

ANALISIS PENJUALAN PRODUK MAKANAN DARIPADA LAMAN WEB E-DAGANG DENGAN MENGGUNAKAN TEKNIK PERLOMBONGAN DATA

Tan Kai Shen
Sabrina Tiun

Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

Dalam dunia yang berkembang dengan pesat ini, E-dagang telah berkembang pada kadar yang tidak dapat dijangkakan di seluruh dunia disebabkan oleh penggunaan Internet dan aplikasi laman Web yang semakin banyak. E-dagang telah menyediakan satu platform yang baik untuk masyarakat memulakan perniagaan. Ia hanya memerlukan kos permulaan yang sangat rendah untuk masyarakat membina perniagaan sendiri. Di Malaysia, rakyat Malaysia telah dibahagikan kepada tiga golongan, iaitu Kumpulan Tertinggi 20% (T20), Kumpulan Pertengahan 40% (M40) dan Terendah 40% (B40) berdasarkan kaji selidik Penyiasatan Pendapatan Isi Rumah dan Kemudahan Asas (PPIR & KA) oleh Jabatan Perangkaan Malaysia. Walaupun rakyat Malaysia telah dibahagikan kepada tiga kumpulan yang berbeza, kerajaan Malaysia tetap menggalakkan keusahawanan antara tiga kumpulan ini. Objektif projek ini adalah untuk membangunkan sebuah sistem yang meramalkan produk makanan yang berpotensi sebagai laris jualan untuk dijual melalui analisis data siri masa jualan produk makanan dengan menggunakan teknik perlombongan data. Produk makanan yang diramalkan akan bertindak sebagai cadangan produk makanan kepada pengusaha B40 untuk mereka memulakan perniagaannya di laman web E-dagang. Teknik pengikisan web digunakan untuk mengumpulkan maklumat daripada laman web E-dagang. Manakala, teknik perlombongan data akan digunakan untuk membersihkan, mengubah, mengaruh dan mengekstrak maklumat yang penting daripada maklumat yang telah dikumpulkan. Kesimpulannya, projek ini akan membantu pengusaha B40 dalam pemilihan produk makanan yang sesuai bagi mereka memulakan perniagaan di laman web E-dagang.

1 PENGENALAN

Pada tahun 2020, ekonomi negara Malaysia telah terjejas akibat COVID-19. Sektor pelancongan Malaysia dan sektor-sektor berkaitan terjejas akibat sekatan kemasukan pelancongan asing dan kawalan pergerakan yang menjejaskan aktiviti pelancongan domestik. Perniagaan fizikal di Malaysia juga terjejas secara langsung akibat daripada arahan kawalan pergerakan penduduk dan pekerja-pekerja yang terkena jangkitan. Pada masa yang sama,

perniagaan digital di laman web E-dagang telah berkembang dengan cepat dan mampu menjadi 'penyelamat' kepada ekonomi negara Malaysia.

Sejak kewujudan Internet, telah berlaku satu lonjakan paradigma yang besar pada perniagaan. Segala urusan jual beli barangan dan perkhidmatan boleh dilakukan secara atas talian dengan menggunakan laman web E-dagang. E-dagang boleh didefinisikan sebagai satu proses membeli dan menjual produk dengan cara elektronik seperti aplikasi mudah alih dan Internet (Muhammad Hafiz Bin Mustafar 2016). Sebenarnya, E-dagang telah memberi banyak peluang kepada para rakyat untuk memulakan perniagaannya kerana ia telah menyediakan satu platform yang mudah dan senang diakses. E-dagang juga merupakan salah satu alternatif jalan untuk memulakan perniagaan selain daripada perniagaan fizikal terutamanya pada tahun 2020.

Rakyat Malaysia dibahagikan kepada tiga golongan iaitu Kumpulan tertinggi 20% (T20), Kumpulan pertengahan 40% (M40) dan Terendah 40% (B40) berdasarkan hasil kaji selidik Penyiasatan Pendapatan Isi Rumah dan Kemudahan Asas (PPIR & KA) oleh Jabatan Perangkaan Malaysia. Menurut hasil penyelidikan Kaji Selidik Kemudahan Asas dan Pendapatan Isi Rumah 2019 dari JSM (CompareHero.my 2020), kumpulan T20 mempunyai had pendapatan isi rumah yang lebih daripada RM10,959 dan had pendapatan isi rumah untuk kumpulan M40 adalah antara RM4,850 sehingga RM10,959 dan B40 adalah kurang daripada RM4,850. Walaupun rakyat Malaysia telah dibahagikan kepada tiga golongan, kerajaan Malaysia tetap menggalakkan keusahawanan pada semua golongan; antaranya dengan mengadakan kempen dan program berkaitan dengan keusahawanan. Malaysia Digital Economy Corporation (MDEC) telah diberi dana oleh kerajaan Malaysia sebanyak RM20 juta untuk program Perkhidmatan e-Dagang Setempat (PeDAS) bagi menaik taraf Pusat Internet Desa kepada hab E-dagang. Program ini telah memberikan peluang kepada rakyat Malaysia untuk mengasah kemahiran perniagaan dalam E-dagang.

Perlombongan data sering digunakan oleh syarikat-syarikat untuk mengubah data mentah kepada maklumat yang berguna. Dengan menggunakan perisian, corak set data besar dapat dilihat dan boleh digunakan oleh syarikat untuk menentukan keputusan. Ini adalah salah satu kegunaan perlombongan data dalam perniagaan untuk mengembangkan strategi pemasaran dengan efektif. Perlombongan data juga banyak digunakan dengan pelbagai cara seperti pemasaran pangkalan data, pengesanan penipuan dan lain-lain (ALEXANDRA TWIN 2020). Selain itu, perlombongan data juga sering digunakan untuk melakukan ramalan dalam penjualan produk untuk meningkatkan jualan. Perlombongan data juga mempunyai empat

teknik utama iaitu pengelompokan dan petua sekutuan sering digunakan untuk mengesan pola, manakala klasifikasi dan regresi sering digunakan untuk ramalan (Jiawei Han, Micheline Kamber 2012). Dalam kertas kerja ini, sebuah sistem akan dibangunkan dengan menggunakan teknik perlombongan data yang dapat meramalkan produk makanan yang berpotensi sebagai laris jualan untuk dijual melalui analisis data siri masa jualan produk makanan. Produk makanan yang diramal akan bertindak sebagai cadangan produk makanan kepada pengusaha B40 untuk memulakan perniagaannya di laman web E-dagang.

2 PENYATAAN MASALAH

Pada masa kini, memulakan perniagaan bukan satu perkara yang susah tetapi yang susah adalah memulakan perniagaan yang betul. Tiga perkara yang penting untuk memulakan perniagaan adalah belanjawan yang dapat ditanggung, tempat menjual dan apa yang patut dijual. Setakat ini, banyak analisis jualan produk daripada laman web E-dagang telah diadakan (Dimira Teneva 2016) tetapi analisis jualan produk yang membantu pengusaha B40 untuk memulakan perniagaan amat terhad.

Menurut hasil penyelidikan Kaji Selidik Kemudahan Asas dan Pendapatan Isi Rumah 2019 dari JSM, had pendapatan isi rumah kumpulan B40 adalah kurang daripada RM4,850. Belanjawan pengusaha B40 yang sangat terhad untuk memulakan perniagaannya telah menghalang banyak golongan B40 memulakan perniagaannya. Sumber yang terhad telah melemahkan semangat mereka untuk memulakan perniagaan.

Memulakan sebuah perniagaan juga memerlukan satu tempat untuk menjual produk. Lokasi yang strategi seperti pusat bandar adalah pilihan utama bagi menjual produk kerana mempunyai orang ramai. Tidak dapat dibantah bahawa lokasi-lokasi yang strategi memerlukan harga sewa yang tinggi. Ini juga akan jadi satu tanggungan kepada golongan B40 untuk memulakan perniagaan.

Selain itu, memilih produk yang betul juga adalah perkara yang amat penting untuk memulakan perniagaan. Produk-produk elektronik seperti telefon pintar dan komputer riba adalah produk yang paling laris pada musim norma baru ini. Tetapi produk-produk ini tidak realistik untuk diambil sebagai produk dijual oleh golongan B40 bagi memulakan perniagaan. Sebuah telefon pintar yang paling murah telah berharga sekurang-kurangnya RM400. Golongan B40 tidak mempunyai modal untuk mengambil stok yang banyak untuk memulakan

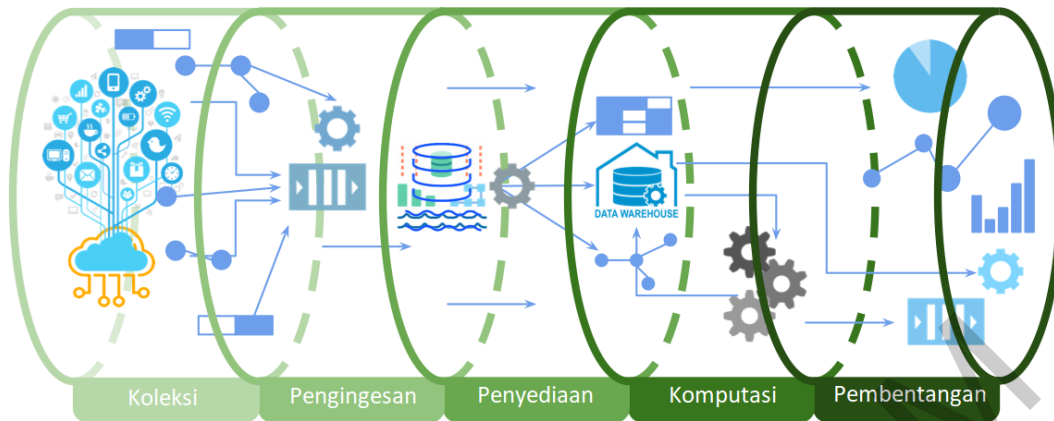
perniagaannya. Oleh itu, produk yang lebih murah dan produk yang berpotensi sebagai laris jualan seperti produk makanan boleh dijadikan sebagai sasaran produk jualan dan dijangka dapat keuntungan daripada jualan produk.

3 OBJEKTIF KAJIAN

Objektif projek ini adalah untuk membangunkan sebuah sistem yang meramalkan produk makanan yang berpotensi sebagai laris jualan untuk dijual melalui analisis data siri masa jualan produk makanan dengan menggunakan teknik perlombongan data. Produk makanan yang diramalkan akan bertindak sebagai cadangan produk makanan kepada pengusaha B40 untuk mereka memulakan perniagaannya di laman web E-dagang. Produk makanan yang diramalkan dijangka berpotensi untuk mendapatkan keuntungan daripada jualan produk kerana produk makanan diramal berdasarkan jualan sebelumnya dan ia adalah laris jualan berbanding dengan produk makanan yang lain daripada data siri masa jualan produk makanan yang diguna.

4 METOD KAJIAN

Dalam projek ini, metodologi yang akan digunakan adalah seni bina saluran data (*data pipeline architecture*) untuk membina sistem analisis. Seni bina saluran data adalah satu sistem yang dapat menangkap, mengatur dan mengarahkan data sehingga mendapat pandangan (snapLogic 2020). Seni bina saluran data mengatur data dan digunakan untuk menghasilkan laporan, dan analisis. Kombinasi teknologi dan protokol perisian yang sesuai akan digunakan untuk menguruskan visualisasi, transformasi dan pergerakan data dari pelbagai sumber mengikut tujuan perniagaan. Seni bina saluran data juga banyak digunakan dalam perniagaan kecerdasan (BI) dan analitis. Rajah 1 menunjukkan fasa-fasa metodologi bagi seni bina saluran data.



Rajah 1 Fasa-fasa dalam metodologi (Satish Chandra Gupta 2020)

4.1 Koleksi (*Collection*)

Dalam fasa ini, sumber data seperti laman web, aplikasi web, aplikasi telefon pintar akan diinstrumenkan untuk mengumpul data mentah.

4.2 Pengingesan (*Ingestion*)

Fasa ini akan mengepamkan data yang diinstrumenkan ke dalam saluran data. Data dibentuk dalam dua jenis iaitu gumpalan dan aliran. Semua data yang dikumpul akan disimpan ke dalam pangkalan data.

4.3 Penyediaan (*Preparation*)

Pengekstrakan, transformasi, load (ETL) akan beroperasi didalam fasa ini untuk membersihkan, menyesuaikan, membentuk, mengubah dan katalogkan gumpalan dan aliran data yang ada di dalam pangkalan data. Data yang melalui proses ETL akan dipersiapkan untuk digunakan pada fasa seterusnya.

4.4 Komputasi (*Computation*)

Dalam fasa ini, data boleh digunakan untuk analisis, sains data dan pembelajaran mesin. Dalam projek ini, kaedah yang sesuai akan ditentukan untuk menganalisis data yang ada.

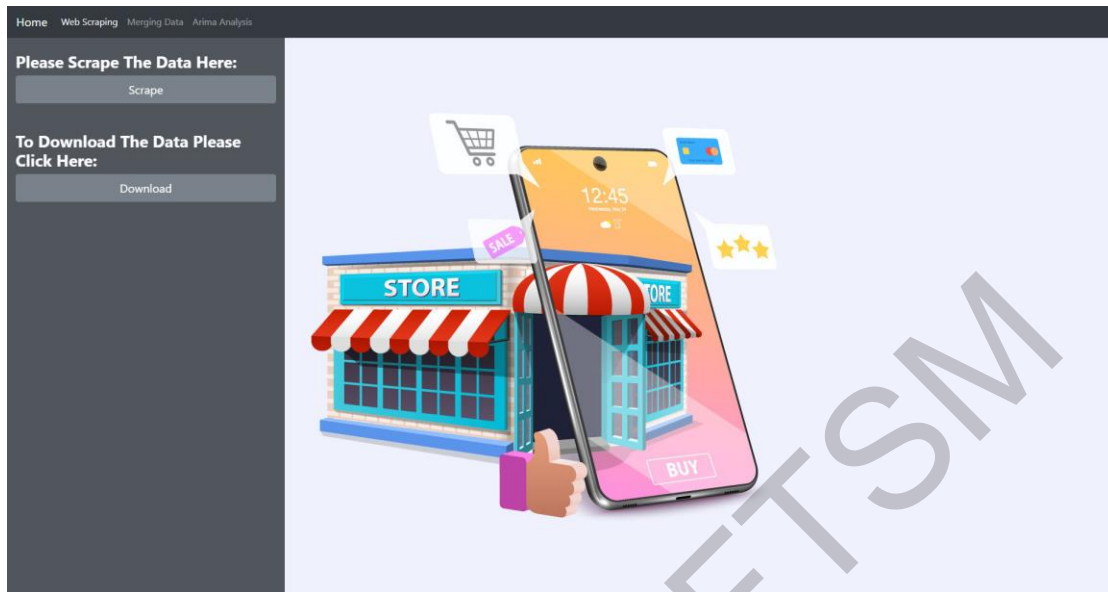
4.5 Pembentangan (*Presentation*)

Dalam fasa ini, model dan analisis boleh divisualisasikan sebagai papan semuka, graf dan lain-lain untuk dibentangkan kepada pengguna.

6 HASIL KAJIAN

Bahagian ini membincang hasil daripada proses pembangunan sistem yang bertujuan untuk meramalkan produk makanan yang berpotensi sebagai laris jualan untuk dijual sebagai cadangan produk kepada pengguna. Spesifikasi reka bentuk juga diterangkan dalam terperinci. Spesifikasi reka bentuk menerangkan bagaimana sistem melaksanakan keperluan yang telah digariskan dalam keperluan fungsian (*Ofni Systems*). Spesifikasi reka bentuk amat penting dalam pembangunan sistem kerana ia memberikan arahan yang jelas dalam mengenai pembinaan dan prestasi. Dengan ini, permasalahan dapat dikurangkan dalam sepanjang proses pembangunan sistem ini. Bagi mengelakkan pengguna keliru atau menjalankan langkah yang kompleks untuk mengoperasi sistem ini, antara muka sistem yang direka mudah difahami dan senang diguna. Seterusnya, sistem yang dibangunkan diuji berdasarkan keperluan fungsian yang telah dirancang dengan menggunakan kaedah kotak hitam (*Black Box Testing*). Tujuan pengujian dijalankan adalah untuk memastikan segala perancangan yang telah dilakukan di dalam sistem tersebut boleh berjalan dengan lancar. Proses ini dijalankan supaya pepijat atau kesilapan di dalam sistem ini dapat dikesan dan dapat bertambah baik.

5.1 Reka Bentuk Antara Muka



Rajah 2 Reka Bentuk Antara Muka *Web Scraping*



Rajah 3 Reka Bentuk Antara Muka *Merging Data*



Rajah 4 Reka Bentuk Antara Muka *Arima Analysis*

Rajah 2, Rajah 3 dan Rajah 4 menunjukkan reka bentuk antara muka bagi sistem yang akan dibangunkan. Reka bentuk adalah mudah difahami dan senang diguna oleh pengguna. Tiga antara muka yang direka untuk sistem ini adalah “*Web Scraping*”, “*Merging Data*” dan “*Arima Analysis*”. Butang “*Scrape*” di antara muka “*Web Scraping*” boleh ditekan oleh pengguna untuk mengikis data daripada laman web E-dagang dan data yang telah dikikis akan dipaparkan kepada pengguna. Butang “*Download*” di antara muka “*Web Scraping*” berfungsi untuk memuat turunkan data yang telah dikikis. Dua butang “*Browse*” di antara muka “*Merging Data*” berfungsi untuk memilih data yang ingin dimuat naikkan manakala butang “*Upload*” memuat naikkan data yang telah dipilih. Data yang telah berjaya dimuat naikkan akan dipaparkan kepada pengguna. Selepas data telah dimuat naikkan, pengguna boleh menekan butang “*Merge*” di antara muka “*Merging Data*” untuk menggabungkan dua data yang telah dimuat naikkan untuk menghasilkan data siri masa. Data yang berjaya digabungkan juga akan dipaparkan kepada pengguna. Butang “*Download*” di antara muka “*Merging Data*” juga boleh ditekan untuk memuat turunkan data siri masa. Butang “*Browse*” dan “*Upload*” di antara muka “*Arima Analysis*” mempunyai fungsi yang sama dengan butang “*Browse*” dan “*Upload*” di antara muka “*Merging Data*” tetapi data yang akan dimuat naikkan di antara muka “*Arima Analysis*” adalah data siri masa. Butang “*Analyse*” di antara muka “*Arima Analysis*” boleh ditekan oleh pengguna untuk menjalankan pemprosesan dan analisis terhadap

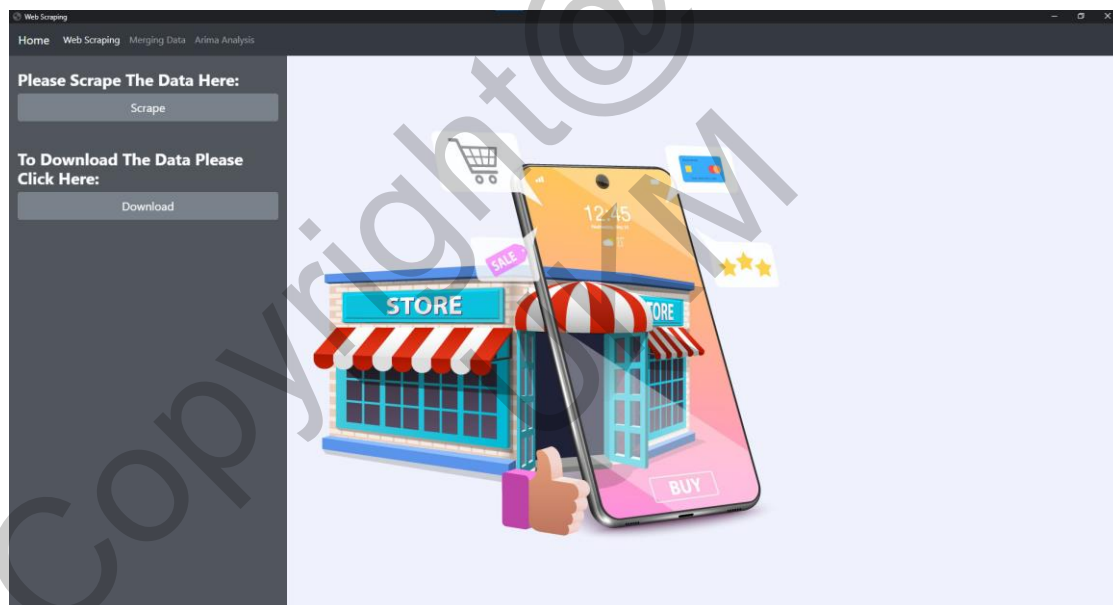
data yang telah dimuat naikan. Akhirnya, data yang telah diproses dan keputusan yang dihasilkan akan dipaparkan kepada pengguna.

5.2 Pengujian Dan Keputusan Pengujian

Pengujian dijalankan untuk memastikan segala perancangan yang telah dilakukan di dalam sistem tersebut boleh berjalan dengan lancar. Keputusan pengujian akan ditunjukkan dalam bentuk rajah.

5.2.1 Pengujian di antara muka “Web Scraping”

Pengujian yang dijalankan di antara muka “*Web Scraping*” adalah mengikis data daripada laman web E-dagang, memuat turunkan data dan memaparkan data atau keputusan.



Rajah 5 Antara muka “*Web Scraping*”

Selepas butang “*Scrape*” ditekan, data yang telah dikikis telah dipaparkan. Ini menunjukkan bahawa data telah berjaya dikikis dan berjaya dipaparkan kepada pengguna seperti yang ditunjukkan dalam Rajah 6.



The screenshot shows a web scraping application window with a dark theme. On the left, there are two sections: "Please Scrape The Data Here:" with a "Scrape" button, and "To Download The Data Please Click Here:" with a "Download" button. The main area displays a table with 17 rows of data. The table has two columns: "Item Name" and "Item sold".

	Item Name	Item sold
0	Signature Market Dried Strawberries (110g)	8900
1	Kleenex Bath Tissue Clean Care Regular 2005 x 20R - 3PLY	3000
2	A1 Crab Flavour Rice Vermicelli White Soup (100g) 螃蟹米粉	23100
3	Dutch Lady Pure Farm Milk - Kurma/Banana (1L)	6800
4	Air Wick Freshmatic Starter Air Refreshener / Air Care	20000
5	[LIMITED EDITION] F&N Sparkling Gift Pack	2400
6	Dutch Lady UHT Milk Pack Of 3 (1L)	1800
7	Ribena Concentrate - Regular (2L)	5300
8	Vinda Deluxe Travel Pack Tissue (50's x 4)	9600
9	NESCAFE Blend & Brew Original (19g x 28g)	3600
10	STARBUCKS Caramel Latte Premium Coffee Mix, 86g Box of 4 x 21.5g Sticks	9700
11	Mister Potato Chips - Assorted Flavor (150g)	18500
12	Vinda Prestige 4D Deco Embossed Toilet Tissue (16 rolls x 200g) (3200 sheets)	1500
13	MAGGI 2-Min Curry 5 Packs, 79g Each	10800
14	Kleenex Bath Tissue Clean Care Regular - 3 Ply (200's x 20R)	27500
15	Maggi 2-Min Chicken (5 Packs/77g Each)	16800
16	Vinda Deluxe Soft Pack Facial Tissue Large (120s x 4)	31400

Rajah 6 Pengikisan data dan paparan data

Untuk memuat turunkan data yang telah dikikis, butang “Download” telah ditekan. Data yang dikikis yang bernama “export” akan dimuat turun dalam format .csv seperti yang ditunjukkan dalam Rajah 7.



The screenshot shows the same web scraping application window as in Rajah 6. The table of items is visible. At the bottom left, a file named "export.csv" is shown, indicating that the data has been successfully downloaded. The "Download" button is highlighted.

	Item Name	Item sold
0	Signature Market Dried Strawberries (110g)	8900
1	Kleenex Bath Tissue Clean Care Regular 2005 x 20R - 3PLY	3000
2	A1 Crab Flavour Rice Vermicelli White Soup (100g) 螃蟹米粉	23100
3	Dutch Lady Pure Farm Milk - Kurma/Banana (1L)	6800
4	Air Wick Freshmatic Starter Air Refreshener / Air Care	20000
5	[LIMITED EDITION] F&N Sparkling Gift Pack	2400
6	Dutch Lady UHT Milk Pack Of 3 (1L)	1800
7	Ribena Concentrate - Regular (2L)	5300
8	Vinda Deluxe Travel Pack Tissue (50's x 4)	9600
9	NESCAFE Blend & Brew Original (19g x 28g)	3600
10	STARBUCKS Caramel Latte Premium Coffee Mix, 86g Box of 4 x 21.5g Sticks	9700
11	Mister Potato Chips - Assorted Flavor (150g)	18500
12	Vinda Prestige 4D Deco Embossed Toilet Tissue (16 rolls x 200g) (3200 sheets)	1500
13	MAGGI 2-Min Curry 5 Packs, 79g Each	10800
14	Kleenex Bath Tissue Clean Care Regular - 3 Ply (200's x 20R)	27500
15	Maggi 2-Min Chicken (5 Packs/77g Each)	16800
16	Vinda Deluxe Soft Pack Facial Tissue Large (120s x 4)	31400

Rajah 7 Data yang dikikis berjaya dimuat turunkan

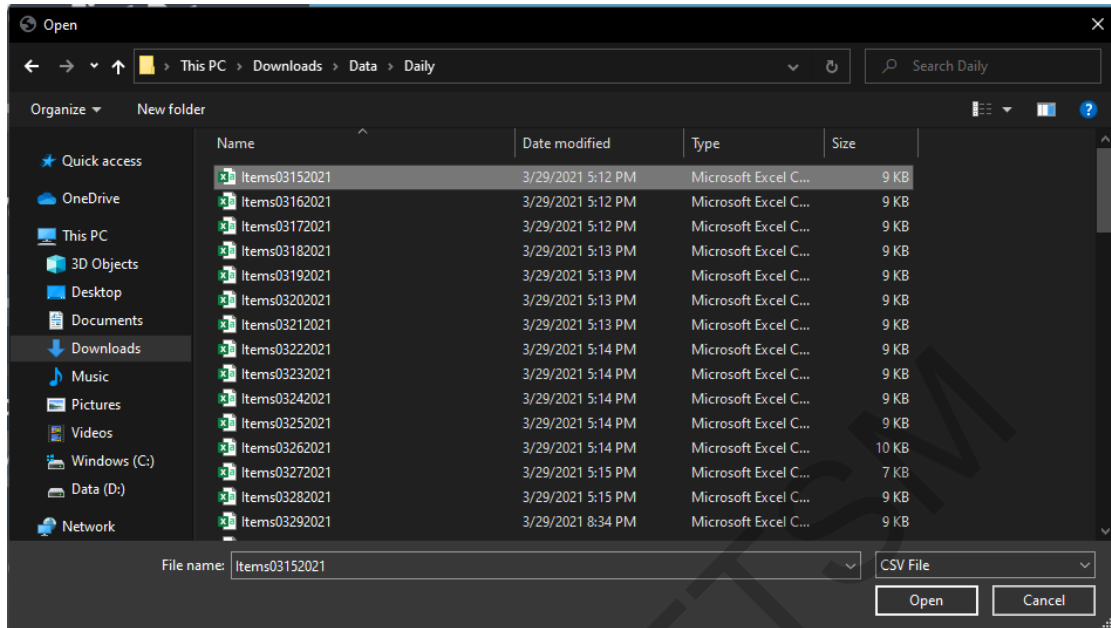
5.2.2 Pengujian di antara muka “*Merging Data*”

Pengujian yang dijalankan di antara muka “*Merging Data*” adalah memuat turunkan data, memuat naikkan data, mengabung data dan memaparkan data atau keputusan.

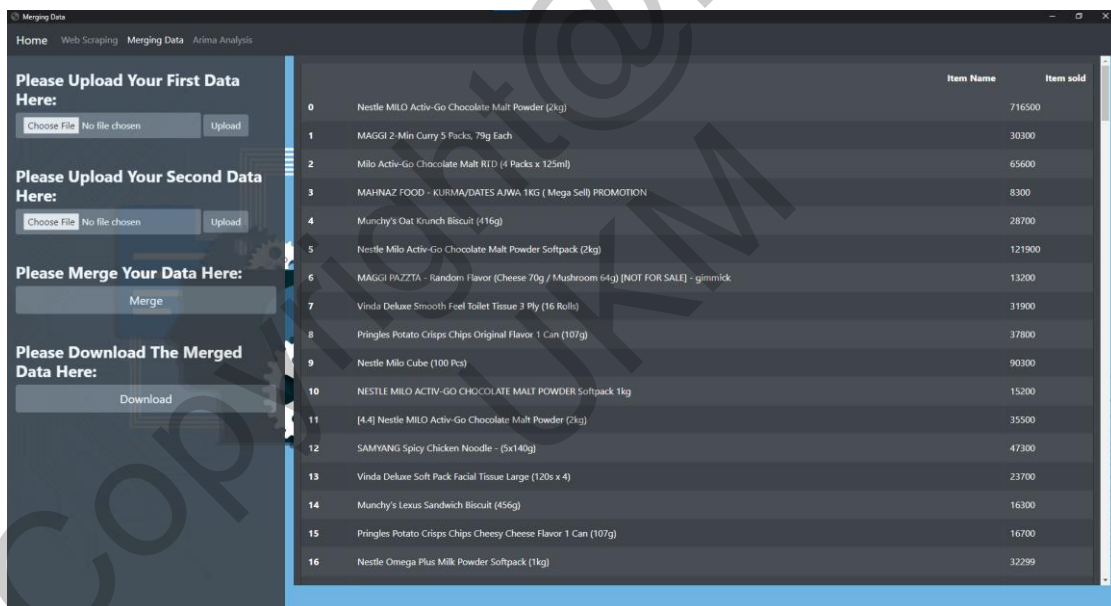


Rajah 8 Antara muka “*Merging Data*”

Untuk memilih data yang ingin dimuat naikkan, butang “*Choose File*” perlu ditekan. Selepas butang “*Choose File*” ditekan, pengguna boleh memilih data yang ingin dimuat naikkan dalam format .csv seperti yang ditunjukkan dalam Rajah 9. Selepas memilih data, butang “*Upload*” perlu ditekan untuk memuat naikkan data dan data yang berjaya dimuat naikkan akan dipaparkan seperti yang ditunjukkan dalam Rajah 10 dan Rajah 11. Langkah bagi memuat naikkan kedua-dua data adalah sama.



Rajah 9 Memilih data



Rajah 10 Paparan data bagi pemilihan data yang pertama

	Item Name	Item sold
0	Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	716800
1	MAGGI 2-Min Curry 5 Packs, 79g Each	30400
2	Milo Activ-Go Chocolate Malt RTD (4 Packs x 125ml)	65800
3	MAHNAZ FOOD - KURMA/DATES AJWA 1KG (Mega Sell) PROMOTION	8500
4	Munchy's Oat Krunch Biscuit (416g)	28800
5	Nestle Milo Activ-Go Chocolate Malt Powder Softpack (2kg)	122000
6	MAGGI PAZZIA - Random Flavor (Cheese 70g / Mushroom 64g) [NOT FOR SALE] - gimmick	13200
7	Vinda Deluxe Smooth Feel Toilet Tissue 3 Ply (16 Rolls)	32000
8	Pringles Potato Crisps Chips Original Flavor 1 Can (107g)	37900
9	Nestle Milo Cube (100 Pcs)	90400
10	NESTLE MILO ACTIV-GO CHOCOLATE MALT POWDER Softpack 1kg	15300
11	Himalaya Salt Sports Candy - Extra Cool Lemon (15g)	77600
12	[4-4] Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	35500
13	SAMYANG Spicy Chicken Noodle - (5x140g)	47300
14	Vinda Deluxe Soft Pack Facial Tissue Large (120s x 4)	23900
15	Munchy's Lexus Sandwich Biscuit (456g)	16400
16	Vinda Deluxe Soft Pack Facial Tissue Large 3ply - 120's x 4	56800

Rajah 11 Paparan data bagi pemilihan data yang kedua

Bagi menggabungkan kedua-dua data yang telah dimuat naikkan, butang “*Merge*” perlu ditekan dan data yang telah digabungkan akan dipaparkan kepada pengguna. Data yang berjaya digabungkan telah dipaparkan kepada pengguna seperti yang ditunjukkan dalam Rajah 12. Untuk memuat turunkan data yang telah digabungkan, butang “*Download*” telah ditekan. Data yang digabungkan yang bernama “*export*” akan dimuat turun dalam format .csv seperti yang ditunjukkan dalam Rajah 13.

	Item Name	1	2
0	Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	716500	716800
1	MAGGI 2-Min Curry 5 Packs, 79g Each	30300	30400
2	Milo Activ-Go Chocolate Malt RTD (4 Packs x 125ml)	65600	65800
3	MAHNAZ FOOD - KURMA/DATES AJWA 1KG (Mega Sell) PROMOTION	8300	8500
4	Munchy's Oat Krunch Biscuit (416g)	28700	28800
5	Nestle Milo Activ-Go Chocolate Malt Powder Softpack (2kg)	121900	122000
6	MAGGI PAZZIA - Random Flavor (Cheese 70g / Mushroom 64g) [NOT FOR SALE] - gimmick	13200	13200
7	Vinda Deluxe Smooth Feel Toilet Tissue 3 Ply (16 Rolls)	31900	32000
8	Pringles Potato Crisps Chips Original Flavor 1 Can (107g)	37800	37900
9	Nestle Milo Cube (100 Pcs)	90300	90400
10	NESTLE MILO ACTIV-GO CHOCOLATE MALT POWDER softpack 1kg	15200	15300
11	[4-4] Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	35500	35500
12	SAMYANG Spicy Chicken Noodle - (5x140g)	47300	47300
13	Vinda Deluxe Soft Pack Facial Tissue Large (120s x 4)	23700	23900
14	Munchy's Lexus Sandwich Biscuit (456g)	16300	16400
15	Pringles Potato Crisps Chips Cheesy Cheese Flavor 1 Can (107g)	16700	16700
16	Nestle Omega Plus Milk Powder Softpack (1kg)	32299	32400

Rajah 12 Penggabungan data dan paparan data

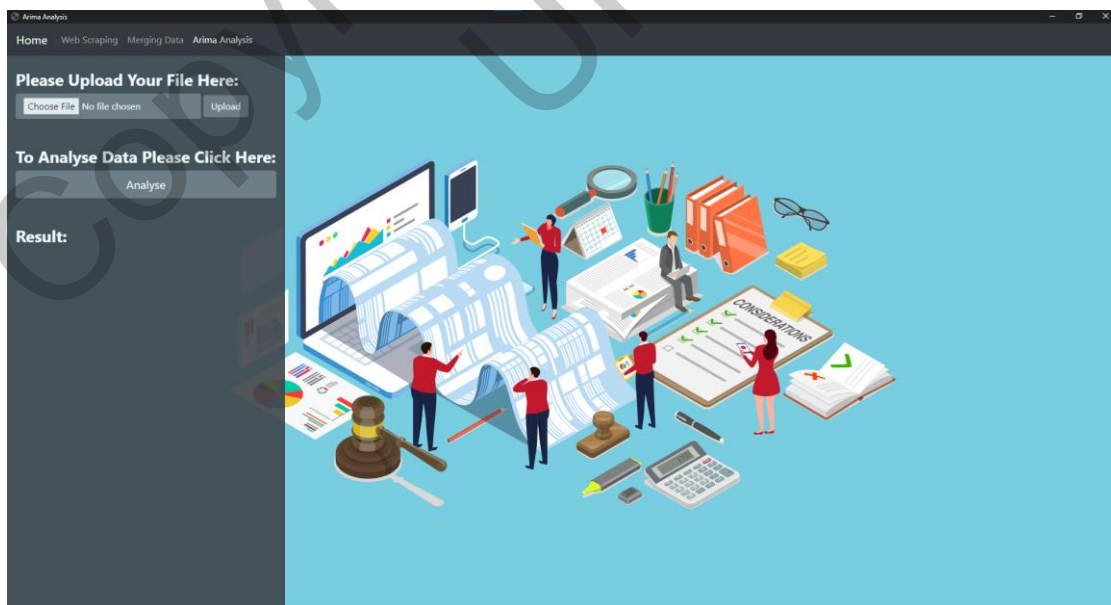


	Item Name	1	2
0	Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	716500	716800
1	MAGGI 2-Min Curry 5 Packs, 75g Each	30300	30400
2	Milo Activ-Go Chocolate Malt RTD (4 Packs x 125ml)	65600	65800
3	MAHNAZ FOOD - KURMA/DATES A/WA 1KG (Mega Sell) PROMOTION	8300	8500
4	Munchy's Oat Krunch Biscuit (416g)	28700	28800
5	Nestle Milo Activ-Go Chocolate Malt Powder Softpack (2kg)	121900	122000
6	MAGGI PAZZTA - Random Flavor (Cheese 70g / Mushroom 64g) [NOT FOR SALE] - gimmick	13200	13200
7	Vinda Deluxe Smooth Feel Toilet Tissue 3 Ply (16 Rolls)	31900	32000
8	Pringles Potato Crisps Chips Original Flavor 1 Can (107g)	37800	37900
9	Nestle Milo Cube (100 Pcs)	90300	90400
10	NESTLE MILO ACTIV-GO CHOCOLATE MALT POWDER Softpack 1kg	15200	15300
11	[4-4] Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	35500	35500
12	SAMYANG Spicy Chicken Noodle - (5x140g)	47300	47300
13	Vinda Deluxe Soft Pack Facial Tissue Large (120s x 4)	23700	23900
14	Munchy's Lexus Sandwich Biscuit (456g)	16300	16400
15	Pringles Potato Crisps Chips Cheesy Cheese Flavor 1 Can (107g)	16700	16700
16	Nestle Omega Plus Milk Powder Softpack (1kg)	32299	32400

Rajah 13 Data yang digabungkan berjaya dimuat turunkan

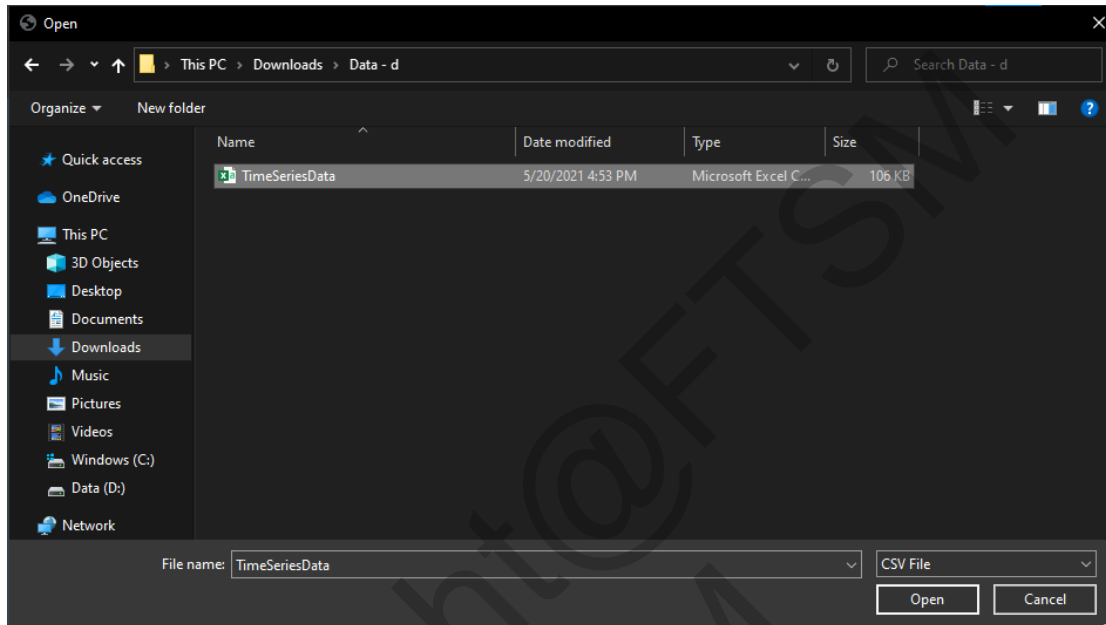
5.2.3 Pengujian di antara muka “Arima Analysis”

Pengujian yang dijalankan di antara muka “Arima Analysis” adalah memuat naikkan data, memaparkan data atau keputusan, pemprosesan data dan analisis data.



Rajah 14 Antara muka “Arima Analysis”

Untuk memilih data siri masa yang ingin dimuat naikkan, butang “Choose File” perlu ditekan. Selepas butang “Choose File” ditekan, pengguna boleh memilih data yang ingin dimuat naikkan dalam format .csv seperti yang ditunjukkan dalam Rajah 15. Selepas memilih data, butang “Upload” perlu ditekan untuk memuat naikkan data siri masa dan data risi masa yang berjaya dimuat naikkan akan dipaparkan seperti yang ditunjukkan dalam Rajah 16.



Rajah 15 Memilih data siri masa

The screenshot shows the Arima Analysis software interface. On the left, there are sections for 'Please Upload Your File Here:' (with a 'Choose File' button), 'To Analyse Data Please Click Here:' (with an 'Analyse' button), and 'Result:'. The main area displays a data table with the following columns: Item Name, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13. The table contains 12 rows of data for various products, including Nestle MILO, MAGGI, Munchy's, and Pringles.

Item Name	1	2	3	4	5	6	7	8	9	10	11	12	13
0 Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	716500	716800	717200	717600	717900	718400	718800	719300	719800	720500	722400	724800	725600
1 MAGGI 2-Min Curry 5 Packs, 79g Each	36300	30400	30600	30700	30900	30900	31200	31300	31800	32299	32800	33200	33300
2 Milo Activ-Go Chocolate Malt RTD (4 Packs x 125ml)	65600	65800	65900	66100	66300	66400	66500	66700	67000	67500	68100	68400	68500
3 MAHNAZ FOOD - KURMA/DATES AJWA 1KG (Mega Sell) PROMOTION	8300	8500	8800	9000	9100	9300	9500	9700	9800	10000	10300	10700	11000
4 Munchy's Oat Krunch Biscuit (416g)	28700	28800	29000	29100	29200	29300	29400	29500	29700	29900	30300	30400	30500
5 Nestle Milo Activ-Go Chocolate Malt Powder Softpack (2kg)	121900	122000	122200	122400	122500	122600	0	122900	123200	123600	124100	124400	124600
6 MAGGI PAZZIA - Random Flavor (Cheese 70g / Mushroom 64g) [NOT FOR SALE] - gimmick	13200	13200	13300	13300	13300	13400	13500	13600	13700	13800	13800	13800	0
7 Vinda Deluxe Smooth Feel Toilet Tissue 3 Ply (16 Rolls)	31900	32000	32100	32200	0	32600	32800	32900	33100	33400	33900	34000	34100
8 Pringles Potato Crisps Chips Original Flavor 1 Can (107g)	37800	37900	37900	38000	38100	38200	38300	38300	38500	38900	39300	39500	39600
9 Nestle Milo Cube (100 Pcs)	90300	90400	90500	90500	90600	90700	90700	90800	91100	91300	91500	91600	91600
10 NESTLE MILO ACTIV-GO CHOCOLATE MALT POWDER Softpack 1kg	15200	15300	15400	15500	15500	15600	15600	15700	15800	16000	16200	16300	16400
11 [4.4] Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	35500	35500	35600	35600	35700	35900	36000	36000	36300	36600	36900	37200	37300

Rajah 16 Paparan data siri masa

Bagi memulakan analisis data siri masa, butang “*Analyse*” telah ditekan. Data siri masa akan diproses dan seterusnya dianalisis. Data siri masa yang telah diproses akan dipaparkan dan hasil analisis data siri masa iaitu nama produk yang dicadangkan juga akan dipaparkan seperti yang ditunjukkan dalam Rajah 17.

The screenshot shows the Arima Analysis software interface. On the left, there is a sidebar with the following sections:

- Please Upload Your File Here:** Includes a 'Choose File' button (showing 'No file chosen') and an 'Upload' button.
- To Analyse Data Please Click Here:** Includes an 'Analyse' button.
- Result:** Displays the text: "The recommended food product is Himalaya Salt Sports Candy - Extra Cool Lemon (15g) with the predicted sales of 6425."

The main area displays a table with the following data:

Item Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0 Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	300	400	400	300	500	400	500	500	700	1900	2400	800	700	1400	1600	700
1 MAGGI 2-Min Curry 5 Packs, 79g Each	100	200	100	200	0	300	100	500	499	501	400	100	300	300	400	200
2 Milo Activ-Go Chocolate Malt RTD (4 Packs x 125ml)	200	100	200	200	100	100	200	300	500	600	300	100	200	400	500	200
3 MAHINAZ FOOD - KURMA/DATES AJWA 1KG (Mega Sell) PROMOTION	200	300	200	100	200	200	200	100	200	300	400	300	300	300	400	300
4 Munchy's Oat Krunch Biscuit (416g)	100	200	100	100	100	100	100	200	200	400	100	100	200	300	300	100
5 Nestle Milo Activ-Go Chocolate Malt Powder Softpack (2kg)	100	200	200	100	100	0	300	300	400	500	300	200	200	500	500	200
6 MAGGI PAZZIA - Random Flavor (Cheese 70g / Mushroom 64g) [NOT FOR SALE] - gimmick	0	100	0	0	100	100	100	100	100	0	0	0	0	0	100	0
7 Vinda Deluxe Smooth Feel Toilet Tissue 3 Ply (16 Rolls)	100	100	100	0	400	200	100	200	300	500	100	100	200	400	400	200
8 Pringles Potato Crisps Chips Original Flavor 1 Can (107g)	100	0	100	100	100	100	0	200	400	400	200	100	200	300	300	100
9 Nestle Milo Cube (100 Pcs)	100	100	0	100	100	0	100	300	200	200	100	0	100	100	100	0
10 NESTLE MILO ACTIV-GO CHOCOLATE MALT POWDER Softpack 1kg	100	100	100	0	100	0	100	100	200	200	100	100	100	100	200	100
11 [4.4] Nestle MILO Activ-Go Chocolate Malt Powder (2kg)	0	100	0	100	200	100	0	300	300	300	300	100	100	300	300	200

Rajah 17 Paparan data siri masa yang diproses dan nama produk yang dicadangkan

5.2.4 Pengujian model ARIMA

Untuk mengenal pasti parameter yang paling optimum untuk membina model ARIMA yang sesuai, fungsi `auto_arima()` telah digunakan. Untuk menguji prestasi fungsi `auto_arima()`, nilai AIC telah diambil kira. Rajah 18 telah menunjukkan parameter ARIMA yang dikembalikan, nilai AIC dan parameter yang paling optimum ditunjukkan di “Best model:”


```
Performing stepwise search to minimize aic
ARIMA(0,1,0)(0,1,0)[7] : AIC=517.703, Time=0.01 sec
ARIMA(1,1,0)(1,1,0)[7] : AIC=500.727, Time=0.11 sec
ARIMA(0,1,1)(0,1,1)[7] : AIC=inf, Time=0.20 sec
ARIMA(1,1,0)(0,1,0)[7] : AIC=502.301, Time=0.04 sec
ARIMA(1,1,0)(2,1,0)[7] : AIC=496.355, Time=0.23 sec
ARIMA(1,1,0)(3,1,0)[7] : AIC=489.479, Time=0.37 sec
ARIMA(1,1,0)(4,1,0)[7] : AIC=inf, Time=0.92 sec
ARIMA(1,1,0)(3,1,1)[7] : AIC=490.330, Time=0.73 sec
ARIMA(1,1,0)(2,1,1)[7] : AIC=inf, Time=0.66 sec
ARIMA(1,1,0)(4,1,1)[7] : AIC=inf, Time=1.25 sec
ARIMA(0,1,0)(3,1,0)[7] : AIC=503.154, Time=0.14 sec
ARIMA(2,1,0)(3,1,0)[7] : AIC=489.856, Time=0.50 sec
ARIMA(1,1,1)(3,1,0)[7] : AIC=484.763, Time=1.12 sec
ARIMA(1,1,1)(2,1,0)[7] : AIC=491.402, Time=0.60 sec
ARIMA(1,1,1)(4,1,0)[7] : AIC=486.704, Time=1.89 sec
ARIMA(1,1,1)(3,1,1)[7] : AIC=486.693, Time=0.99 sec
ARIMA(1,1,1)(2,1,1)[7] : AIC=inf, Time=0.73 sec
ARIMA(1,1,1)(4,1,1)[7] : AIC=inf, Time=1.47 sec
ARIMA(0,1,1)(3,1,0)[7] : AIC=inf, Time=0.45 sec
ARIMA(2,1,1)(3,1,0)[7] : AIC=inf, Time=0.78 sec
ARIMA(1,1,2)(3,1,0)[7] : AIC=inf, Time=0.90 sec
ARIMA(0,1,2)(3,1,0)[7] : AIC=484.815, Time=0.65 sec
ARIMA(2,1,2)(3,1,0)[7] : AIC=inf, Time=1.61 sec
ARIMA(1,1,1)(3,1,0)[7] intercept : AIC=inf, Time=0.94 sec

Best model: ARIMA(1,1,1)(3,1,0)[7]
Total fit time: 17.312 seconds
```

Rajah 18 Parameter Model ARIMA

	aic
ARIMA(1,1,1)(3,1,0)[7]	484.763
ARIMA(0,1,2)(3,1,0)[7]	484.815
ARIMA(1,1,1)(3,1,1)[7]	486.693
ARIMA(1,1,1)(4,1,0)[7]	486.704
ARIMA(1,1,0)(3,1,0)[7]	489.479
ARIMA(2,1,0)(3,1,0)[7]	489.856
ARIMA(1,1,0)(3,1,1)[7]	490.330
ARIMA(1,1,1)(2,1,0)[7]	491.402
ARIMA(1,1,0)(2,1,0)[7]	496.355
ARIMA(1,1,0)(1,1,0)[7]	500.727
ARIMA(1,1,0)(0,1,0)[7]	502.301
ARIMA(0,1,0)(3,1,0)[7]	503.154
ARIMA(0,1,0)(0,1,0)[7]	517.703
ARIMA(1,1,0)(2,1,1)[7]	inf
ARIMA(1,1,0)(4,1,0)[7]	inf
ARIMA(0,1,1)(0,1,1)[7]	inf
ARIMA(1,1,1)(2,1,1)[7]	inf
ARIMA(1,1,1)(4,1,1)[7]	inf
ARIMA(0,1,1)(3,1,0)[7]	inf
ARIMA(2,1,1)(3,1,0)[7]	inf
ARIMA(1,1,2)(3,1,0)[7]	inf
ARIMA(1,1,0)(4,1,1)[7]	inf
ARIMA(2,1,2)(3,1,0)[7]	inf

Rajah 19 Parameter ARIMA and nilai AIC

Parameter ARIMA yang dikembalikan dan nilai AIC telah direkodkan dalam satu jadual mengikut nilai AIC dari rendah hingga tinggi seperti yang ditunjukkan dalam Rajah 18. Nilai “inf” di kolum AIC adalah bermaksud infiniti iaitu nilai yang sangat besar. Nilai AIC

yang lebih rendah menunjukkan model yang lebih sesuai. Selepas menyusun nilai AIC dari rendah hingga tinggi, parameter ARIMA yang dikembalikan yang mempunyai nilai AIC yang paling rendah dalam kes ini adalah sama dengan “Best model:” yang dipilih melalui fungsi `auto_arma()`. Ini telah menunjukkan bahawa fungsi `auto_arma()` telah mengembalikan parameter yang optimum untuk membina model ARIMA yang sesuai untuk melakukan ramalan.

6 KESIMPULAN

Kesimpulannya, sistem yang dibangunkan dalam projek ini dapat meramalkan produk makanan yang berpotensi sebagai laris jualan untuk dijual melalui analisis data siri masa jualan produk makanan. Produk makanan yang diramal akan bertindak sebagai cadangan produk makanan kepada pengusaha B40 untuk memulakan perniagaannya di laman web E-dagang.

Kelebihan sistem adalah sistem yang dibangunkan mempunyai tiga fungsi utama iaitu pengikisan web, penggabungan data dan analisis data. Pengguna boleh menggunakan pengikisan web untuk mengumpulkan data jualan produk makanan dari semasa ke semasa dan menggabungkan data yang dikumpul untuk menghasilkan data siri masa jualan produk makanan. Data siri masa jualan produk makanan boleh dijalankan analisis dengan menggunakan fungsi analisis data sistem ini. Selain daripada data siri masa jualan produk makanan yang dikumpul melalui sistem ini, pengguna juga boleh menggunakan data siri masa sendiri untuk mengadakan analisis data dengan mengikut format yang ditetapkan.

Manakala, had sistem adalah data mesti dalam bentuk format yang ditentukan seperti mempunyai atribut nama produk dan jualan produk supaya dapat diproses oleh algoritma ARIMA. Selain itu, had untuk sistem ini adalah data set yang dikikis dengan menggunakan teknik pengikisan web terhadap pada jenis platform tertentu sahaja.

Peningkatan masa depan sistem ini adalah menambah skop data. Data set yang dikikis dalam sistem ini adalah produk makanan. Pada masa hadapan, data seperti produk elektronik dan lain-lain boleh ditambah ke dalam sistem ini. Selain itu, produk makanan yang diramal boleh dipaparkan dalam bentuk peringkat. Sebagai contoh, memaparkan sepuluh makanan yang paling laris dalam peringkat.

7 RUJUKAN

- ALEXANDRA TWIN. 2020. Data Mining.
<https://www.investopedia.com/terms/d/datamining.asp#:~:text=Data%20mining%20is%20a%20process,increase%20sales%20and%20decrease%20costs>
- CompareHero.my. 2020. T20, M40 And B40 Income Classifications in Malaysia.
 Information Group. 1998. <https://www.comparehero.my/budgets-tax/articles/t20-m40-b40-malaysia>
- Dimira Teneva. 2016. Data-driven marketing for ecommerce.
<https://www.metrilo.com/blog/data-driven-marketing-for-ecommerce>
- imperva. 2021. Black Box Testing. <https://www.imperva.com/learn/application-security/black-box-testing/>
- Jiawei Han, Micheline Kamber, Jian Pei. 2012. DATA MINING Concepts and Techniques Third Edition.
- Muhammad Hafiz Bin Mustafar. 2016. KONSEP E- DAGANG MENYELURUH.
<http://hafizmustafar123.blogspot.com/2016/05/konsep-e-dagang-menyeluruh.html>
- Mustapha Ismail, Mohammed Mansur Ibrahim, Muesser Cemal Nat. 2015. Data Mining in Electronic Commerce: Benefits and Challenges.
- Ofni Systems. Design Specification. <http://www.ofnisystems.com/services/validation/design-specification/>
- Rob J Hyndman and George Athanasopoulos. 2018. 8.6 Estimation and order selection.
<https://otexts.com/fpp2/arma-estimation.html>
- Satish Chandra Gupta. 2020. Architecture for High-Throughput Low-Latency Big Data Pipeline on Cloud. <https://towardsdatascience.com/scalable-efficient-big-data-analytics-machine-learning-pipeline-architecture-on-cloud-4d59efc092b5>
- snapLogic. 2020. Data Pipeline Architecture. <https://www.snaplogic.com/glossary/data-pipeline-architecture#:~:text=A%20data%20pipeline%20architecture%20is,analysis%2C%20and%20using%20data%20easier>
- Staff Writer. 2019. Black Box Testing – Understanding the Basics.
<https://reqtest.com/testing-blog/black-box-testing/#:~:text=The%20black%20box%20is%20a,the%20output%20that%20is%20produced.>