

KEBOLEHTERIMAAN MASYARAKAT TERHADAP APLIKASI E-COMMERCE (LAZADA) MENGGUNAKAN ANALISIS SENTIMEN

SHAATISH A/L SELVAM

Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

E-dagang menjadi semakin luas dalam era teknologi ini. Lazada merupakan salah satu contoh aplikasi yang membenarkan pengguna untuk membeli barang atas talian. Namun, kualiti syarikat tersebut boleh dilihat melalui komen serta maklum balas daripada pengguna aplikasi itu sendiri. Oleh itu, analisis sentimen merupakan satu proses untuk menentukan sama ada sesuatu tulisan itu ialah positif, negatif ataupun neutral. Analisis Sentimen dilakukan terhadap aplikasi Lazada supaya ia memberikan gambaran yang lebih jelas terhadap perkhidmatan yang diberikan oleh syarikat itu. Set data untuk ulasan terhadap Lazada dikumpul melalui aplikasi Twitter kerana ia merupakan aplikasi yang paling kerap digunakan oleh masyarakat untuk menyatakan sesuatu. Set data dibersihkan menggunakan kaedah pra-pemprosesan data seperti mengyingkirkan URL, simbol, tanda baca dan selebihnya. Seterusnya, dataset tersebut melalui pengekstrakan ciri untuk dilabel. Kemudian, proses analisis sentimen dilaksanakan diikuti dengan kaedah pembelajaran mesin *Naive Bayes Classifier* dan *Support Vector Machine* untuk menentukan polariti teks serta ketepatan algoritma. Akhirnya, hasil tersebut diterjemahkan dalam bentuk graf untuk pemahaman lebih.

1. PENGENALAN

Dalam era pemodenan ini, masyarakat merupakan kelompok sebilangan manusia yang tinggal bersama-sama serta saling berinteraksi antara satu sama lain. Mengikut Kamus Dewan Edisi Keempat, kebolehterimaan membawa maksud keadaan boleh diterima atau mencari jalan untuk membuat penyelesaian. Jelaslah bahawa kebolehterimaan masyarakat bermaksud perihal yang boleh diterima atau membuat penyelesaian oleh orang ramai terhadap sesuatu.

Perniagaan atas talian melibatkan pembelian dan penjualan barang serta servis melalui penggunaan Internet (Dr. Shahid Amin Bhat, 2016). Perniagaan atas talian ini boleh dilakukan menggunakan pelbagai jenis peranti seperti komputer, tablet dan telefon pintar. Dalam era ini, pelbagai jenis pembelian boleh dibeli atas talian seperti barangan kesihatan, perabot rumah, pakaian dan alat elektrik. Perniagaan atas talian menjadi lebih penting kini semenjak pandemik Covid-19. Perniagaan atas talian menolong perniagaan dan menubuhkan pasaran yang lebih luas secara lebih efisien (Reynolds J,2000). Dalam kajian ini, aplikasi Lazada yang kerap digunakan oleh masyarakat Malaysia akan menjadi sasaran.

Lazada di Malaysia pula ditubuhkan di Malaysia kira-kira 8 tahun dahulu, iaitu pada tahun 2013. Aplikasi Lazada dilancarkan untuk sistem Android pada tahun 2013 dan untuk sistem IOS pada tahun seterusnya. Semenjak itu, statistik menunjukkan sebanyak \$1.3 bilion Amerika Syarikat dijana oleh Lazada Group melalui pengguna alat mudah alih. Kini, sebanyak 14 juta pengguna bulanan menggunakan aplikasi Lazada di Malaysia. Lazada juga merupakan aplikasi perniagaan atas talian kedua paling popular di Malaysia. (Joschka Muller, 2021).

Analisis Sentimen pula ialah satu proses untuk mengenal pasti sentimen positif ataupun negatif dalam teks (Cornelius Puschmann, 2018). Lazimnya, analisis sentimen digunakan dalam perniagaan untuk mengenalpasti sentimen dalam reputasi serta untuk memahami perasaan pengguna dengan lebih mudah. Proses ini adalah penting untuk dijalankan dalam setiap perniagaan terutama perniagaan atas talian untuk menambahbaikkan penggunaan harian. Secara ringkas, analisis sentimen digunakan untuk pemprosesan Bahasa semula jadi, linguistik pengiraan, analisis teks untuk mengenal pasti, mengekstrak dan mengkaji keadaan afektif serta maklumat subjektif secara sistematik.

Jadi, kebolehterimaan masyarakat terhadap Lazada tentulah menjadi sasaran hangat kini. Perbincangan serta komen dalam media sosial seperti Twitter dan termasuklah dalam aplikasi Lazada sendiri menyebabkan perubahan ketara dalam topik dan sentimen dari masa ke masa agar dapat memahami persepsi serta emosi orang ramai. Oleh itu, kajian ini adalah untuk mengenal pasti kebolehterimaan masyarakat terhadap aplikasi Lazada menggunakan analisis sentimen.

2. PENYATAAN MASALAH

Semenjak pandemik Covid-19 dilanda, perniagaan atas talian menjadi tumpuan orang ramai untuk membeli barang keperluan. Perniagaan atas talian ataupun e-dagang merupakan kaedah paling efektif untuk membeli barang. Masyarakat boleh menggunakan aplikasi Lazada untuk tujuan membeli belah. Mengikut statistik daripada GlobalData, perniagaan atas talian ataupun e-dagang telah meningkat sebanyak 24.7% dalam tahun 2020 dan dijangka akan terus meningkat ke 50% dalam tempoh 4 tahun akan datang. Terdapat banyak penjual dalam aplikasi ini yang memberi perkhidmatan kepada pengguna namun, adakah servis dan perkhidmatan yang disediakan adalah sempurna? Selain itu, adakah pembelian atas talian melalui aplikasi

Lazada ini memuaskan? Akhirnya, adakah pandangan masyarakat Malaysia terhadap aplikasi Lazada ini positif?

3. OBJEKTIF KAJIAN

Antara tujuan kajian ini adalah seperti berikut:

1. Mengetahui dan mengklasifikasikan kandungan teks berkenaan aplikasi Lazada sama ada kandungannya adalah sentimen positif, neutral ataupun negatif.
2. Mengetahui kaedah analisis sentimen yang paling tepat.
3. Mengetahui sama ada masyarakat setuju dan menerima perkhidmatan yang disediakan dalam aplikasi Lazada ataupun tidak.

4. METOD KAJIAN

Kajian ini menggunakan teknik analisis sentimen untuk mendapatkan persepsi masyarakat terhadap perkhidmatan dalam aplikasi Lazada. Untuk melaksanakan teknik ini, setiap fasa harus dilengkapkan sebelum fasa seterusnya untuk mendapatkan output yang tepat. Di penghujung setiap fasa, penilaian dibuat untuk memastikan kajian ini dilaksanakan seperti yang dirancang.

4.1 Fasa Pengumpulan Data

Fasa pertama yang dijalankan adalah pengumpulan set data iaitu yang melibatkan tweets daripada aplikasi Twitter yang mengandungi ulasan mengenai aplikasi Lazada. Antara contoh kata kunci yang digunakan dalam projek ini adalah *Lazada Service*, *Lazada Quality* dan *Lazada Malaysia*. Untuk mendapatkan set data mentah daripada aplikasi Twitter melalui Bahasa pengaturcaraan Python, perpustakaan *Snsrape* dan *Tweepy* digunakan. Untuk menggunakan perpustakaan ini, ia hendaklah dimuat turun ke dalam perisian Python dahulu. Kemudian, API yang digunakan dalam perisian mestilah mempunyai kata kunci yang berkenaan dengan skop iaitu perkhidmatan dalam aplikasi Lazada bagi memperincikan tweets yang dikumpulkan adalah sesuai serta tepat. Sebanyak 8866 tweets (data mentah) yang terkini berjaya dikumpulkan untuk projek ini. Selain itu, tweets yang dikumpulkan juga ditapis dalam Bahasa Inggeris sahaja dan dimuat turunkan ke dalam satu fail csv.

4.2 Fasa Pra-pemrosesan Data

Fasa kedua dalam projek ini adalah pra pemrosesan di mana tweets (data mentah) yang telah dikumpulkan dalam fail csv akan dibersihkan untuk menghasilkan set data yang tepat untuk analisis sentimen dan pengklasifikasian. Antara kaedah pra pemrosesan yang digunakan termasuklah penyingkiran ruang kosong, menukar perkataan kepada huruf kecil, penapisan karakter seperti “@pengguna”, “RT@pengguna”, dan URL serta penyingkiran karakter yang tidak membawa makna seperti tanda baca. Selepas pra pemrosesan data dijalankan, didapati bahawa hanya 7125 data yang tinggal. Seterusnya, senarai perkataan yang unik di dalam set data divisualisasikan dalam bentuk awan perkataan (*WordCloud*). Penyampaian kata kunci dalam bentuk gambar yang berwarna seperti awan supaya perkataan yang bermakna dapat diberi perhatian dalam fasa seterusnya dan perkataan yang tidak bermakna boleh disingkirkan untuk memudahkan proses pengekstrakan ciri dan pengelasan kategori. Pakej *wordcloud* dalam Python digunakan untuk menghasilkan awan perkataan yang menarik. Awan perkataan ini juga memberi gambaran ringkas dan mudah untuk difaham bagi mengetahui kandungan set data yang dikumpulkan secara keseluruhan.

4.3 Fasa Pengekstrakan Ciri

Fasa ketiga dalam projek ini melibatkan pengekstrakan ciri dan proses analisis sentimen. Dalam proses ini, terdapat dua teknik utama yang akan digunakan untuk mengenal pasti sentimen pengguna terhadap perkhidmatan aplikasi Lazada iaitu TextBlob dan VADER. Terdapat tiga label yang akan dipaparkan untuk menunjukkan sama ada positif, neutral ataupun negatif bagi tweets yang telah dibersihkan.

a) TextBlob

TextBlob merupakan sebuah pakej bagi Bahasa pengaturcaraan Python untuk *Natural Language Processing* (NLP). Kaedah TextBlob ini menggunakan *Natural Language Toolkit* (NLTK) untuk mendapatkan sentimen dalam sesuatu ayat. Terdapat dua sifat yang digunakan untuk mengenal pasti sentimen dalam sesuatu ayat iaitu polariti dan subjektiviti. Secara ringkas, polariti terletak di antara julat [-1,1] di mana -1 menunjukkan sentimen negatif dan +1 menunjukkan sentimen positif manakala subjektiviti pula mengukur dari segi pendapat peribadi atau maklumat fakta yang terkandung dalam sebuah teks.

b) VADER (*Valence Aware Dictionary for Sentiment Reasoning*)

VADER pula merupakan teknik yang digunakan dalam analisis sentimen untuk mengesan polariti iaitu positif ataupun negatif dan intensiti iaitu kekuatan emosi dalam teks yang sensitif. Analisis sentimen VADER menggunakan kamus yang ditakrifkan ciri leksikal kepada intensiti emosi yang dikenali sebagai skor sentimen. Kaedah ini juga sangat sesuai digunakan untuk data yang diperolehi daripada media sosial seperti Twitter.

4.4 Fasa Pembelajaran Mesin

Fasa keempat dalam projek ini melibatkan pengklasifikasian *tweets* yang telah dilabel menggunakan TextBlob dan juga VADER. Pengklasifikasian ini dilakukan menggunakan dua jenis teknik pembelajaran mesin iaitu pengklasifikasian *Naïve Bayes* dan pengklasifikasian *Support Vector Machine*. Teknik pembelajaran mesin ini dilakukan untuk membuat ramalan sama ada teks itu adalah positif, neutral ataupun negatif. Untuk menggunakan teknik pembelajaran mesin ini, set data yang telah dilabel menggunakan TextBlob dan VADER akan diguna pakai. Sebelum melakukan teknik pembelajaran mesin, teknik TF-IDF akan digunakan untuk menjadikan sifat-sifat dalam set data sebagai vektor. Kemudian, set data akan dibahagikan kepada set data latihan dan set data uji iaitu 80% peratus set data latihan dan 20% set data uji. Selepas itu, model pengklasifikasian pembelajaran mesin *Naïve Bayes* akan diguna pakai untuk mengelaskan dan membuat ramalan pada set data uji. Proses ini akan diulang untuk kedua-dua teknik analisis sentimen iaitu TextBlob dan VADER. Ketepatan pengklasifikasian *Naïve Bayes* dan *Support Vector Machine* akan dikira untuk mengenal pasti model yang paling tepat dan sesuai untuk digunakan.

4.5 Fasa Hasil & Diskusi

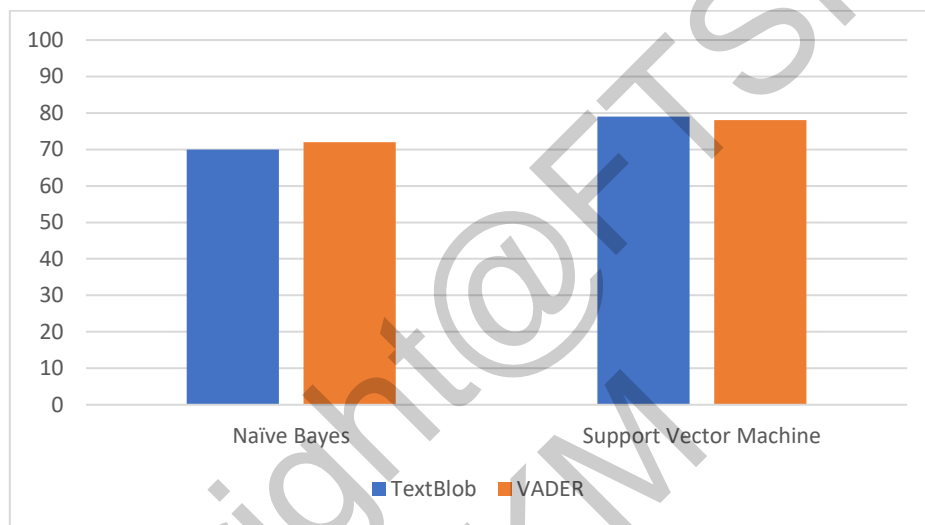
Dalam fasa hasil dan diskusi, teknik analisis sentimen yang paling tepat dan sesuai akan dipilih berdasarkan ketepatan pengklasifikasian pembelajaran mesin.

5. HASIL KAJIAN

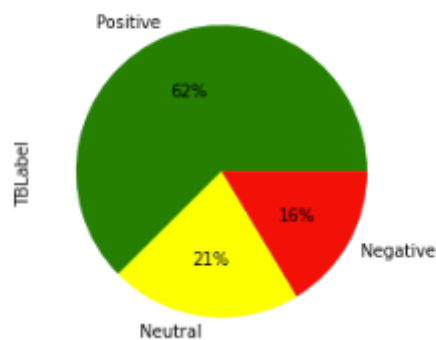
Kaedah analisis sentimen yang paling tepat dan sesuai akan dipilih berdasarkan ketepatan pengklasifikasian pembelajaran mesin.

Teknik Pembelajaran Mesin/ Teknik analisis sentimen	<i>Naïve Bayes</i>	<i>Support Vector Machine</i>
TextBlob	70%	79%
VADER	72%	78%

Jadual 1 Ketepatan Pengklasifikasian Pembelajaran Mesin menggunakan TextBlob dan VADER



Rajah 1 Graf Bar Perbandingan Ketepatan Pengklasifikasian Pembelajaran Mesin



Rajah 2 Carta Pai TextBlob

Berdasarkan jadual serta graf bar yang dijana atas, teknik pembelajaran mesin *Naïve Bayes* memberi ketepatan sebanyak 70% menggunakan TextBlob dan 72 bagi VADER manakala teknik pembelajaran mesin *Support Vector Machine* pula memberi ketepatan sebanyak 79% menggunakan TextBlob dan 78 bagi VADER. Di sini, didapati bahawa ketepatan pengklasifikasian *Support Vector Machine* adalah lebih tinggi berbanding pengklasifikasian *Naïve Bayes*. Selain itu, untuk teknik analisis sentimen pula, didapati bahawa ketepatan TextBlob adalah lebih tinggi berbanding kepada VADER. Jadi, merujuk kepada carta pai TextBlob, carta menunjukkan bahawa terdapat 62% teks yang dilabel sebagai positif, diikuti dengan 21% teks neutral dan 16% teks negatif. Ini menunjukkan bahawa majoriti penduduk di Malaysia mempunyai sentimen yang positif terhadap perkhidmatan serta kualiti yang disediakan dalam aplikasi Lazada. Sebagai konklusi, masyarakat Malaysia boleh dikatakan menerima dan menyukai penggunaan aplikasi Lazada.

6. KESIMPULAN

Secara keseluruhannya, kajian ini adalah untuk menganalisis sentimen tentang perkhidmatan aplikasi Lazada. Proses ini dimulakan dengan pengumpulan data daripada aplikasi Twitter dan Lazada dalam bentuk fail csv. Kemudian, data akan dibersihkan dan dilabelkan terlebih dahulu dalam fasa pra-pemprosesan dan fasa pengekstrakan ciri. Analisis sentimen dibuat menggunakan kaedah TextBlob dan *VADER Sentiment* kerana sesuai untuk mengenalpasti nada sentimen sama ada positif, neutral ataupun negatif. Kemudian, teknik pembelajaran mesin yang digunakan adalah pengelas *Naïve Bayes* dan pengelas *Support Vector Machine* untuk mendapatkan sentimen bagi set data uji.

Hasil Output bagi kajian telahpun mencapai objektif kajian serta dipaparkan dalam bentuk visual seperti graf. Objektif yang dicapai adalah seperti berikut:

1. Mengenal pasti dan mengklasifikasikan kandungan teks berkenaan aplikasi Lazada sama ada kandungannya adalah sentimen positif, neutral ataupun negatif
2. Mengenal pasti kaedah analisis sentimen yang paling tepat
3. Mengenal pasti sama ada masyarakat setuju dan menerima perkhidmatan yang disediakan dalam aplikasi Lazada ataupun tidak.

Kajian ini menggunakan teknik analisis sentimen serta pembelajaran mesin. Oleh itu, dalam menjalankan kajian ini, diharap dapat membantu meningkatkan pengetahuan dan kemahiran dalam pembelajaran mesin yang kini menjadi tumpuan dalam industri. Selain itu, kajian ini juga diharap dapat meningkatkan lagi kemahiran koding menggunakan Bahasa pengaturcaraan *Python*

7. RUJUKAN

- Tha'Er Majali, Malek Alsoud, Dirar Abdelaziz Almaaitah, Tamara Maaitah. (2021). *Journal of Contemporary Issues in Business and Government* Vol. 27, No. 1
- Dr. Shahid Amin Bhat, Keshav Kansana. (2016). *A Review Paper on E-Commerce*.
- Reynolds, J. (2000), "eCommerce: a critical review", *International Journal of Retail & Distribution Management*, Vol. 28 No. 10, pp. 417-444.
<https://doi.org/10.1108/09590550010349253>
- Wikipedia*. (n.d). Retrieved from Masyarakat:
<https://ms.wikipedia.org/wiki/Masyarakat>
- Dewan Bahasa dan Pustaka. Retrieved from kebolehterimaan:
<https://prpm.dbp.gov.my/cari1?keyword=kebolehterimaan>
- Wikipedia*. (n.d). Retrieved from Lazada Malaysia:
https://ms.wikipedia.org/wiki/Lazada_Malaysia
- GlobalData. (2020, September 8).
<https://www.globaldata.com/covid-19-accelerates-e-commerce-growth-malaysia-says-globaldata/>
- Joschka Muller. (2021, September 29). Top 10 e-commerce sites in Malaysia Q2 2021, by monthly traffic.
<https://www.statista.com/statistics/869640/malaysia-top-10-e-commerce-sites/>
- The Writing Center. (2021). University of North Carolina at Chapel Hill.
<https://writingcenter.unc.edu/tips-and-tools/literature-reviews/>
- Reeves and Sons Limited. (2021). Apa itu E-dagang. 86-90 Paul Street, London, EC2A 4NE, UK. ecommerce-platforms.com
- Aqlan, Ameen & Bairam, Dr. Manjula & Naik, R Lakshman. (2019). A Study of Sentiment Analysis: Concepts, Techniques, and Challenges. 10.1007/978-981-13-6459-4_16.

- Ligthart, A., Catal, C. & Tekinerdogan, B. Systematic reviews in sentiment analysis: a tertiary study. *Artif Intell Rev* 54, 4997–5053 (2021).
<https://doi.org/10.1007/s10462-021-09973-3>
- Yang, Li & Li, Ying & Wang, Jin & Sherratt, R.. (2020). Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning. *IEEE Access*. 8. 1-1. 10.1109/ACCESS.2020.2969854.
- Puspita Kencana Sari. (2018). Measuring e-Commerce service quality from online customer review using sentiment analysis.
- MonkeyLearn Inc. (2022). Sentiment Analysis: A Definitive Guide
<https://monkeylearn.com/sentiment-analysis/>
- Aditya Beri. (2020, May 28). Sentiment Analysis Using Vader.
<https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664>
- Hardikkumar Dhaduk. (2021, July 13). Performing Sentiment Analysis With Naive Bayes Classifier!
<https://www.analyticsvidhya.com/blog/2021/07/performing-sentiment-analysis-with-naive-bayes-classifier/>
- Manish Sharma. (2020, May11). Sentiment Analysis: An Introduction to Naive Bayes Algorithm.
<https://towardsdatascience.com/sentiment-analysis-introduction-to-naive-bayes-algorithm-96831d77ac91>
- C. Troussas, M. Virvou, K. J. Espinosa, K. Llaguno and J. Caro, "Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning," *IISA 2013*, 2013, pp. 1-6, doi: 10.1109/IISA.2013.6623713.
- M. Wongkar and A. Angdresey, "Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter," *2019 Fourth International Conference on Informatics and Computing (ICIC)*, 2019, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985884.
- Tan S., Cheng X., Wang Y., Xu H. (2009) Adapting Naive Bayes to Domain Adaptation for Sentiment Analysis. In: Boughanem M., Berrut C., Mothe J., Soule-Dupuy C. (eds) *Advances in Information Retrieval. ECIR 2009. Lecture Notes in Computer Science*, vol 5478. Springer, Berlin, Heidelberg.
https://doi.org/10.1007/978-3-642-00958-7_31
- H. Parveen and S. Pandey, "Sentiment analysis on Twitter Data-set using Naive Bayes algorithm," *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, 2016, pp. 416-419, doi: 10.1109/ICATCCT.2016.7912034.
- Jason Brownlee. (2016, April 11). Naive Bayes for Machine Learning.
<https://machinelearningmastery.com/naive-bayes-for-machine-learning/>

- Shubham.jain Jain (2018, Februari 11). Natural Language Processing for Beginners: Using TextBlob.<https://www.analyticsvidhya.com/blog/2018/02/natural-language-processing-for-beginners-using-textblob/>
- Aryan Bajaj. (2021, Jun 17). Can Python understand human feelings through words? – A brief intro to NLP and VADER Sentiment Analysis.
<https://www.analyticsvidhya.com/blog/2021/06/vader-for-sentiment-analysis/>
- Zoumana Keita. (2021, Mac 2). Social Media Sentiment Analysis In Python With VADER No Training Required! <https://towardsdatascience.com/social-media-sentiment-analysis-in-python-with-vader-no-training-required-4bc6a21e87b8>
- Kanakaraddi, Suvarna & Chikaraddi, Ashok & Aivalli, Nandeesh & Maniyar, Jawaad & Singh, Neelendra. (2022). Sentiment Analysis of Covid-19 Tweets Using Machine Learning and Natural Language Processing. 10.1007/978-981-16-7330-6_28.
- Kaur, Sukhnandan & Mohana, Rajni. (2015). A roadmap of sentiment analysis and its research directions. International Journal of Knowledge and Learning. 10. 296. 10.1504/IJKL.2015.073485.