aplikasi Foodpanda, Grab dan Uber merupakan satu platform digital yang menawarkan pekerjaan gig tetapi khusus untuk penghantaran makanan dan e-hailing. Seiring dengan perkembangan gig ekonomi, pekerjaan gig bukanlah difokuskan kepada kerja penghantaran sahaja, malah kerja-kerja berkaitan kontrak profesional seperti pembinaan web, reka bentuk projek dan kontraktor persendirian juga banyak dilakukan secara gig. Hal ini menyukarkan sesetengah pihak untuk mendapatkan khidmat pekerja gig kerana tiada platform yang khusus dalam penawaran kerja.

Di samping itu, kadar kebarangkalian pekerja gig untuk menemukan pekerjaan dalam bidang mereka adalah rendah. Perkara ini disebabkan oleh kurangnya pasaran kerja yang menawarkan pelbagai bidang kerja dan mereka perlu melakukan carian kerja melalui pelbagai platform. Hal ini menyebabkan carian kerja menjadi kurang efektif. Di samping itu, pekerja gig juga mempunyai kekurangan informasi dan maklumat berkaitan pekerjaan yang dicari kerana kebanyakan platform tidak menyediakan perincian data kepada pengguna. Perkara ini menyebabkan pekerja gig terpaksa melakukan kajian dan perincian data tersendiri di mana ia dapat menjejaskan kadar keberkesanan dalam pencarian kerja.

3 OBJEKTIF KAJIAN

Objektif projek ini ialah membangunkan portal yang mempunyai fungsi untuk pengguna menawar dan mencari pekerjaan secara gig yang menggunakan analisis sentimen untuk membuat perolehan data mengenai gig ekonomi. Secara umumnya, portal ini dibangunkan untuk memudahkan pengguna membuat carian dan penawaran kerja-kerja secara gig. Selain itu, terdapat juga paparan analitik data yang telah dibuat bagi meningkatkan lagi informasi pengguna portal berkaitan gig ekonomi. Paparan utama pada portal ini adalah sentimen mengenai gig ekonomi dan ia dipaparkan dalam bentuk carta 'pie' di mana pengguna boleh menafsirkan data dengan melihat peratusan setiap sentimen iaitu positif, neutral dan negatif. Menggunakan pembelajaran mesin yang diselia, sentimen bagi gig ekonomi diramal menggunakan dataset yang telah diambil dari Twitter.

4 METOD KAJIAN

Kajian ini menggunakan kaedah 'Agile' bagi membangunkan portal gig ekonomi. Kaedah Agile ini digunakan untuk membangunkan sistem dalam jangka masa yang pendek dan kurang berisiko dibandingkan dengan kaedah yang lain. Tiada pra-perancangan dalam pembinaan projek membuatkan tugas boleh ditambahbaik pada bila-bila masa. Justeru, semua keperluan baharu boleh ditambah pada fasa reka bentuk kerana mempunyai struktur kitaran yang membenarkan penambahan ciri. Dalam pada itu, implementasi analisis sentimen sebagai ciri utama portal ini memerlukan beberapa proses yang khusus dalam usaha penggalian pendapat. Dalam projek ini, pendekatan berdasarkan pembelajaran mesin digunakan dalam menjayakan teknik analisis sentimen ini.

4.1 Fasa Keperluan

Fasa ini merupakan fasa yang melibatkan proses perancangan dan pengenalpastian pelan projek bagi membantu proses pembangunan yang dilakukan. Semua spesifikasi keperluan telah disenaraikan dalam fasa ini. Perkara ini dilakukan bagi melancarkan lagi struktur pembangunan projek kerana ia amat memerlukan perancangan yang teliti bagi membina sebuah sistem yang baik. Sebuah carta Gantt telah dihasilkan sebagai panduan untuk pembinaan projek mengikut garis masa yang ditetapkan supaya projek dibangunkan selari dengan pelan asal. Pada fasa ini juga pernyataan masalah, skop, objektif dan metodologi kajian telah disenaraikan.

4.2 Fasa Reka Bentuk

Pada fasa ini, segala lakaran dan rajah yang berkaitan dengan reka bentuk sistem telah dibina bagi menjelaskan lagi struktur sistem yang dibangunkan secara logik dan mengikut algoritma yang bersesuaian. Aliran pangkalan data juga telah dibina bagi menjelaskan bagaimana sistem ini berfungsi dibalik tabir bersesuaian dengan teknologi yang digunakan. Dalam sistem utama, aliran pengguna ditafsirkan dalam bentuk kes guna bagi menunjukkan kefungian sistem dan bagaimana ia mengendalikan pengguna secara dasarnya. Manakala implementasi analisis sentimen dilakukan menggunakan kaedah pembelajaran mesin secara diselia untuk membuat ramalan terhadap dataset Twitter.

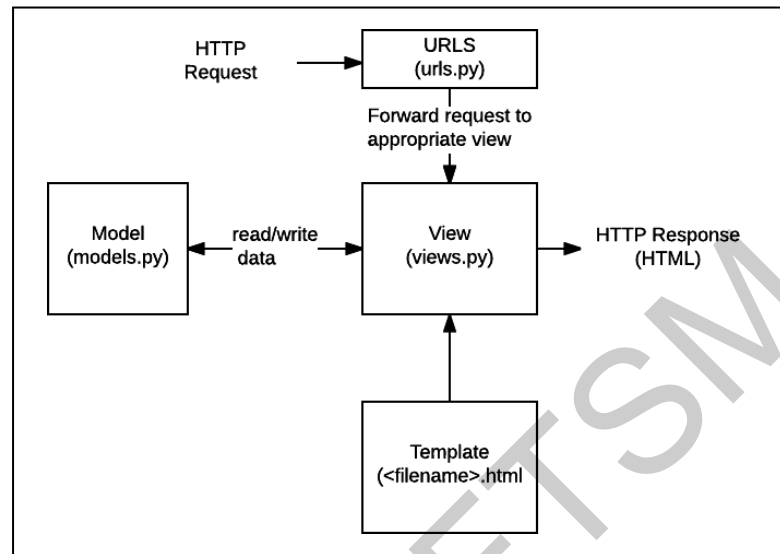
4.3 Fasa Pembangunan

Dalam fasa pembangunan ini, cara implementasi bagi projek yang telah dibina menggunakan pelbagai pendekatan, perisian dan teknologi bagi melengkapkan segala keperluan dan rancangan yang dicadangkan dalam fasa lepas telah dikemukakan. Pelbagai teknologi yang digunakan sama ada diaplikasikan melalui pengalaman dari subjek yang telah dipelajari mahupun pembelajaran sendiri. Dalam implementasi projek, penting untuk menyatakan pendekatan dan teknologi yang terlibat sepanjang pembinaan dan pembangunan bagi menjelaskan lagi bagaimana projek ini dibangunkan secara teknikal.

Proses pembangunan portal ini dibangunkan menggunakan pelbagai perkakasan dan perisian. Perisian dan perkakasan yang bertepatan dengan projek yang dibangunkan amat penting bagi mendapatkan hasil seperti yang dirancang. Terdapat banyak perisian yang digunakan dan perkakasan yang sedia ada telah digunakan bagi mengelakkan perbelanjaan yang melebihi kemampuan, berikut adalah spesifikasi perkakasan yang digunakan untuk membangunkan projek:

- i. Sistem Pengoperasian: Windows 10 Home
- ii. Pemprosesan: Intel Core i5 8300H @ 2.30GHz
- iii. Cakera Keras: 512GB ADATA SSD, 1024 HDD
- iv. Cakera Rawak: 8 GB DDR4
- v. Kad Grafik: Nvidia GTX 1050 4GB

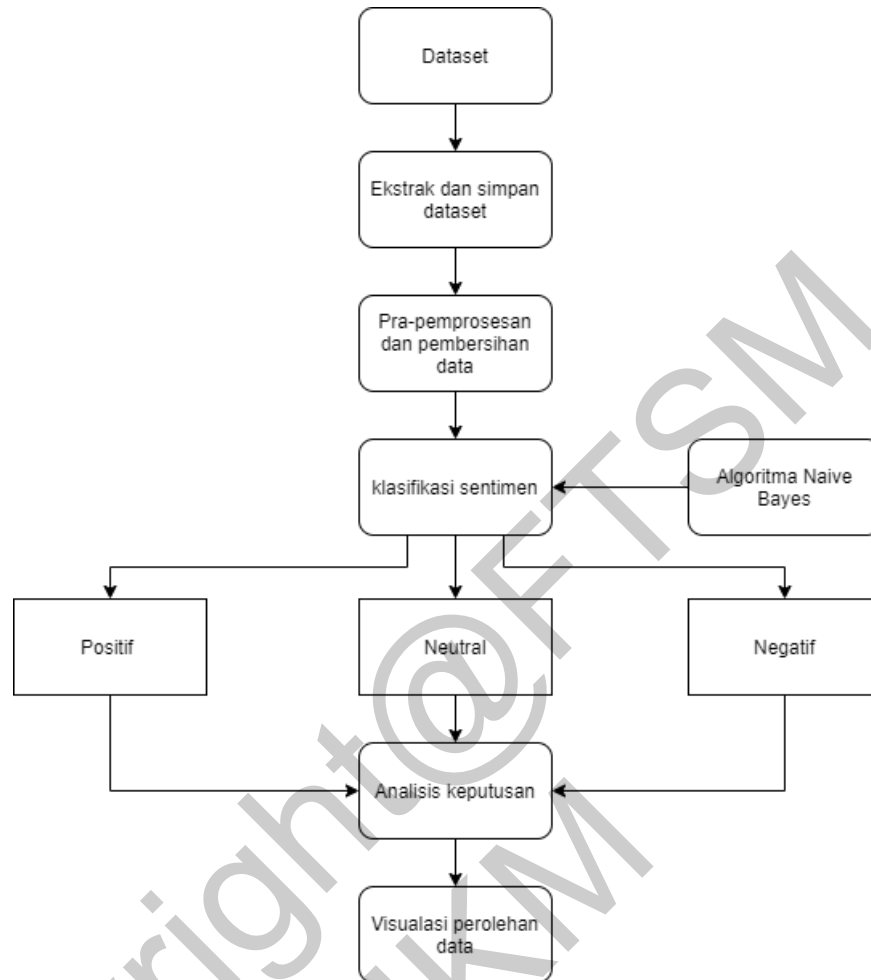
Proses pembangunan Portal Gig Ekonomi dan Implementasi Analisis Sentimen atau ‘Gigger’ ini menggunakan kerangka Django bagi menghasilkan sebuah sistem yang mempunyai ciri keselamatan yang baik dan mudah untuk diselenggara. Bahasa aturcara yang digunakan adalah Python, HTML, CSS dan Javascript bagi membina portal ini. Kerangka Django juga mudah untuk dipelajari membuatkan pengkaji selesa untuk membangunkan projek menggunakan kerangka ini. Kerangka Django mempunyai seni bina MVC di mana ia sesuai untuk dibangunkan dalam masa yang singkat dan menghasilkan kualiti yang terbaik.



Rajah 1 Model Kerangka Web Django

Rajah 1 menunjukkan model kerangka web bagi Django. Secara dasarnya, terdapat empat modul yang digunakan semasa pembangunan web iaitu *model*, *urls*, *view* dan *template*. *Model* mempunyai maklumat-maklumat pangkalan data seperti jenis dan saiz data. *Urls* mengawal alamat web yang bakal dimasukkan oleh pengguna. Jika pengguna memasukkan sesuatu alamat, *urls* akan berinteraksi dengan *View* untuk tindakan seterusnya. *View* mempunyai segala kod yang membenarkan Django membuat paparan antaramuka portal. *Template* pula merupakan fail-fail HTML yang dibangunkan untuk memaparkan antaramuka sesuatu halaman pada portal.

Bagi proses penghasilan model analisis sentimen, suasana integrasi pembangunan yang digunakan adalah menggunakan Google Colaboratory di mana seluruh proses dalam pembinaan model analisis sentimen menggunakan bahasa aturcara Python. Hasil model yang dibina digunakan untuk meramal dataset ciapan Twitter dan divisualisasikan mengikut kadar peratusan sentimen. Terdapat beberapa proses yang dijalankan bagi menghasilkan model analisis sentimen menggunakan pembelajaran mesin secara diselia. Model yang telah dilatih disimpan dalam format *pickle* dan dipanggil semula untuk membuat ramalan pada dataset yang belum dilabel.



Rajah 2 Seni Bina Analisis Sentimen

Rajah 2 menunjukkan proses yang dijalankan sebelum model analisis sentimen dihasilkan. Proses perolehan dataset dijalankan dengan membuat pensekerapan data dari Twitter dan kemudiannya melalui proses pra-pemrosesan data. Dalam proses pra-pemrosesan data, data-data ciapan Twitter dibersihkan sebelum membuat pelabelan data secara manual. Setelah itu, data yang dibersihkan dilabel bagi menandakan setiap sentimen bagi setiap ciapan. Dataset yang telah siap dilabel dibawa ke proses seterusnya iaitu proses pengekstrakan ciri menggunakan teknik *Count Vectorizer*. Menggunakan teknik ini, kekerapan sesebuah perkataan diambil kira dan disimpan serta ditukarkan dari bentuk teks kepada bentuk vektor yang boleh difahami oleh mesin untuk membuat analisis.

CountVectorizer membina matriks di mana setiap perkataan unik mewakili lajur matriks, dan setiap sampel teks dari dokumen adalah baris dalam matriks. Nilai setiap sel tidak lain adalah jumlah perkataan dalam sampel teks tertentu.

Selepas proses pengekstrakan ciri dijalankan, model teks yang divektor, disimpan di dalam file pickle / .pkl. Modul Python pickle digunakan untuk serializing dan deserializing objek Python. Pickle menserialisasi objek terlebih dahulu sebelum menulisnya ke dalam fail. *Pickling* adalah cara untuk menukar objek python (list, dict, dll) menjadi aliran karakter. Secara dasarnya, aliran karakter ini mengandungi semua maklumat yang diperlukan untuk membina semula objek dalam skrip python yang lain atau model ini boleh dipanggil semula untuk membuat ramalan pada dataset yang belum dilabel.

Selain itu, untuk menambahkan lagi maklumat untuk paparan pengguna portal, analisis data telah dilakukan dan dipaparkan pada bahagian papan pemuka portal. Data yang dianalisis diperoleh daripada laman sesawang carian kerja seperti Glassdoor, Jora, Jobstreet dan Indeed. Data-data seperti jenis pekerjaan, lokasi pekerjaan dan gaji yang ditawarkan di dalam laman sesawang disekrap dan dianalisis. Data-data yang dikaji berkhususkan kepada kerja-kerja yang ditawarkan secara gig atau di bawah carian 'freelancer'. Hal ini kerana kata kunci 'freelancer' yang digunakan dalam laman sesawang carian kerja adalah merupakan kerja-kerja yang berkonsepkan gig ekonomi di mana ia sesuai dengan fokus domain kajian. Dataset yang disekrap disimpan dalam fail .csv dan menggunakan kod Python dan 'library' *panda* untuk membuat pengelasan terhadap setiap atribut data dengan memanggil fail .csv dan membuat pengiraan kekerapan yang terbanyak bagi setiap atribut data.

Melalui dapatan analisis, atribut data yang dikategorikan dan dipaparkan di dalam portal adalah dalam bentuk graf dan statistik yang berinteraktif yang dijana menggunakan *chart.js* dalam bahasa aturcara Javascript. Maklumat seperti jenis pekerjaan yang terbanyak ditawarkan, lokasi yang menawarkan kerja secara gig dan syarikat-syarikat yang menawarkan kerja secara gig dapat dipaparkan hasil daripada analisis yang dijalankan.

Manakala dapatan data seperti komen atau 'review' hasil dari komen pekerja di ruangan 'company review' laman sesawang carian kerja, pengkaji dapat mengumpulkan hasil komen ini dan membuat analisis sentimen terhadap komen-komen yang dikumpul. Model analisis sentimen yang telah dibina pada proses analisis sentimen untuk data ciapan Twitter

pada awal pembangunan digunakan untuk membuat ramalan sentimen pada dataset ini. Selain itu, paparan perbualan terhangat menggunakan fungsi perkataan awan telah dibuat pada fasa pembangunan dan analisis data. Hasil daripada dataset ulasan syarikat yang menawarkan kerja secara gig, paparan sentimen terhadap syarikat yang diulas dipaparkan di papan pemuka portal pada bahagian 'company analysis'.

4.4 Fasa Pengujian

Dalam fasa pengujian, sistem yang dibangunkan, dipastikan agar bebas daripada *bug* dan segala ralat kritikal serta selari dengan semua fungsi yang telah dirancang pada pelan perancangan awal. Ujian dijalankan bagi memastikan kod yang ditulis berjalan dengan lancar dan mematuhi syarat yang digariskan pada fasa pembangunan. Pengujian juga telah dibuat dengan mencuba semua fungsi utama portal dan membuat senarai periksa agar tiada ralat yang kritikal terjadi dalam proses penghasilan.

Setelah model analisis sentimen dibina, proses validasi model akan dilakukan untuk mendapatkan perincian tentang prestasi sesuatu model itu. Pada dasarnya, dataset telah dibahagikan secara rawak semasa proses melatih model. Alasan untuk melakukan perkara ini adalah untuk melihat keberkesanan model untuk menilai data yang belum diketahui sebelumnya. Dataset dibahagikan kepada 80% untuk dilatih dan 20% digunakan untuk membuat pengujian ke atas model. Kelebihan cara ini adalah model pembangun dapat melihat bagaimana model ini memberi kesan dalam menilai data yang belum diketahuinya. Kemudian data ini akan diramal menggunakan dua pengkelas yang berbeza bagi membuat perbandingan skor.

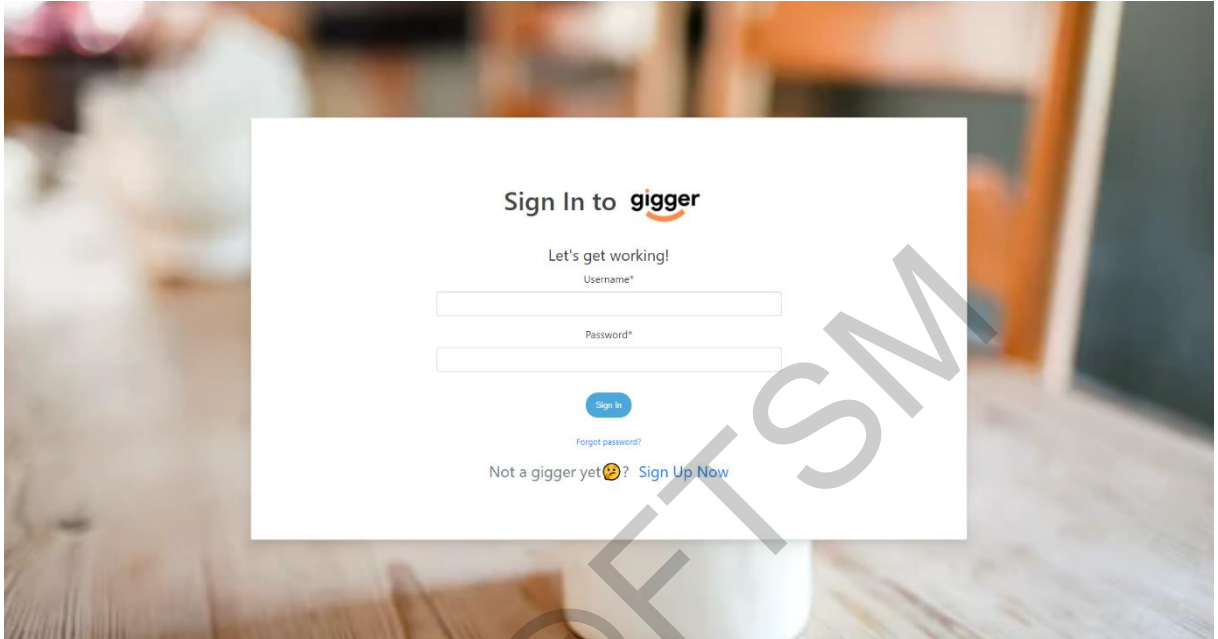
Pada bahagian ini, hasil daripada proses pembangunan portal Gigger dibincangkan bagi membuktikan hasil akhir pembangunan projek. Portal Gigger ini menggunakan HTML, CSS dan Javascript dalam pembinaan antaramuka dan mengintegrasikan Bootstrap 4 untuk menghasilkan sebuah portal yang menarik dan minimal. Manakala bahagian ‘backend’ pula kerangka Django memainkan peranan penting dalam mengawal dan menyelia semua kehendak sistem. Django menggunakan sepenuhnya bahasa aturcara python.



How does  work?

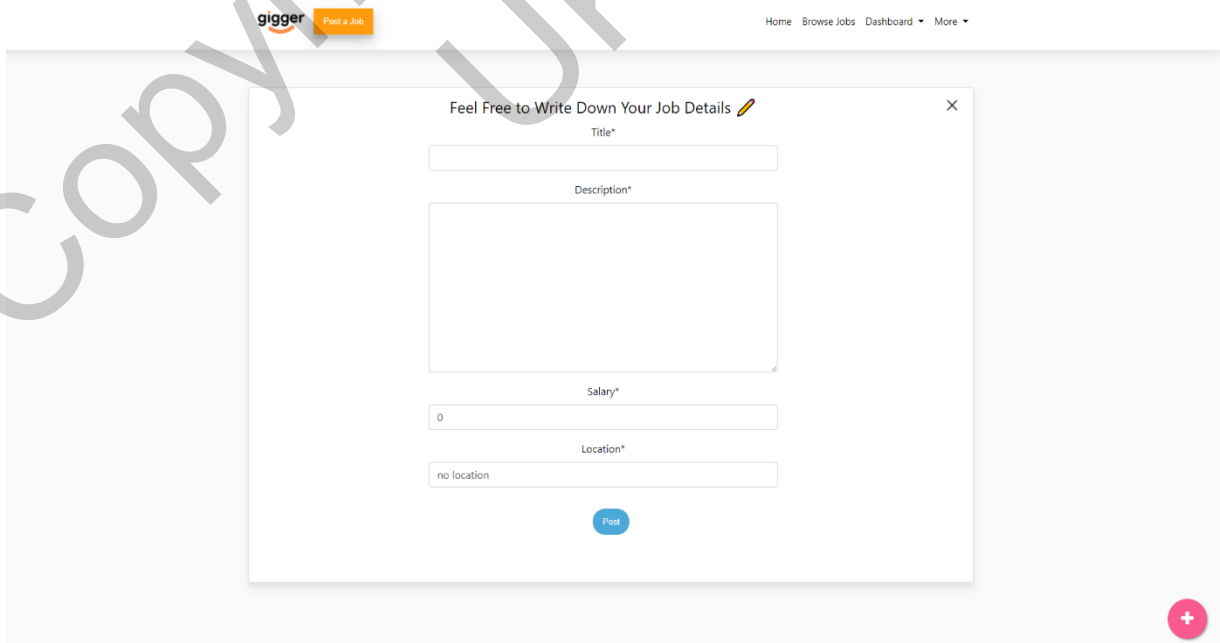
Rajah 3 Antaramuka halaman utama

Rajah 3 menunjukkan rekabentuk antaramuka halaman utama apabila pengguna membuat navigasi ke <http://gigger-fyp.herokuapp.com>. Pengguna boleh log masuk dan membuat pendaftaran jika belum mendaftar dengan menekan pilihan pada bahagian ‘navigation bar’.



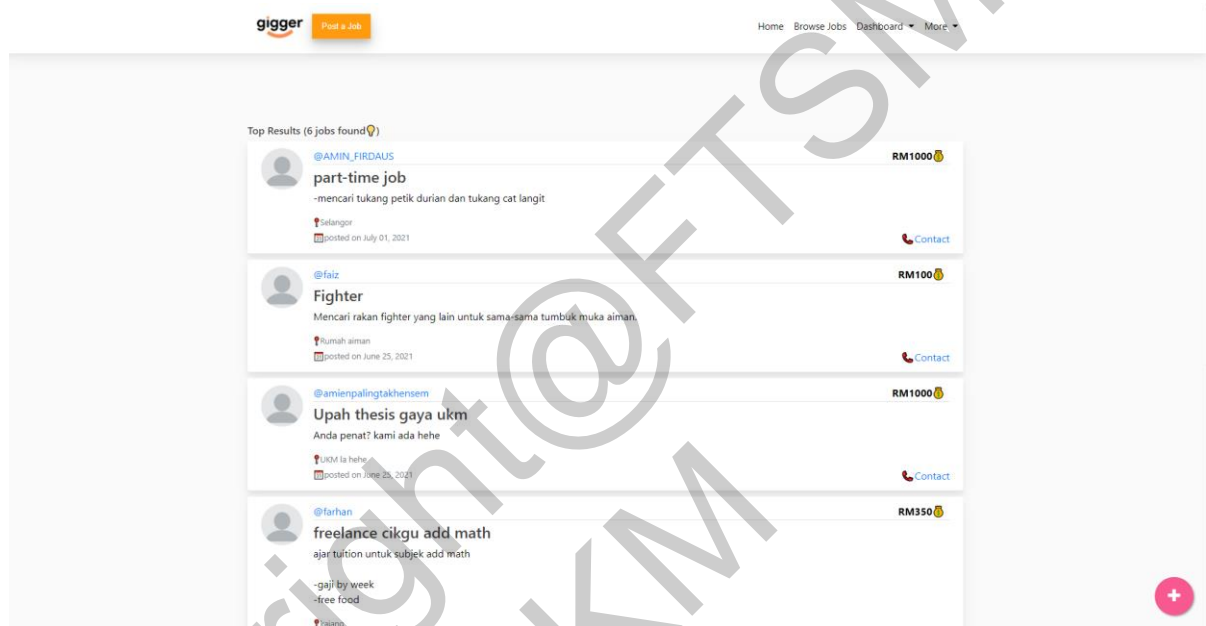
Rajah 4 Antaramuka halaman utama

Rajah 4 menunjukkan antaramuka log masuk portal gigger. Pengguna perlu memasukkan 'username' dan kata laluan seperti yang didaftarkan dan sistem akan membuat semakan jika pengguna sudah mendaftar jika belum ralat akan dipaparkan.



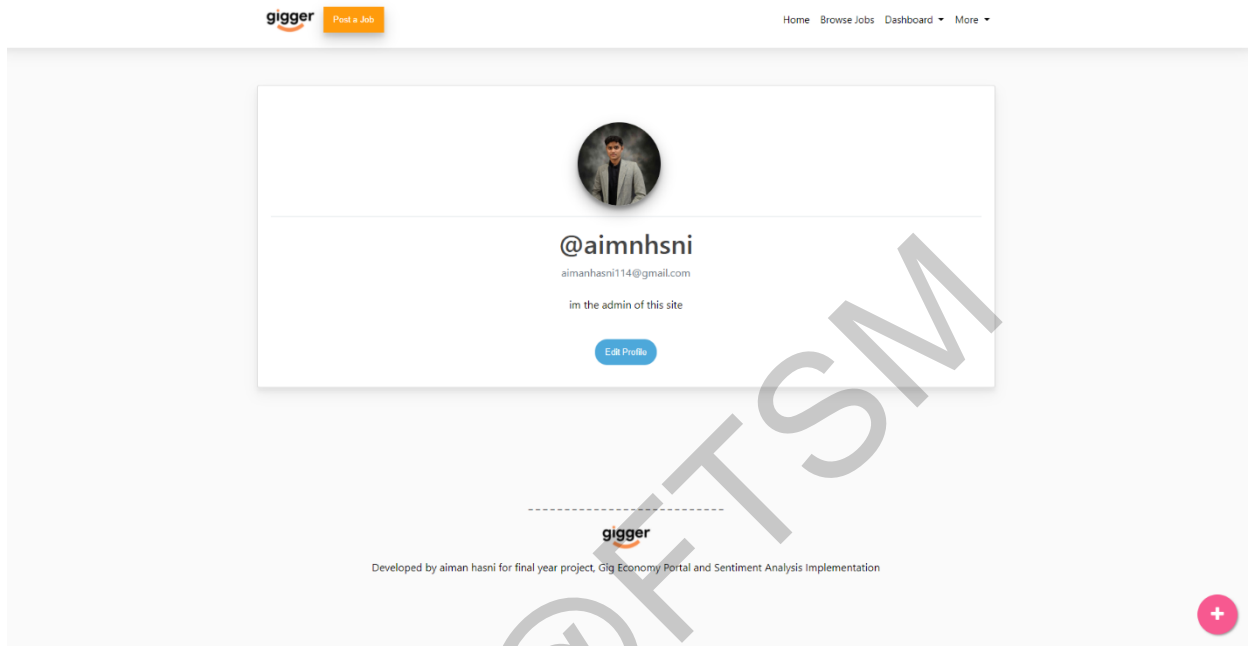
Rajah 5 Antaramuka penawaran kerja

Rajah 5 menunjukkan antaramuka untuk pengguna membuat penawaran kerja. Pengguna perlu mengisi perincian mengikut format yang ditetapkan oleh sistem jika tidak ralat akan diaparkan. Pengguna jugak tidak dibenarkan untuk meninggalkan ruangan kosong atau sistem akan paparkan ralat.



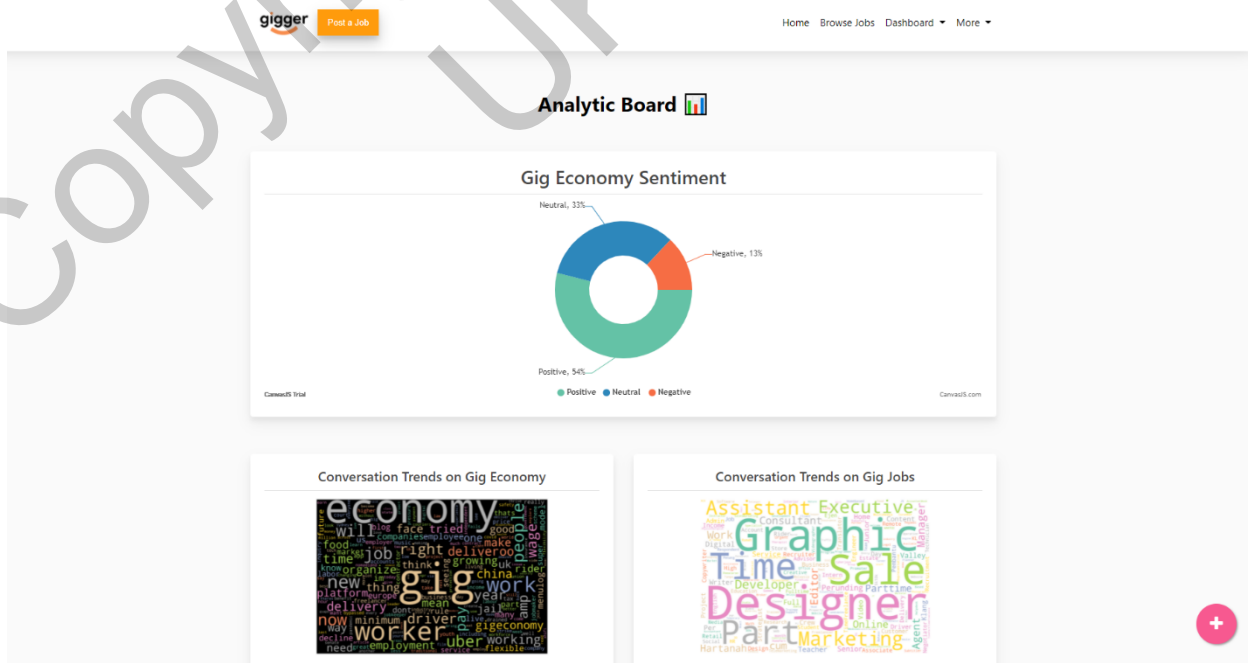
Rajah 6 Antaramuka senarai kerja

Rajah 6 menunjukkan antaramuka untuk pengguna mencari senarai kerja yang ditawarkan oleh pengguna lain. Sistem akan memuat naik setiap kerja yang dimuatnaik oleh pengguna lain dan terdapat butang 'contact' untuk para pencari kerja menghubungi pengguna yang menawarkan kerja melalui email.

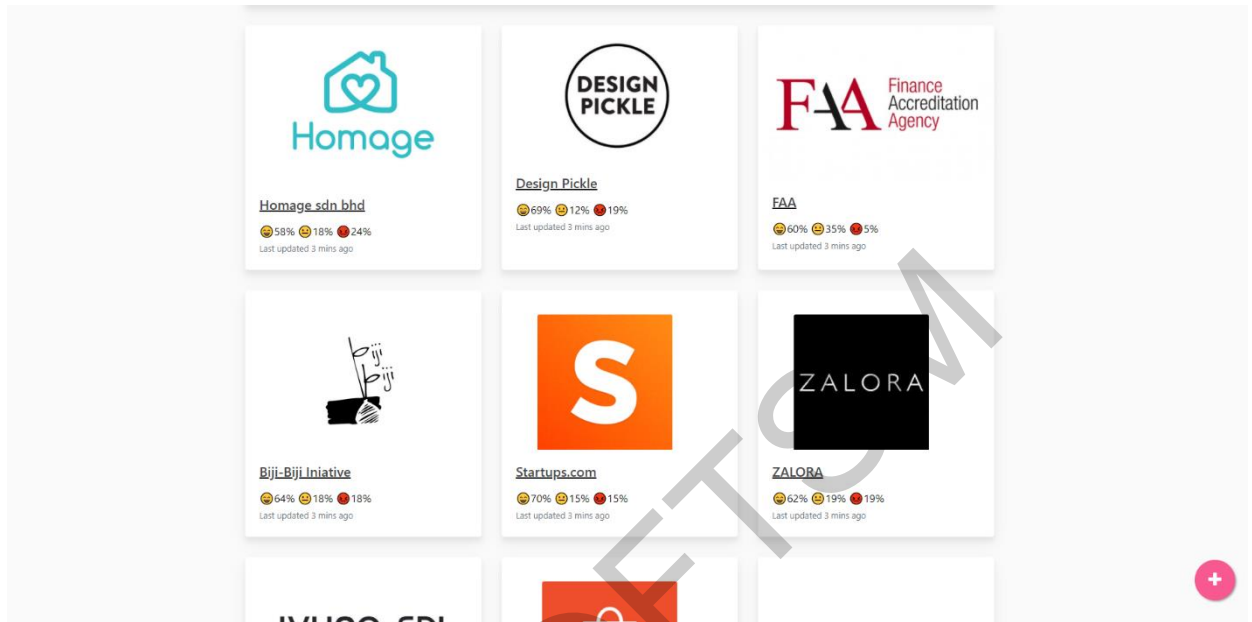


Rajah 7 Antaramuka profil pengguna

Rajah 7 menunjukkan antaramuka profil pengguna di mana setiap pengguna yang berdaftar boleh melihat profil masing-masing dan membuat kemaskini terhadap maklumat yang disediakan sistem secara 'default'.



Rajah 8.1 Antaramuka papan pemuka



Rajah 9.2 Antaramuka 'company analysis'

Rajah 9.1 dan 9.2 menunjukkan antaramuka 'company analysis' di mana paparan sentimen terhadap syarikat yang menawarkan kerja secara gig dipaparkan bagi memberi informasi kepada pengguna berkenaan sentimen komen terhadap syarikat-syarikat yang dikomen oleh pekerja.

6 KESIMPULAN

Kesimpulannya, proses pembangunan Portal Gig Ekonomi dan Implementasi Analisis Sentimen dapat dibangunkan dengan lancar. Model analisis sentimen yang dibina berjaya membuat ramalan pada dataset dengan ketepatan yang memuaskan dan diterima. Selain itu, hasil analisis data juga telah berjaya dibuat dan berjaya dipaparkan pada portal yang dibangunkan.

Semasa pembangunan Portal Gig Ekonomi dan Implementasi Analisis Sentimen ini, beberapa masalah telah dikenalpasti sepanjang pembangunan. Antara masalah yang terbesar adalah kekurangan dataset yang dipercayai (*unreliable*). Dataset latihan yang digunakan untuk membentuk model analisis sentimen adalah daripada ciapan Twitter dari peringkat global. Hal ini kerana kata kunci *gig economy* digunapakai ekoran daripada hasil ciapan dalam Bahasa Melayu yang kurang memuaskan dan tidak mencapai kehendak jumlah dataset untuk dikaji. Oleh itu, hasil sentimen tidak mewakili rakyat Malaysia

sebaliknya mempunyai skop yang tidak bertumpu kerana skop global digunakan. Selain itu, dataset untuk kerja-kerja gig untuk proses analisis data juga tidak mendapat sambutan yang memuaskan kerana kebanyakan data yang disekrap adalah bercampur dengan penawaran kerja yang tidak berunsurkan ekonomi gig. Walaupun begitu, proses-proses analisis data masih dapat dijalankan ekoran pengumpulan data secara berperingkat di mana data disekrap selang beberapa minggu bagi mendapatkan hasil yang tidak bercampur dan lebih banyak dataset yang boleh diperoleh.

Akhir sekali, projek yang dirancang telah berjaya disiapkan dan dihasilkan dengan wujudnya portal ini di internet dan para pengguna boleh menggunakannya secara nyata di <http://gigger-fyp.herokuapp.com> .

7 RUJUKAN

Ela Pustulka-Hunt, Thomas Hanne, Eliane Blumer, Manuel Frieder, 2018.

Multilingual Sentiment Analysis for a Swiss Gig. Conference: 6th International Symposium on Computational and Business Intelligence, 10.1109/ISCBI.2018.00028.

Haseena Rahmath, Tanvir Ahmad, July 2014. Sentiment Analysis Techniques - A Comparative Study . International Journal of Computational Engineering & Management, Vol. 17 Issue 4,.

Olga Kolchyna, Th´arsis T. P. Souza, Philip C. Treleaven and Tomaso Aste, 2015. Methodology for Twitter Sentiment Analysis, Department of Computer Science, UCL, Gower Street, London, UK.

Copyright@FTSM
UKM