

PERSEPSI MASYARAKAT TERHADAP FILEM DAN DRAMA BERSIRI NETFLIX MENGGUNAKAN ANALISIS SENTIMEN

Leshuan a/l Sivakumar, Mohd Juzaidin Ab Aziz

Fakulti Teknologi dan Sains Maklumat

Universiti Kebangsaan Malaysia

Emel: a177388@siswa.ukm.edu.my, juzaidin@ukm.edu.my

ABSTRAK

Umum mengetahui bahawa, Penggunaan hiburan internet semakin meluas di Malaysia, dan salah satu perkhidmatan hiburan dalam talian yang paling banyak digunakan ialah Netflix. Terdapat banyak filem dalam pelbagai genre yang tersedia di aplikasi Netflix, tetapi sebahagian daripadanya melanggar keputusan Lembaga Penapisan Filem Malaysia dan tidak tertakluk kepada syarat tayangan yang ditentukan oleh FINAS. Oleh sebab itu, tujuan kajian ini adalah untuk mengkaji, berdasarkan tiga kategori utama agama, budaya, dan nilai moral kandungan filem dan drama bersiri yang terdapat di Netflix yang membincangkan subjek yang bertentangan dengan norma masyarakat Malaysia. Untuk mengkategorikan siri drama dan filem Netflix yang mematuhi garis panduan kandungan yang dibenarkan FINAS, pembelajaran mesin digunakan. Apabila bercakap tentang sebuah filem atau drama bersiri, ulasan kritis memainkan peranan penting dalam meluahkan pendapat atau perasaan tentang filem atau drama bersiri tertentu. Berdasarkan algoritma ini, kita dapat menentukan bagaimana ulasan Netflix yang disiarkan di media sosial seperti Twitter boleh mempengaruhi mood sesebuah filem ataupun drama bersiri, yang seterusnya akan mempengaruhi penonton Netflix di Malaysia. Hasilnya, kedua-dua sektor pemasaran hiburan dan komuniti sains data akan mendapat manfaat yang besar daripada sumbangan berharga kajian ini kepada pembangunan analisis teks.

1 PENGENALAN

Populariti platform OTT, juga dirujuk sebagai media hiburan dalam talian, semakin meningkat dari semasa ke semasa. Pada zaman moden, perkhidmatan penstriman video ini telah menyaksikan peningkatan mendadak dalam pelanggan, terutamanya sejak Perintah Kawalan Pergerakan (PKP) dilaksanakan berikutan pandemik Covid-19. Kajian metro menunjukkan bahawa, PKP yang berkuat kuasa pada Mac 2020, meningkatkan penggunaan OTT di kebanyakan kawasan, termasuk industri filem. Oleh itu, orang ramai beralih kepada perkhidmatan hiburan dalam talian seperti Amazon Prime, Astro on the Go, Netflix, iFlix dan Viu. Sejak dilancarkan beberapa tahun yang lalu, orang ramai telah memberi respons positif kepada platform ini kerana penggunaannya yang lebih selesa, kebolehcapaian yang lebih mudah, kos yang berpatutan dan akses kepada pelbagai saluran selaras dengan perkembangan rangkaian Internet di Malaysia.

Netflix ialah sebuah syarikat penerbitan berasaskan langganan dan pengedar perkhidmatan penstriman mediayang menyediakan penstriman dalam talian daripada perpustakaan filem, drama bersiri dan rancangan televisyen, menurut Wikipedia. Di California pada tahun 1997, Marc Randolph dan Reed Hastings melancarkan Netflix. Bermula pada tahun 2007, Netflix melancarkan perkhidmatan penstriman media dalam taliannya, yang menjadi global pada tahun 2010. Pada tahun 2016, Netflix membuat debutnya di Malaysia dengan pelbagai pakej pakej langganan yang memenuhi piawaian yang telah ditetapkan. Sebaik sahaja Netflix tiba di Malaysia, perkhidmatan penstriman ini diterima dengan baik oleh pelanggan. Pelanggan boleh menonton kandungan yang diberikan pada bila-bila masa mereka memilih dengan pelbagai program televisyen, drama bersiri, dokumentari dan filem.

Pelanggan juga boleh menonton Netflix kerana ia sangat mesra pengguna pada semua platform dan peranti. Kos langganan juga adil. Beberapa rancangan Netflix semasa yang sering dibatalkan hanya tersedia dalam talian kepada pelanggan. Netflix juga tidak menapis sebarang kandungan siaran. Tamadun kita di Malaysia terdiri daripada orang yang berbeza bangsa, kepercayaan, budaya dan bangsa. Rakyat Malaysia hidup mengikut

set nilai sosial, tradisi budaya dan konvensyen mereka sendiri. Standard masyarakat Malaysia mungkin tidak dipenuhi oleh semua kandungan Netflix. Dengan mematuhi piawaian masyarakat Malaysia dan garis panduan kandungan yang dibenarkan oleh Perbadanan Kemajuan Filem Nasional Malaysia (FINAS), Netflix digunakan di negara tersebut secara sah. Kebimbangan yang boleh dikaitkan kepada kandungan Netflix yang menyimpang daripada norma sosial Malaysia termasuk sensitiviti agama, sensitiviti politik, unsur-unsur tidak senonoh seperti lucah dan pedofilia, LGBTQ, keganasan melampau, unsur-unsur yang mengganggu keharmonian sosiobudaya rakyat Malaysia, dan amalan ilmu hitam.

Sebagai cara menangani masalah ini, teknik pembelajaran mesin boleh dilakukan ke atas data-data yang berkaitan Netflix di media sosial seperti Twitter. Proses analisis sentimen dapat mengenal pasti corak dalam sesuatu data dan memproses data tersebut ke dalam bentuk yang lebih mudah dibaca dan difahami. Proses ini biasanya mengenali emosi di sebalik teks. Emosi terbahagi kepada tiga jenis iaitu positif, negatif dan neutral. Apabila melakukan analisis sentimen pada data berkaitan kandungan Netflix, anda boleh mempertimbangkan model pendapat umum untuk memahami perasaan penonton tentang filem itu. Oleh kerana orang pada masa kini menggunakan platform media sosial secara meluas, kebanyakannya lebih cenderung untuk mengkritik atau memberi pendapat tentang filem atau siri TV di Twitter mahupun Facebook. Oleh itu, dalam kajian ini, sentimen teks media sosial berkaitan filem Netflix akan diklasifikasikan mengikut kandungan yang sesuai untuk penonton di Malaysia.

2 PENYATAAN MASALAH

Pada era ini, hiburan dalam talian adalah platform yang sangat popular. Walaupun saiz industri hiburan, perkhidmatan OTT seperti Netflix dan lain-lain digunakan secara meluas, terutamanya pada masa wabak Covid-19 melanda negara. Banyak kajian lepas menawarkan kesimpulan umum mengenai topik bahan Netflix yang melibatkan subjek yang bertentangan dengan norma hidup sosial Malaysia. Terdapat beberapa variasi yang ketara antara cara hidup di Malaysia dan negara-negara Barat seperti Amerika Syarikat.

Model atau kaedah yang digunakan untuk menganalisis sentimen ulasan filem dalam talian, seperti IMDB dan Rotten Tomatoes, juga semuanya sangat generik. Siaran Netflix harus mengambil kira garis panduan kandungan yang diluluskan oleh Perbadanan Kemajuan Filem Nasional Malaysia (FINAS), yang selaras dengan norma sosial Malaysia. Pada masa kini, media sosial seperti Twitter adalah medium yang sangat popular dengan masyarakat. Akibatnya, tweet sering digunakan untuk menyampaikan pelbagai idea dan kritikan. Maka, pemeriksaan pendapat pengguna terhadap bahan Netflix berdasarkan keutamaan mereka adalah perlu untuk memastikan sentimen.

3 OBJEKTIF KAJIAN

Terdapat banyak model yang boleh menentukan sentimen teks kepada nada positif, negatif atau neutral. Dalam cicapan yang bertujuan untuk mengkritik filem daripada Netflix, banyak perkataan khusus harus dipertimbangkan untuk memahami konteks teks secara keseluruhan. Oleh itu, tujuan utama kajian ini adalah untuk mengklasifikasikan kandungan cicapan filem Netflix dengan bersesuaian mengikut kandungan yang bersesuaian dengan penonton yang terdiri daripada masyarakat Malaysia. Untuk mencapai matlamat ini, objektif berikut digariskan:

- a. Untuk Mendapatkan tweets ataupun cicapan daripada Twitter melalui Twitter API
- b. Untuk mengenal pasti sentimen filem dan drama bersiri dari Netflix dengan menggunakan analisis sentimen.
- c. Untuk mengenal pasti kategori filem dan drama bersiri dari Netflix tersebut berdasarkan kategori yang telah ditetapkan oleh FINAS menerusi pendekatan Naïve Bayes.

4 METOD KAJIAN

Pembangunan model yang sesuai penting untuk memastikan kajian berjalan dengan baik dan mendapatkan hasil keputusan yang tepat dan berkualiti. Model metodologi yang digunakan untuk membangunkan algoritma ini adalah Model Klasifikasi Sentimen. Model ini mudah untuk difahami dan digunakan dalam proses analisis sentimen. Fasa pembangunan algoritma senang untuk dipantau dan risiko kegagalan semasa pembangunan algoritma adalah kurang. Rajah 1 merujuk kepada model metodologi kajian ini. Beberapa model sistem bagi projek ini dirujuk daripada kajian salah seorang pelajar Universiti Kebangsaan Malaysia iaitu Tan Fui Ken, (2019).



Rajah 1 Model Klasifikasi Sentimen

4.1 Fasa Perancangan

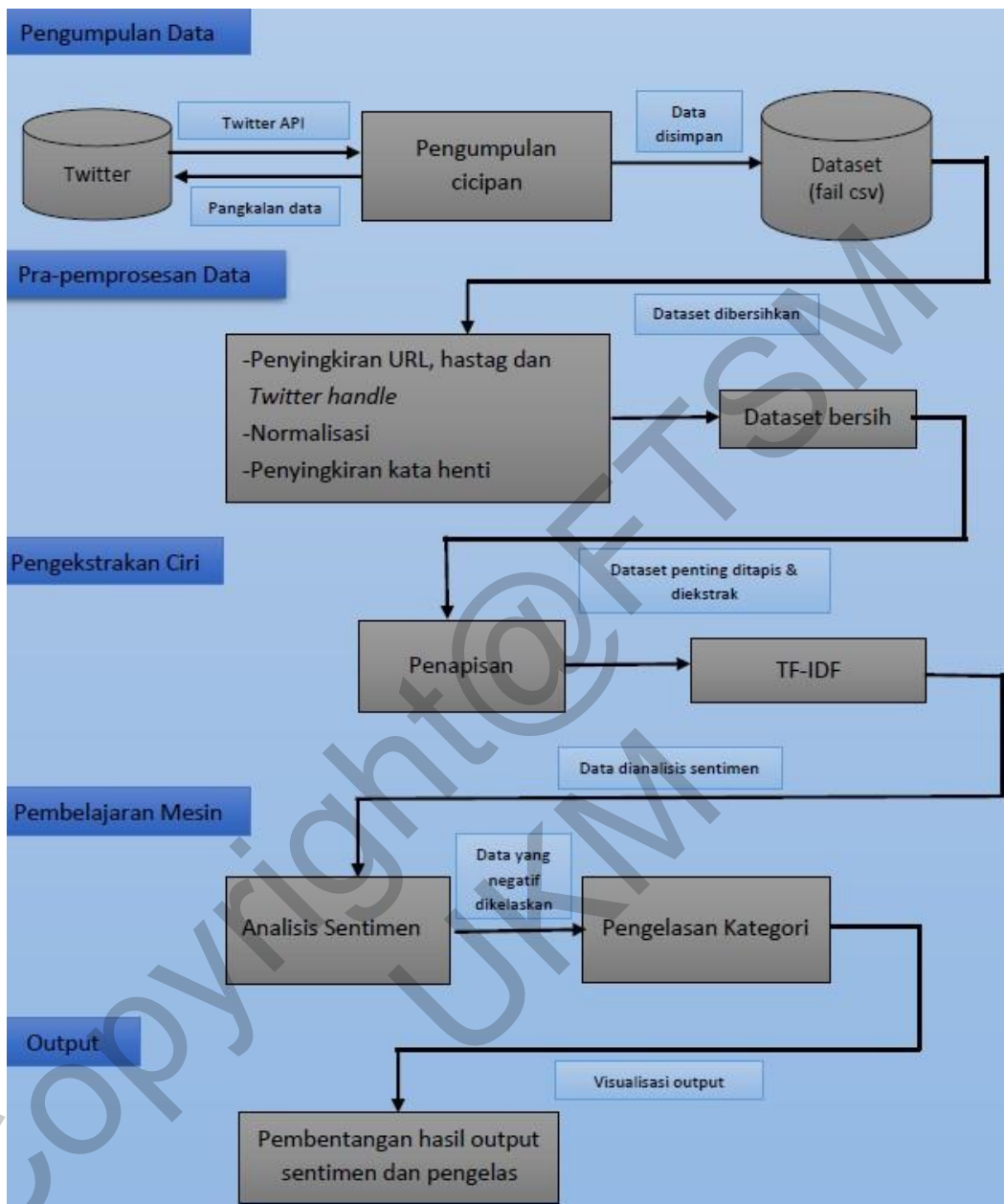
Fasa perancangan melibatkan proses pengumpulan data dan analisis keperluan kajian seperti set data yang berkaitan dengan objektif kajian dan permasalahan kajian serta menganalisis sorotan kajian susastera yang melibatkan pengumpulan maklumat, pencarian jurnal, pembacaan jurnal dan kajian lepas sebagai rujukan. Antara topik yang dikaji adalah berkaitan pengesanan sentimen ulasan filem daripada *IMDB*, *Rotten Tomatoes* dan media sosial menggunakan pembelajaran mesin sedia ada. Set data yang berkaitan dengan kajian daripada Twitter dikumpulkan.

4.2 Fasa Analisis

Fasa ini melibatkan analisis dan tafsiran terhadap data yang dikumpulkan. Hal ini bertujuan untuk memastikan data yang dikumpulkan adalah sesuai dan relevan dengan topik kajian serta diselaraskan kepada Bahasa Inggeris sahaja. Kemudian, kesemua data yang ditapis dibersihkan lagi ke dataset yang lebih bermakna melalui pra-pemrosesan data dan pengekstrakan ciri. Sebanyak 2,316 set data iaitu cicipan dari Twitter berjaya dikumpul dan dilabelkan secara manual kepada tiga kategori yang menjadi skop kajian ini iaitu nilai moral, kebudayaan dan keagamaan (Wan Amizah , Chang & Jamaluddin 2009). Ini bertujuan untuk memudahkan data untuk dianalisis. Analisis data dijalankan bagi memastikan tiada data yang hilang, tiada data yang berulang, mengetahui jenis data, statistik set data dan visualisasi taburan data bagi mendapatkan set data yang berkualiti.

4.3 Fasa Reka Bentuk

Fasa ini melibatkan reka bentuk seni bina yang secara umumnya memberi gambaran awal secara keseluruhan bagi pembangunan system kajian ini. Reka bentuk ini juga menunjukkan hubung kait antara sub-sistem secara terperinci. Merujuk kepada rajah 2 di bawah, reka bentuk kajian ini terbahagi kepada lima fasa utama iaitu pengumpulan data, pra-pemrosesan data, pengekstrakan ciri, pembelajaran mesin dan pembentangan output.



Rajah 2 Reka Bentuk Seni Bina Kajian

- i) Dalam kajian ini, fasa pertama melibatkan pengumpulan data iaitu cicipan daripada media sosial Twitter yang mengandungi komen-komen filem dan drama bersiri dari Netflix. Perpustakaan Tweepy digunakan bagi mendapatkan set data mentah daripada aplikasi Twitter menggunakan Twitter API. 'Kunci API' dan 'kunci rahsia API' dari akaun Twitter Developer digunakan untuk mencipta akses kepada token. Kemudian, API yang digunakan dalam perisian mempunyai kata kunci yang tertentu seperti 'nama filem'

atau ‘nama drama bersiri’ berserta dengan ‘Netflix ’ bagi memperincikan cicipan yang dikumpulkan adalah sesuai dengan projek ini. Sebanyak 3000-4000 cicipan (data mentah) yang terkini berjaya dikumpulkan sepanjang perancangan projek ini iaitu bermula bulan Disember 2022 sehingga April 2023. Cicipan yang dikumpulkan disimpan dalam beberapa fail *csv* mengikut tajuk filem atau drama bersiri Netflix tetapi kemudian disatukan dalam satu fail *csv* secara keseluruhan selepas ditapis secara manual. Data yang dikumpulkan juga ditapis dan diselaraskan kepada satu Bahasa sahaja iaitu Bahasa Inggeris. Jumlah data mentah yang telah dikumpulkan adalah sebanyak 3001 data.

ii) Pra-pemprosesan cicipan yang dikumpulkan bagi menghasilkan dataset yang bersih untuk dianalisis sentimen dan diklasifikasi. Antara kaedah pra pemprosesan yang digunakan dalam projek ini termasuklah menyingkirkan ruang kosong, menukar cicipan kepada huruf kecil, penapisan karakter seperti “@pengguna”, “RT@pengguna”, dan URL serta penyingkiran karakter yang tidak membawa makna seperti nombor dan tanda baca serta penyingkiran kata henti. Pakej NLTK diimport untuk proses normalisasi. Pakej ini adalah rangkaian perpustakaan dan program untuk pra pemprosesan bahasa semula jadi simbolik dan statistik (NLP) untuk perisian Python dalam bahasa Inggeris. Pakej NLTK ini secara kebiasaannya digunakan untuk menyokong fungsi klasifikasi, tokenisasi, *stemming*, *tagging*, penghuraian, dan penaakulan semantik. Rajah 3 menunjukkan output data sebelum dan selepas diproses.

	Text	Processed
0	@SodoMe1 @GOJOT1TTIES eh, it's netflix. same people who thought cuties was a great idea	eh netflix people thought cuties great idea
1	@spideysnews @netflix Even Spider-Man won't get me to use Netflix. Not until they remove Cuties and apologize.	even spider man get use netflix remove cuties apologize
2	@IGN Damn! Sounds like Sony supports pedo. Cuties is still on Netflix. I canceled Netflix. I do not support pedo. #RemoveCuties	damn sounds like sony supports pedo cuties still netflix canceled netflix support pedo
3	@spideysnews @netflix Bleep you, @netflix I canceled Netflix due to Cuties! I refuse to return till you remove Cutá https://t.co/TH7RcsvRRV	bleep you canceled netflix due cuties refuse return till remove cut
4	@fridaybobb Does not make it true. Netflix aired Cuties sooo?	make true netflix aired cuties sooo
5	@fridaybobb I don't think they care tbh. Netflix aired "Cuties" what does that tell you about Netflix?	think care tbh netflix aired cuties tell netflix
6	@MiriaAkagi5 @Memelin_sequel I don't know, Miria, you look like the type of person who hates lolicon, but loves Netflix's Cuties.	know miria look like type person hates lolicon loves netflix cuties
7	Because Netflix is the same company made Cuties. https://t.co/CIAMQfU2hD	netflix company made cuties
8	@GameSpot ugh so the only way I will see spiderman is rental or theaters I WILL NOT WATCH ANYTHING on netflix untilá https://t.co/nw67W1DhaF	ugh way see spiderman rental theaters willnot watch anything netflix until
9	if you enjoyed @netflix movie #cuties" it means you're a chomo. #StopChildAbuse #stopsexualtrafficing #ProtectTheKids	enjoyed movie means chomo
10	@TheePill @VitoGesualdi You're arguing with the guy that defended Netflix Cuties for days on end. He does it to beá https://t.co/qcpZdu4NcB	arguing guy defended netflix cuties days end be
11	hold on all these baseball tweets I thought we were boycotting MLB like we all gave up our Netflix accounts for cuties?	hold baseball tweets thought boycotting mlb like gave netflix accounts cuties

Rajah 3 Data Sebelum dan Selepas Diproses

iii) Analisis sentimen dilakukan dengan menggunakan Textblob. Fasa ini penting bagi mengenalpasti data yang mendapat nilai sentimen yang positif, negatif dan neutral. Textblob akan mengabaikan kata-kata yang tidak diketahui, ia akan mempertimbangkan perkataan dan frasa yang dapat memberikan polariti dan purata untuk memperoleh skor akhir. Pengiraan ciri sentimen cicipan dibuat dalam bentuk polariti dan subjektiviti. Kemudian, perlu mengklasifikasikan polariti sebagai positif, neutral atau negatif. Cicipan dilabel mengikut keputusan sentimen yang diperolehi. Bagi kajian ini, hanya data yang mendapat nilai sentimen yang negatif sahaja akan dianalisis dengan lebih terperinci. Data yang mendapat nilai sentimen yang negatif menunjukkan ketidaksesuaian kandungan filem atau drama bersiri tersebut serta peranan negatif netizen terhadap kandungan Netflix yang dibincangkan. Oleh itu, analisis sentimen ini penting bagi mendapatkan data yang bermakna bagi memenuhi objektif kajian ini.

iv) Fasa pengekstrakan ciri dilakukan bagi mendapatkan ciri penting yang akan digunakan sebagai perwakilan data. Ciri utama yang diekstrak ialah *Term Frequency-Inverse Document Frequency (TF-IDF)*. TF-IDF adalah ukuran statistik yang menilai kadar relevan sesuatu perkataan dalam sesebuah dokumen dengan kumpulan dokumen tersebut. Ini dilakukan dengan mendarabkan dua metrik iaitu, frekuensi istilah dan frekuensi dokumen songsang. Frekuensi Istilah (TF) merujuk kepada berapa kali sesuatu perkataan muncul dalam dokumen dan Frekuensi Dokumen Songsang (IDF) merujuk kepada kekerapan songsang perkataan tersebut dalam sekumpulan dokumen. TF dan IDF tidak perlu dikira secara manual, sebaliknya *sklearn* mempunyai fungsi yang boleh mendapatkannya secara langsung iaitu menggunakan *TfidfVectorizer* dalam bahasa pengaturcaraan Python. Vektor akan mengira berat setiap perkataan dalam korpus dan akan mengembalikan TF-IDF.

iv) Fasa pengelasan data dilakukan dengan menggunakan teknik *Multinomial Naïve Bayes*. Teknik ini dipilih kerana secara kebiasannya digunakan untuk data set media sosial dan sangat mudah dan pantas untuk meramalkan kelas set data ujian. Ia juga menunjukkan prestasi yang baik dalam ramalan *multi-class* sepertimana yang digunakan dalam kajian ini juga. *Naïve Bayes* menggunakan kebarangkalian bersyarat. Algoritma yang digunakan

dalam pengelasan *Naïve Bayes* dalam kajian ini menggunakan semua ciri vector TF-IDF dan pengelas menganalisis ciri secara bebas. Data dibahagi kepada data set latihan sebanyak 80% dan data set pengujian sebanyak 20%. Data akan dikategorikan dalam tiga kelas utama iaitu keagamaan, nilai moral dan kebudayaan.

4.4 Fasa Pengujian

Fasa ini melibatkan pengujian teknik analisis sentimen dan model pengelas yang dipilih bagi kajian ini. Terdapat banyak teknik analisis sentiment yang boleh digunakan bagi menganalisis sentimen di sebalik cicipan yang dikumpulkan. Oleh itu, kajian ini telah menjalankan analisis awal (*preliminary analysis*) bagi membandingkan dan memilih teknik analisis sentimen yang bersesuaian, iaitu sama ada *VADER Sentiment Analysis* atau *TextBlob Sentiment*. Analisis awal ini dijalankan dengan menggunakan 200 cicipan daripada data mentah. 200 cicipan ini dilabel terlebih dahulu mengikut pemerhatian pengkaji sama ada positif, negatif atau neutral. Kemudian, kesemua cicipan tersebut menjalani kaedah pra pemprosesan seperti yang diterangkan di fasa reka bentuk. Seterusnya, cicipan yang sudah diproses dimasukkan pula ke dalam *VADER* dan *TextBlob* bagi mendapatkan output sentiment mengikut teknik masing-masing. Output dari kedua-dua teknik ini dibandingkan dengan label sentimen yang sudah dibuat secara manual oleh pengkaji terlebih dahulu tadi. Hasil analisis awal mendapati *TextBlob Sentiment Analysis* adalah lebih tepat, maka teknik ini yang digunakan bagi pembangunan algoritma yang selanjutnya. Pengujian yang seterusnya bagi projek ini melibatkan pengklasifikasian cicipan yang bersentimen negatif kepada tiga kategori utama yang sudah ditetapkan iaitu nilai moral, keagamaan dan kebudayaan menggunakan vektor TFIDF. Bagi penilaian pengelas, ketepatan pengelas *Naïve Bayes* bagi latihan dan pengujian adalah seperti dalam rajah 4 di bawah.

```
nb = BernoulliNB()
nb.fit(X,y)
train_acc = nb.score(X, y)
test_acc = nb.score(X_test, y_preds)
print("Training Accuracy for Naïve Bayes:",train_acc)
print("Testing Accuracy for Naïve Bayes:",test_acc)
```

```
Training Accuracy for Naïve Bayes: 0.9726651480637813
Testing Accuracy for Naïve Bayes: 1.0
```

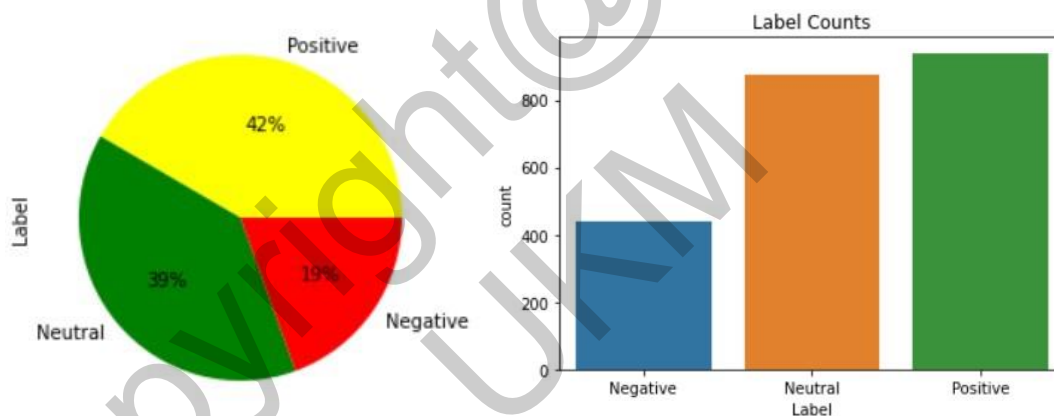
Rajah 4 Ketepatan Pengelas *Naïve Bayes*

5 HASIL KAJIAN

Untuk memudahkan pengguna menilai ulasan pengguna Twitter tentang siri drama dan filem Netflix yang tidak sesuai untuk tayangan di Malaysia mengikut peraturan yang ditetapkan oleh FINAS, hasilnya divisualisasikan sebagai graf bar dan carta pai. Keputusan analisis sentimen menggunakan teknik *TextBlob* dikeluarkan dalam bentuk carta pai dalam Rajah 5 di mana carta tersebut menggambarkan peratusan cicipan yang bersentimen positif, negatif dan neutral. Selamat sejahtera Dr, Di sini saya sertakan bersama D7 saya untuk rujukan Dr.

Terima kasih Dr.

. Rajah 5 juga merujuk kepada graf bar yang menunjukkan jumlah cicipan yang bersentimen positif, negatif dan neutral. Didapati jumlah cicipan yang bersentimen positif adalah 910, jumlah cicipan yang bersentimen negatif adalah 454 dan jumlah cicipan yang bersentimen neutral adalah 857.



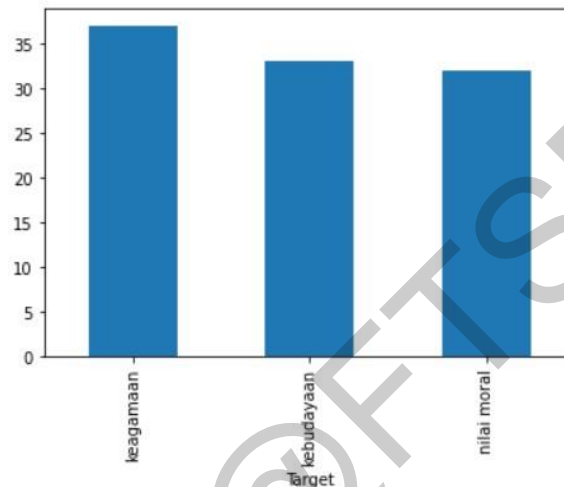
Rajah 5 Output Analisis Sentimen

Processed	Polarity	Subjectivity	Label
nooooo i love big mouth especially lola she is the reason i watch the show lol	0.325	0.6	Positive
big mouth is one show that always makes my mood better	0.25	0.3	Positive
pedophilia is wrong cuties was a terrible film	-0.75	0.95	Negative
i just started watching the keepers on netflix and i just wanted to say i am disgusted men cant be trusted	-1	1	Negative
our feisty indian aunty watches and believes that dna testing has nothing on indian matchmaking	0	0	Neutral
elite season is out on netflix overrated show	0	0	Neutral

Rajah 6 Hasil Output *TextBlob*

Hasil keputusan pengelas *Naive Bayes* dikeluarkan dalam bentuk graf bar bagi pandangan keseluruhan output model pengelas iaitu mengklasifikasikan cicipan yang bersentimen

negatif kepada 3 kategori yang tidak sesuai ditayangkan di Malaysia selaras dengan garis panduan yang ditetapkan oleh FINAS.



Rajah 7 Output Pengelas *Naive Bayes*

Berikut adalah contoh cicipan yang yang telah dikenalpasti sebagai kategori keagamaan.

Tweet	Target
i watched messiah on netflix yesterday i not religious but it was not nice	keagamaan
messiah was a clever thought provoking series about contemporary faith deceit mystery very disappoi	keagamaan
jesus chilling out in the rock tomb watching messiah on netflix	keagamaan
hating messiah season hate how jesus is portrayed as problematic but also portrayed inaccurately as u	keagamaan
please let us dispose this christian leftist idea of making castes taboo bollywood and netflix are	keagamaan
i watching lucifer really sad devil guy	keagamaan

Rajah 8 Contoh Cicipan Kategori Keagamaan

Berikut adalah contoh cicipan yang yang telah dikenalpasti sebagai kategori kebudayaan.

Tweet	Target
if you want bad bdsm watch bonding on netflix	kebudayaan
bonding season one on netflix is sadly a flaccid look at the world of bdsm skip it	kebudayaan
randomly started watching indian matchmaking on netflix and it actually way more creepy than what i thought	kebudayaan
netflix needs to stop making all the gays act likes whores its so frustrating	kebudayaan
stopped watching elite when the girl took off her hijab for a white man i hate netflix	kebudayaan
elite its a spanish show on netflix but i watch it illegally a warning the show has a lot of sex so	kebudayaan

Rajah 9 Contoh Cicipan Kategori Kebudayaan

Berikut adalah contoh cicipan yang telah dikenalpasti sebagai kategori nilai moral.

Tweet	Target
isnt cuties a film about little girl strippers yall are so disgusting	nilai moral
people still watch netflix i canceled mine soon as they put that disgusting cuties	nilai moral
i quit netflix when they started catering to pedo with that cuties series shit	nilai moral
i am being accused for endorsing child porn apparently i think cuties is bad and sh	nilai moral
sorry to burst your misguided bubble my guy but he didn say anyone who still use	nilai moral
and unfortunately one of them shows is the pedo attraction movie cuties	nilai moral

Rajah 10 Contoh Cicipan Kategori Nilai Moral

6 KESIMPULAN

Kesimpulannya, kajian ini diharap dapat mewujudkan satu sistem yang boleh mengekstrak dan mengumpul data dari Twitter tentang ulasan atau komen filem atau drama bersiri Netflix dan menganalisis sentimen data tersebut di samping mengelaskan kandungan yang didapati negatif kepada 3 kategori yang dikelaskan terlebih dahulu. Data daripada Twitter API dikumpulkan dalam sebagai fail csv akan dibersihkan dan dilabelkan terlebih dahulu dalam pra-pemprosesan dan pengekstrakan ciri. Pembelajaran mesin adalah Teknik yang digunakan dalam kajian ini. Oleh kerana *TextBlob Sentiment Analysis* sesuai digunakan untuk dataset daripada Twitter untuk mengenalpasti data yang bernada positif dan negatif, ianya digunakan bagi proses analisis sentimen. Kemudian, data bernada negative dikelaskan kepada tiga kategori yang berkaitan menggunakan *Naive Bayes Classifier*.

Hasil kajian ini harus memenuhi objektifnya dan boleh divisualasikan secara berstruktur, seperti graf. Sebagai output pertama terdapat output analisis sentimen untuk membezakan antara data yang bernada positif dan negatif. Seterusnya terdapat output pengelas yang digunakan untuk menentukan sama ada kandungan Netflix yang dianalisis sesuai untuk

ditonton oleh rakyat Malaysia berdasarkan kategori nilai moral, kebudayaan dan keagamaan.

Kaedah pengelasan dan analisis sentimen digunakan dalam kajian ini. Penciptaan sistem ini dijangka menyumbang kepada kemajuan lanjut pengetahuan dalam pembelajaran mesin, yang kini digunakan dengan lebih kerap oleh orang awam dan perusahaan besar. Kajian ini juga dijangka dapat memajukan kebolehan pengaturcaraan Python. Selain itu, ada kemungkinan penyelidikan masa depan akan meningkatkan algoritma kajian ini untuk mencipta sistem atau aplikasi yang boleh mencapai matlamat yang sama. Ini berikutan sekiranya sistem atau program seperti ini diwujudkan, ia akan mendapat sambutan yang lebih tinggi terutama dalam industri hiburan. Akibatnya, dapat mengerahkan usaha ke arah kemajuan analisis teks dalam kedua-dua sektor saintifik dan pemasaran dan hiburan

Copyright@FTSM
UKM

7 RUJUKAN

- Tay Fui Kien. 2019. *Analisis Sentimen Twitter Mengenai Peristiwa-Persitiwa penting yang Berlaku di Sekitar UKM, Bangi*. Bangi: Universiti Kebangsaan Malaysia.
- Lembaga Penapis Filem. 2010. *Garis Panduan Penapisan Filem*. W.P Putrajaya: Bahagian Kawalan Penapisan Filem dan Penguatkuasaan Kementerian Dalam Negeri.
- Wan Amizah Wan Mahmud, Chang Peng Kee & Jamaluddin Aziz. 2009. *Film Censorship in Malaysia: Sanctions of Religious, Cultural and Moral Values*. Jurnal Komunikasi, Malaysian Journal of Communication Vol, 25: 42-49.
- Lina L. Dhande & Dr. Prof. Girish K. Patnaik. 2014. *Analyzing Sentiment of Movie Review Data using Naive Bayes Neural Classifier*. International Journal of Emerging Trends & Technology in Computer Science, Vol 3, Issue 4.
- Palak Baid, Apoorva Gupta & Neelam Chaplot. 2017. *Sentiment Analysis of Movie Reviews using Machine Learning Techniques*. International Journal of Computer Applications. Vol 179, No.7.
- Kuat Yessenov & Sasa Misailovic. 2009. *Sentiment Analysis of Movie Review Comments*, 6.863 Spring 2009 final project.
- Rucha Jadhavar & Agastya Kumar Komarraju. 2018. *Sentiment Analysis of Netflix and Competitor Tweets to Classify Customer Opinions*. Paper 2708-2018.
- Lum Choi Kian. 2019. *Analisis Sentimen Dalam Bitcoin Tweets*. Bangi: Universiti Kebangsaan Malaysia.
- Ahmad Imam Fahrur Rozi, Elok Nur Hamdana & Muhammad Balya Iqbal Alfahmi. 2018. *Pengembangan Aplikasi analisis Sentimen Twitter Menggunakan Metode Naive Bayes Classifier (Studi Kasus SAMSAT Kota Malang*. Jurnal Informatika Polinema 4(2):149.
- Asma Md. Isa, Wan Amizah Wan Mahmud, Wan Idros Wan Sulaiman & Muhammad Adnan Pitchan. 2019. *Netflix and Dilemma of Content Regulation in Malaysia*. International Journal of Advanced Science and Technology, Vol. 28, No. 16.

- Jiena G. McLellan. 2017. *Sentiment Analysis of Real-Time Twitter Data*. The Beef Cattle Institute, Kansas State University.
- Alexandra Balahur. 2013. *Sentiment Analysis in Social Media Texts*. European Commission Joint Research Centre, Italy.
- Ahmad Faezal. 2020. Filem Cuties: Netflix Cuba Normalisasikan Pedophilia?. *The Patriots*, 14 September. <https://www.thepatriots.asia/filem-cuties-netflix-cubanormalisasikanpedophili a/>
- Afiezy Azaman. 2019. Penapisan kandungan Netflix: Bincang dalam-dalam agar hak pengguna tidak terjejas. *Astro Awani*, 23 November. <https://www.astroawani.com/beritamalaysia/penapisan-kandungan-netflix-bincang-dalamdalam-agar-hak-pengguna-tidakterjejas-223455>
- Norhafzan Jaafar. 2019. Kawal kandungan program siaran Netflix elak kebimbangan ibu bapa. *Berita Harian Online*, 16 November. <https://www.bharian.com.my/berita/nasional/2019/11/629027/kawal-kandungan-program-siaran-netflix-elak-kebimbangan-ibu-bapa>
- Ida Lim. 2020. MCO: Malaysians stuck at home give Netflix, Tonton biggest traffic growth rates against competitors. *Malay Mail*, 1 April. <https://www.malaymail.com/news/malaysia/2020/04/01/mco-malaysians-stuck-at-home-give-netflix-tonton-biggest-traffic-growth-rat/1852432>
- Rahimi Yunus. 2020. Increasing streaming subscribers signals viewing preferences. *The Malaysian Reserve*, 3 August. <https://themalaysianreserve.com/2020/08/03/increasingstreamingsubscriberssignal-s-viewing-preferences/>
- Nur Haziqah A Malek. 2020. Malaysians turn to social media, streaming services to fill time during MCO. *The Malaysian Reserve*, 6 April. <https://themalaysianreserve.com/2020/04/06/malaysians-turn-to-socialmediastreamingservices-to-fill-time-during-mco/>
- Bernamea. 2019. Finas calls for law to regulate contents on Netflix and VOD. *Free Malaysia Today*, 16 November. <https://www.freemalaysiatoday.com/category/nation/2019/11/16/finas-calls-for-law-to-regulate-contents-on-netflix-and-vod/>

Mix. 2020. Here are the 9 titles Netflix purged due to government censorship. *The next Web*, 7 February.

<https://thenextweb.com/insights/2020/02/07/netflix-government-censorship/>

Harris Andrea. 2020. What are the Advantages and Disadvantages of Netflix?. *Tech 21 Century*. <https://www.tech21century.com/pros-and-cons-of-netflix/>

Julius. 2019. Sekatan Dalam Menonton Filem. *Kata Malaysia*, 30 Oktober.

<https://katamalaysia.my/social/sekatan-dalam-menonton-filem/>

Artem Oppermann. 25 February 2019. Sentiment Analysis with Deep Learning of Netflix Reviews. *towardsdatascience*.

<https://towardsdatascience.com/sentiment-analysis-with-deep-learning-62d4d0166ef6>

1 Oktober 2019. Pra-Pemrosesan Tweet untuk Analisis Sentimen. *Ichi.Pro*.

<https://ichi.pro/id/pra-pemrosesan-tweet-untuk- analisis-sentimen- 183953141635390>

Shubham Jain. 2018. Ultimate guide to deal with Text Data (using Python) – for Data Scientists and Engineers. *Analytics Vidhya*, 27

Februari. <https://www.analyticsvidhya.com/blog/2018/02/the-different-methods-deal-text-datapredictive-python/>

Deepak Das. 2019. Social Media Sentiment Analysis using Machine Learning : Part — I. *towardsdatascience*, 7 September.

<https://towardsdatascience.com/social-media-sentiment-analysis-49b395771197>

Avinash Navlani. 2018. Naive Bayes Classification using Scikit-learn?. *datacamp*, 4 December. <https://www.datacamp.com/community/tutorials/naive-bayes-scikit-learn>

AFAF5. (2021, Jan 6). Analyticsvidhya. Retrieved from Sentiment Analysis: VADER or TextBlob?: <https://www.analyticsvidhya.com/blog/2021/01/sentiment-analysis-vader-or-textblob/>

Baris Hasdemir. (2020, Aug 16). Retrieved from Sentiment analysis on the tweets about distance learning with TextBlob:

<https://towardsdatascience.com/sentiment-analysis-on-the-tweets-about-distance-learning-with-textblob-cc73702b48bc>

Netflix. 10 2016. <https://ms.wikipedia.org/wiki/Netflix>