

SISTEM PEMROSESAN DATA DESKRIPTIF SEBAGAI SAAAS

Nur Najihah Binti Nizaruddin

Assoc. Prof. Dr. Mohd Zakree Ahmad Nazri

Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

Selari dengan perkembangan teknologi maklumat, Data Raya telah menjadi satu elemen penting dalam strategi pelbagai perusahaan untuk memperoleh keuntungan dalam dunia perniagaan yang semakin kompetitif. Pelbagai data yang diterima setiap hari melalui telefon bimbit dan komputer riba telah menyebabkan pelbagai perusahaan memerlukan ruang storan yang besar untuk menyimpan data. Data-data yang diterima mampu ditafsirkan kepada sesuatu maklumat atau pengetahuan yang baru. Pelbagai sistem pemrosesan data analitik telah dihasilkan namun masih terdapat ruang untuk ditambah baik. Sistem yang bakal dibangunkan ini akan memfokuskan kepada data deskriptif iaitu pemahaman tentang setiap atribut yang dimiliki oleh set data. Konsep perisian sebagai perkhidmatan (SaaS) akan digunakan untuk menyelesaikan masalah penyimpanan data yang kecil dan kos operasi perisian. Hal ini demikian kerana SaaS akan berperanan sebagai penyedia awan maka pengguna dan pembangun sistem hanya memerlukan internet sahaja dan tidak perlu membeli perkakasan atau memuat turun sebarang sistem. Sistem yang akan dibangunkan ini akan menggunakan Python, HTML dan MySQL dengan rangka kerja Flask. Pengguna yang menggunakan sistem ini iaitu pengguna umum akan mampu memahami kriteria atribut dan melihat gambaran set data melalui graf manakala admin akan dapat melihat jumlah pengguna umum yang berdaftar.

1 PENGENALAN

Seiring dengan perkembangan dalam menuju industri keempat, salah satu teras utama bagi mencapai revolusi ini ialah Teknologi Pemrosesan Awan. Teknologi awan ini dapat didefinisikan sebagai perkhidmatan perkomputeran yang mempunyai pelayan, penyimpanan, pangkalan data, rangkaian, perisian, analisis dan kepintaran secara Internet (“Awan”) (Microsoft 2019). Salah satu perkhidmatan yang ditawarkan oleh teknologi awan ialah perisian sebagai perkhidmatan (SaaS). SaaS ialah satu alternatif untuk menyampaikan aplikasi perisian melalui Internet yang memerlukan penggunaanya untuk melanggan. Melalui SaaS, penyedia awan bertanggungjawab menguruskan

aplikasi perisian, infrastruktur asas dan mengendalikan sebarang penyelenggaraan, seperti peningkatan perisian dan pengawal keselamatan.

Selain itu, analisis data juga merupakan alternatif penting bagi mencapai revolusi industri 4.0. Analisis data adalah proses memeriksa set data besar untuk memberi gambaran dan menyediakan maklumat yang berguna daripada data yang dikumpulkan (Matthew, 2014). Salah satu kaedah analisis data ialah analisis data deskriptif. Data deskriptif ialah data yang dapat difahami dalam bentuk rata-rata (mean), jumlah (sum), simpangan baku (standard deviation), varians (variance), rentang (range), nilai maksimum, nilai minimum dan lain-lain (Coladarci A.P. & Coladarci A.T, 1980). Melalui pemprosesan data berbentuk deskriptif, pengguna dapat mengenalpasti kriteria utama atribut yang terlibat. Kaedah ini digunakan dalam pelbagai industri sebagai contoh perbankan dan bisnes. Dengan ini, pengguna akan mudah untuk menganalisa dan membuat keputusan bagi meningkatkan kualiti perniagaan atau sesebuah perkhidmatan. Pengguna juga mampu untuk meramal keputusan sesuatu keadaan yang bakal berlaku melalui proses analisis data.

Berdasarkan perkara-perkara diatas, SaaS sangat sesuai digunakan bagi membangunkan sebuah sistem pemprosesan data berbentuk deskriptif. Melalui kombinasi ini, sistem pemprosesan data deskriptif boleh digunakan secara atas talian dan membantu pengguna untuk memahami data yang telah diproses melalui konteks visual secara corak, trend atau korelasi dengan kadar segera. Oleh itu, pengguna mampu membuat keputusan yang lebih baik untuk menjimatkan masa dan tenaga.

2 PERNYATAAN MASALAH

Menurut Dr Nurfadhlina, kemajuan teknologi penyimpanan digital dan capaian rangkaian yang kian meluas (Adee, 2018) telah memberi peluang kepada syarikat perniagaan dan agensi luar untuk memajukan perniagaan mereka. Misalnya peniaga dapat menambah naik teknik pembungkusan produk mereka dan memperoleh keuntungan hasil daripada maklum balas yang diterima daripada pelanggan melalui media sosial dan saluran promosi media digital.

Namun begitu, terdapat beberapa masalah yang dihadapi oleh pengguna dalam menguruskan data digital ini. Antara masalah yang dihadapi adalah seperti berikut:

1. Kos operasi lesen dan penyelenggaraan perisian yang mahal.

Bagi meningkatkan keuntungan dan kepuasan pelanggan, syarikat perniagaan telah membelanjakan wang dalam kadar yang tinggi untuk membayar lesen dan menyelenggara perisian yang sedia ada. Syarikat perniagaan juga perlu membeli perkakasan atau pelayan untuk menyimpan semua maklumat pengguna. Kesukaran menampung kos operasi perniagaan akan menyebabkan syarikat tersebut tertinggal di belakang dan berkemungkinan bankrap.

2. Ruang storan yang tidak mencukupi.

Pengguna atau syarikat memerlukan ruang storan yang besar pada harga yang mahal untuk menyimpan data pelanggan-pelanggan mereka. Selain itu, ruang fizikal yang besar juga perlu disediakan untuk menampung maklumat tersebut.

3. Sistem pemprosesan data deskriptif sedia ada susah dianalisa dan mahal.

Gambaran data melalui carta dan graf yang dihasilkan susah difahami dan dianalisa. Perisian profesional yang sedia ada seperti Tableau dan Power BI juga agak mahal. Ini menyukarkan pengguna misalnya pelajar untuk mencuba dan menggunakan sistem analisis data.

3 OBJEKTIF KAJIAN

Tujuan utama projek ini adalah untuk membangunkan sebuah sistem perisian yang mampu memproses data berbentuk deskripsi secara atas talian. Bagi mencapai objektif ini, terdapat dua matlamat yang perlu dicapai:

- i. Membangunkan sebuah enjin pemprosesan data deskripsi menggunakan bahasa pengaturcaraan Python.
- ii. Menyediakan pelantar perisian untuk memproses data deskripsi kepada SaaS.

4 METOD KAJIAN

Dalam pelaksanaan projek pembangunan Sistem Pemprosesan Data Deskriptif sebagai SaaS, metodologi kajian yang dipilih adalah Pembangunan Tangkas. Ini demikian kerana Teknik pecutan tunggakan kerja (sprint backlog) yang boleh disusun mengikut keutamaan dan kemampuan fleksibel yang membenarkan sebarang perubahan dilakukan sepanjang projek dijalankan (Charles, 2015).

Menurut Ghule (2014), jangkaan jumlah masa yang diperlukan untuk menyempurnakan keperluan yang diberikan bagi menghasilkan sesebuah perisian adalah sangat buruk. Melalui kaedah pembangunan tangkas, sebarang modifikasi dapat dilakukan terhadap sesuatu keperluan dapat diatasi secara efektif tanpa perlu memanjangkan jangka masa untuk membangunkan sistem ini.

Bagi memastikan semua keperluan perisian dapat disempurnakan dalam jangka masa yang ditetapkan, pembangunan sistem ini akan dibahagikan kepada beberapa fasa. Pembahagian beberapa fasa ini akan membenarkan pihak berkepentingan untuk terlibat dan memberi sebarang maklum balas mengenai penghasilan sistem secara langsung (Charles, 2015). Rajah 1.1 memaparkan model pembangunan tangkas yang digunakan dalam membangunkan sistem ini.



Rajah 1 Teknik Pembangunan Tangkas

4.1 Fasa Keperluan

Fasa ini merangkumi aspek mengenalpasti masalah yang ingin diselesaikan (Asyraf & Amirah, 2017). Apabila pernyataan masalah sudah diperolehi, aspek keperluan yang penting juga dapat dikenalpasti. Selain itu, pengumpulan data daripada kajian lepas seperti jurnal, buku dan laman sesawang yang boleh dipercayai perlu dijalankan bagi menepati objektif dan skop kajian.

4.2 Fasa Perancangan

Fasa ini pula melibatkan sorotan kesusasteraan di mana semua jurnal dan kajian lepas yang dikumpulkan akan dikaji untuk meningkatkan dan menambahbaik nilai sesuatu sistem. Idea penyelesaian masalah akan disusun mengikut kepentingan dan diasingkan dalam Scrum. Sekali gus dapat memastikan keperluan projek dapat dipenuhi dalam jangka masa yang ditetapkan (Charles, 2015).

4.3 Fasa Reka Bentuk

Fasa ini akan memberi fokus terhadap seni bina sistem dan perisian yang akan dibangunkan berdasarkan spesifikasi keperluan yang telah dikenalpasti. Rekabentuk sistem perlu memenuhi kehendak pengguna. Antara aktiviti yang akan dilakukan ialah menyediakan rajah kes guna, rajah urutan, rajah kelas, model-papar-kawal (MVC) dan antara muka (IU) untuk memberi gambaran mudah mengenai sistem ini.

4.4 Fasa Perlaksanaan

Proses pengekodan dalam rangka kerja Flask dan bahasa pengaturcaraan Python akan digunakan dalam fasa ini. Proses Scrum juga akan kerap dijalankan. Oleh itu, spesifikasi keperluan yang besar akan dibahagikan dalam struktur yang kecil supaya kesemua spesifikasi ini dapat disiapkan dalam kadar waktu yang ditetapkan.

4.5 Fasa Pengujian

Setiap komponen akan diuji untuk memastikan fungsi sistem berjalan dengan baik berdasarkan objektif dan kehendak pengguna. Kesalahan sistem juga akan dicari sehingga tahap keyakinan dapat dicapai.

4.6 Fasa Penilaian

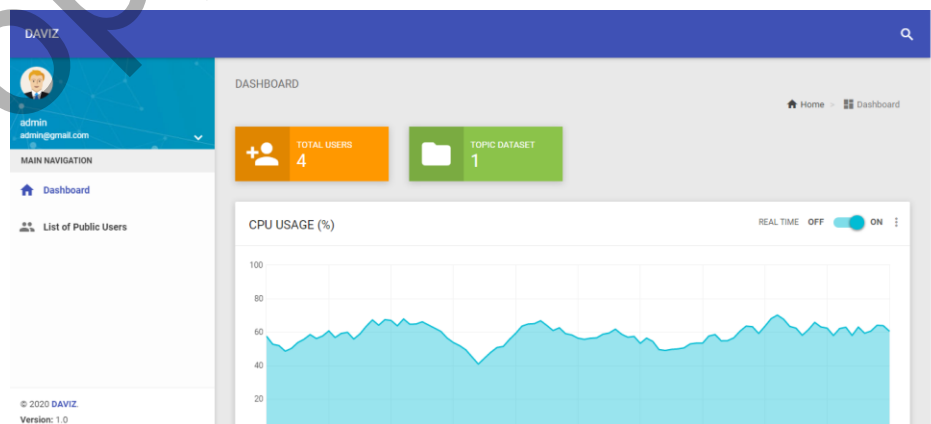
Fasa ini pula membenarkan pengguna sasaran untuk mencuba sistem yang dihasilkan. Proses penilaian ini akan dilakukan sehingga mencapai tahap kualiti yang diperlukan dan dipersetujui oleh pihak berkepentingan.

5 HASIL KAJIAN

Projek pembangunan Sistem Pemrosesan Data Deskriptif sebagai SaaS ini secara keseluruhannya telah berjaya dibangunkan mengikut spesifikasi yang ditetapkan dan mencapai objektif utama untuk membangunkan sebuah enjin pemrosesan data deskripsi menggunakan bahasa pengaturcaraan Python dan menyediakan pelantar perisian menggunakan AWS. Rajah 2 dan Rajah 3 menunjukkan antaramuka bagi dua pengguna sistem iaitu pengguna umum dan admin.



Rajah 2 Laman utama Pengguna Umum



Rajah 3 Laman utama Admin

(C)

Rajah 4 Antaramuka laman utama (A), pendaftaran akaun (B) dan log masuk (C)

Selepas proses membangunkan Rajah 4 selesai, sistem ini akan membangunkan 4 fungsian utama bagi pengguna umum iaitu proses memuat naik set data, paparan kriteria atribut bagi data, penghasilan graf dan proses penyimpanan lembaran kerja. Proses memuat naik set data hanya membenarkan fail format .csv sahaja. ID bagi setiap set data yang dimuat naik akan digunakan untuk paparan kriteria atribut dan proses memuat naik set data ke pangkalan data. Apabila kriteria atribut set data telah dipaparkan, pengguna umum boleh menghasilkan graf berdasarkan atribut set data dan memuat turun hasil lembaran kerja ke komputer pengguna umum.

ID	DATA NAME	ACTION
4	Iris.csv	VIEW DELETE

(A)

DAVIZ

ATTRIBUTES DETAILS Graph

Attributes Name	
0	sepal.length
1	sepal.width
2	petal.length
3	petal.width
4	variety

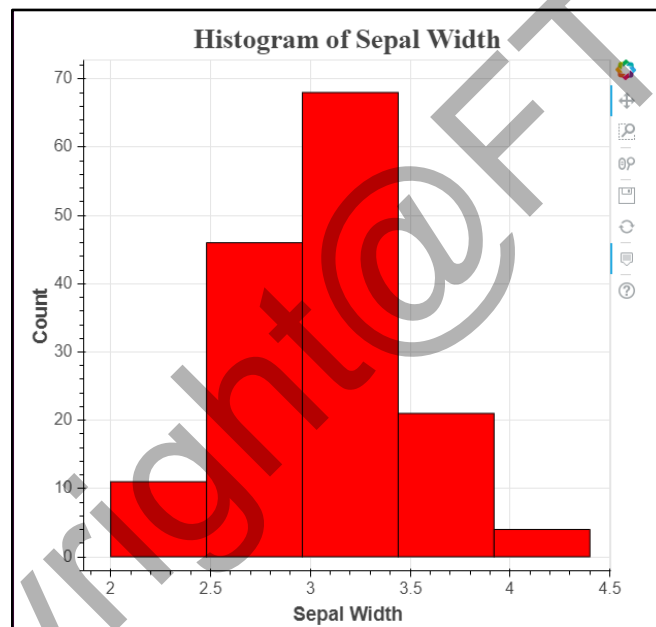
Data Type	
sepal.length	float64
sepal.width	float64
petal.length	float64
petal.width	float64
variety	object

Missing Values	
sepal.length	0
sepal.width	0
petal.length	0
petal.width	0
variety	0

Any Missing Values? **Total**
False 0

Statistic

(B)



(C)

Rajah 5 Antaramuka bagi proses memuat naik set data (A), paparan kriteria atribut (B) dan contoh penghasilan graf berdasarkan atribut dan simbol muat turun lembaran kerja (C)

Setakat ini, sistem pemprosesan data deskriptif ini hanya membenarkan Admin untuk melihat butiran diri pengguna umum sahaja. Berikut ialah Rajah 6 bagi fungsian melihat butiran diri pengguna umum.

ID	Username	Email	Career
1	jiha	jiha@gmail.com	student
3	meor	meor@gmail.com	data analyst

Rajah 6 Butiran diri pengguna umum

Pengguna bagi sistem ini akan mudah memahami fungsi bagi sistem ini kerana reka bentuk antara muka yang cara Sistem yang dibangunkan menggunakan reka bentuk yang mudah difahami dan menggunakan bahasa Inggeris sebagai bahasa pengantar supaya diharapkan pengguna antarabangsa juga mampu untuk memahami fungsi sistem yang dibangunkan.

Sistem yang telah siap dibangunkan ini akan diuji dengan menggunakan teknik kotak hitam sahaja dan diubahsuai untuk memenuhi kehendak pihak berkepentingan dan mengikut objektif spesifikasi sistem yang ditetapkan.

6 KESIMPULAN

Sistem yang telah dibangunkan ini dijangka dapat membantu memudahkan pengguna untuk menganalisa set data yang mereka perolehi atau kumpul untuk menghasilkan informasi yang berguna sekali gus membantu proses operasi syarikat atau sesebuah organisasi.

Proses menganalisa set data menggunakan perpustakaan-perpustakaan Python juga memudahkan pembangun sistem memproses set data yang dimuat naik oleh pengguna. Manakala, penggunaan SaaS dalam membangunkan sistem ini juga dapat mengurangkan isu storan data dan lain-lain. Diharapkan sistem yang dibina dapat dimanfaatkan dan ditambahbaik pada masa hadapan.

7 RUJUKAN

- Adee. 2018. Sihat Sejahtera Melalui Data Analitis Data Raya. <https://www.majalahsains.com/sihat-melalui-analitis-data-ray/> [23 September 2019].
- Amirul Shafiq Ahmad Zuperi. 2015. *Chapter 1 sistem maklumat dalam dunia global MIS*. Slaid. Malaysia: LinkedIn.
- Brown, J.R, Earnshaw, R, Jern, M & Vince, J. 1995. *Visualization Using Computer Graphics to Explore Data and Present Information*. ISBN 10471129917. New York: Penerbit John Wiley & Sons, Inc.
- Charles Edeki. 2015. Agile Software Development Methodology. *European Journal of Mathematics and Computer Science* 2(1): 1-6.
- Coladarci, A.P & Coladarci, T.P. 1980. *Elementary Descriptive Statistics*. ISBN 100534007821. Amerika Syarikat: Penerbit Wadsworth, Inc.
- Eric, E.s, Linderman, M.D., Sorenson, J, Lee, L & Nolan, G.P. 2010. Computational solutions to large-scale data management and analysis. *National Institutes of Health*. 11(9): 647-657.
- Matthew Wall. 2014. Big Data: Are you ready for blast off? <https://www.bbc.com/news/business-26383058> [20 September 2019].
- Mbarek Marwan, Ali Kartit & Hassan Ouahmane. 2018. A Secured Data Processing Technique for Effective Utilization of Cloud Computing. *Journal of Data Mining and Digital Humanities*, Episciences.org. hal-01466986v2f (1): 1-13.
- Microsoft. 2019. What is cloud computing?. <https://azure.microsoft.com/en-us/overview/what-is-cloud-computing/> [23 September 2019].
- Muhammad Asyraf Zulkaple, Amirah Ismail. 2017. Media Sosial Interaktif Alumni. Degree. Tesis, Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia.
- Nor Aslamiah Ali. 2013. Cloud Computing: Revolusi Perkomputeran Masa Kini. *DimensiKOOP* Disember: 10-15.