



**JOURNAL OF INFORMATION
SYSTEM AND TECHNOLOGY
MANAGEMENT
(JISTM)**
www.jistm.com



VISUALISASI DATA E-DAGANG MENGGUNAKAN AUGMENTASI REALITI

VISUALIZATION OF E-COMMERCE DATA USING AUGMENTED REALITY

Azzan Amin¹, Haslina Arshad^{2*}, Ummul Hanan Mohamad³

¹ Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia, Malaysia
Email: a172209@siswa.ukm.edu.my

² Institut IR4.0, Universiti Kebangsaan Malaysia, Malaysia
Email: haslinarshad@ukm.edu.my

³ Institut IR4.0, Universiti Kebangsaan Malaysia, Malaysia
Email: ummulhanan@ukm.edu.my

* Corresponding Author

Article Info:

Article history:

Received date: 15.10.2020

Revised date: 05.11.2020

Accepted date: 17.11.2020

Published date: 07.12.2020

To cite this document:

Amin, A., Arshad, H., & Mohamad, U. H. (2020). Visualisasi Data E-Dagang Menggunakan Augmentasi Realiti. Journal of Information System and Technology Management, 5 (19), 104-122.

DOI: 10.35631/JISTM.519009.

This work is licensed under [CC BY 4.0](#)



Abstrak:

Visualisasi data dapat dilihat sebagai elemen yang signifikan dalam analisis data dan komunikasi. Apabila penglibatan data menjadi semakin banyak dan kompleks, persembahan data secara visual sememangnya membantu pengguna bagi memudahkan pemahaman terhadap penyampaian data tersebut. Pada masa kini, paparan data secara dua dimensi (2D) kerap digunakan untuk proses visualisasi data, namun ketiadaan dimensi kedalaman menyebabkan pemahaman terhadap sesuatu data itu menjadi terhad dan tidak efisien. Oleh itu, keberkesanan AR dalam visualisasi data dikaji melalui pembangunan sebuah aplikasi AR Visualisasi Data menggunakan Data E-Dagang. Model pembelajaran mesin juga terlibat dalam pembangunan aplikasi AR ini bagi penyediaan data untuk fungsi analisis ramalan. Untuk menyediakan Data E-Dagang yang berkualiti dan model pembelajaran mesin yang optimum, proses sains data dijalankan dengan menggunakan bahasa pengaturcaraan Python. Data E-Dagang yang dipilih bagi kajian ini merupakan data terbuka yang diambil melalui laman sesawang Kaggle. Set data ini mempunyai 9994 bilangan data dan 21 attribut. Aplikasi AR Visualisasi Data ini akan memudahkan pengguna untuk memahami Data E-Dagang dengan mendalam melalui penggunaan teknologi AR dan berkemampuan untuk membuat visualisasi ramalan bagi keuntungan jualan berdasarkan algoritma model "Auto Regressive Integrated Moving Average" (ARIMA).

Kata Kunci:

E-Dagang, Ramalan Keuntungan, Visualisasi Data

Abstract:

Data visualization is viewed as a significant element in data analysis and communication. As the data engagement becomes more and more complex, visual presentation of data does help users understand the data. So far, two-dimensional (2D) data visuals are often used for the data visualization process, but the lack of depth dimension leads to inefficient and limited understanding of the data. Therefore, the effectiveness of augmented reality (AR) in data visualization was studied through the development of an AR Data Visualization application using E-commerce data. Machine learning models are also involved in the development of this AR application for the provision of data using predictive analysis functions. To provide quality E-commerce data and an optimal machine learning model, the data science process is carried out using the python programming language. The E-commerce data selected for this study is open data taken through the Kaggle Website. This database has 9994 data numbers and 21 attributes. This AR data visualization application will make it easier for users to understand the E-commerce data in-depth through the use of AR technology and be able to visualize the forecasts for sales profit based on the algorithm model "Auto-Regressive Integrated Moving Average" (ARIMA).

Keywords:

E-Commerce, Profit Forecasting, Data Visualization

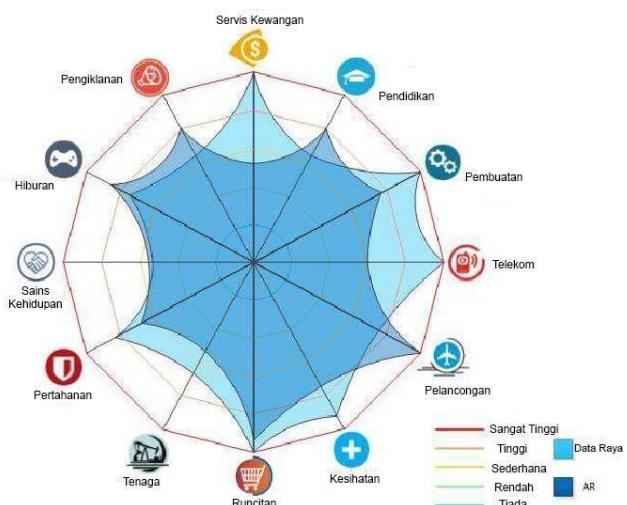
Pengenalan

Visualisasi data adalah satu teknik untuk mempersempahkan data dan informasi dalam bentuk grafik. Grafik yang digunakan dalam data visualisasi lebih dikenali sebagai statistik deskriptif seperti graf, rajah, carta, plot dan sebagainya. Melalui visualisasi data, data yang kompleks ditambah dengan lebih banyak dimensi bagi memudahkan pengguna untuk memahami data tersebut (Huang, Hui, & Peylo, 2014). Sehingga kini, terlalu banyak data yang terkumpul sejajar dengan pertambahan bilangan pengguna internet dari hari ke hari. Data yang terkumpul ini adalah sangat berharga dan boleh digunakan untuk menghasilkan suatu pengetahuan mahupun membantu dalam membuat keputusan. Oleh itu, teknik visualisasi data bertujuan untuk memudahkan pengguna untuk memahami data dengan lebih jelas dan efisyen.

Data Raya atau lebih dikenali sebagai "*Big Data*" adalah satu istilah yang menggambarkan sebuah data dalam skala atau jumlah yang sangat besar dan kompleks (Cielen, Meysman, & Ali, 2016). Data Raya ini boleh diklasifikasikan sebagai berstruktur dan tidak berstruktur. Data berstruktur adalah data yang tersimpan di dalam pangkalan data atau jadual manakala data yang tidak berstruktur adalah data yang terdapat dalam artikel, jurnal, gambar, video dan sebagainya. Data Raya boleh dikhususkan kepada beberapa domain yang tertentu, mengikut spesifikasi pengetahuan yang digunakan, bagi tujuan analisis dan kajian. Data yang lengkap dan huraian kefahaman terhadap data dapat dijustifikasi dengan lebih baik sekiranya teknologi Augmentasi Realiti (AR) digunakan (Yim, Chu, & Sauer, 2017).

Menurut Furht (2011), Augmentasi Realiti (AR) adalah teknologi yang boleh menggabungkan kandungan informasi digital dengan dunia nyata dalam masa sebenar pengguna. AR membolehkan pengguna melihat objek maya 2D atau 3D yang dipaparkan di dunia nyata secara interaktif, menarik dan efektif (Haller, 2006). Antara bidang yang mengaplikasikan teknologi AR pada masa kini adalah bidang Pendidikan (Mahayuddin & Saif, 2020), khususnya untuk melaksanakan proses pengajaran dan pembelajaran (Akçayır & Akçayır, 2017). Meng et al. (2019) berpendapat dengan menggunakan AR, pembelajaran di sekolah menjadi lebih menarik sekaligus memudahkan pelajar untuk memahami sesuatu perkara yang diajar oleh guru. Selain itu, bidang lain yang turut mengaplikasikan teknologi AR adalah pelancongan (Arshad, Ahmad Shukri, Zainal, & Tan, 2019), peruncitan (Tredinnick, 2018), kesihatan (Herron, 2016), perubatan (Rahman, Arshad, Mahmud, Mahayuddin, & Obeidy, 2017), hiburan (García-Crespo, González-Carrasco, López-Cuadrado, Villanueva, & González, 2016), pengiklanan dan e-dagang (Kumar, Chandra, Bharati, & Manava, 2016).

Dalam era Revolusi Industri ke-4 ini, visualisasi data e-dagang menggunakan AR dilihat sangat penting. Secara definisinya, E-dagang adalah membeli dan menjual sesuatu produk atau servis melalui media elektronik seperti radio, televisyen, komputer dan telefon pintar (Barenji, Wang, Li, & Guerra-Zubiaga, 2019). Kebanyakan E-dagang ini menggunakan medium rangkaian internet bagi menghubungkan pembeli dan penjual. Seperti yang dilaporkan oleh Huang et al. (2014), Rajah 1 menunjukkan bahawa penggunaan AR terhadap domain Runcitan adalah antara yang tertinggi. Ini disokong oleh penggunaan sumber Data Raya terbanyak dari domain tersebut berbanding domain yang lain untuk tujuan kajian. Secara tidak langsung, kajian ini yang melibatkan penggabungan AR dan Data Raya bagi tujuan visualisasi terhadap Data Runcitan atau secara khususnya, E-dagang, adalah signifikan kerana ia dapat mengatasi limitasi kognitif pengguna terhadap representasi data tersebut. Seterusnya, kajian ini dapat membantu sesebuah syarikat perniagaan dalam membuat keputusan yang efektif untuk mengoptimumkan sumber dan meningkatkan keuntungan.



Rajah 1: Pengaruh Data Raya dan Augmentasi Realiti Terhadap Sesuatu Domain
Modifikasi daripada Sumber: Huang et al. (2014)

Kajian Literasi

Kajian literasi ini akan membincangkan tentang visualisasi data dan sejauh mana Augmentasi Realiti serta penggunaan algoritma pembelajaran mesin yang optimum berpotensi untuk memvisualisasikan Data E-Dagang yang berkualiti.

Pada masa kini, beberapa alat visualisasi yang digunakan oleh organisasi atau syarikat adalah bertujuan untuk menganalisis data yang terkumpul untuk mendapatkan keputusan yang cepat dan efektif. Contoh alat visualisasi adalah seperti Tableau, ReportPlus, Charted dan Highcharts. Namun, kebanyakan alat visualisasi yang sedia ada ini hanya berfokus untuk menghasilkan visualisasi gambar abstrak dan grafik yang ringkas bagi sesuatu kumpulan data. Ini menyebabkan beberapa data yang signifikan diketepikan dan hubungan yang penting antara data itu disembunyikan daripada penglihatan pengguna. Keadaan ini dikatakan sebagai kehilangan darjah penglihatan. Perkara ini juga boleh terjadi disebabkan masalah paparan resolusi yang melibatkan kualiti representasi data kerana ia bergantung kepada jumlah ketumpatan piksel skrin (Olshannikova, Ometov, Koucheryavy, & Olsson, 2015).

Selain itu, manusia sememangnya mempunyai limitasi kemampuan untuk membuat persepsi dan kognitif yang tepat terhadap sesuatu perkara. Justeru, visualisasi data memainkan peranan yang amat penting dalam menghasilkan suatu maklumat yang bermakna buat pengguna. Terdapat beberapa cara untuk memvisualisasikan data. Antaranya adalah visualisasi dalam bentuk tiga-dimensi (3D) atau dua-dimensi (2D). Namun begitu, Gustafsson & Odd (2018) mengetengahkan bahawa terdapatnya limitasi untuk memaparkan data dengan cara begini kerana kekurangan kedalaman dalam skrin 2D. Hal ini juga menyebabkan timbulnya isu dalam visualisasi data seperti kesempitan sudut visual, navigasi, perskalaan dan sebagainya.

Visualisasi Data

Visualisasi data adalah teknik yang digunakan untuk membantu pemahaman sesebuah data dengan menggunakan kaedah konteks visual (Qin, Luo, Tang, & Li, 2020). Penggunaan visualisasi data akan memudahkan lagi proses menganalisis corak data sekaligus menambahbaik suatu keputusan untuk menjadi lebih efektif. Data visualisasi berkeupayaan dalam menyediakan fungsi untuk memanipulasikan data mengikut kehendak dan mempersempahkan data dalam medium grafik yang mudah difahami. Oleh itu, visualiasi data bagi Data Raya amat signifikan dalam usaha mengeksplorasi dan memahami akan data yang dikaji kerana ia membantu mengatasi limitasi kognitif manusia dalam mengolah sesuatu perkara. Antara tujuan utama visualisasi data diperkenal adalah untuk menambahbaik representasi data dalam pelbagai bentuk imej, rajah dan animasi. Terdapat beberapa teknik visualisasi data yang kerap digunakan seperti tag awam, klustergram, carta gerakan dan papan pemuka (PS & Chintala, 2020).

Kebanyakan syarikat pada masa kini menggunakan alat visualisasi data yang sedia ada dalam membantu mereka membuat keputusan untuk meningkatkan keuntungan serta mengurangkan kos dengan hanya menggunakan sumber yang minimum. Terdapat banyak alat visualisasi yang tersedia dalam menghasilkan visualisasi yang bermakna dan menarik. Sebagai contoh, alat visualisasi yang terbuka untuk analisis dan visualisasi data adalah seperti ‘Highcharts’, ‘Weave’, ‘Infogram’, ‘DataWrapper’ dan sebagainya. Visualisasi data memainkan peranan yang sangat penting dalam menterjemahkan sesebuah data dengan menggunakan teknik

visualisasi yang terbaik. Keberkesanannya suatu visualisasi data bergantung kepada cara pengimplementasian untuk mengimbangi persepsi (pandangan) dan kognisi (fikiran) manusia bagi memudahkan untuk memahaminya (Gustafsson & Odd, 2018). Oleh itu, pengetahuan yang jelas berkenaan teknik visualisasi amatlah penting dalam menghasilkan satu aplikasi visualisasi data yang berguna dan berkesan buat pengguna.

Analisis Ramalan

Untuk membuat ramalan keuntungan jualan melalui Data E-Dagang, permodelan data perlu dibuat dengan menggunakan teknik analisis ramalan seperti teknik analisis regresi atau '*Moving Average*' (Purata Bergerak). Bagi Purata Bergerak, ia adalah lebih efektif bersama dengan data yang pegun yang mana ia berterusan dari masa ke masa. Untuk data siri masa yang mempunyai trend atau corak bermusim, teknik ramalan yang boleh dipertimbang untuk digunakan adalah '*Auto Regressive Integrated Moving Average*' (ARIMA) (Nochai & Nochai, 2006).

Menurut Ho & Xie (1998), ARIMA merupakan salah satu algoritma pembelajaran mesin yang paling mudah dan efektif dalam melaksanakan ramalan siri masa. Penggunaan model ARIMA telah pun diaplikasikan untuk membuat ramalan harga komoditi seperti minyak (Xiang & Zhuang, 2013) dan ramalan musim kemarau (Khan, Muhammad, & El-Shafie, 2020). Perbezaan siri yang muncul di dalam persamaan ramalan adalah dipanggil sebagai process '*Auto Regressive*'. Oleh kerana data E-Dagang memiliki tren corak pembelian pengguna berdasarkan siri masa, model ARIMA dilihat sesuai digunakan untuk tujuan ramalan titik data berdasarkan data lampau.

Visualisasi data pada era ini perlu mengimplementasikan konsep dan cara yang baru iaitu dengan menggunakan teknologi AR di dalam visualisasi data. Teknologi AR mampu menyediakan visualisasi yang lebih mendalam berbanding dengan teknik visualisasi yang biasa digunakan. Penggunaan AR mampu menambahkan lagi dimensi terhadap data yang dipaparkan dengan melihat data melalui pelbagai sudut pandangan (Hune & Müller, 2018). Menurut Donalek et al. (2014), semakin banyak dimensi yang dapat kita bayangkan, semakin tinggi kemungkinan kita mengenali corak dan hubungan sesebuah data.

Metodologi

Kajian ini bertujuan untuk mengimplementasikan Augmentasi Realiti dalam visualisasi Data Raya bagi membantu menginterpretasi data dengan lebih bermakna dan efektif. Dua jenis metodologi yang digunakan di dalam kajian ini ialah pembangunan aplikasi AR Visualisasi Data E-Dagang dan metodologi O.S.E.M.N bagi proses sains data untuk penyediaan data dan model pembelajaran mesin.

Fungsi utama di dalam aplikasi AR Visualisasi Data E-Dagang ini ialah Visualisasi Data menggunakan AR yang interaktif dan informatif dan Fungsi Ramalan Keuntungan bagi Data E-Dagang. [1] Fasa yang penting ialah fasa rekabentuk dan implementasi. Fasa Rekabentuk termasuk rekabentuk Seni Bina Aplikasi, Reka Bentuk Pangkalan Data, Reka Bentuk Antara Muka [2] dan Reka Bentuk Algoritma. [3] Aplikasi AR Visualisasi Data ini memiliki algoritma yang melibatkan AI iaitu model pembelajaran mesin. Model pembelajaran mesin yang digunakan di dalam aplikasi ini adalah untuk meramal keuntungan perniagaan berdasarkan Data E-Dagang. Terdapat beberapa proses yang melibatkan pengimplementasian

ilmu pengetahuan sains data bagi menghasilkan data yang berkualiti dan analisis yang efektif. Penghasilan algoritma model pembelajaran mesin bagi penyediaan data E-Dagang ke dalam aplikasi akan menggunakan rangka kerja O.S.E.M.N. Fasa tersebut adalah pengumpulan data, pembersihan data, explorasi data, permodelan data dan penginterpretasian data melalui visualisasi.

Rangka Kerja O.S.E.M.N

Pengumpulan Data

Pengumpulan data ialah proses mengumpul dan mengukur maklumat berdasarkan pembolehubah ditetapkan dalam perancangan sesebuah sistem atau kajian yang membolehkan menjawab soalan relevan serta menilai keputusan yang berkemungkinan.

Data E- Dagang yang dipilih bagi kajian ini merupakan data terbuka yang diambil melalui laman web Kaggle. Data ini dinamakan sebagai “Superstore Sales”. Rajah 2 menunjukkan data tersebut di dalam format CSV di dalam Microsoft Excel.

1	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	State
2	1	CA-2016-152156	8/11/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky
3	2	CA-2016-152156	8/11/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky
4	3	CA-2016-138688	12/6/2016	16/6/2016	Second Class	DV-13045	Darrin Van H.	Corporate	United States	Los Angeles	California
5	4	US-2015-108966	11/10/2015	18/10/2015	Standard Class	SO-20335	Sean O'Donn	Consumer	United States	Fort Lauderdale	Florida
6	5	US-2015-108966	11/10/2015	18/10/2015	Standard Class	SO-20335	Sean O'Donn	Consumer	United States	Fort Lauderdale	Florida
7	6	CA-2014-115812	9/6/2014	14/6/2014	Standard Class	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
8	7	CA-2014-115812	9/6/2014	14/6/2014	Standard Class	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
9	8	CA-2014-115812	9/6/2014	14/6/2014	Standard Class	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
10	9	CA-2014-115812	9/6/2014	14/6/2014	Standard Class	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
11	10	CA-2014-115812	9/6/2014	14/6/2014	Standard Class	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
12	11	CA-2014-115812	9/6/2014	14/6/2014	Standard Class	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
13	12	CA-2014-115812	9/6/2014	14/6/2014	Standard Class	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
14	13	CA-2017-114412	15/4/2017	20/4/2017	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord	North Carolina
15	14	CA-2016-161389	5/12/2016	10/12/2016	Standard Class	IM-15070	Irene Maddo	Consumer	United States	Seattle	Washington
16	15	US-2015-118983	22/11/2015	26/11/2015	Standard Class	HP-14815	Harold Pawla	Home Office	United States	Fort Worth	Texas
17	16	US-2015-118983	22/11/2015	26/11/2015	Standard Class	HP-14815	Harold Pawla	Home Office	United States	Fort Worth	Texas
18	17	CA-2014-105893	11/11/2014	18/11/2014	Standard Class	PK-19075	Pete Kriz	Consumer	United States	Madison	Wisconsin
19	18	CA-2014-167164	13/5/2014	15/5/2014	Second Class	AG-10270	Alejandro Grc	Consumer	United States	West Jordan	Utah
20	19	CA-2014-143336	27/8/2014	1/9/2014	Second Class	ZD-21925	Zuschuss Dor	Consumer	United States	San Francisco	California
21	20	CA-2014-143336	27/8/2014	1/9/2014	Second Class	ZD-21925	Zuschuss Dor	Consumer	United States	San Francisco	California
22	21	CA-2014-143336	27/8/2014	1/9/2014	Second Class	ZD-21925	Zuschuss Dor	Consumer	United States	San Francisco	California
23	22	CA-2016-137330	9/12/2016	13/12/2016	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska
24	23	CA-2016-137330	9/12/2016	13/12/2016	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska
25	24	US-2017-156909	16/7/2017	18/7/2017	Second Class	SF-20065	Sandra Flana	Consumer	United States	Philadelphia	Pennsylvania
26	25	CA-2015-106320	25/9/2015	30/9/2015	Standard Class	EB-13870	Emily Burns	Consumer	United States	Orem	Utah
27	26	CA-2016-121755	16/1/2016	20/1/2016	Second Class	EH-13945	Eric Hoffman	Consumer	United States	Los Angeles	California
28	27	CA-2016-121755	16/1/2016	20/1/2016	Second Class	EH-13945	Eric Hoffman	Consumer	United States	Los Angeles	California

Rajah 2: Data E-Dagang

Set data ini mempunyai 9994 bilangan data dan 21 attribut. Terdapat beberapa peringkat pengukuran data yang dapat dikuantifikasikan untuk setiap attribut Data E-Dagang ini. Peringkat data tersebut adalah nominal, ordinal, sela dan nisbah. Data nominal adalah data berbentuk kategori (Non Parametrik). Data ordinal pula adalah ukuran berdasarkan urutan, keutamaan dan kedudukan. Data yang mempunyai urutan dan wujud perbezaan antara jeda tetapi tiada mutlak kosong dinamakan sebagai data sela. Manakala data nisbah pula mempunyai nilai mutlak kosong.

Pembersihan Data

Proses pembersihan data adalah sangat signifikan bagi membantu pemprosesan data untuk memastikan ia lebih efisyen dan menyediakan keputusan yang efektif. Jadual 1 menunjukkan teknik pra-pemprosesan yang dijalankan bagi Data E- Dagang untuk mewujudkan data yang berkualiti untuk kegunaan proses seterusnya.

Jadual 1: Teknik Pra-Pemprosesan bagi Data E- Dagang

Teknik Pra Pemprosesan	Example	Example
Kumpulkan data yang mempunyai tarikh yang sama. [1]	Data duplikasi dapat dielakkan dan penyelarasan data dapat dilaksanakan. Jumlahkan kesemua keuntungan jualan mengikut rekod tarikh yang sama.	Kumpulkan rekod mengikut tarikh dan tambahkan nilai keuntungan data tersebut mengikut kumpulan:[1]ecom_sales_new = ecom_sales.groupby('Order Date')[['Profit']].sum().reset_index()
Menetapkan siri masa sebagai index data .	Rekod Data E-Dagang akan berpandukan tarikh sebagai index terbaru mereka.	Penetapan attribut ‘Order Date’ sebagai index terbaru:[1]ecom_sales_new = ecom_sales_new.set_index('Order Date') ecom_sales_new.index
Penyelerasian Data melalui teknik ‘Resample’	Membuat pengusunan semula rekod data E-Dagang mengikut tarikh permulaan bagi setiap bulan disusuli dengan purata keuntungan jualan pada bulan tersebut.	Penyusunan semula Data E-Dagang menggunakan ‘MS’: [1]ecom_sales_new = ecom_sales_new[['Profit']].resample('MS').mean()

Explorasi Data

Fasa eksplorasi data ini adalah peringkat yang sangat penting dalam kitaran hidup sains data kerana melalui dengan cara ini, ciri-ciri data dapat dikenalpasti dengan lebih baik dan memudahkan proses dijalankan untuk peringkat seterusnya. Antara perkara yang dilihat dan dikaji bagi Data E-Dagang dalam proses eksplorasi data ini adalah:

- Jumlah keuntungan jualan dari segi bulanan dan tahunan [1]
- Perbandingan keuntungan jualan berdasarkan kategori barang. [1]

Permodelan Data

Proses ini juga berfokus untuk meramalkan sifat kejadian yang tidak pernah dilihat sebelumnya dan perkara ini boleh dibuat melalui teknik klasifikasi, regresi dan pengesanan anomal. Bagi model pembelajaran mesin yang dipilih, ia perlu mencari ‘hyperparameter’ yang paling optimum untuk mewujudkan nilai ketepatan yang bagus. Ralat sesebuah model juga

boleh dilihat melalui nilai '*Mean Squared Error*' dan '*Root Mean Squared Error*' (Jadual 2). Untuk kajian ini, model ARIMA akan digunakan bagi meramal nilai keuntungan berdasarkan Data E-Dagang.

Jadual 2: Cara Pengukuran Ralat Model ARIMA

Teknik Pengukuran Ralat	Penerangan	Formula
'Root Mean Square Error' (RMSE) 	Mengukur sisisian piawai model ramalan dengan label yang sebenar.	$RMSE(X, h) = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(x^{(i)}) - y^{(i)})^2}$
'Mean Absolute Error' (MAE) 	Mengukur purata magnitud ralat di dalam set ramalan. Ia adalah purata berdasarkan ujian sampel perbezaan mutlak antara ramalan dengan pemerhatian sebenar yang mana kesemua perbezaan individu memiliki berat atau ' <i>weight</i> ' yang sama.	$MAE(X, h) = \frac{1}{m} \sum_{i=1}^m h(x^{(i)}) - y^{(i)} $

Penilaian bagi kedua-dua teknik ini perlu dilakukan bagi mengenalpasti ketepatan model ramalan yang dihasilkan itu adalah efektif atau tidak. Model ini juga mengaplikasikan teknik carian grid bagi mendapatkan nilai parameter model yang paling optimum yang mampu menghasilkan ramalan yang terbaik.

Interpretasi dan Visualisasi Data

Peringkat akhir untuk proses sains data adalah interpretasi model dan data. Kekuatan sesebuah ramalan itu terletak kepada model yang berkemampuan untuk melihat corak umum bagi sesuatu set data. Proses interpretasi data ini adalah merupakan satu proses untuk mempersempit data berdasarkan penemuan yang dijumpai melalui proses sains data daripada peringkat awal hingga ke peringkat yang akhir.

Salah satu inisiatif untuk membawa interpretasi Data E-Dagang ke peringkat yang lebih tinggi adalah dengan memvisualisasikan data tersebut dengan menggunakan teknologi AR. Dengan penghasilan Aplikasi AR Visualisasi Data E-Dagang, ia mampu meningkatkan lagi kefahaman dari segi persepsi dan kognitif pengguna terhadap data tersebut dengan adanya fungsi untuk berinteraksi dengan data bagi mengeluarkan maklumat yang lebih informatif. Visualisasi data dalam aplikasi ini membolehkan pengguna untuk melihat data dari pelbagai sudut pandangan.

Hasil Kajian

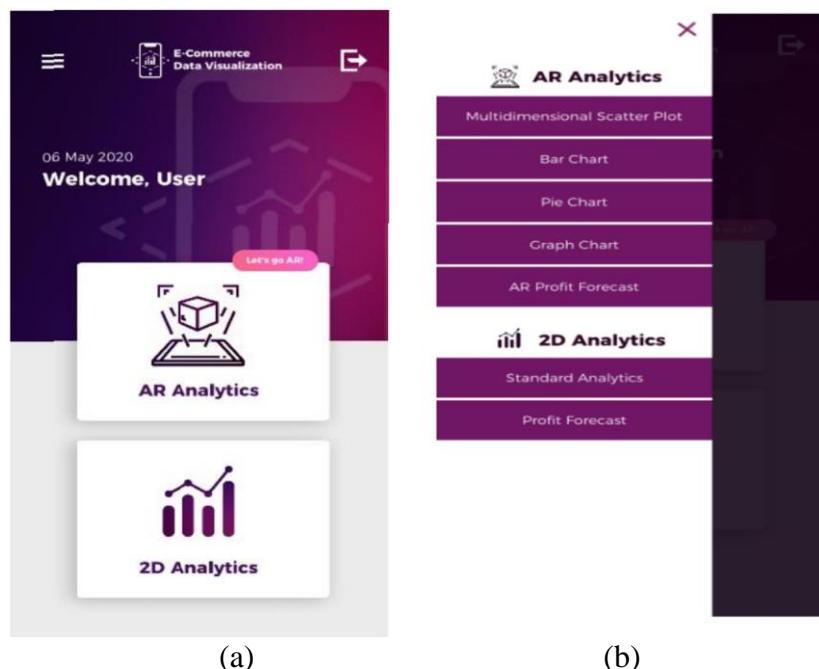
Aplikasi AR Visualisasi Data E-Dagang

Aplikasi AR Visualisasi Data E-Dagang merupakan aplikasi yang bertujuan menyediakan fungsi visualisasi data dalam bentuk yang lebih interaktif dan informatif dalam bentuk AR dan juga visualisasi data 2D. Permasalahan seperti ketiadaan dimensi kedalaman dan juga limitasi yang lain dalam penyampaian visualisasi data 2D telah mencetuskan idea untuk

mengimplementasikan pendekatan yang baharu iaitu dengan menggunakan teknologi AR dalam visualisasi data. Penggunaan AR sememangnya sangat membantu dalam mewujudkan visualisasi yang lebih unik dan menarik perhatian pengguna dalam masa yang sama menyediakan maklumat yang lengkap dan jelas.

Menu Utama

Pada bahagian menu utama, nama pengguna dan tarikh terkini dipaparkan di bahagian atas. Terdapat dua butang utama iaitu “AR Analytics” dan juga “2D Analytics”. Di bahagian atas kiri skrin, terdapat butang ke menu navigasi manakala di bahagian kanan skrin tersebut adalah butang untuk log keluar akaun. Rajah 3(a) menunjukkan antara muka menu utama aplikasi. Rajah 3(b) pula menunjukkan antara muka menu navigasi aplikasi ini.



Rajah 3: (a) Menu Utama Aplikasi (b) Menu Navigasi Aplikasi 14

Analitik AR

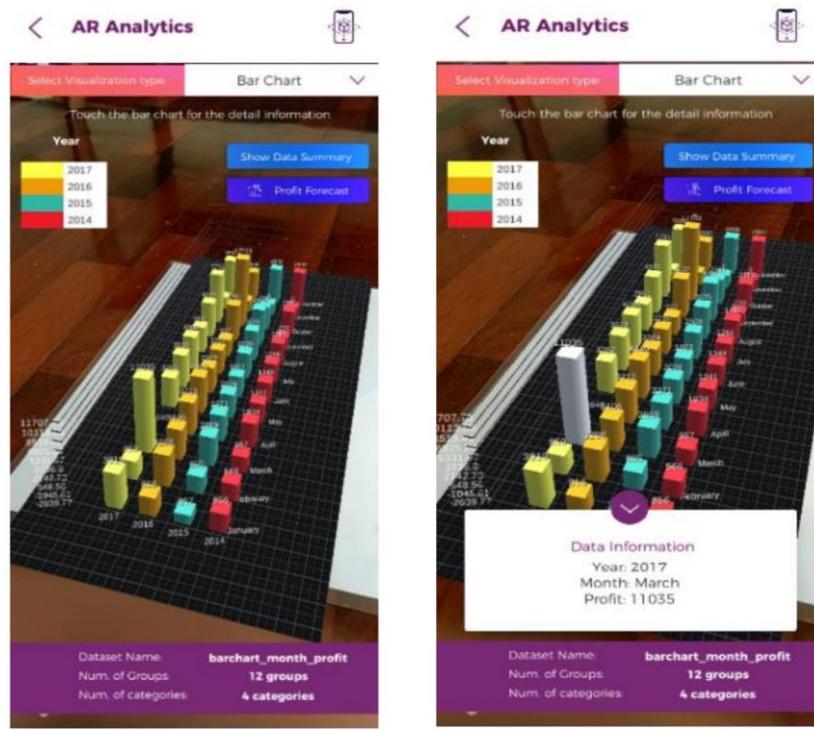
Analitik AR merupakan fungsi utama aplikasi yang memaparkan visualisasi data E-Dagang dalam bentuk AR. Aplikasi ini menggunakan AR berpenanda seperti dalam Rajah 4 sebagai imej sasaran untuk mengeluarkan objek atau paparan visualisasi data secara AR.



Rajah 4: Sasaran Imej Visualisasi Data AR

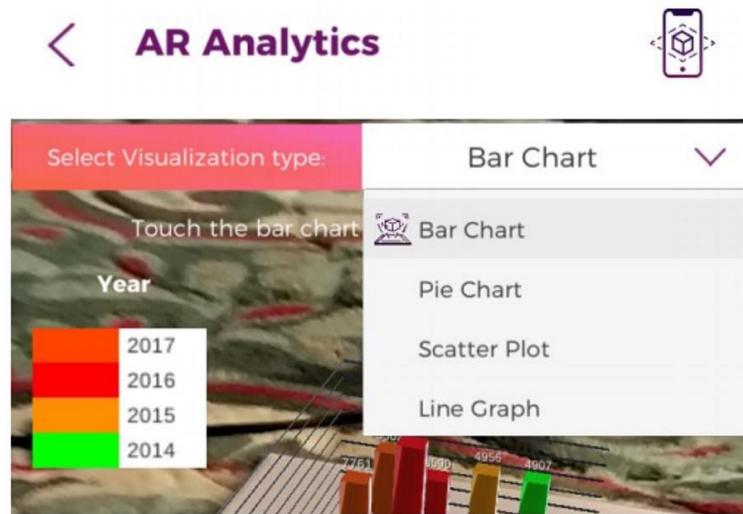
Carta Bar AR

Carta bar AR bagi aplikasi ini memaparkan hasil implementasi penggunaan aset Graph and Chart Unity (Rajah 5(a)). Pengguna boleh berinteraksi dengan carta bar tersebut untuk mengeluarkan informasi berkaitan perincian sesuatu data itu. Pengguna hanya perlu menekan mana-mana bar pada carta bar AR tersebut dan informasi tentang data itu dapat dipaparkan. Rajah 5(b) menunjukkan gambaran perincian informasi yang dikeluarkan sekiranya salah satu bar ditekan oleh pengguna.



Rajah 5: (a) Carta Bar AR (b) Perincian Data bagi Carta Bar

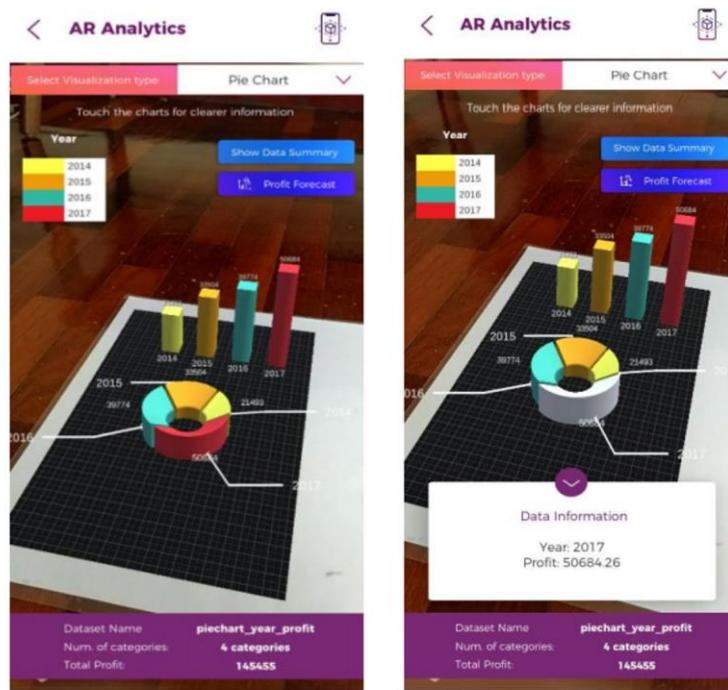
Di bahagian atas skrin terdapat satu kotak kombo yang membantu pengguna untuk mengakses ke skrin visualisasi data AR yang lain (lihat Rajah 6).



Rajah 6: Kotak Kombo Visualisasi Data AR

Carta Pai AR

Bagi carta pai ini juga, pengguna boleh berinteraksi dengan visualisasi data tersebut untuk mengeluarkan informasi berkaitan bahagian pecahan carta pai (Rajah 7 (a)) tersebut. Rajah 7 (b) menunjukkan gambaran perincian informasi yang dikeluarkan apabila carta pai itu disentuh oleh pengguna.



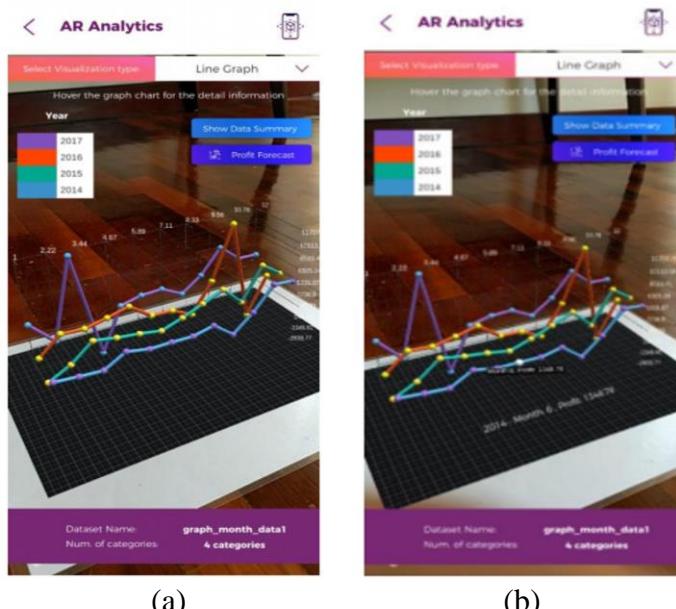
(a)

(b)

Rajah 7: (a) Carta Pai AR (n) Perincian Data bagi Carta Pai

Carta Garisan AR

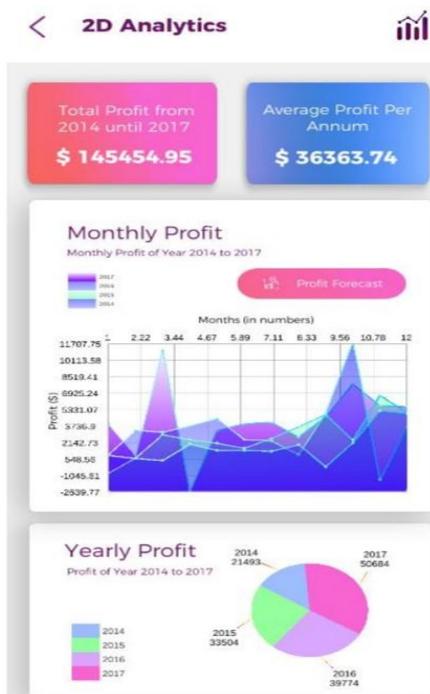
Bagi carta garisan AR, perincian informasi yang dikeluarkan adalah sedikit berbeza berbanding carta-carta yang lain kerana ia tidak dipaparkan dalam bentuk kad. Informasi yang dikeluarkan adalah berbentuk tulisan sahaja beserta warna yang kontras dibelakang tulisan tersebut (Rajah 8 (a)). Tulisan di sebelah kiri mewakili bulan manakala di sebelah kanan mewakili nilai keuntungan yang dicapai pada bulan tersebut. Rajah 8(b) menunjukkan ringkasan data bagi carta garisan AR tersebut.



Rajah 8 : (a) Carta Garisan AR (b) Perincian Data bagi Carta Garisan AR

Analitik 2D

Bagi fungsi analitik 2D, hanya dua visualisasi data iaitu carta pai dan carta garisan dalam berbentuk 2D yang dipaparkan. Di dalam skrin ini juga terpapar jumlah keuntungan bagi kesemua tahun untuk jualan barang teknologi dan juga purata keuntungan tahunan. Rajah 9 menunjukkan antara muka analitik 2D bagi aplikasi ini.



Rajah 9: Analitik 2D

Ramalan Keuntungan

Fungsi ramalan keuntungan merupakan juga bahagian utama aplikasi bagi memaparkan ramalan jumlah keuntungan jualan pada tahun-tahun akan datang. Aplikasi ini menyediakan dua jenis ramalan keuntungan iaitu secara visualisasi 2D dan AR. Pendekatan yang berbeza ini mampu membantu memahami data tersebut dengan lebih mendalam serta menarik.

Visualisasi Ramalan Keuntungan Menggunakan AR

Rajah 10 menunjukkan visualisasi AR bagi data ramalan keuntungan E-Dagang. Data tersebut boleh diinteraksi bagi memaparkan informasi yang lebih banyak. Di bahagian atas skrin, terdapat kombo yang membolehkan pengguna untuk memilih tahun untuk diramal dan seterusnya tekan pada butang “*FORECAST*”.



Rajah 10: Visualisasi AR bagi Fungsi Ramalan Keuntungan

Visualisasi Ramalan Keuntungan Secara 2D

Untuk bahagian ini, paparan data tersebut adalah sama seperti data pada ramalan keuntungan AR. Ia hanya berbeza di bahagian pengimplementasi teknik visualisasi yang mana bahagian ini menggunakan teknik 2D. Rajah 11 menunjukkan visualisasi 2D bagi ramalan keuntungan E-Dagang. Di bahagian tengah skrin, terdapat juga kotak kombo yang membenarkan pengguna untuk memilih tahun yang ingin diramal keuntungannya dan tekan pada butang “*FORECAST*” untuk memaparkan data ramalan.



Rajah 11: Visualisasi AR bagi Fungsi Ramalan Keuntungan

Prestasi Model Pembelajaran Mesin (ARIMA)

Untuk mencari ketepatan data siri masa bagi ramalan tersebut menggunakan ARIMA, matlamat utama bagi langkah yang pertama model adalah dengan mencari nilai metrik minat yang optimum bagi fungsi:

$$\text{ARIMA}(p,d,q)(P,D,Q)s$$

Teknik ini menggunakan carian grid bagi mengeksplorasi data dengan gabungan pelbagai parameter sehingga nilai matlamat itu tercapai. Bagi setiap kombinasi parameter, model ARIMA akan disesuaikan menggunakan fungsi SARIMAX() daripada modul “statsmodel” and menilai kualiti parameter itu secara keseluruhannya. Rajah 12 menunjukkan contoh output yang dihasilkan untuk carian nilai metrik optimum.

```

ARIMA(1, 1, 0)x(1, 0, 1, 12)12 - AIC:1406.170280886325
ARIMA(1, 1, 0)x(1, 1, 0, 12)12 - AIC:316.09413416974314
ARIMA(1, 1, 1)x(0, 0, 0, 12)12 - AIC:581.1841479513987
ARIMA(1, 1, 1)x(0, 0, 1, 12)12 - AIC:5940.171626109858
ARIMA(1, 1, 1)x(0, 1, 0, 12)12 - AIC:450.25406642822605
ARIMA(1, 1, 1)x(1, 0, 0, 12)12 - AIC:448.5680221538713
ARIMA(1, 1, 1)x(1, 0, 1, 12)12 - AIC:2534.330574615973
ARIMA(1, 1, 1)x(1, 1, 0, 12)12 - AIC:305.1359719413757

```

Rajah 12: Keputusan Parameter Bagi Nilai Metrik Minat

Berdasarkan Rajah 12, berikut adalah nilai parameter yang terpilih yang memiliki jumlah skor AIC yang terendah:-

ARIMA(1, 1, 1)x(1, 1, 0, 12)12 - AIC:305.1359719413757

Setelah parameter yang paling optimum dikenalpasti, parameter tersebut akan disesuaikan ke dalam model ARIMA. Perkara ini akan membawa kepada model yang paling sesuai untuk data siri masa E-Dagang. Rajah 13 menunjukkan keputusan prestasi model apabila diproses bersama Data E-Dagang tersebut.

	coef	std err	z	P> z	[0.025	0.975]
ar.L1	-0.2990	0.189	-1.581	0.114	-0.670	0.072
ma.L1	-0.9297	0.484	-1.921	0.055	-1.878	0.019
ar.S.L12	-0.6029	0.493	-1.222	0.222	-1.570	0.364
sigma2	4.039e+04	2.08e+04	1.940	0.052	-420.581	8.12e+04

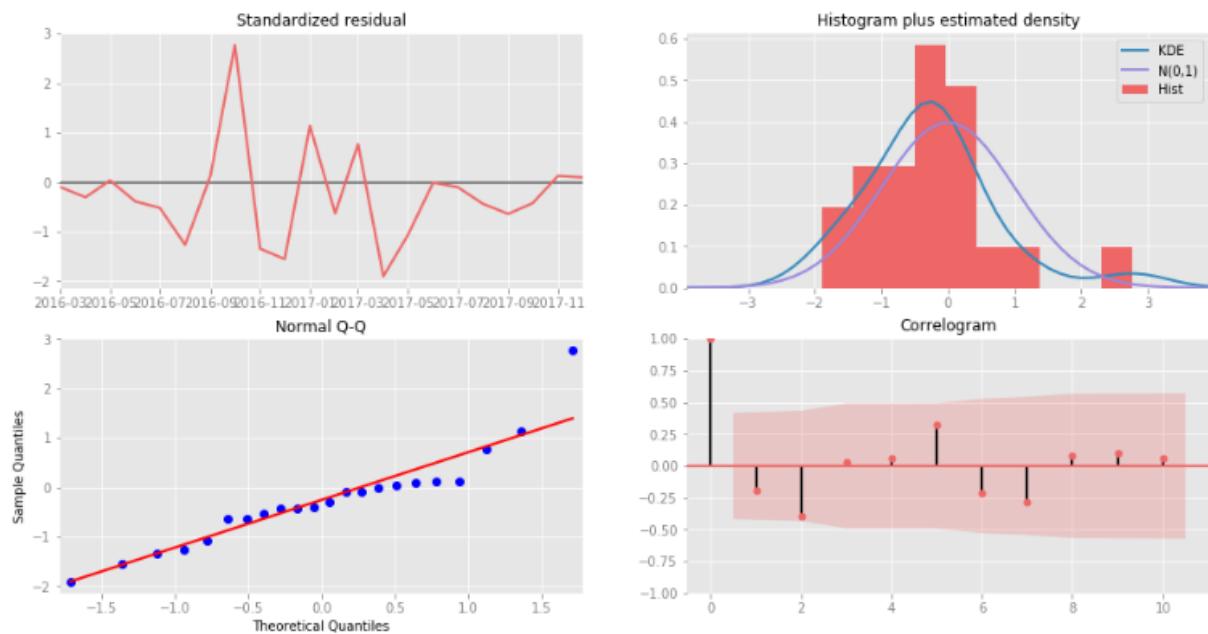
Statespace Model Results						
Dep. Variable:	Profit	No. Observations:	48			
Model:	SARIMAX(1, 1, 1)x(1, 1, 0, 12)	Log Likelihood	-148.568			
Date:	Thu, 28 May 2020	AIC	305.136			
Time:	11:13:08	BIC	309.500			
Sample:	01-01-2014	HQIC	306.164			
	- 12-01-2017					
Covariance Type:	opg					

	coef	std err	z	P> z	[0.025	0.975]
ar.L1	-0.2990	0.189	-1.581	0.114	-0.670	0.072
ma.L1	-0.9297	0.484	-1.921	0.055	-1.878	0.019
ar.S.L12	-0.6029	0.493	-1.222	0.222	-1.570	0.364
sigma2	4.039e+04	2.08e+04	1.940	0.052	-420.581	8.12e+04

Ljung-Box (Q):	nan	Jarque-Bera (JB):	9.74
Prob(Q):	nan	Prob(JB):	0.01
Heteroskedasticity (H):	0.38	Skew:	1.12
Prob(H) (two-sided):	0.23	Kurtosis:	5.37

Rajah 13: Prestasi Model

Untuk memvisualisasikan prestasi data dengan lebih jelas, fungsi plot_diagnostic() boleh digunakan untuk menyelidik ciri-ciri data. Rajah 14 menunjukkan prestasi model apabila diproses bersama Data E-Dagang tersebut.

**Rajah 14: Keputusan Diagnostik Model**

Berdasarkan rajah tersebut, ia menunjukkan prestasi model bagi menjalankan ramalan adalah memuaskan dan tidaklah terlalu tepat. Perkara ini bergantung kepada corak data yang terdahulu bagi Data E-Dagang. Semakin banyak data dan corak yang sama dalam sesuatu siri masa, semakin bagus dan tepat ramalan itu dihasilkan. Rajah 15 menunjukkan graf ramalan keuntungan bagi Data E-Dagang yang dihasilkan berdasarkan rajah di atas tersebut.

**Rajah 15: Graf Ramalan Keuntungan**

Berpandukan Rajah 15, ia menunjukkan bahawa semakin jauh ramalan itu dibuat, semakin tinggi ralat terhadap ramalan tersebut. Perkara tersebut boleh dilihat daripada pengembangan sempadan atas dan sempadan bawah ramalan tersebut. Jadual 3 menunjukkan nilai MSE dan RMSE bagi model ARIMA berdasarkan Data E-Dagang tersebut.

Jadual 3: Nilai MSE dan RMSE Model ARIMA

Teknik Pengukuran Ralat	Keputusan
'Root Mean Square Error' (RMSE) [SEP]	26928.47
'Mean Absolute Error' (MAE) [SEP]	165.1

Berdasarkan Jadual 3, nilai-nilai pengukuran ralat menunjukkan jumlah yang agak banyak dan oleh sebab itulah ralat yang dihasilkan juga agak besar. Walaubagaimanapun, melalui keputusan ini juga ia membantu para pengguna untuk mengetahui corak siri masa yang dihasilkan untuk masa akan datang berdasarkan trend keuntungan yang masa yang lalu.

Kesimpulan

Aplikasi AR Visualisasi Data E-Dagang adalah berfokus untuk menggabungkan elemen AR, Kecerdasan Buatan dan Data Raya. Melalui kajian yang dijalankan, dapat dikenalpasti bahawa AR dan visualisasi data adalah sangat signifikan untuk diintegrasikan bersama kerana kedua-dua mempunyai matlamat yang sama. Dengan adanya elemen AR ini, visualisasi data akan menjadi lebih efektif dan mampu mengatasi limitasi kognitif manusia. Perkara ini dapat diperjelaskan lagi dengan kebolehannya yang membenarkan pengguna untuk menganalisis data yang divisualisasikan dari pelbagai sudut pandangan. Aplikasi AR Visualisasi Data E-Dagang yang telah dibangunkan ini memiliki kelebihan yang membolehkan pengguna untuk berinteraksi dengan sesuatu titik data dan mampu meramalkan keuntungan perniagaan mengikut tahun yang dipilih. Aplikasi AR Visualisasi Data E-Dagang adalah berada pada tahap yang sangat memuaskan kerana ia merupakan salah satu inisiatif yang unik dalam masa yang sama berguna untuk meningkatkan lagi kefahaman pengguna terhadap sesuatu data itu.

Rujukan

- Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, 20, 1–11.
- Arshad, H., Ahmad Shukri, S. A., Zainal, R., & Tan, S. Y. (2019). User experience assessment of mobile augmented reality for tourism Control Systems. *Journal of Advanced Research in Dynamical and Control Systems*, 11(08), 1059–1064.
- Barenji, A. V., Wang, W. M., Li, Z., & Guerra-Zubiaga, D. A. (2019). Intelligent E-commerce logistics platform using hybrid agent based approach. *Transportation Research Part E: Logistics and Transportation Review*, 126, 15–31.
- Cielen, D., Meysman, A., & Ali, M. (2016). *Introducing data science: big data, machine learning, and more, using Python tools*. Manning Publications Co.
- Donalek, C., Djorgovski, S. G., Cioc, A., Wang, A., Zhang, J., Lawler, E., ... Drake, A. (2014). Immersive and collaborative data visualization using virtual reality platforms. *2014 IEEE International Conference on Big Data (Big Data)*, 609–614. IEEE.
- Furht, B. (2011). *Handbook of augmented reality*. Springer Science & Business Media.
- García-Crespo, Á., González-Carrasco, I., López-Cuadrado, J. L., Villanueva, D., & González, Á. (2016). CESARSC: Framework for creating cultural entertainment systems with augmented reality in smart cities. *Computer Science and Information Systems*, 13(2), 395–425.
- Gustafsson, M., & Odd, O. (2018). *Virtual reality data visualization: concepts, technologies and more*.

- Haller, M. (2006). *Emerging technologies of augmented reality: Interfaces and design: Interfaces and design*. IGI Global.
- Herron, J. (2016). Augmented reality in medical education and training. *Journal of Electronic Resources in Medical Libraries*, 13(2), 51–55.
- Ho, S. L., & Xie, M. (1998). The use of ARIMA models for reliability forecasting and analysis. *Computers & Industrial Engineering*, 35(1–2), 213–216.
- Huang, Z., Hui, P., & Peylo, C. (2014). When augmented reality meets big data. *ArXiv Preprint ArXiv:1407.7223*.
- Hune, N., & Müller, M. (2018). *The Data in Your Hands: Exploring Novel Interaction Techniques and Data Visualization Approaches for Immersive Data Analytics*.
- Khan, M. M. H., Muhammad, N. S., & El-Shafie, A. (2020). Wavelet based hybrid ANN-ARIMA models for meteorological drought forecasting. *Journal of Hydrology*, 590, 125380.
- Kumar, K. N., Chandra, S., Bharati, S., & Manava, S. (2016). Factors influencing adoption of augmented reality technology for e-commerce. *Pacific Asia Conference on Information Systems (PACIS)*. Association For Information System.
- Mahayuddin, Z. R., & Saif, A. F. M. S. (2020). Augmented Reality based AR Alphabets towards Improved Learning Process in Primary Education System. *Journal of Critical Reviews*, 7(19), 514–521.
- Meng, C. L., Abd Majid, N. A., Muhammad Nizam, S. S., Suwadi, N. A., & Sadik, M. J. (2019). Towards Interactive Mobile Augmented Reality Fusion. *National Symposium Human Computer Interaction*.
- Nochai, R., & Nochai, T. (2006). ARIMA model for forecasting oil palm price. *Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications*, 13–15.
- Olshannikova, E., Ometov, A., Koucheryavy, Y., & Olsson, T. (2015). Visualizing Big Data with augmented and virtual reality: challenges and research agenda. *Journal of Big Data*, 2(1), 22.
- PS, M. G., & Chintala, B. R. (2020). Big Data Challenges and Opportunities in Agriculture. *International Journal of Agricultural and Environmental Information Systems (IJAEIS)*, 11(1), 48–66.
- Qin, X., Luo, Y., Tang, N., & Li, G. (2020). Making data visualization more efficient and effective: A survey. *The VLDB Journal*, 29(1), 93–117.
- Rahman, H., Arshad, H., Mahmud, R., Mahayuddin, Z. R., & Obeidy, W. K. (2017). A Framework to Visualize 3D Breast Tumor Using X-Ray Vision Technique in Mobile Augmented Reality. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 9(2–11), 145–149.
- Tredinnick, L. (2018). Augmented reality in the business world. *Business Information Review*, 35(2), 77–80.
- Xiang, Y., & Zhuang, X. H. (2013). Application of ARIMA model in short-term prediction of international crude oil price. *Advanced Materials Research*, 798, 979–982. Trans Tech Publ.
- Yim, M. Y.-C., Chu, S.-C., & Sauer, P. L. (2017). Is augmented reality technology an effective tool for e-commerce? An interactivity and vividness perspective. *Journal of Interactive Marketing*, 39, 89–103.