PENGESANAN TUTURAN MENGHINA DAN TUTURAN BENCI BERINTEGRASI EMOJI

Tracy Chai Yee May & Wandeep Kaur a/p Ratan Singh

[1,2]*Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia, 43600 UKM Bangi,, Selangor Darul Ehsan, Malaysia*

Abstrak

Penggunaan media sosial telah meningkat dengan pesat, terutamanya pasca pandemik COVID-19. Peningkatan ini juga menyebabkan peningkatan kes tuturan benci atau kes siber buli, di mana pengguna media sosial tinggalkan komen atau berkongsi kandungan yang mempromosikan sikap tidak bertoleransi, diskriminasi atau perasaan yang tidak baik terhadap seseorang, sesuatu organisasi atau sekumpulan orang berkaitan dengan kewarganegaraan, agama, etnik, usia, jantina dan sebagainya. Oleh itu, model pengesanan tuturan menghina dan tuturan benci sangat diperlukan untuk mengenalpasti kandungan mana merupakan konteks tuturan menghina dan tuturan benci. Terdapat kajian yang dilakukan untuk mengesan tuturan menghina dan tuturan benci menggunakan sama ada kaedah leksikon berasaskan peraturan, pembelajaran mesin, pembelajaran mendalam dan sebagainya. Dalam kajian ini, set data penanda aras yang mengandungi tuturan menghina dan tuturan benci daripada Twitter telah digunakan. Kajian ini merangkumi analisis emoji dalam data tersebut untuk mengesan tuturan menghina dan tuturan benci. Emoji2Vec akan digunakan untuk mendapatkan pembenaman emoji dan ia kemudiannya digabungkan dengan data teks. Pelbagai variasi model BERT telah dieksperimenkan untuk menentukan model pembelajaran mesin yang paling berkesan dalam penyelidikan ini. DistilBERT mendapat skor FI yang tertinggi sebanyak 0.958 berbanding dengan variasi model BERT lain.

Kata Kunci: Emoji2Vec, Offensive Language, Hate Speech, BERT

Pengenalan

Rangkaian sosial telah wujud selama lebih sedekad dan ia masih berkembang. Ia adalah sebahagian daripada kehidupan seharian kita sekarang. Masyarakat telah banyak mendapat manfaat daripada persekitaran maya ini sama ada dari segi persahabatan atau perhubungan dengan orang lain atau keluarga (Bhaskar et al. 2022). Media sosial menggalakkan usahawan, *influencer*, dan perniagaan untuk berkembang namun terdapat juga kelemahan yang boleh didapati daripada platform-platform media sosial. Walaupun media sosial membolehkan orang ramai berhubung antara satu sama lain, meluaskan rangkaian dan menyediakan ruang untuk kebebasan bersuara (Fortunas et al. 2020), kebebasan bersuara boleh bertukar menjadi tuturan menghina dan tuturan benci dengan mudah apabila komen yang disiarkan itu mempunyai cenderung untuk memudaratkan atau mengkritik seseorang atau sekumpulan orang tertentu (Krupalija et al, 2022).

Tuturan menghina ditakrifkan sebagai apa-apa jenis bahasa atau ungkapan yang mungkin dianggap tidak sopan, kesat atau tidak sesuai. Ia boleh melibatkan penggunaan kata-kata cabul, penghinaan atau kata-kata menghina, menyasarkan individu atau kumpulan berdasarkan ciri-ciri seperti bangsa, jantina, agama atau penampilan (Ahmed et al. 2022). Tuturan benci melampaui tuturan menghina dan digambarkan sebagai kandungan yang membawa sikap tidak bertoleransi, diskriminasi, prasangka atau sebarang perasaan yang tidak baik terhadap seseorang, organisasi atau sekumpulan orang berkenaan dengan kewarganeraraan, agama, etnik, umur, jantina dan sebagainya (Krupalija et al. 2022).

Seperti yang dinyatakan sebelum ini, media sosial sentiasa berkembang, dan rangkaian yang luas menyebabkan pengesanan dan pengklasifikasi setiap kandungan di Internet secara manual adalah sesuatu misi yang mustahil. Justeru, satu sistem yang boleh mengenal pasti dan mengklasifikasikan komen di Twitter diperlukan supaya ia boleh digunakan sebagai usaha pencegahan untuk mengurangkan diskriminasi.

Kajian in berfokuskan Twitter kerana Twitter Twitter mempunyai 368.4 juta pengguna dan 7.2% pengguna internet mengakses Twitter sekurang-kurangnya sekali dalam sebulan[1]. Jumlah pengguna Twitter pada tahun 2023 adalah tertinggi berbanding tahun 2019 hingga 2021. Walaupun terdapat kajian lepas yang dijalankan untuk mengesan tuturan menghina atau tuturan benci (Tolba et al. 2021; Sadiq et al. 2019; Balakrishnan et al. 2019), penyelidikan ini mencadangkan untuk mencari variasi BERT yang paling sesuai yang mampu mengenal pasti tuturan menghina dan tuturan benci.

Pada masa kini, orang ramai menggunakan emoji atau emotikon dan bukannya teks panjang dalam siaran, ulasan atau mesej teks mereka untuk menyampaikan hal yang ingin mereka nyatakan (Bai et al. 2019). Secara intuitif, pengaruh emoji atau emotikon boleh menjadi besar kerana ia mewakili keadaan emosi sebenar yang tersembunyi dalam teks (Wu et al. 2018). Oleh itu, memasukkan emoji sebagai sebahagian daripada ciri teks adalah sangat penting supaya makna emoji yang digunakan dapat diekstrak juga. Kajian ini bertujuan untuk menggunakan gabungan ciri teks termasuk teks itu sendiri dan emoji untuk mengesan tuturan menghina dan tuturan benci.

Menurut Messaoudi et al. (2022), kejituan penggunaan teks dan emoji ialah 83.7%, manakala kejituan penggunaan teks sahaja ialah 78.4%. Ini telah membuktikan bahawa emoji mempengaruhi klasifikasi teks. Terdapat beberapa kaedah yang boleh digunakan untuk memasukkan emoji dalam teks iaitu menukar emoji dengan nama Unikod yang sepadan (Pota et al. 2020), menggabungkan skor sentimen yang dikira untuk teks dan menggabungkannya dengan skor sentimen emoji tersebut (Fortunas et al. 2020) dan Emoji2Vec (Aquino et al. 2021). Dalam penyelidikan ini, Emoji2Vec digunakan untuk ciri emoji dan kemudiannya akan digabungkan dengan teks. Ia kemudiannya akan digunakan untuk melatih model BERT untuk mencari variasi BERT terbaik untuk tugasan ini.

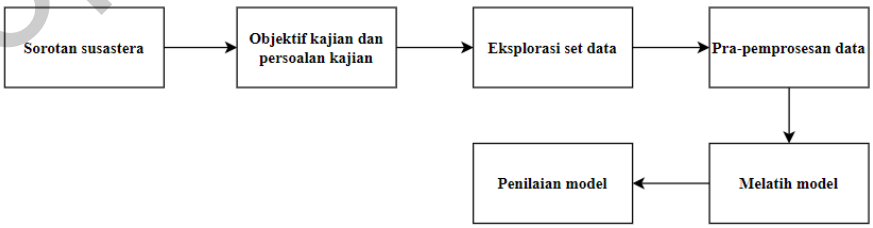
Objektif kajian ini adalah seperti berikut:

- Untuk mengenal pasti ciri dalam teks yang boleh digunakan untuk pengesanan tuturan menghina dan tuturan benci.

[1] https://www.statista.com/statistics/303681/twitter-users-worldwide/

- Untuk mencadang model pembelajaran mesin untuk mengesan tuturan menghina dan tuturan benci.
- Untuk menilai model pengesanan tuturan menghina dan tuturan benci menggunakan metrik penilaian.

Terdapat skop dalam menjalankan penyelidikan ini. Satu ialah data yang tidak disiarkan di Twitter tidak dapat dianalisis kerana kajian ini menggunakan set data yang disediakan oleh (Davidson et al. 2017). Kedua, hanya data sehingga tahun 2017 boleh dianalisis kerana twit yang dikumpul adalah sebelum 2018. Tambahan pula, hanya twit bahasa Inggeris dimasukkan dalam penyelidikan ini kerana komen kecuali komen bahasa Inggeris dengan emoji akan dialih keluar semasa pra-pemprosesan data. Selain itu, kajian ini mencadangkan untuk menggunakan Emoji2Vec untuk mendapatkan pembenaman emoji, dan kemudiannya akan digabungkan dengan pembenaman teks. Walau bagaimanapun, Emoji2Vec terdiri daripada 300 vektor untuk satu emoji, yang mungkin menyebabkan komplikasi apabila menggabungkannya dengan pembenaman teks. Akhir sekali, set data mempunyai 3 kelas: tuturan menghina, tuturan benci dan bukan kedua-duanya. Walau bagaimanapun, penyelidikan ini memfokuskan pada teks dengan emoji sahaja, dan disebabkan bilangan contoh yang terhad dalam kelas benci, kelas tuturan benci dan tuturan menghina mungkin digabungkan menjadi satu, menghasilkan hanya 2 kelas: tuturan menghina dan tuturan benci serta bukan kedua-duanya.



Rajah 1 Ringkasan metodologi untuk pengesanan tuturan menghina dan tuturan benci

Rajah 1 menunjukkan ringkasan metodologi untuk kajian ini. Sorotan susastera dijalankan pada peringkat pertama di mana kajian telah dijalankan ke atas pelbagai domain dan akhirnya topik kajian telah ditentukan. Seterusnya, peringkat ini bertujuan untuk mengenal pasti objektif kajian dan persoalan kajian. Pada peringkat ketiga, eksplorasi set data telah dijalankan untuk menentukan set data yang sesuai untuk digunakan dalam kajian ini. Seterusnya, pra-pemprosesan data termasuk

pembersihan dan penormalan emoji dan teks akan dijalankan. Akhir sekali, model pembelajaran mesin akan dilatih dan dinilai.

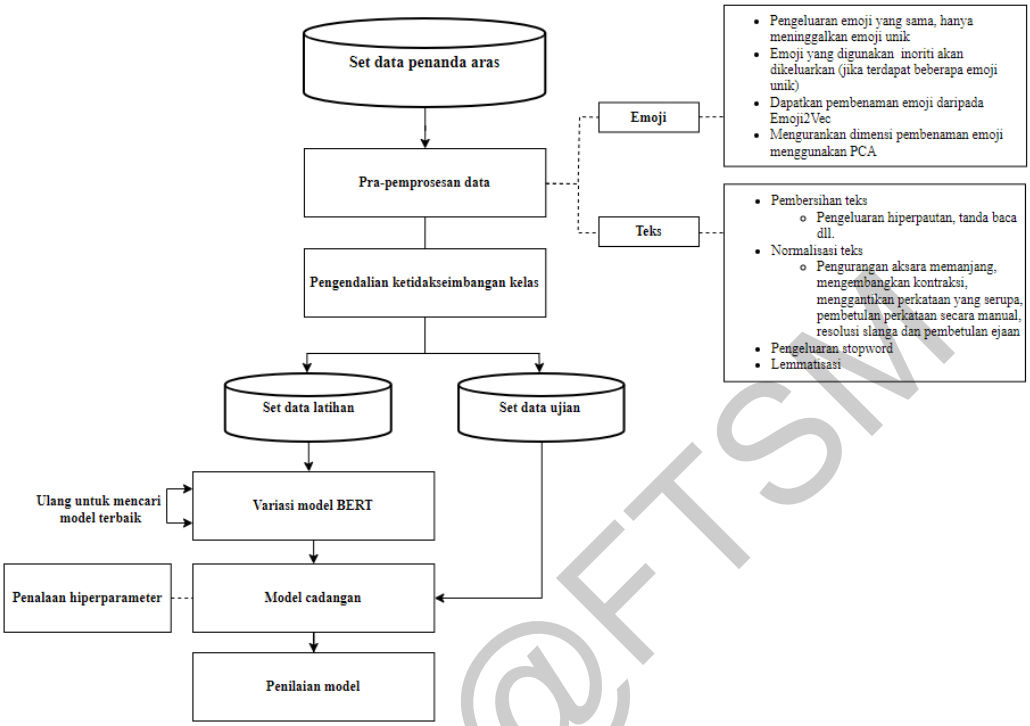
Jadual 1 menunjukkan kajian literatur tentang bahasa yang menyinggung perasaan dan pengesanan ucapan benci. Dalam penyelidikan terdahulu, Fortunas et al. (2020) menggunakan kaedah berasaskan peraturan leksikon pada set data Facebook, mencapai recall 0.96 dan precision 0.79. Walau bagaimanapun, had pendekatan mereka terletak pada bilangan sampel latihan yang kecil, yang mungkin memberi kesan kepada kebolehgeneralisasiannya. Risch dan Krestel (2018) menggunakan ensemble *Logistic Regressing* dan model rangkaian saraf dengan perwakilan ciri seperti pembenaman perkataan, *character n-grams*, *word n-grams*, ciri sintaksis pada set data Facebook, memperoleh skor F1 0.61. Penambahbaikan masa hadapan mungkin melibatkan penerokaan pembenaman perkataan lain seperti CBoW, GloVe untuk menilai potensi peningkatan prestasi. Mozafari et al. (2019) *fine-tuned* model BERT base dan CNN models pada set data daripada Waseem dan Hovey (2016) dan Davidson et al. (2017), mencapai skor F1 masing-masing 0.88 dan 0.92, respectively. Prapemprosesan mereka menggantikan emoji dan emotikon dengan <emoji> dan <emotikon>, yang memerlukan penyiasatan lanjut tentang kesannya terhadap prestasi model. Pratiwi et al. (2018) menggunakan data daripada Instagram dan menggunakan FastText n-gram dan char n-gram, memperoleh skor F1 0.66 untuk bigram dengan klasifikasi FastText. Skor F1 yang rendah mungkin disebabkan oleh set data yang kecil kerana FastText tidak sesuai dan pengalihan keluar emoji semasa prapemprosesan boleh mempengaruhi hasil. Machova et al. (2022) mencipta leksikon baharu (Slovak) dan menggunakan model seperti NB, SVM, Bagging dan RF pada set data Facebook dan Instagram, yang mencapai *accuracy* SVM sebanyak 0.89. Walau bagaimanapun, leksikon mereka mungkin terlepas perkataan toksik, yang berpotensi menjejaskan ketepatan pengesanan. Khan et al. (2022) memfokuskan pada topik khusus yang berkaitan dengan hashtag Twitter, menggunakan TFIDF, BoW, dan pembelajaran ensemble dengan model seperti LR, MNB, DT, SVM, Bagging, Adaboost dan SGB. Pendekatan mereka mencapai ketepatan 0.98 dengan SGB tetapi mungkin mempunyai kebolehgunaan terhad

kerana sifat khusus topiknya. Ilma et al. (2021) menggunakan dataTwitter, GloVE dan model BiLSTM untuk mencapai skor F1 0.82. langkah prapemprosesan mereka melibatkan mengalih keluar emoji, yang mungkin memberi kesan kepada keupayaan model untuk mengesan tuturan menghina yang melibatkan emoji. Aroyehun and Gelbukh (2018) menggunakan CNN, LSTM, BiLSTM, dan *data augmentation* dengan pelabelan pseudo pada set data Facebook, memperoleh Skor F1 0.65 dengan LSTM. Emoji telah dinyahkodkan menjadi teks semasa prapemprosesan dan mempertimbangkan emoji dalam bentuk asalnya mungkin meningkatkan prestasi model.

Jadual 1 Kajian pengesanan tuturan menghina dan tuturan benci

Penulis	Set Data	Metodologi	Hasil	Had/Jurang
Fortunas et al. (2020)	Facebook	Kaedah berasaskan peraturan leksikon (SentiWordNet, kamus slanga dsb.)	*Recall*: 0.96 *Precision*: 0.79	Set data yang kecil.
Risch and Krestel (2018)	Facebook	Perwakilan ciri: Pembenaman perkataan, *char n-gram*, *word n-gram*, ciri sintaksis Model: Ensemble LR dan NN	Skor F1: 0.61	Pembenaman perkataan lain (CBoW, GloVe dll) boleh dieksperimen.
Mozafari et al. (2019)	D1: Waseem dan Hovey (2016) D2: Davidson et al (2017)	BERT *base fine-tuning*	BERT base+CNN Skor F1: 0.88 (D1), 0.92 (D2)	Emoji dan emotikon telah digantikan dengan <emoji>, <emoticon>
Pratiwi et al. (2018)	Instagram	*FastText n-gram* dan *char n-grams* Klasifikasi: FastText, RF, DT, LR	Skor F1: 0.66 untuk FastText (*bigram*)	FastText tidak sesuai untuk set data kecil. Emoji telah dialih keluar semasa langkah prapemprosesan.
Machova et al. (2022)	Facebook dan Instagram	Mencipta leksikon baharu (Slovak) Model: NB, SVM, Bagging, RF)	*Accuracy*: 0.89 (SVM)	Leksikon yang baru dicipta mungkin tiada perkataan toksik tertinggal.
Khan et al. (2022)	Twitter #CoronaJihad, #CoronaTerrorism, #MuslimCorona	Pengekstrakan ciri: TFIDF, BoW Model: LR, MNB, DT, SVM Pembelajaran ensemble: Bagging, Adaboost, SGB	*Accuracy*: 0.98 (SGB)	Fokus pada topik tertentu sahaja.
Ilma et al. (2021)	Twitter	GloVE, BiLSTM	Skor F1: 0.82	Emoji telah dialih keluar semasa langkah prapemprosesan.
Aroyehun and Gelbukh (2018)	Facebook	CNN, LSTM, BiLSTM etc. Penggunaan data augmentation dan pelabelan pseudo dalam DNN	Skor F1: 0.65 (LSTM)	Emoji telah dinyahkodkan menjadi teks

Metodologi Kajian



Rajah 2 Carta alir algoritma pengesanan tuturan menghina dan tuturan benci

Rajah 2 menggambarkan carta alir algoritma untuk model tuturan menghina dan tuturan benci. Data yang dipilih adalah set data penanda aras yang disediakan oleh penyelidik semasa menjalankan kajian untuk mengesan tuturan menghina dan tuturan benci dalam twit (Davidson et al, 2017). Set data mempunyai 6 lajur, diterangkan dalam Jadual 2. Twit dan lajur kelas akan digunakan dalam kajian ini untuk melatih model untuk mengesan tuturan menghina dan benci.

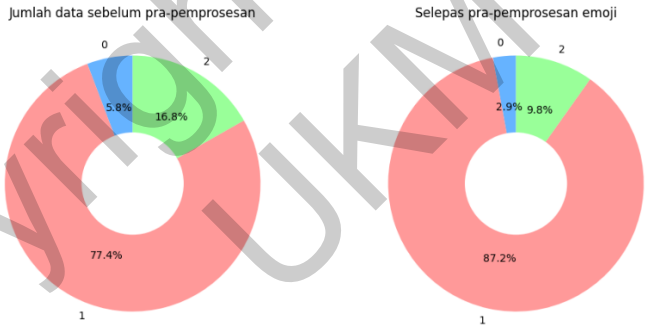
Jadual 2 Senarai lajur dalam set data

Feature	Description	Data Type	Feature for Training
count	number of users who coded each tweet	integer	
hate_speech_count	number of users who judged the tweet to be hate speech	integer	
offensive_language_count	number of users who judged the tweet to be offensive	integer	
neither_count	number of users who judged the tweet to be neither offensive nor hate speech	integer	
class	class label for majority count	integer	✓
tweet	-	string	✓

Bilangan data bagi setiap kelas: tuturan benci, tuturan menghina dan bukan kedua-duanya digambarkan dalam Jadual 3. Selepas pra-pemprosesan emoji telah dilakukan, disebabkan pengagihan entri yang tidak seimbang antara ketiga-tiga kelas dengan bilangan yang agak kecil dalam kategori tuturan benci, kelas tuturan benci dan tuturan menghina akan digabungkan menjadi satu kelas. Penggabungan ini membolehkan penciptaan kategori yang lebih luas yang merangkumi contoh tuturan benci dan tuturan menghina, memberikan saiz sampel yang lebih besar untuk analisis dan tafsiran seperti yang ditunjukkan dalam Jadual 4. Pra-pemprosesan data akan dibentangkan dalam bahagian seterusnya.

Jadual 3 Jumlah data dalam set data sebelum dan selepas pra-pemprosesan data

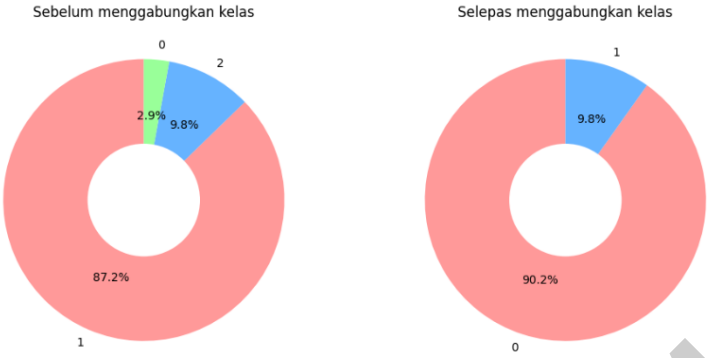
Kelas	Penerangan	Jumlah data	Selepas pra-pemprosesan emoji
0	Hate speech	1430	105
1	Offensive	19190	3154
2	Neither	4163	356



Rajah 3 Perbandingan carta pai sebelum dan selepas pra-pemprosesan emoji

Jadual 4 Jumlah data selepas pra-pemprosesan emoji dan teks serta menggabungkan kelas tuturan benci dan tuturan menghina

Class	Penerangan	Jumlah data	Selepas pra-pemprosesan emoji dan teks
0	Hate speech/Offensive	3259	3145
1	Neither	356	318



Rajah 4 Perbandingan carta pai sebelum dan selepas menggabungkan kelas

Selepas mendapatkan set data, prapemprosesan data akan bermula dengan memisahkan emoji daripada tweet ke lajur lain. Berikutan ini, emoji menjalani prapemprosesan emoji, manakala teks tweet tertakluk kepada proses prapemprosesan berasingan.

Pra-pemprosesan emoji bertujuan untuk mengurangkan pertindihan emoji yang sama dalam baris yang sama. Jadual 5 menggambarkan langkah pra-pemprosesan yang dijalankan pada lajur emoji. Hasil akhir prapemprosesan emoji ialah hanya mempunyai satu emoji untuk setiap baris. Dalam kajian ini, Emoji2Vec[2] akan digunakan untuk mendapatkan pembenaman emoji untuk watak emoji yang sepadan. Walau bagaimanapun, setiap emoji dalam Emoji2Vec menghasilkan vektor benam sepanjang 300 yang mungkin menyebabkan komplikasi dalam memasukkannya dalam melatih model pengesanan pertuturan benci. Oleh itu, *Principal Component Analysis* (PCA) digunakan pada benam untuk mengurangkan dimensi vektor.

Jadual 5 Modul pra-pemprosesan emoji

Pra-pemprosesan	**Penerangan**
Mengeluarkan emoji yang berulang	Untuk memastikan setiap baris mempunya emoji yang berbeza Contoh: 😂☝🔪
Menapis emoji yang jarang digunakan dalam set data	Untuk mengutamakan emoji yang digunakan majoriti Contoh: 😂☝
Hanya menyimpan emoji yang digunakan majoriti (jika terdapat pelbagai emoji dalam baris yang sama)	Untuk memastikan setiap baris hanya mempunyai satu emoji Eg: 😂

[2] https://github.com/uclnlp/emoji2vec

Dengan andaian emoji yang banyak digunakan: '😂'; emoji yang paling jarang digunakan: '🔪'.

Twit mentah biasanya mempunyai tahap redundansi dan hangar yang tinggi, seperti terdapat hashtag, pautan dan lain-lain. Untuk memastikan teks dapat dikendalikan dengan betul, model pra-pemprosesan teks yang diterangkan dalam Jadual 6 dilaksanakan.

Jadual 6 Modul pra-pemprosesan tekss

Kaedah pra-pemprosesan	**Penerangan**
Pengurangan aksara memanjang	Kurangkan aksara berulang ke maksimum dua aksara sahaja Contoh: 'goooooddd' kepada 'goodd'
Pengeluaran pautan, *hashtag*, *mentions*	Mengeluarkan perkataan yang bermula dengan https://, @ dan # Contoh: https://fyp.com, @namapengguna, #fyp
Pengembangan kontraksi	Mengembangkan perkataan dengan tanda apostrof Contoh: I'm kepada I am
Penggantian manual perkataan yang mengandungi aksara asterisk	Gantikan perkataan yang mempunyai asterisk di dalamnya Contoh: n**ga kepada nigga
Penggantian perkataan menggunakan *FuzzyWuzzy*	Gantikan perkataan yang berkemungkinan besar tidak dikenal pasti oleh penyemak ejaan Contoh: 'lmaaoo' kepada 'lmao'
Resolusi slanga	Selesaikan slanga yang biasa digunakan seperti 'lol' kepada '*laughing out loud*'
Pembersihan teks menggunakan *clean text*	Bersihkan keseluruhan teks menggunakan perpustakaan *cleantext*
Semakan ejaan	Menyemak ejaan setiap perkataan menggunakan perpustakaan *pyspellcheck*
Pengeluaran *stopword*	Keluarkan semua *stopword* kecuali perkataan *negator* seperti *no*, *not*
Lematisasi	Jalankan lematisasi ke atas kata nama sahaja

Rajah 5 menunjukkan *Word Cloud* bagi kedua-dua kelas dan dapat dilihat bahawa perkataan yang majoriti didapati dalam kelas tuturan menghina dan tuturan benci ini ialah '*bitch*', '*fuck*', '*nigga*', '*hoe*' dan lain-lain. Word Cloud untuk kelas bukan kedua-duanya mempunyai majoriti perkataan '*love*', '*got*', '*laughing*', '*loud*', dan lain-lain.

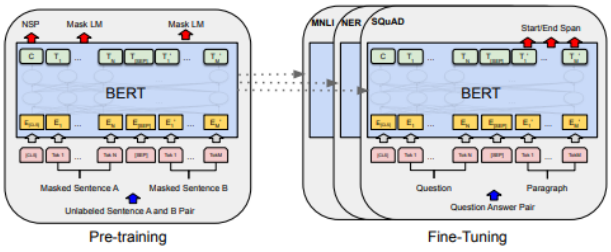
Rajah 5 Word Cloud untuk kedua-dua kelas

Menurut Jadual 4, berikutan pra-pemprosesan emoji dan teks, kelas yang dilabelkan sebagai '0' (mewakili bukan kedua-duanya) mempunyai 318 data, yang sepadan dengan kira-kira 11% daripada jumlah kejadian. Sebaliknya, kelas lain terdiri daripada 3145 data, menyumbang kira-kira 89% daripada set data. Ini menunjukkan ketidakseimbangan yang sangat ketara. Untuk menangani isu ini, pensampelan berlebihan akan dijialankan ke atas set data menggunakan kaedah SMOTE. SMOTE diaplikasikan untuk menambah bilanan data kelas minoriti. Selepas mengaplikasikan SMOTE, kedua-dua kelas akan mempunyai bilangan data yang sama, iaitu 3145 data untuk setiap kelas, seperti yang ditunjukkan di Jadual 7.

Jadual 7 Jumlah data selepas SMOTE

Class	Description	After Emoji and Text Preprocessing		After SMOTE	
0	Tuturan menghina dan tuturan benci	3259	89%	3145	50%
1	Bukan kedua-duanya	356	11%	3145	50%

Variasi model BERT yang berlainan akan digunakan untuk mencari model yang berprestasi terbaik. Model BERT adalah model pra-latihan yang diperkenalkan oleh Google pada 2018 di mana BERT adalah berdasarkan seni bina Transformer dan keseluruhan prosedur pra-latihan dan penalaan halus ditunjukkan dalam Rajah 6. Transformer adalah sejenis rangkaian saraf yang cemerlang dalam menangkap hubungan kontekstual dalam bahasa. Titik kemenangan BERT ialah dwiarahnya, bermaksud ia boleh mempertimbangkan koteks dari kiri dan kanan perkataan tertentu (Devlin et al. 2019). Pendekatan dua hala ini membolehkan BERT menangkap pemahaman yang lebih mendalam tentang bahasa dan mengendalikan tugas. Model terlatih tidak semestinya memerlukan pengekstrakan ciri tambahan. Menutut Subakti et al. (2022), BERT boleh mengekstrak ciri tetap daripada model BERT pra-latihan, dan ini juga dikenali sebagai pembenaman perkataan kontekstual.



Rajah 6 Keseluruhan prosedur pra-latihan dan penalaan halus untuk BERT (Devlin et al. 2019)

BERT base-uncased adalah salah satu varian BERT di mana ia merupakan model pra-latihan pada teks bahasa Inggeris dan varian BERT khusus ini adalah ‘uncased’, bermakna ia menukarkan semua teks kepada huruf kecil untuk generalisasi yang lebih baik (Tsai et al. 2019). Model ini telah dilatih dengan data teks bahasa Inggeris yang tidak berlabel, membolehkannya mempelajari representasi perkataan yang kaya dan bermakna serta konteksnya.

DistilBERT ialah versi suling BERT yang bertujuan untuk mengurangkan saiz model dan kos pengiraan sambil mengekalkan prestasinya. Menurut Adel et al. (2022), DistilBERT menggunakan pengetahuan penyulingan untuk meminimumkan parameter model asas BERT sebanyak 40%. Idea utama penyulingan adalah untuk mengangarkan taburan keluaran penuh model BERT menggunakan model yang lebih kecil seperti DistilBERT.

LAMBERT menggunakan seni bina BERT, dilatih dengan teknik pengoptimuman kelompok besar penyesuaian lapisan iaitu LAMB. You et al. (2020) mengatakan bahawa LAMBERT bertujuan untuk mengurangkan masa pengiraan. Sama seperti BERT, model ini telah dilatih terlebih dahulu untuk bahasa Inggeris di Wikipedia dan BookCorpus. Input teks telah ditulis dengan huruf kecil sebelum tokenisasi menjadi kepingan perkataan.

ALBERT dengan bilangan parameter berkurangan diterbitkan oleh Lan et al. (2019). ALBERT menggabungkan dua teknik pengurangan parameter yang menyelesaikan halangan utama dalam penskalaan model pra-latihan. Yang pertama dalam parameterisasi pembenaman berfaktor di mana metrik pembenaman perbendaharaan kata yang besar diuraikan kepada dua metrik kecil. Teknik ini memudahkan untuk mengembangkan saiz tersembunyi tanpa meningkatkan saiz parameter pembenaman perbendaharaan kata dengan ketara. Teknik kedua ialah perkongsian parameter rentas lapisan di mana teknik ini menghalang parameter daripada pengembangan BERT tanpa menyebabkan kesan besar kepada prestasi BERT, sekali gus meningkatkan kecekapan parameter.

ELECTRA ialah kaedah untuk pembelajaran perwakilan bahasa yang terselia sendiri. Model ELECTRA dilatih untuk membezakan token input ‘sebenar’ berbanding token input ‘palsu’ yang dijana oleh rangkaian saraf lain (Clark et al. 2020).

adual 8 menunjukkan butiran bagi setiap variasi BERT lapisan tersembunyi, unit tersembunyi, parameter, prepocessor dan pengekod yang digunakan dalam kajian ini.

Jadual 8 Lapisan dan unit tersembunyi, parameter, dan prapemproses dan pengekod digunakan

Model	Lapisan tersembunyi	Unit tersembunyi	Parameter	Prapemproses	Pengekod
BERT (en_uncased)	12	768	110M	[3]	[4]
DISTILBERT	6	768	66M	[5]	[6]
LAMBERT	24	1024	-	[7]	[8]
ALBERT	12	768	11M	[9]	[10]
Electra	12	768	110M	[11]	[12]

Metrik penilaian dalam projek ini termasuklah skor F1, lengkungan AUC-ROC dan *confusion marix*. Semua ini dikira daripada *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN) (Sadiq et al. 2021). Data tuturan menghina dan tuturan benci yang diklasifikasikan dengan betul sebagai tuturan menghina dan tuturan benci ialah TP, dan yang salah klasifikasi ialah FN manakala data bukan tuturan menghina dan tuturan benci yang dikelaskan dengan betul ialah TN manakala twit yang salah klasifikasi ialah FP.

- ***Precision*** mengira bilangan *true positives*, dan berapa banyak contoh yang dikelaskan dengan betul dalam kelompok yang sama (Palacio-Niño et al. 2019). Dalam projek ini,

[3] https://tfhub.dev/tensorflow/bert_en_uncased_preprocess/3
[4] https://tfhub.dev/tensorflow/bert_en_uncased_L-12_H-768_A-12/4
[5] https://tfhub.dev/jeongukjae/distilbert_en_uncased_preprocess/2
[6] https://tfhub.dev/jeongukjae/distilbert_en_uncased_L-6_H-768_A-12/1
[7] https://tfhub.dev/tensorflow/bert_en_uncased_preprocess/3
[8] https://tfhub.dev/tensorflow/lambert_en_uncased_L-24_H-1024_A-16/2
[9] http://tfhub.dev/tensorflow/albert_en_preprocess/3
[10] https://tfhub.dev/tensorflow/albert_en_base/3
[11] https://tfhub.dev/tensorflow/bert_en_uncased_preprocess/3
[12] https://tfhub.dev/google/electra_base/2

ketepatan mengukur bilangan data tuturan menghina dan tuturan benci yang diklasifikasikan dengan betul di antara semua data yang diklasifikasikan sebagai tuturan kebencian.

$$Precision = \frac{TP}{TP+FP} \qquad (2)$$

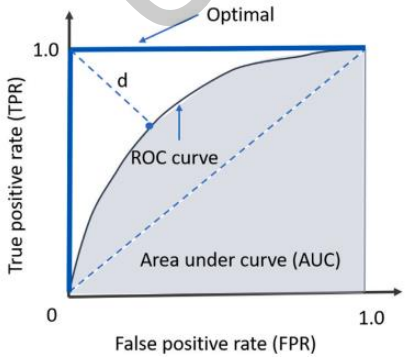
- ***Recall*** menilai peratusan elemen yang dimasukkan ke dalam kelompok yang sama dengan betul.

$$Recall = \frac{TP}{TP+FN} \qquad (3)$$

- **Skor F1** ditakrifkan sebagai min harmonik bagi metrik *precision* dan *recall*. Ia juga membantu dalam menilai *accuracy* model (Muñoz and Iglesias, 2022).

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision+Recall} (4)$$

- **Lengkungan AUC-ROC (*Area Under the ROC Curve*)** ialah pengukuran prestasi untuk masalah pengelasan pada pembagai tetapan ambang. ROC ialah lengkung kebarangkaliaan dan AUC mewakili tahap atau ukuran kebolehpisahan di mana ia memberitahu sejauh mana model mampu membezakan antara kelas. Semakin tinggi nilai AUC, semakin baik model tersebut membezakan antara kelas positif dan negatif. Berdasarkan Rajah 3.9, lengkungan ROC diplot terhadap FPR dimana TPR berada pada paksi-y dan FPR berada pada paksi-x.



Rajah 7 Lengkungan AUC-ROC (Liu et al. 2020)

$$TPR = Recall \qquad (5)$$

$$FPR = 1 - \frac{FP}{TN+FP} \qquad (6)$$

- **Confusion matrix** mewakili istilah seperti sensitiviti, *false positive rate* dan lain-lain. Ia menggambarkan metrik algoritma tertentu dan memberi gambaran tentang prestasi model yang dicadangkan dalam setiap kelas (Handelman et al. 2019). True positives dan true negatives dalam data ujian dengan *predicted positive* dan *predicted negative* boleh dilihat dalam *confusion matrix* dan maklumat ini boleh digunakan untuk mengira pelbagai statistik seperti *recall*, *precision*, skor F1, *accuracy* dan lain-lain.

Keputusan dan Perbincangan

Jadual 9 membentangkan ringkasan perbandingan metrik penilaian untuk variasi berbeza model BERT. Rajah 8 menggambarkan perbandingan metrik yang berbeza untuk variasi model BERT manakala Rajah 10 menunjukkan lengkung ROC untuk model yang berbeza.

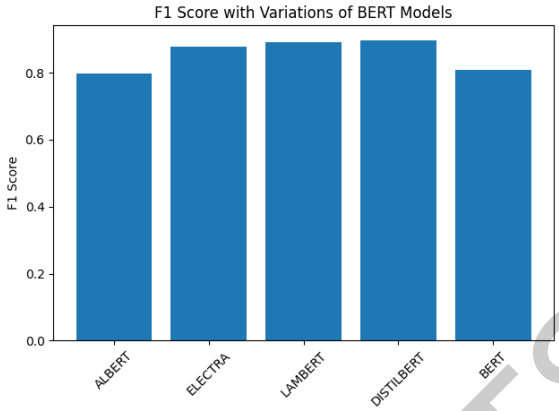
Berdasarkan analisis prestasi, model DistilBERT memenangi model BERT lain dalam tugas pengesanan tuturan menghina dan tuturan benci. Ia mencapai kadar skor F1 yang mengagumkan sebanyak 87.2% dan skor AUC yang kukuh sebanyak 95.2%.

Table 9 Evaluation metrics for variations of BERT

Model	**Metrik penilaian**			
	Recall	***Precision***	**Skor F1**	**AUC**
BERT	0.698	0.748	0.735	0.905
DISTILBERT	0.862	0.874	0.872	0.952
Electra	0.855	0.836	0.839	0.911
ALBERT	0.669	0.754	0.73	0.909
LAMBERT	0.846	0.862	0.86	0.958

Dalam beberapa keadaan, *precision* boleh membawa kepada hasil yang salah. Dengan mengambil kira metrik lain, skor F1 juga dinilai kerana ia mungkin menawarkan hasil yang lebih tepat. Rajah 8 menggambarkan DistilBERT mempunyai skor F1 tertinggi di kalangan semua, di mana ini bermakna model itu mencapai keseimbangan yang baik antara *precision* dan *recall*, yang

membawa kepada prestasi yang berkesan dalam membezakan antara tuturan menghina dan tuturan benci atau bukan kedua-duanya.



Rajah 8 Perbandingan nilai skor F1 antara variasi BERT

Rajah 9 menggambarkan bahawa DistilBERT mempunyai skor AUC tertinggi di kalangan semua, di mana ini bermakna model tersebut mempunyai kuasa diskriminasi yang lebih baik dan boleh membezakan secara berkesan antara kelas positif dan negatif.

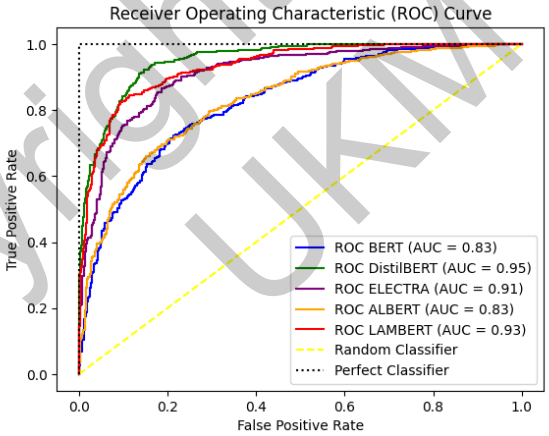
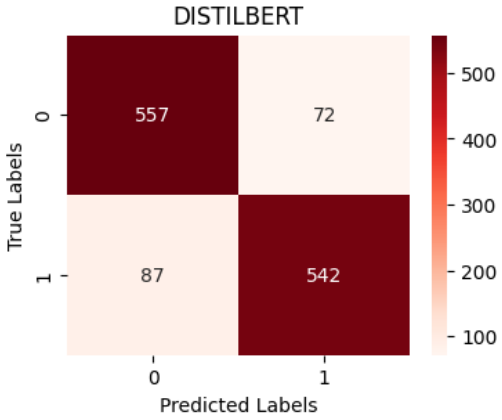


Figure 9 ROC curves for variations of BERT

Berdasarkan Rajah 10, model dengan betul mengklasifikasikan 542 kes sebagai bukan kedua-duanya dan 557 kes sebagai tuturan menghina dan tuturan benci. Walau bagaimanapun, ia salah mengklasifikasikan 72 contoh sebagai tuturan menghina dan tuturan benci apabila ia sepatutnya menjadi bukan kedua-duanya, dan 87 contoh sebagai bukan kedua-duanya apabila ia sepatutnya menjadi tuturan menghina dan tuturan benci.

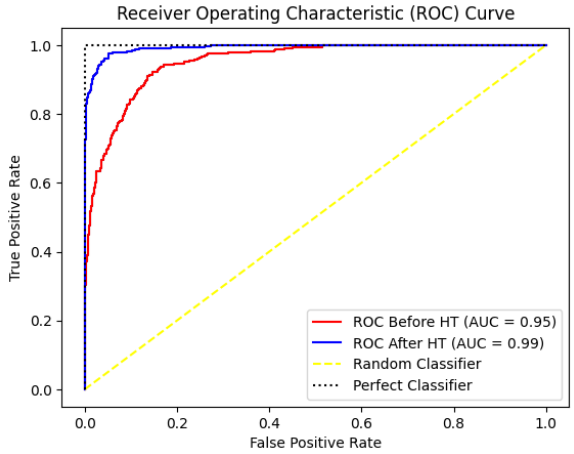


Rajah 10 Confusion matrix untuk DistilBERT

Jadual 10 membentangkan keputusan untuk model DistilBERT sebelum dan selepas penalaan hiperparameter. Selepas menala hiperparameter, terdapat peningkatan yang ketara dalam nilai recall, yang telah meningkat kira-kira 14%, daripada 86.2% kepada 98.4%. Nilai AUC selepas penalaan hiperparameter juga telah meningkat sebanyak 4% berbanding dengan AUC sebelum penalaan hiperparameter.

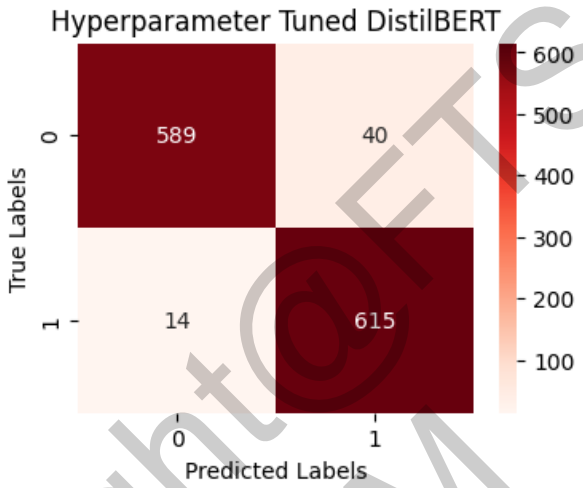
Jadual 10 Ringkasan perbandingan metrik penilaian model DistilBERT sebelum dan selepas penalaan hiperparameter

Model DistilBERT	**Metrik penilaian**			
	Recall	***Precision***	**Skor F1**	**AUC**
Sebelum penalaan hiperparameter	0.862	0.874	0.872	0.952
Selepas penalaan hiperparameter	0.984	0.942	0.958	0.993



Rajah 11 Lengkungan ROC bagi model DistilBERT sebelum dan selepas penalaan hiperparameter

Berdasarkan Rajah 12, model dengan betul mengklasifikasikan 615 kes sebagai bukan kedua-duanya dan 589 kes sebagai tuturan menghina dan tuturan benci. Walau bagaimanapun, ia salah mengklasifikasikan 40 contoh sebagai tuturan menghina dan tuturan benci apabila ia sepatutnya menjadi bukan kedua-duanya dan 14 contoh sebagai bukan kedua-duanya apabila ia sepatutnya menjadi tuturan menghina dan tuturan benci. Membandingkan prestasi keseluruhan sebelum dan selepas penalaan hiperparameter, model menunjukkan peningkatan yang ketara.



Rajah 12 Confusion matrix untuk model DistilBERT selepas penalaan hiperparameter

Model penanda aras yang digunakan dalam penyelidikan oleh Mozafari et al. (2019) melibatkan penggunaan BERT dengan strategi penalaan halus yang berbeza untuk membina model pengesanan tuturan menghina dan tuturan benci. Dalam kajian mereka, emoji yang terdapat dalam twit telah ditukar kepada token pemegang tempat '<EMOTICON>'. Jadual 11 menunjukkan prestasi model penanda aras.

Satu cabaran penting ialah menyesuaikan kod model penanda aras untuk menggunakan set data yang dipraproses dalam kajian ini dengan berkesan. Kod penanda aras yang disediakan mempunyai kekangan dan kod khususnya untuk data pembersihan mereka, yang menyukarkan penyepaduan yang lancar. Disebabkan oleh kekangan masa, kod tidak dapat disesuaikan sepenuhnya set data kajian ini.

Satu lagi kekangan ialah masa yang terhad untuk penilaian. Proses menyesuaikan model penanda aras, menyahpepijat isu dan menjalankan eksperimen komprehensif terbukti memakan masa. Akibatnya, penilaian menyeluruh tidak dapat dijalankan dalam tempoh masa yang ada.

Jadual 11 Results for benchmark model by Mozafari et al. (2019)

Model DistilBERT	**Metrik Penilaian**		
	Recall	***Precision***	**Skor F1**
BERT base	0.91	0.91	0.91
BERT base + nonlinear layers	0.78	0.76	0.77
BERT base + LSTM	0.92	0.91	0.92
BERT base + CNN	0.92	0.92	0.92

Conclusion

Kesimpulannya, model DistilBERT dengan penalaan hiperparameter telah mencapai Skor F1 sebanyak 0.958.

Objektif yang dinyatakan sebelum ini telah berjaya dicapai:

1. Untuk mengenal pasti ciri dalam teks yang boleh digunakan untuk pengesanan tuturan menghina dan tuturan benci.

Ciri teks yang penting untuk tugasan ini telah dikenal pasti, dan ini memanfaatkan pembenaman kontekstual berasaskan BERT untuk menangkap makna semantik, susunan perkataan dan pola sintaksis dengan berkesan. Ini telah membawa kepada kejayaan membina model.

2. Untuk mencadangkan model pembelajaran mesin untuk pengesanan tuturan menghina dan tuturan benci.

Model pembelajaran mesin yang dicadangkan dalam kajian ini ialah variasi BERT yang berbeza, dan selepas eksperimen dijalankan, DistilBERT dengan menggabungkan benam teks berasaskan BERT dan benam emoji, memaparkan prestasi yang terbaik dalam pengesanan tuturan menghina dan tuturan benci.

3. Untuk menilai model pengesanan tuturan menghina dan tuturan benci.

Prestasi model dinilai menggunakan metrik penilaian seperti skor F1, lengkungan AUC-ROC dan *confusion matrix* dan ia menghasilkan keputusan yang baik.

Model tersebut dibina menghadapi beberapa had dan had-had ini perlu diambil kira pada masa hadapan:

1. Ketidakseimbangan kelas.

   Dalam kajian ini, kelas 'bukan kedua-duanya' mempunyai data yang sangat sedikit, menjadikannya mencabar untuk model untuk belajar dan meramal kelas ini dengan berkesan.

2. Data hanya termasuk tweet yang dikumpul sehingga tahun 2017.

   Had ini boleh menjejaskan kebolehan model untuk mengendalikan trend dan corak terkini dalam twit Twitter. Akibatnya, model mungkin tidak menangkap sepenuhnya perubahan terkini dalam aliran bahasa dan ramalannya mungkin tidak tepat apabila digunakan pada data masa nyata.

3. Mengumpul penggunaan emoji masa nyata di Twitter.

   Sebelum ini, bilangan emoji yang digunakan dalam masa nyata (real-time) boleh dijejaki dengan laman web EmojiTracker , namun disebabkan polisi Twitter, penggunaan emoji masa nyata di Twitter tidak dapat diperoleh, yang mana pengumpulan data emoji tersebut pada mulanya dirancang untuk digunakan untuk menapis dan mengutamakan emoji yang sama dan terdapat dalam set data. Kaedah semasa yang digunakan dalam kajian ini mengehadkan setiap baris hanya mempunyai satu emoji. Apabila terdapat pelbagai emoji dalam baris yang sama, emoji yang sering digunakan akan dipilih dan emoji lain akan dikeluarkan. Walaupun pendekatan ini boleh memudahkan proses pemodelan, ia mungkin menyebabkan kehilangan maklumat kerana beberapa emoji yang penting mungkin boleh diabaikan.

Beberapa cadangan untuk pembinaan model masa hadapan adalah untuk mengumpul data yang lebih pelbagai. Isu ketidakseimbangan kelas adalah sangat biasa dilihat dan ia boleh diatasi menggunakan teknik pensampelan atau melaraskan pemberat kelas. Walau bagaimanapun, teknik pensampelan mungkin membawa kepada overfitting, di mana model menjadi terlalu khusus untuk

data latihan dan gagal digeneralisasikan dengan baik kepada data baharu yang tidak kelihatan. Sebaliknya, pensampelan rendah kelas majoriti boleh mengakibatkan kehilangan maklumat berharga dan mengurangkan keupayaan model untuk mengenali corak dalam kelas majoriti.

Selain itu, daripada mengehadkan setiap baris untuk mengandungi hanya satu emoji, prestasi model boleh dipertingkatkan dengan menggunakan pendekatan yang boleh mengendalikan berbilang emoji berturut-turut dengan berkesan kerana emoji yang berbeza sudah pasti menyampaikan makna yang berbeza.

Untuk menangani ketoksikan dalam media sosial atau komuniti dalam talian, penyelidik boleh mempertimbangkan untuk mengumpul data daripada platform yang diketahui mempunyai insiden tingkah laku toksik yang lebih tinggi. Berdasarkan Statista, Facebook telah mengambil tindakan dalam mengalih keluar 10.7 juta kandungan tuturan benci , dan Instagram telah mengalih keluar 5.1 juta kandungan tuturan benci , kedua-duanya pada suku pertama tahun 2023. Data ini menunjukkan bahawa Facebook mungkin mempunyai kadar kandungan toksik yang lebih tinggi, tetapi ia juga boleh menunjukkan bahawa Facebook lebih berkesan dalam mengesan dan mengalih keluar kandungan tersebut. Oleh itu, penyelidik perlu meneliti dengan teliti ciri-ciri setiap platform untuk memilih platform yang paling sesuai untuk memperoleh data bagi membina model bahasa yang menyinggung perasaan dan ucapan benci.

Penghargaan

Saya ingin merakamkan setinggi-tinggi penghargaan kepada penyelia saya, Dr. Wandeep Kaur a/p Ratan Singh, atas bimbingan dan sokongan yang tidak berbelah bahagi sejak cadangan projek ini dimulakan. Kepakaran dan dorongan beliau amat berharga sepanjang perjalanan ini. Saya mengucapkan terima kasih kepada pensyarah yang berdedikasi di FTSM, yang sentiasa membantu dan membolehkan saya menyiapkan projek tahun akhir dengan jayanya. Saya juga amat berterima kasih kepada ibu bapa dan rakan-rakan saya atas sokongan dan dorongan yang berterusan.

Kepercayaan mereka dan sokongan mental mereka yang berterusan telah memainkan peranan penting dalam memastikan saya sentiasa bermotivasi dan fokus pada pelajaran saya.

RUJUKAN

Bhaskar, R. & Bansal, A. 2022. Implementing Prioritized-Breadth-First-Search for Instagram Hashtag Recommendation. *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence),* pp.66-70.

Fortunatus, M. Anthony, P. & Charters, S. 2020. Combining textual features to detect cyberbullying in social media posts. *Procedia Computer Science* 176: 612-621.

Krupalija, E. Đonko, D. & Šupić, H. 2022. Usage of user hate speech index for improving hate speech detection in Twitter posts. *2022 XXVIII International Conference on Information, Communication and Automation Technologies (ICAT),* pp.1-6.

Ahmed, I., Abbas, M., Hatem, R., Ihab, A. & Fahkr, M. W. 2022. Fine-tuning Arabic Pre-Trained Transformer Models for Egyptian-Arabic Dialect Offensive Language and Hate Speech Detection and Classification. 2022 20th International Conference on Language Engineering (ESOLEC), pp.170-174.

Tolba, M., Ouadfel, S. & Meshoul, S. 2021. Hybrid ensemble approaches to online harassment detection in highly imbalanced data. *Expert Systems with Applications* 175: 114751.

Sadiq, S., Mehmood, A., Ullah, S., Ahmad, M., Choi, G. S. & On, B.-W. 2021. Aggression detection through deep neural model on Twitter. Future Generation Computer Systems 114: 120-129.

Balakrishnan, V., Khan, S., Fernandez, T. & Arabnia, H. R. 2019. Cyberbullying detection on twitter using Big Five and Dark Triad features. Personality and Individual Differences 141: 252-257.

Bai, Q. Dan, Q. Mu, Z. & Yang, M. 2019. A Systematic Review of Emoji: Current Research and Future Perspectives. *Frontiers in Psychology* 10:

Wu, Y. Kang, X. Matsumoto, K. Yoshida, M. Xielifuguli, K. & Kita, K. 2018. Sentence Emotion Classification for Intelligent Robotics Based on Word Lexicon and Emoticon Emotions. *2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE),* pp.38-41.

Messaoudi, C. Guessoum, Z. & Romdhane, L. B. 2022. A Deep Learning Model for Opinion mining in Twitter Combining Text and Emojis. *Procedia Computer Science* 207: 2628-2637.

Pota, M. Ventura, M. Catelli, R. & Esposito, M. 2021. An Effective BERT-Based Pipeline for Twitter Sentiment Analysis: A Case Study in Italian. *Sensors.* 21 (1):

Aquino, M. Ortiz, Y. Rashid, A. Tumlin, A. M. Artan, N. S. Dong, Z. & Gu, H. 2021. Toxic Comment Detection: Analyzing the Combination of Text and Emojis. *2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS),* pp.661-662.

Davidson, T., Warmsley, D., Macy, M. & Weber, I. 2017. Automated Hate Speech Detection and the Problem of Offensive Language. Proceedings of the International AAAI Conference on Web and Social Media 11:

Risch, J. & Krestel, R. 2018. Aggression Identification Using Deep Learning and Data Augmentation.

Mozafari, M., Farahbakhsh, R. & Crespi, N. 2019. A BERT-Based Transfer Learning Approach for Hate Speech Detection in Online Social Media. International Workshop on Complex Networks & Their Applications,

Waseem, Z., Davidson, T., Warmsley, D. & Weber, I. 2017. Understanding Abuse: A Typology of Abusive Language Detection Subtasks. Proceedings of the First Workshop on Abusive Language Online,Association for Computational Linguistics. Vancouver, BC, Canada. 2017/8.

Pratiwi, N. I., Budi, I. & Alfina, I. 2018. Hate Speech Detection on Indonesian Instagram Comments using FastText Approach. 2018 International Conference on Advanced Computer Science and Information Systems (ICACSIS), pp.447-450.

Machova, K., Mach, M. & Adamišín, K. 2022. Machine Learning and Lexicon Approach to Texts Processing in the Detection of Degrees of Toxicity in Online Discussions. Sensors 22: 6468.

Khanday, A. M. U. D., Rabani, S. T., Khan, Q. R. & Malik, S. H. 2022. Detecting twitter hate speech in COVID-19 era using machine learning and ensemble learning techniques. International Journal of Information Management Data Insights 2(2): 100120.

Ilma, R. A., Hadi, S. & Helen, A. 2021. Twitter's Hate Speech Multi-label Classification Using Bidirectional Long Short-term Memory (BiLSTM) Method. 2021 International Conference on Artificial Intelligence and Big Data Analytics, pp.93-99.

Aroyehun, S. T. & Gelbukh, A. 2018. Aggression Detection in Social Media: Using Deep Neural Networks, Data Augmentation, and Pseudo Labeling. pp.90-97

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. ArXiv abs/1810.04805:

Subakti, A., Murfi, H. & Hariadi, N. 2022. The performance of BERT as data representation of text clustering. Journal of Big Data 9:

Tsai, H., Riesa, J., Johnson, M., Arivazhagan, N., Li, X. & Archer, A. 2019. Small and Practical BERT Models for Sequence Labeling. ArXiv abs/1909.00100:

Khalid, U., Beg, M. & Arshad, M. 2021. RUBERT: A Bilingual Roman Urdu BERT Using Cross Lingual Transfer Learning.

Adel, H., Dahou, A., Mabrouk, A., Elsayed Abd Elaziz, M., Kayed, M., El-Henawy, I., Alshathri, S. & Ali, A. 2022. Improving Crisis Events Detection Using DistilBERT with Hunger Games Search Algorithm. Mathematics 10: 447.

You, Y., Li, J., Reddi, S., Hseu, J., Kumar, S., Bhojanapalli, S., Song, X., Demmel, J., Keutzer, K. & Hsieh, C.-J. 2020. Large Batch Optimization for Deep Learning: Training BERT in 76 minutes. arXiv [cs.LG]:

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P. & Soricut, R. 2020. ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. arXiv [cs.CL]:

Clark, K., Luong, M.-T., Le, Q. V. & Manning, C. D. 2020. ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators. arXiv [cs.CL]:

Handelman, G. S., Kok, H. K., Chandra, R. V., Razavi, A. H., Huang, S., Brooks, M., Lee, M. J. & Asadi, H. 2018. Peering Into the Black Box of Artificial Intelligence: Evaluation Metrics of Machine Learning Methods. *American Journal of Roentgenology* 212(1): 38-43.

Palacio-Niño, J.-O., & Berzal, F. 2019. Evaluation metrics for unsupervised learning algorithms. *arXiv.org*.

Muñoz, S. & Iglesias, C. A. 2022. A text classification approach to detect psychological stress combining a lexicon-based feature framework with distributional representations. *Information Processing & Management* 59(5): 103011.

Liu, Y., Pang, Z., Karlsson, M. & Gong, S. 2020. Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. Building and Environment 183: 107212.

Tracy Chai Yee May (A180813)
Dr. Wandeep Kaur a/p Ratan Singh
Fakulti Teknologi & Sains Maklumat,
Universiti Kebangsaan Malaysia