

ANALISIS SENTIMEN TERHADAP ULASAN MAKANAN

AKMAL NADZIRAH ASWAMI FADILLAH

NOOR FARIDATUL AINUN ZAINAL

Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

Dengan kebangkitan era rangkaian sosial, Internet telah menjadi platform yang menjanjikan pembelajaran dalam talian, bertukar idea dan berkongsi pendapat. Media sosial mengandungi jumlah data yang besar dalam bentuk tweet, blog, status dan dataset. Kebanyakan manusia suka meluahkan pendapat mereka di laman sosial setiap hari. Dalam sekelip mata, kita juga boleh mengetahui latar belakang seseorang berdasarkan data yang diperolehi. Analisis sentimen boleh menentukan sama ada seseorang itu menyokong (suka) atau menentangnya (tidak suka). Tujuan analisis sentimen terhadap ulasan makanan dibuat untuk membantu restoran kecil yang tidak beroperasi dengan baik dan juga menyedarkan pemilik restoran tentang kepentingan ulasan restoran. Model yang akan digunakan untuk mendapat ketepatan yang berhampiran dengan 1 adalah Gaussian NB, Multinomial NB, Random Forest dan Logistic Regression. Setelah memperolehi ketepatan, model yang mempunyai ketepatan yang berhampiran dengan 1 akan diuji untuk mengklasifikasikan teks. Ia akan mengenal pasti dan mengkategorikan teks sama ada positif atau negatif. Jumlah data yang akan digunakan adalah 1000 data. Oleh kerana keputusan analisis akan berdasarkan ketepatan juga, di mana 4 kaedah akan digunakan untuk meningkatkan ketepatan. Akhir sekali, 3 kaedah untuk meningkatkan ketepatan akan diaplikasikan kepada 4 model bagi memperoleh klasifikasi yang baik.

1. PENGENALAN

Analisis Sentimen ialah proses pengiraan menentukan sama ada sesuatu tulisan itu adalah positif atau negatif. Ini juga dikenali sebagai perlombongan pendapat iaitu memperoleh pendapat atau sikap penutur. Terdapat tiga cara analisis sentimen boleh digunakan dalam situasi kehidupan iaitu dalam perniagaan, politik dan tindakan awam. Pertama, dalam perniagaan, syarikat pemasaran boleh menggunakan analisis sentimen bagi membangun strategi mereka dalam memahami minat manusia terhadap sesuatu produk atau jenama. Selain itu, analisis sentimen boleh melihat cara manusia bertindak balas terhadap kempen atau pelancaran produk bagi mengetahui punca sesetengah manusia tidak membeli produk tersebut. Dari sini strategi yang dilaksanakan boleh disusun semula dan menambah baik produk mereka agar sesuai dengan kesukaan manusia. Kedua, dalam politik, kita tidak perlu mengadakan hari mengundi kerana analisis sentimen boleh digunakan untuk meramalkan keputusan pilihan raya. Akhir sekali, dalam tindakan awam, ini boleh digunakan untuk memantau dan menganalisis fenomena sosial untuk mengesan situasi yang berpotensi berbahaya. Dari sini, kita boleh mengelak keadaan berbahaya daripada menjadi teruk.

2 PENYATAAN MASALAH

Berdasarkan kajian literatur di mana sebilangan besar tindakan pelanggan berpunca daripada emosi mereka dan pengalaman yang mereka terima. Perkhidmatan pelanggan yang positif boleh membangkitkan emosi positif, yang membawa kepada tindakan positif. Sebaliknya, perkhidmatan yang buruk boleh mencetuskan emosi dan tindakan negatif (Pauly 2022). Maklum balas pelanggan yang baik dilihat sebagai penyumbang utama untuk peningkatan kualiti, kadar kejayaan produk baharu dan keuntungan korporat. Tetapi banyak syarikat masih bergelut atau gagal dalam proses menjadikan data kepada maklumat kerana mereka telah mengumpul data tetapi tidak menjadikannya sebagai sumber peningkatan produk.

Maklum balas pelanggan adalah penting untuk membimbing dan memaklumkan keputusan serta mempengaruhi inovasi dan perubahan pada produk atau perkhidmatan. Ia juga penting untuk mengukur kepuasan pelanggan di kalangan pelanggan semasa. Pemilik restoran memandang ringan akan ulasan makanan dan tidak mengambil perhatian sepenuhnya bahawa dengan maklum balas pelanggan, perniagaan boleh mengubah 360 darjah daripada kegagalan kepada kejayaan atau sebaliknya. Ulasan dalam talian untuk perniagaan ialah salah satu elemen terpenting dalam kes analisis pemasaran. Dalam perniagaan, ulasan pelanggan dalam talian menjadi sangat penting

untuk produk dan perkhidmatan, oleh itu, kami boleh mengesan ulasan buruk dan baik dengan bantuan yang mana kami boleh menganalisis kualiti produk dan piawaiannya, juga berguna dalam membuat kaedah dan teknik baharu untuk menambah baik. kualiti produk (Ahmed et al. 2021).

3 OBJEKTIF KAJIAN

Terdapat dua objektif kajian iaitu:

- i. Membuat pengelasan ulasan restoran sama ada positif atau negatif.
- ii. Membuat analisis data yang bertepatan paling hampir kepada nilai 1.0 dengan menggunakan empat jenis algoritma pengelas.

4 KAEDAH KAJIAN

Kajian ini menggunakan Model Air Terjun yang juga disebut sebagai model kitaran hayat berjujukan linear. Ia sangat mudah untuk difahami dan digunakan. Dalam model air terjun, setiap fasa mesti diselesaikan sebelum fasa seterusnya boleh dimulakan dan tidak ada pertindihan dalam fasa.

4.1 Fasa Perancangan

Fasa ini merupakan fasa terpenting dalam pembangunan sistem. Pelan projek didokumenkan, penghantaran dan keperluan projek ditakrifkan dan jadual projek dibuat. Ia melibatkan mewujudkan satu set rancangan untuk membantu membimbing melalui fasa pelaksanaan dan penutupan projek analisis sentimen terhadap ulasan makanan. Dalam fasa ini di mana ia menentukan sebab projek analisis sentimen terhadap ulasan makanan dibangunkan dan dapat mengetahui secara terperinci tentang projek ini. Juga, ia menyatakan matlamat untuk dicapai pada akhir fasa nanti.

4.2 Fasa Analisis

Fasa ini mentakrifkan keperluan pembangunan, bebas daripada cara keperluan ini akan dicapai. Fasa ini mentakrifkan masalah yang cuba diselesaikan. Dalam fasa ini di mana dapat mencari perisian dan perkakasan yang sesuai yang akan digunakan untuk membangunkan projek analisis sentimen tanpa sebarang masalah dengan menyatakan spesifikasi yang betul.

4.3 Fasa Reka Bentuk

Fasa ini menentukan seni bina pembangunan yang akan digunakan. Fasa ini juga membantu untuk menggambarkan aliran dan bagaimana pembangunan analisis sentimen terhadap ulasan makanan akan kelihatan. Dalam fasa ini di mana ia akan menentukan kod aliran projek analisis sentimen. Terdapat lima langkah iaitu Ulasan Produk, Pengenalpastian Sentimen, Pengekstrakan Ciri, Klasifikasi Sentimen dan terakhir Kekutuban Sentimen. Selain itu, dalam fasa ini juga di mana dapat mengenal pasti *python library* yang diperlukan dan sesuai untuk digunakan dalam projek ini.

4.4 Fasa Implementasi

Fasa ini membincangkan tentang aspek pembangunan dan implementasi pembangunan yang dibangunkan. Semua kajian kecil akan digabungkan untuk menjadi pembangunan lengkap yang akan mencapai objektif projek. Dalam fasa ini di mana proses analisis sentimen dibahagikan kepada 3 langkah iaitu prapemprosesan, analisis dan keputusan. Prapemprosesan akan dijalankan dahulu diikuti oleh yang lain. Ini membantu untuk memahami kod dengan lebih baik dan dapat mengenal pasti ralat dengan mudah jika ia berlaku. Selain daripada itu, dalam fasa ini juga di mana akan mengenal pasti dan menggunakan 3 kaedah untuk meningkatkan ketepatan selepas berjaya membangunkan analisis sentimen. Kaedah itu penting supaya mesin dapat mengenal pasti teks positif dan negatif dengan lebih tepat. Kaedah tersebut adalah tambahkan lebih banyak sampel data, eksperimen dengan algoritma yang berbeza dan tambahkan beberapa konteks pada data.

5 HASIL KAJIAN

Pembangunan ini dibangunkan dengan menggunakan perisian Spyder dan bahasa Python dengan sistem operasi terkini dan perkakasan spesifikasi tinggi.

5.1 Perolehan Keputusan Dari Kaedah – Kaedah Yang Digunakan

Terdapat empat kaedah yang digunakan dan hasil setiap kaedah adalah seperti yang dibincangkan dalam sub topik ini.

i. Kaedah 1: Tambahkan lebih banyak sampel data

Pada asalnya pembangunan menggunakan 1000 data sahaja. Tetapi salah satu penemuan fakta mendapati bahawa dengan menambah lebih banyak data di mana ia dapat meningkatkan ketepatan. Mempunyai lebih banyak data adalah idea yang baik. Ia membenarkan "data memberitahu sendiri," dan bukannya bergantung pada andaian dan korelasi yang lemah. Kehadiran lebih banyak data menghasilkan model yang lebih baik dan tepat (Ray 2020). 1000 data lagi telah ditambah kepada pembangunan dan menjadi 2000 data. Ia telah meningkatkan ketepatan. 1000 data tersebut telah di dimasukkan dari satu fail lain ke fail yang digunakan untuk pembangunan. Rajah 1 menunjukkan data dalam fail Excel yang belum ditambah iaitu mengandungi 1000 data. Manakala Rajah 2 pula menunjukkan fail Excel yang telah ditambah 1000 lagi data baru dan data sebelumnya untuk menjalankan kaedah 1 ini.

	A	B
987	The problem I have is that they charge \$11.99 for a sandwich that is no bigger than a Subway sub (which offers better	0
988	Shrimp- When I unwrapped it (I live only 1/2 a mile from Brushfire) it was literally ice cold.	0
989	It lacked flavor, seemed undercooked, and dry.	0
990	It really is impressive that the place hasn't closed down.	0
991	I would avoid this place if you are staying in the Mirage.	0
992	The refried beans that came with my meal were dried out and crusty and the food was bland.	0
993	Spend your money and time some place else.	0
994	A lady at the table next to us found a live green caterpillar In her salad.	0
995	the presentation of the food was awful.	0
996	I can't tell you how disappointed I was.	0
997	I think food should have flavor and texture and both were lacking.	0
998	Appetite instantly gone.	0
999	Overall I was not impressed and would not go back.	0
1000	The whole experience was underwhelming, and I think we'll just go to Ninja Sushi next time.	0
1001	Then, as if I hadn't wasted enough of my life there, they poured salt in the wound by drawing out the time it took to bring	0
1002		
1003		

Rajah 1 Sebelum data ditambah

	A	B	C
1985	Best resturant of Dhaka in terms of food taste, quantity, price, customer service and hospitality. They are best of the best. Keep it up Hakka Dh	1	
1986	Each and every steak they serve is just perfect and no doubt the best restaurant for steak in Dhaka	1	
1987	From my wandering of tasting a good, decent steak, I found my end to this journey in Bangladesh. They serve you the steak full of juicy, tende	1	
1988	Went there last Friday with my friend, got surprised a bit seeing such a small place. Upon arrival they waiter attended us with details of their m	1	
1989	Amazing steak! The price is on the higher end, but the taste of the steaks did not disappooint. Would have preferred that the place was a bit big	1	
1990	Excellent taste of Steak comes up with Woodhouse! A place where you can go again n again.	1	
1991	Food is really good but don't like the place It's just too congested.	1	
1992	Good place , their process of aging the steak, makes it delicious.	1	
1993	Undeniably one of the best steakhouses around(minus the small space)....the meat was perfectly cooked & was flavorful.	1	
1994	Food is good, both in quality and quantity. A little bit overpriced though. Gave 0 star less because the interior needs to be changed.The place i	1	
1995	One of the best steak in Dhaka.... Place is small but cozy... Service is good.... Steak lovers should definitely enjoy...	1	
1996	Best steak... really good and delicious... Juicy and tender... Loved it...	1	
1997	The only redeeming quality of Foodbuzz is the friendly and polite service. Unfortunately the first experince for the wife and I was terrible. I pers	0	
1998	Unfortunately I didn't like the behaviour of manager. I'm leaving my number for the owner 00711-717171. Please call me so that I can give you	0	
1999	Horrible experince I had there. No more foodbuzz.	0	
2000	Yummy food... with the nice behavior...loved the chicken wings	1	
2001	Awesome foods... I enjoyed a lot.... try sub-sandwich....	1	

Rajah 2 Selepas Data ditambah

ii. Kaedah 2: Eksperimen dengan algoritma yang berbeza

Pada asalnya pembangunan menggunakan Gaussian Naive Bayes sahaja. Tetapi salah satu penemuan fakta mendapati bahawa dengan menggunakan algoritma yang berbeza, boleh dikenal pasti yang mana satu algoritma yang sesuai untuk digunakan pada projek ini. Sesetengah algoritma lebih sesuai untuk jenis set data tertentu daripada yang lain (Ray 2020). Empat algoritma telah digunakan untuk menganalisis set data semakan restoran adalah *GaussianNB*, *Multinomial NB*, *Random Forest* dan *Logistic Regression*. Ini kerana algoritma tersebut adalah sesuai digunakan untuk pengelasan. Setelah menjalankan algoritma tersebut, *Multinomial NB* didapati memperoleh ketepatan tertinggi antara yang lain seperti yang ditunjukkan dalam Jadual 1.

Jadual 1 Ketepatan analisis ulasan makanan berdasarkan algoritma yang berbeza

Algoritma	Ketepatan
GaussianNB	73%
Multinomial NB	78%
Random Forest	70%
Logistic Regression	72%

iii. Kaedah 3: Tambahkan beberapa konteks pada data

Konteks adalah penting dalam apa jua keadaan dan melatih model pembelajaran mesin tidak berbeza. Kadangkala, satu titik data tidak boleh menceritakan kisah, jadi perlu menambah lebih banyak konteks untuk mana-mana algoritma yang ingin gunakan pada data ini untuk mempunyai prestasi yang baik (Metwalli 2021). Setelah menjalankan kaedah 3, didapati ketepatan tidak meningkat. Seseengah data telah diperincikan supaya memberi banyak maklumat kepada mesin untuk memahami ayat lebih mendalam tetapi ketetapan masih sama dan tiada perubahan.

iv. Kaedah 4: Gabungan Kaedah 1, Kaedah 2 dan Kaedah 3

Setelah menggabungkan ketiga-tiga kaedah tersebut iaitu gabungan kaedah1, kaedah2 dan kaedah3 didapati ketepatan telah meningkat dan Logistic Regression mendapati ketepatan paling tinggi antara yang lain. Jadual 2 menunjukkan ketepatan ulasan makanan sebelum menggunakan gabungan ketiga-tiga kaedah. Jadual 3 pula menunjukkan ketepatan ulasan makanan selepas menggabungkan ketiga-tiga kaedah.

Jadual 2 Ketepatan ulasan makanan sebelum menggabungkan ketiga-tiga kaedah.

Algoritma	Ketepatan
GaussianNB	73%
Multinomial NB	78%
Random Forest	70%
Logistic Regression	72%

Jadual 3 Ketepatan ulasan makanan selepas menggabungkan ketiga-tiga kaedah.

Algoritma	Ketepatan
GaussianNB	60%
Multinomial NB	84.5%
Random Forest	85.5%
Logistic Regression	88%

Logistic Regression dengan ketepatan yang paling tinggi antara yang lain telah memberikan analisis yang lebih baik untuk membezakan perkataan positif dan negatif. Ini menunjukkan bahawa dengan ketepatan 0.88 iaitu hampir 1.0 memberikan lebih tepat dalam mengklasifikasikan teks. Rajah 3 menunjukkan data yang telah dimasukkan ke dalam sistem dan sistem menunjukkan output yang bertepatan dengan data yang dimasukkan iaitu samada ulasan yang dimasukkan adalah positif atau negatif.

Rajah 3 Contoh ulasan yang dimasukkan dan output yang dihasilkan oleh sistem.

```
Enter your text : The cashier had no care what so ever on what I had to say it still ended up being wayyy overpriced.  
This is a NEGATIVE review.  
Do You Want To Continue? (y/n) y  
Enter your text : Also there are combos like a burger, fries, and beer for 23 which is a decent deal.  
This is a POSITIVE review.
```

6 KESIMPULAN

Lazimnya, tujuan analisis sentimen adalah untuk mengetahui sikap pelanggan terhadap sesuatu produk atau perkhidmatan. Maklum balas pelanggan adalah penting untuk perniagaan dan media sosial sebagai platform yang berkuasa, boleh digunakan untuk menambah baik dan meningkatkan peluang perniagaan jika maklum balas pada media sosial dapat dianalisis tepat pada masanya (Zahoor et al. 2020). Secara keseluruhannya, Analisis sentimen terhadap ulasan makanan berjaya dibangunkan walaupun terdapat sedikit masalah paparan output. Pembangunan analisis sentimen dengan ketepatan yang berhampiran 1.0 dapat dicapai dan akan membantu pemilik restoran menyedarkan bahawa ulasan pelanggan adalah penting untuk meningkatkan perniagaan mereka kerana ia memberikan maklumat keseluruhan sama ada pelanggan menyukai atau tidak menyukai restoran mereka. Akhir sekali, ketiga-tiga kaedah tersebut dipilih kerana merupakan kaedah yang terpenting dalam meningkatkan ketepatan.

RUJUKAN

- Ahmed, H.M., Awan, M.J., Khan, N.S., Yasin, A. & Shehzad, F. 2021. Sentiment Analysis of Online Food Reviews using Big Data Analytics. *İlköğretim Online*, Vol. 20(2) : 827-836.
- Metwalli, S.A. 2021. 6 Ways to Improve Your ML Model Accuracy. <https://towardsdatascience.com/6-ways-to-improve-your-ml-model-accuracy-ec5c9599c436>
- Pauly, P. 2022. How to Conduct Sentiment Analysis to Improve Customer Experience?. <https://www.mediatoolkit.com/blog/how-to-conduct-sentiment-analysis-to-improve-customer-experience/>
- Ray, S. 2020. 8 Proven Ways for improving the “Accuracy” of a Machine Learning Model. *Analytics Vidhya*. <https://www.analyticsvidhya.com/blog/2015/12/improve-machine-learning-results/>
- Zahoor, K., Bawany, N. Z. & Hamid, S. 2020. Sentiment Analysis and Classification of Restaurant Reviews using Machine Learning, 21st International Arab Conference on Information Technology (ACIT), hlmn 1-6.