

PENGESANAN IDEA BUNUH DIRI DALAM FORUM MEDIA SOSIAL MENGUNAKAN INTENSITI EMOSI DAN PEMBENAMAN PERKATAAN

STUART HENG FU YU

WANDEEP KAUR A/P RATAN SINGH

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

ABSTRAK

Pengesanan kandungan berkaitan bunuh diri di platform digital adalah mencabar. Kebanyakan pengguna teknologi dalam talian meluahkan idea bunuh diri pada platform-platform digital. Hal ini meningkatkan keperluan pengesanan tanda-tanda emosi terhadap kecenderungan bunuh diri. Kajian semasa terbatas dalam penerokaan peranan ungkapan emosi bagi ramalan kecenderungan bunuh diri di media sosial. Projek ini bertujuan untuk mengesan kandungan bunuh diri menggunakan kaedah leksikon emosi. Dengan memanfaatkan pengkomputeran bahasa tabii, fitur emosi dapatlah diekstrak daripada kandungan teks media sosial. Sebuah metrik digunakan untuk mengira intensiti hubungan emosi dengan teks. Kemudian, teks-teks tersebut divektorkan menggunakan pembedaan perkataan (*word embeddings*) untuk menonjolkan perhubungan semantik. Kedua-dua metrik intensiti emosi dan vektor pembedaan perkataan (*word embeddings*) disatukan dalam kerangka pembelajaran mendalam (*deep learning*) berdasarkan mekanisme *Attention* untuk mengesan ketidaknormalan emosi yang merupakan petunjuk kepada kecenderungan bunuh diri. Teks-teks media sosial diklasifikasikan berdasarkan petunjuk bunuh diri menggunakan algoritma klasifikasi. *NRC Emotion Intensity Lexicon* (NRC-EIL) menghubungkan perkataan-perkataan Inggeris dengan skor intensiti bagi lapan emosi asas. Bagi setiap teks, metrik intensiti kumulatif bagi setiap emosi diperolehi dengan menggabungkan skor intensiti setiap perkataan. Projek ini menggunakan model Google News Word2Vec yang telah siap dilatih untuk mengekstrak vektor pembedaan perkataan (*word embeddings*). Gabungan vektor-vektor ini dengan metrik intensiti emosi membentuk input bagi model klasifikasi LSTM-Attention-CNN. Keputusan menunjukkan bahawa penambahan skor intensiti emosi sebagai input tidak meningkatkan ketepatan model yang dicadangkan. Model yang dicadangkan mencapai kejaitan 77.99% dan skor F1 78.08%.

Kata kunci: bunuh diri, media sosial, emosi

PENGENALAN

Fikiran untuk bunuh diri sering dikaitkan dengan perasaan putus asa yang kuat, sering kali berpunca daripada keadaan seperti kemurungan dan khayalan diri. Menurut Pertubuhan Kesihatan Sedunia (2023), lebih daripada 700,000 orang meninggal dunia akibat bunuh diri setiap tahun. Lebih-lebih lagi, tidak terhitung jumlahnya percubaan bunuh diri. Statistik menunjukkan bunuh diri ialah penyebab kematian keempat terbanyak dalam kalangan orang muda, dengan kadar 10.5 kes bunuh diri bagi setiap 100,000 individu (Pertubuhan Kesihatan Sedunia 2023). Pada tahun 2019, terdapat sebanyak 5.8 kes bunuh diri bagi setiap 100,000 individu di Malaysia. Anggaran menunjukkan 1841 individu meninggal dunia akibat bunuh diri, iaitu kira-kira 5 kes kematian bunuh diri setiap hari (Lew et al. 2021). Selain itu, kadar bunuh diri bagi lelaki di Malaysia telah meningkat secara ketara dari tahun 2014 hingga 2019 (Lew et al. 2021). Kajian daripada Mohd Fadhli et al. (2022) melaporkan bahawa 17.1% remaja di Semenanjung Malaysia adalah tertakluk kepada tingkah laku bunuh diri. Dalam kumpulan tersebut, 11.9% pernah mengalami fikiran untuk bunuh diri, 10.2% pernah membuat perancangan bunuh diri, dan 8.4% pernah membuat percubaan bunuh diri.

Kajian telah dijalankan untuk menyiasat faktor-faktor yang menyumbang kepada risiko bunuh diri. Menurut Mohd Fadhli et al. (2022), remaja di Malaysia yang merupakan mangsa buli siber lebih cenderung untuk menunjukkan tingkah laku bunuh diri. Macrynikola et al. (2021) menyatakan bahawa ketagihan telefon pintar, penghantaran teks mesej seksual, dan penggunaan media sosial berkaitan bunuh diri mempunyai korelasi positif dengan kemunculan fikiran dan tingkah laku bunuh diri. Kajian oleh Wan Ismail et al. (2020) menunjukkan keputusan yang serupa, di mana pelajar universiti awam di Lembah Klang, Malaysia, yang mengalami ketagihan internet dan ketagihan telefon pintar adalah lebih cenderung untuk mengalami kemurungan, kegelisahan, stres, dan bunuh diri. Tambahan pula, gangguan mental, peristiwa negatif dalam kehidupan, dan percubaan bunuh diri pada masa lepas akan meningkatkan risiko kemunculan idea bunuh diri dalam kalangan pelajar universiti awam di Malaysia (Chan & Chng 2022). Kajian oleh Zhang et al. (2019) melaporkan bahawa terdapat hubungan antara bunuh diri dengan penggunaan media sosial bermasalah. Menurut Sedgwick et al. (2019) remaja yang menggunakan media sosial secara berlebihan atau bermasalah adalah berisiko tinggi untuk mencuba bunuh diri. Pendedahan kepada kandungan pencederaan diri sendiri (self-harm) di Instagram adalah berkorelasi positif dengan kemunculan idea bunuh diri, pencederaan diri sendiri, dan gangguan emosi (Arendt et al. 2019). Penggunaan sosial media secara bermasalah boleh memudaratkan kesihatan mental (Raudsepp & Kais 2019). Sebaliknya, terdapat beberapa faktor yang mempunyai korelasi negatif dengan kemunculan idea bunuh diri. Kesejahteraan rohaniah dan kesejahteraan eksistensi yang baik dapat menurunkan kemunculan idea bunuh diri (Ibrahim et al. 2019). Penerimaan sokongan daripada ahli keluarga dan rakan-rakan juga didapati berkorelasi negatif dengan kemunculan idea bunuh diri.

Kajian semasa telah menunjukkan bahawa media sosial boleh digunakan untuk mengesan tanda-tanda amaran risiko bunuh diri. Fitur emosi mungkin berperanan dalam pengesanan risiko bunuh diri. Dengan menganalisis emosi dalam mesej teks, Glenn et al.

(2020) mendapati bahawa apabila idea untuk bunuh diri semakin meningkat, mesej teks yang dihantar oleh individu tersebut menunjukkan peningkatan dalam perasaan marah dan penurunan bagi emosi positif. Penemuan ini disokong oleh Sarsam et al. (2021) yang mendapati bahawa tweet yang berkaitan dengan bunuh diri menunjukkan tahap ketakutan dan kesedihan yang lebih tinggi secara signifikan. Sebaliknya, perasaan marah dan gembira adalah lebih kerap dalam tweet yang tidak berkaitan dengan bunuh diri. Tambahan pula, individu yang pernah melakukan tindakan bunuh diri menunjukkan penglibatan yang lebih tinggi dalam aktiviti media sosial yang berkaitan dengan bunuh diri seperti memuat naik siaran (post) yang mempunyai unsur bunuh diri dan memberi komen pada kandungan berkaitan bunuh diri (Liu et al. 2020). Justeru, analisis kontekstual terhadap kandungan media sosial untuk mengenal pasti tanda-tanda bunuh diri dapat memudahkan pengesanan risiko bunuh diri. Secara keseluruhan, emosi dan maklumat kontekstual dalam teks media sosial adalah berguna bagi pengesanan risiko bunuh diri. Kajian literasi mendedahkan jurang penyelidikan dalam aplikasi fitur emosi bagi pengesanan idea bunuh diri. Integrasi fitur emosi dengan pembenaman perkataan dalam konteks pengesanan bunuh diri adalah satu kaedah yang kurang diberi perhatian dalam kajian semasa.

Objektif projek ini adalah untuk mengukur intensiti emosi dalam siaran (*post*) Reddit berunsur bunuh diri. Selain itu, untuk mencadangkan model pembelajaran mendalam (*deep learning*) berdasarkan fitur emosi dan sentimen untuk pengesanan automatik idea bunuh diri bagi kandungan teks media sosial. Di samping itu, untuk menilai kaedah penyelesaian yang dicadangkan berdasarkan metrik-metrik penilaian.

METODOLOGI KAJIAN

Set data

Dalam projek ini, set data yang digunakan ialah set data eRisk¹. eRisk ialah satu inisiatif penyelidikan yang menumpukan pada metodologi penilaian, metrik keberkesanan dan aplikasi praktikal, terutamanya yang berkaitan dengan kesihatan dan keselamatan, dalam pengesanan risiko awal di Internet (Clef Initiative 2017). Set data eRisk terdiri daripada tulisan teks, termasuklah siaran (*post*) dan komen yang dikumpulkan daripada sekumpulan pengguna platform media sosial Reddit. Setiap pengguna Reddit, dipanggil “subjek”, dikaitkan dengan satu label sama ada “kemurungan” atau “bukan kemurungan”. Set tulisan bagi setiap subjek dalam set data adalah tersusun dalam susunan kronologi.

Anotasi label bunuh diri

Set data eRisk tersedia dengan label kemurungan tetapi tiada label bunuh diri. Justeru, langkah-langkah berikut telah dilakukan untuk memberikan label bunuh diri kepada semua tulisan dalam set data eRisk.

Pertama sekali, senarai kata kunci yang berkaitan dengan bunuh diri telah dikumpulkan daripada dua sumber. Parrott et al. (2020) telah mengumpulkan satu senarai kata

¹ Losada, D., Crestani, F. & Parapar, J, 2017, hlm. 346-360.

kunci bahasa Inggeris yang sering digunakan dalam perbincangan topik bunuh diri, disemak dan disahkan oleh ahli profesional akan ketepatan senarai itu. Low et al. (2020) telah membina beberapa buah leksikon secara manual untuk menilai kekerapan penggunaan istilah yang berkaitan dengan kesihatan mental dalam siaran (*post*) Reddit. Salah satu leksikon itu adalah berkenaan dengan bunuh diri. Kata-kata kunci yang dikumpul daripada kedua-dua buah sumber ini telah ditapis secara manual untuk mengeluarkan kata-kata yang maknanya tidak semestinya membawa maksud bunuh diri. Senarai kata kunci yang telah ditapis mempunyai sejumlah 29 patah perkataan.

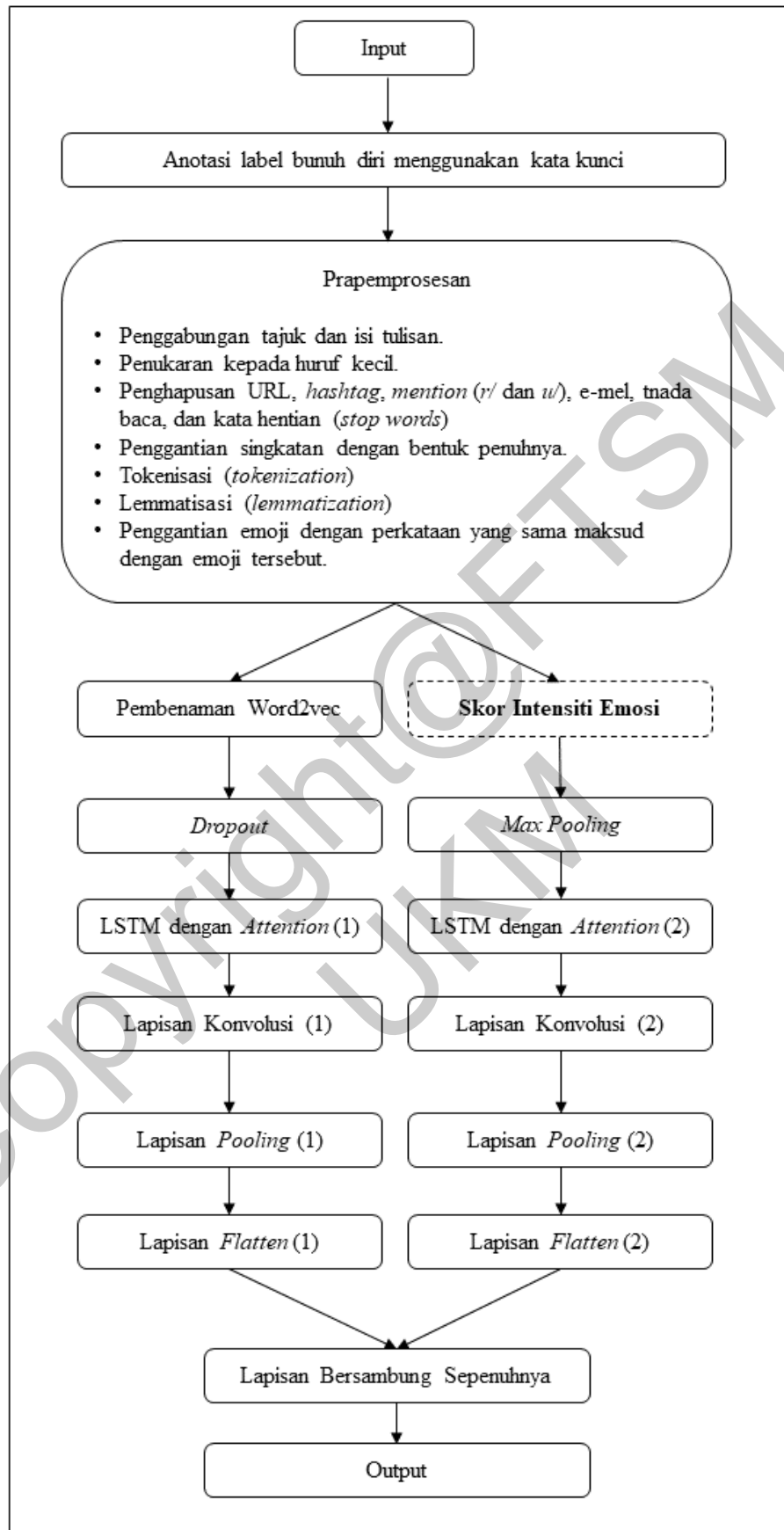
Kemudian, satu model Word2Vec telah digunakan untuk mengembangkan senarai kata kunci tersebut dengan menambahkan perkataan yang mempunyai makna yang hampir serupa. Model Word2Vec digunakan untuk menjana pembenaman perkataan (*word embeddings*) yang membolehkan pengiraan persamaan makna antara dua patah perkataan melalui persamaan kosinus (*cosine similarity*). Model Word2Vec tersebut telah dilatih menggunakan set data Pengesanan Bunuh Diri dan Kemurungan (*Suicide and Depression Detection*, SDD) oleh Komati (2021) yang dimuat turun daripada laman web Kaggle. Set data itu mengandungi 232,074 buah siaran (*post*) Reddit yang dikumpul daripada subreddit r/SuicideWatch, r/depression, dan r/teenagers. Untuk setiap perkataan dalam senarai kata kunci bunuh diri, paling banyak 10 patah perkataan baharu yang mempunyai persamaan kosinus tertinggi (0.5 atau ke atas) telah ditambahkan ke dalam senarai kata kunci. Kemudian, senarai yang dikembangkan itu telah diperiksa secara manual untuk mengeluarkan kata-kata yang tidak berkaitan dengan bunuh diri. Senarai kata kunci bunuh diri kini mengandungi 173 patah kata kunci.

Anotasi label bunuh diri dalam set data eRisk telah dilakukan mengikut subjek, bermakna setiap subjek (iaitu pengguna Reddit) yang menulis siaran Reddit menerima satu label bagi semua siaran mereka, sama ada “bunuh diri” atau “bukan bunuh diri”. Skor bunuh diri telah dikira bagi setiap subjek dalam kedua-dua set data. Skor bunuh diri merupakan jumlah nilai Kekerapan Istilah-Kekerapan Dokumen Songsang (Term Frequency-Inverse Document Frequency, TF-IDF) bagi semua kata kunci bunuh diri yang terdapat dalam siaran subjek, dibahagikan dengan jumlah bilangan token dalam siaran tersebut.

Set data SDD digunakan sebagai satu garis dasar untuk menentukan perbezaan skor bunuh diri antara tulisan bunuh diri dan tulisan bukan bunuh diri. Bagi subjek eRisk dengan skor bunuh diri x , label yang mempunyai skor min paling hampir dengan x akan diberikan kepada subjek tersebut. Selepas itu, pemeriksaan manual dijalankan untuk mengelakkan berlakunya kesilapan label. Sebarang kesilapan label telah dibetulkan.

Seni bina model

Model gabungan LSTM-Attention-CNN yang diperkenalkan oleh Renjith et al. (2022) telah digunakan untuk mengesan idea bunuh diri pada siaran (*post*) Reddit. Model ini berfungsi dengan menggunakan vektor output daripada LSTM sebagai input untuk lapisan *Attention*. Seterusnya, hasil daripada lapisan *Attention* berfungsi sebagai input untuk lapisan konvolusi.



Rajah 1 Model LSTM-Attention-CNN untuk tugas pengesanan idea bunuh diri

Seperti yang ditunjukkan dalam Rajah 1, langkah pertama melibatkan pemberian label “bunuh diri” dan “bukan bunuh diri” kepada setiap subjek eRisk. Seterusnya, setiap tulisan dalam eRisk telah dipraproses. Kemudian, pembenaman perkataan (*word embeddings*) Word2Vec dan skor intensiti emosi *NRC Emotion Intensity Lexicon* (NRC-EIL) telah diekstrak daripada tulisan yang telah dipraproses. Lapisan *dropout* telah dimasukkan untuk mengelakkan *overfitting*. Pembenaman perkataan Word2Vec dan skor intensiti emosi telah dimasukkan ke dalam dua model LSTM-Attention-CNN yang berasingan. Lapisan LSTM dapat menangkap kebergantungan berjujukan pada ayat yang panjang. Lapisan Attention memberikan pemberat (*weights*) kepada setiap token berdasarkan kepentingannya. Selepas itu, lapisan konvolusi mengenal pasti corak dan ciri dalam data input. Lapisan *pooling* telah digunakan untuk mengurangkan dimensi bagi setiap peta fitur rectified (*rectified feature map*). Selepas itu, peta fitur tersebut telah diratakan untuk menjadi satu dimensi. Akhir sekali, lapisan bersambung sepenuhnya (*fully connected*) melakukan klasifikasi dan mengeluarkan output.

Kaedah pengestrakan fitur Word2Vec dan skor intensiti emosi adalah seperti berikut. Pembenaman perkataan (*word embeddings*) menghubungkan perkataan ke ruang vektor berterusan di mana hubungan spatial antara vektor mencerminkan hubungan semantik antara perkataan. Model pra-latih Google News Word2Vec telah digunakan untuk menjana pembenaman perkataan daripada siaran media sosial. Model yang asal mengandungi vektor sebanyak 300 dimensi untuk 3 juta patah perkataan dan frasa. Walau bagaimanapun, model tersebut tidak dapat berjalan kerana kekurangan RAM. Oleh itu, hanya 500 ribu patah perkataan dan frasa yang pertama telah digunakan dalam projek ini.

Fitur sentimen bagi setiap tulisan eRisk telah diekstrak daripada siaran (*post*) mereka menggunakan *NRC Emotion Intensity Lexicon* (NRC-EIL) yang dibangunkan oleh Mohammad (2017). NRC-EIL menghubungkan perkataan Bahasa Inggeris dengan skor intensiti lapan emosi asas iaitu marah (*anger*), jangkauan (*anticipation*), jijik (*disgust*), takut (*fear*), gembira (*joy*), sedih (*sadness*), kejutan (*surprise*), dan percaya (*trust*). Skor intensiti berjalut antara 0 hingga 1. Bagi setiap satu emosi, skor 0 bermakna istilah yang berkenaan tidak menunjukkan sebarang emosi tersebut manakala skor 1 bermakna istilah yang berkenaan menunjukkan intensiti tinggi bagi emosi tersebut. Untuk setiap istilah yang didapati dalam leksikon, skor intensiti istilah tersebut telah dikodkan sebagai vektor lapan dimensi, sepadan dengan lapan emosi asas seperti yang diberikan dalam leksikon. Vektor skor intensiti tersebut berfungsi sebagai perwakilan fitur emosi.

Antara sumbangan rangka kerja ini ialah penambahan analisis emosi sebagai salah satu kaedah pengestrakan fitur. Penambahan analisis emosi dalam pengesanan bunuh diri telah dicadangkan oleh Sarsam et al. (2021). Kajian oleh Sarsam et al. (2021) telah menunjukkan potensi analisis emosi dalam tugas klasifikasi antara *tweet* yang mengandungi unsur bunuh diri dan *tweet* yang tidak mengandunginya.

Persediaan Eksperimen

Objektif model pengesanan bunuh diri adalah untuk menentukan sama ada seseorang pengguna terdapat idea untuk membunuh diri dengan menganalisis siaran (*post*) media sosial pengguna itu. Prestasi model pengesanan bunuh diri telah dinilai berdasarkan metrik penilaian berikut: kejitian (*accuracy*), kepersisan (*precision*), ingatan semula (*recall*), dan skor F1 (*F1-score*).

Seperti yang ditunjukkan dalam Rajah 2, matriks kekeliruan (*confusion matrix*) memaparkan empat nilai: Positif Benar (*True Positive*, TP), Negatif Benar (*True Negative*, TN), Positif Palsu (*False Positive*, FP), Negatif Palsu (*False Negative*, FN) (Rabani et al. 2023). Positif Benar (TP) mewakili situasi model klasifikasi meramalkan kelas positif dengan tepat. Negatif Benar (TN) mewakili situasi model klasifikasi meramalkan kelas negatif dengan tepat. Positif Palsu (FP) mewakili situasi model klasifikasi salah meramalkan kelas negatif sebagai positif. Negatif Palsu (FN) pula mewakili situasi model klasifikasi salah meramalkan kes positif sebagai negatif. Model klasifikasi yang optimum mempunyai nilai TP dan TN yang lebih tinggi.

		Kelas sebenar	
		+	-
Kelas ramalan	+	Positif benar (TP)	Positif palsu (FP)
	-	Negatif palsu (FN)	Negatif benar (TN)

Rajah 2 Matriks kekeliruan

Sumber: Valero-Carreras et al. (2023)

Metrik kejitian (*accuracy*) mengukur nisbah bilangan ramalan yang betul kepada jumlah bilangan ramalan (Birjali et al. 2021). Walaupun kejitian merupakan ukuran yang baik, kejitian menjadi kurang bermakna tentang prestasi model klasifikasi jika set data tidak seimbang (Chatterjee, Kumar, et al. 2022).

$$\text{Kejitian} = \frac{TP + TN}{TP + TN + FP + FN}$$

Kepersisan (*precision*) ialah metrik yang mengukur nisbah bilangan ramalan positif benar (TP) kepada jumlah bilangan ramalan positif (TP + FP) (Chatterjee, Kumar, et al. 2022).

$$\text{Kepersisan} = \frac{TP}{TP + FP}$$

Ingatan semula (*recall*), juga dikenali sebagai sensitiviti (*sensitivity*), ialah nisbah bilangan ramalan positif benar (TP) kepada jumlah bilangan positif sebenar (TP + FN) (Chatterjee, Kumar, et al. 2022). Ingatan semula adalah penting apabila terdapat kesan yang besar dikaitkan dengan negatif palsu. Sebagai contoh, dalam konteks kesihatan mental seperti

mengenal pasti risiko bunuh diri, kesilapan klasifikasi individu berisiko boleh membawa akibat yang teruk. Oleh itu, metrik ingatan semula adalah amat penting (Rabani et al. 2023).

$$\text{Ingatan semula} = \frac{TP}{TP + FN}$$

Skor F1 (*F1-score* atau *F-measure*), ditakrifkan sebagai min harmonik kepersisan dan ingatan semula (Birjali et al. 2021). Skor F1 berfungsi sebagai metrik untuk menilai prestasi model klasifikasi, terutamanya apabila terdapat keperluan untuk mencapai keseimbangan antara kepersisan dan ingatan semula. Selain itu, skor F1 berguna dalam senario ketidakseimbangan set data yang mempunyai lebih banyak data negatif (Rabani et al. 2023).

$$\text{Skor F1} = 2 \times \frac{\text{Kepersisan} \times \text{Ingatan Semula}}{\text{Kepersisan} + \text{Ingatan Semula}}$$

KEPUTUSAN DAN PERBINCANGAN

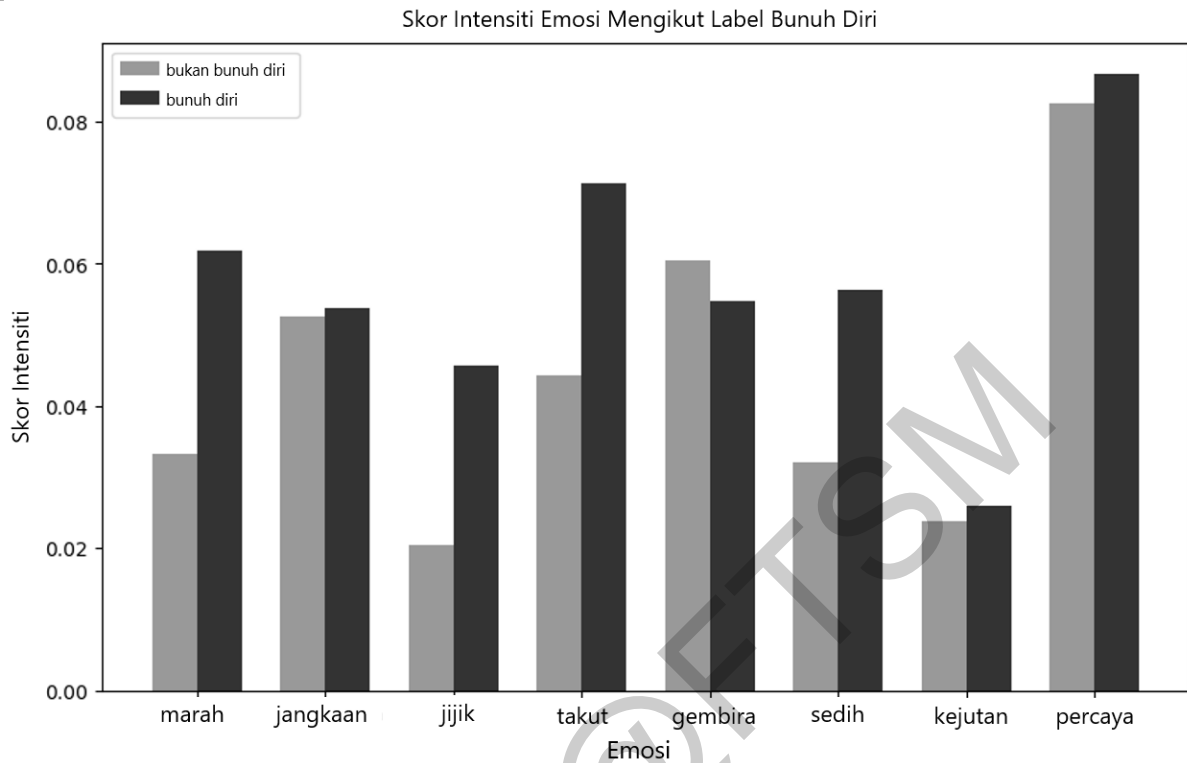
Analisis Set Data

Set data eRisk terdiri daripada 2,348 subjek. 2,184 subjek (93.02%) dilabel bukan kemurungan dan 164 subjek (6.98%) dilabel kemurungan. Set data eRisk merangkumi sejumlah 1,129,425 tulisan, di mana 1,073,752 (95.07%) ditulis oleh subjek bukan kemurungan dan 55,673(4.93%) oleh subjek kemurungan. Bilangan tulisan bagi setiap subjek adalah antara 10 hingga 2,001 tulisan, dengan min (M) 481.02 dan sisihan piawai (SD) 521.00. Jadual 1 menyenaraikan rumusan set data eRisk.

Jadual 1 Rumusan set data eRisk yang dianotasi

Label	Bukan kemurungan	Kemurungan	Jumlah
Bukan bunuh diri	660,344 buah tulisan 2,154 orang subjek	39,679 buah tulisan 156 orang subjek	700,023 buah tulisan 2,310 orang subjek
Bunuh diri	1,436 buah tulisan 22 orang subjek	484 buah tulisan 8 orang subjek	1,920 buah tulisan 30 orang subjek
Jumlah	661,780 buah tulisan 2,176 orang subjek	40,163 buah tulisan 164 orang subjek	701,943 buah tulisan 2,340 orang subjek

Rajah 3 mempersembahkan awan perkataan (*word cloud*) untuk tulisan bukan bunuh diri. Istilah yang paling kerap, tidak termasuk kata hentian (*stop words*), adalah “one” (satu) dan “think” (fikir). Perkataan positif seperti “well” (baik), “love” (cinta), dan “good” (baik) muncul dalam tulisan tersebut. Pemerhatian ini menunjukkan bahawa tulisan bukan bunuh diri mengandungi unsur perasaan dan interaksi yang positif. Penggunaan perkataan seperti “think” (fikir), “say” (kata, cakap), “see” (lihat), dan “look” (lihat) secara kerap menunjukkan bahawa tulisan bukan bunuh diri banyak membincangkan pendapat, pandangan, atau pemerhatian terhadap sesuatu topik. Selain itu, tulisan bukan bunuh diri juga merangkumi pelbagai topik, termasuk kerjaya (“work/kerja”), hiburan (“game/permainan”), dan aktiviti harian (“use/guna”, “go/pergi”, dan “come/datang”). Ringkasnya, kepelbagaian kata-kata dan perasaan positif dalam tulisan bukan bunuh diri menunjukkan komunikasi yang seimbang dan menyeluruh daripada tulisan bukan bunuh diri. Hal ini juga mencerminkan pengalaman harian



Rajah 5 Min skor intensiti emosi bagi tulisan eRisk mengikut label bunuh diri

Jadual 2 Keputusan analisis intensiti emosi

Emosi	Semua Tulisan (N = 680,436)	Tulisan Bukan Bunuh Diri (N = 678,530)	Tulisan Bunuh Diri (N = 1,906)	Signifikan (df = 680,434)
Marah (<i>Anger</i>)	M = 0.0334 SD = 0.0692	M = 0.0333 SD = 0.0692	M = 0.0619 SD = 0.0863	t = -18.0259 p < 0.0001*
Jangkaan (<i>Anticipation</i>)	M = 0.0526 SD = 0.0772	M = 0.0526 SD = 0.0772	M = 0.0537 SD = 0.0768	t = -0.6415 p = 0.5212
Jijik (<i>Disgust</i>)	M = 0.0205 SD = 0.0540	M = 0.0204 SD = 0.0539	M = 0.0457 SD = 0.0783	t = -20.4503 p < 0.0001*
Takut (<i>Fear</i>)	M = 0.0444 SD = 0.0824	M = 0.0443 SD = 0.0823	M = 0.0713 SD = 0.0949	t = -14.2556 p < 0.0001*
Gembira (<i>Joy</i>)	M = 0.0605 SD = 0.0892	M = 0.0605 SD = 0.0892	M = 0.0547 SD = 0.0737	t = 2.8475 p = 0.0044*
Sedih (<i>Sadness</i>)	M = 0.0322 SD = 0.0682	M = 0.0322 SD = 0.0681	M = 0.0563 SD = 0.0834	t = -15.4322 p < 0.0001*
Kejutan (<i>Surprise</i>)	M = 0.0238 SD = 0.0518	M = 0.0237 SD = 0.0518	M = 0.0259 SD = 0.0492	t = -1.8296 p = 0.0673
Percaya (<i>Trust</i>)	M = 0.0825 SD = 0.0968	M = 0.0825 SD = 0.0968	M = 0.0868 SD = 0.0893	t = -1.9134 p = 0.0557

(* perbezaan adalah signifikan)

Skor intensiti emosi marah ($p < 0.0001$), jijik ($p < 0.0001$), takut ($p < 0.0001$), gembira ($p = 0.0044$), dan sedih ($p < 0.0001$) telah menunjukkan perbezaan yang ketara antara tulisan bukan bunuh diri dan tulisan bunuh diri. Tulisan bunuh diri mempamerkan intensiti emosi negatif yang lebih tinggi seperti marah, jijik, takut, dan sedih. Sebaliknya, tulisan bukan bunuh diri mempamerkan intensiti emosi gembira yang lebih tinggi. Keputusan ini konsisten dengan Chatterjee, Samanta, et al. (2022) yang melaporkan bahawa individu yang mempunyai kecenderungan untuk bunuh diri mengekspresikan lebih banyak emosi negatif berbanding dengan individu biasa. Renjith et al. (2022) dan Zhang et al. (2021) juga mendapati penggunaan kata-kata beremosi negatif yang lebih kerap dalam teks bunuh diri berbanding teks bukan bunuh diri melalui analisis kekerapan istilah. Di samping itu, Sarsam et al. (2021) mendapati bahawa terdapat hubungan antara teks bunuh diri dengan tahap ketakutan dan kesedihan yang tinggi. Walau bagaimanapun, Sarsam et al. (2021) mendapati bahawa tahap kemarahan dan kegembiraan dalam teks bukan bunuh diri adalah lebih tinggi daripada teks bunuh diri. Penemuan tersebut berbeza dengan keputusan ini yang mendapati intensiti emosi marah lebih rendah dalam teks bukan bunuh diri daripada teks bunuh diri.

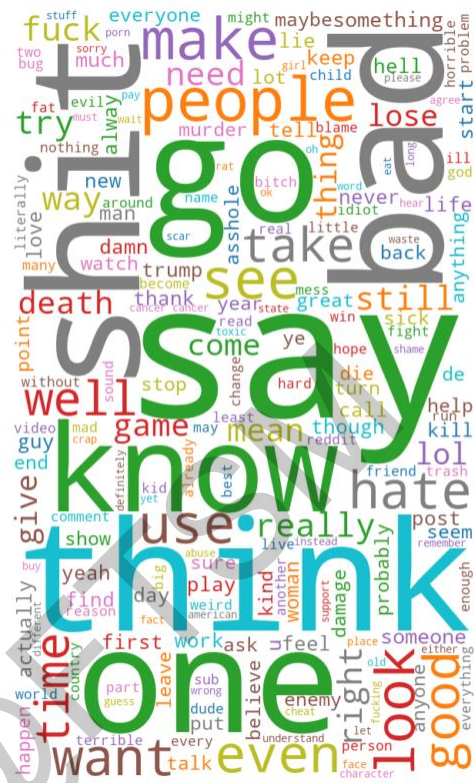
Rajah 6 menggambarkan awan perkataan untuk 10% tulisan teratas dengan skor intensiti tertinggi bagi emosi marah, jijik, takut, dan sedih. Dalam tulisan yang menunjukkan kandungan emosi marah, terdapat banyak istilah ganas seperti "*fight*" (lawan), "*hit*" (pukul), "*attack*" (serang), dan "*kill*" (bunuh). Hal ini mencerminkan sifat agresif dalam tulisan tersebut. Persamaan dengan tulisan bunuh diri ini menunjukkan hubungan yang kuat antara emosi marah dengan bunuh diri.

Tulisan yang tinggi dalam emosi jijik memaparkan istilah yang menunjukkan emosi negatif seperti "*bad*" (buruk), "*shit*", dan "*hate*" (benci). Istilah ini menunjukkan perasaan negatif dan benci yang mendalam. Memandangkan kata-kata tersebut juga sering dijumpai dalam tulisan bunuh diri, malah terdapat hubungan antara emosi jijik dengan kemunculan idea bunuh diri.

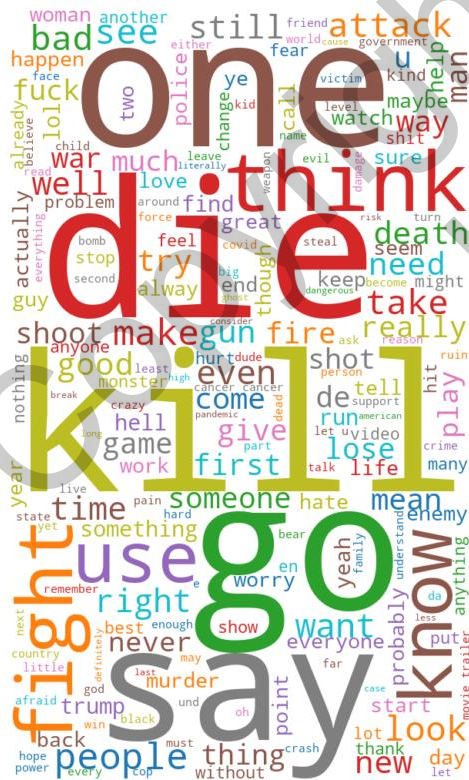
Bagi tulisan dengan intensiti emosi takut yang tinggi, perkataan yang berkaitan dengan kematian seperti "*die*" (mati) dan "*kill*" (bunuh) mempunyai kekerapan yang tertinggi. Hal ini mendedahkan rasa ngeri yang terdapat dalam tulisan bunuh diri. Sama seperti tulisan tinggi dalam emosi marah, istilah ganas seperti "*fight*" (lawan), "*attack*" (serang), "*war*" (peperangan), "*shoot*" (tembak), dan "*gun*" (senjata) juga banyak dijumpai dalam tulisan tinggi dalam emosi takut. Hal ini menunjukkan bahawa topik kematian dan peperangan boleh menyebabkan ketakutan. Kata-kata tersebut juga banyak didapati dalam tulisan bunuh diri, menunjukkan bahawa kebimbangan yang dikaitkan dengan emosi takut mungkin menyumbang kepada kemunculan idea bunuh diri, kerana kebimbangan adalah salah satu tanda kemunculan idea bunuh diri (Grant et al. 2023).



marah



jijik



takut



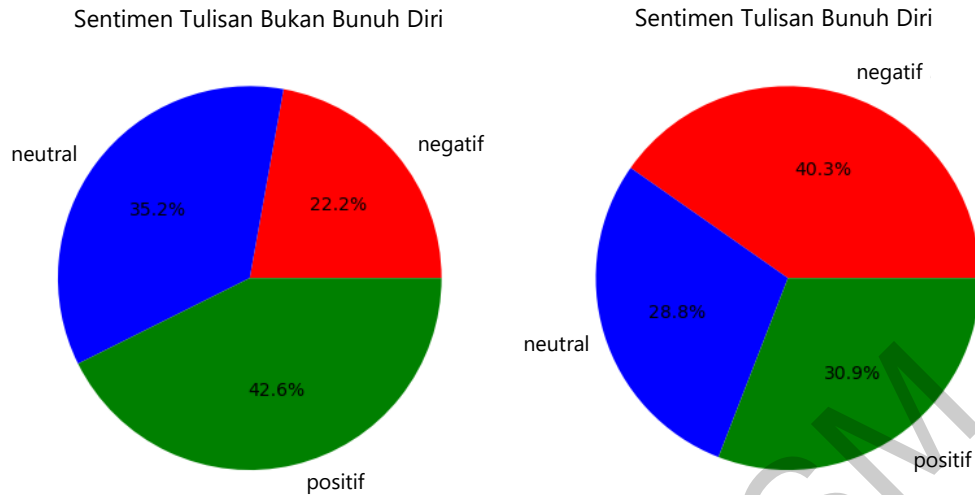
sedih

Rajah 6 Awan perkataan 10% tulisan teratas dengan skor intensiti emosi yang tertinggi bagi emosi marah, jijik, takut, dan sedih

Dalam tulisan yang tinggi dalam emosi sedih, istilah yang berkaitan dengan kematian seperti “*die*” (mati) dan “*kill*” (bunuh) juga banyak dijumpai. Kecepatan istilah “*kill*” (bunuh) dalam tulisan sedih tidak setinggi kekerapannya dalam tulisan takut. Perkataan “*bad*” (buruk), “*lose*” (kalah, hilang), “*hurt*” (sakit), “*hell*” (neraka), dan “*hate*” (benci) menggambarkan emosi negatif yang kuat dalam tulisan sedih. Perkataan-perkataan tersebut menunjukkan bahawa individu yang meluahkan kesedihan sering berhadapan dengan perasaan sakit hati, putus asa, dan benci. Persamaan istilah ini dengan istilah yang didapati dalam tulisan bunuh diri menunjukkan bahawa perasaan sakit hati dan putus asa dalam tulisan sedih mungkin berkait rapat dengan kemunculan idea bunuh diri.

Tuntasnya, kesemua awan perkataan untuk emosi marah, jijik, takut, dan sedih mempunyai persamaan dengan awan perkataan untuk tulisan bunuh diri. Hal ini menunjukkan bahawa emosi negatif seperti marah, jijik, takut, dan sedih mempunyai kaitan rapat dengan keadaan emosi bagi individu yang ingin membunuh diri.

Rajah 7 menunjukkan awan perkataan untuk 10% tulisan teratas dengan skor intensiti emosi gembira yang tertinggi. Awan perkataan tersebut menunjukkan emosi dan pengalaman yang positif. Tulisan yang tinggi dalam emosi gembira mempunyai kekerapan tinggi bagi perkataan positif seperti “*love*” (cinta), “*good*” (baik), “*thank*” (terima kasih), “*well*” (baik), dan “*win*” (menang). Perkataan-perkataan positif itu mencerminkan tema kesejahteraan dan kejayaan. Kelaziman istilah yang bersemangat positif itu membuktikan bahawa tulisan yang tinggi dalam emosi gembira tertumpu pada pengalaman dan emosi yang positif. Bahasa yang bersifat positif ini menonjolkan perbezaan yang ketara dengan tulisan bunuh diri yang lazimnya kaya dengan perkataan beremosi negatif. Contohnya, walaupun tulisan bunuh diri sering memaparkan perkataan yang negatif, tulisan yang penuh kegembiraan ini didominasi oleh perkataan yang menyampaikan kebahagiaan dan kepuasan. Perbezaan ini menyerlahkan perbezaan emosi antara individu yang gembira dengan individu yang menunjukkan tanda-tanda idea bunuh diri. Istilah yang kerap dijumpai dalam tulisan emosi gembira seperti “*well*” (baik), “*love*” (cinta), dan “*good*” (baik) juga mempunyai kekerapan yang sederhana hingga tinggi dalam tulisan bukan bunuh diri. Persamaan ini menyokong kenyataan individu yang beremosi positif kurang berkemungkinan untuk bunuh diri.



Rajah 8 Peratus tulisan eRisk dengan sentimen positif, neutral, dan negatif

Seperti yang digambarkan dalam carta pai Rajah 8, tulisan bukan bunuh diri menunjukkan peratusan positif yang lebih tinggi (42.6%) daripada peratusan negatif (22.2%). Sebaliknya, tulisan bunuh diri mempunyai peratusan negatif yang lebih tinggi (40.3%) berbanding dengan peratusan positif (30.9%). Hal ini membuktikan bahawa terdapat lebih banyak sentimen positif dalam tulisan bukan bunuh diri serta lebih banyak sentimen negatif dalam tulisan bunuh diri. Jadual 3 memaparkan keputusan terperinci analisis sentimen untuk set data eRisk mengikut label bunuh diri.

Jadual 3 Keputusan analisis sentimen

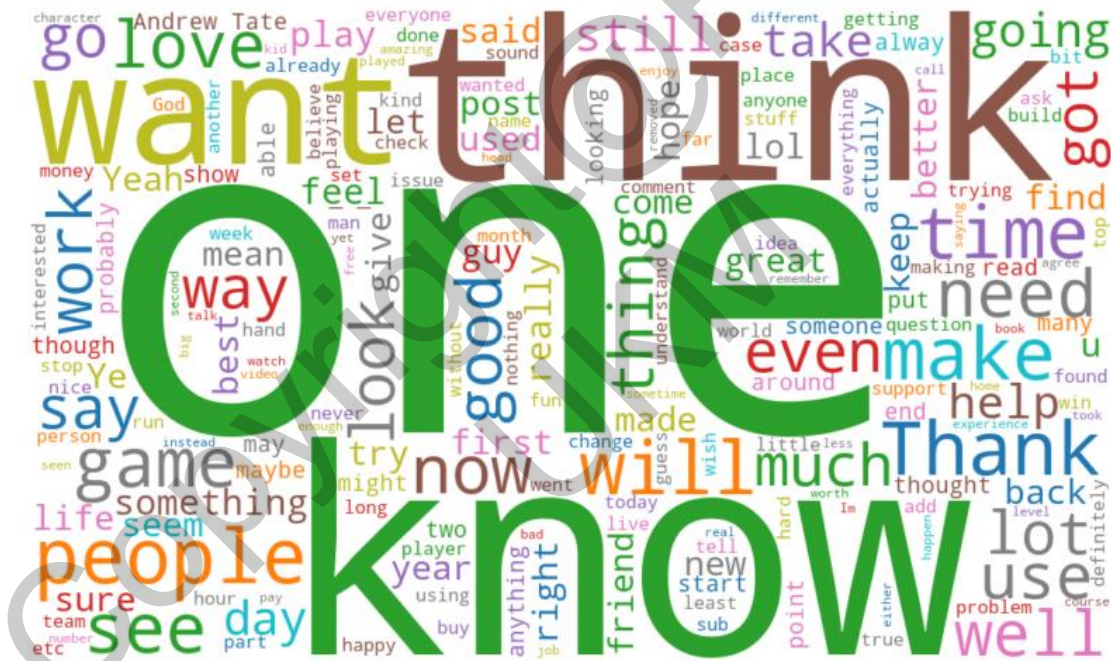
Skor Sentimen	Semua Tulisan (N = 1,127,402)	Tulisan Bukan Bunuh Diri (N = 1,124,320)	Tulisan Bunuh Diri (N = 3,082)	Signifikan (df = 1,127,400)
<i>pos</i> (positif)	M = 0.1367 SD = 0.2002	M = 0.1368 SD = 0.2003	M = 0.1110 SD = 0.1725	t = 7.1393 p < 0.0001*
<i>neu</i> (neutral)	M = 0.7839 SD = 0.2363	M = 0.7839 SD = 0.2364	M = 0.7615 SD = 0.2256	t = 5.2485 p < 0.0001*
<i>neg</i> (negatif)	M = 0.0689 SD = 0.1338	M = 0.0688 SD = 0.1337	M = 0.1229 SD = 0.1684	t = -22.4482 p < 0.0001*
<i>compound</i> (keseluruhan)	M = 0.1299 SD = 0.4502	M = 0.1305 SD = 0.4499	M = -0.0878 SD = 0.5241	t = 26.7598 p < 0.0001*

(* perbezaan adalah signifikan)

Keputusan ujian-t dua sampel bebas telah melaporkan perbezaan yang signifikan dalam skor sentimen *pos*, *neu*, *neg*, dan *compound* antara tulisan bukan bunuh diri dan tulisan bunuh diri. Tulisan bunuh diri mempunyai tahap sentimen negatif yang lebih tinggi manakala tulisan bukan bunuh diri menunjukkan tahap sentimen positif yang lebih tinggi. Hasil dapatan ini sejajar dengan keputusan yang dilaporkan oleh Sarsam et al. (2021).

Rajah 9 mempamerkan awan perkataan bagi tulisan yang mempunyai sentimen positif manakala Rajah 10 bagi tulisan yang mempunyai sentimen negatif. Terdapat persamaan antara tulisan positif dengan tulisan bukan bunuh diri, di mana tulisan positif merangkumi topik kehidupan seharian dengan perkataan seperti “*think*” (berfikir), “*work*” (kerja) dan “*game*” (permainan). Tulisan bersentimen positif juga mengandungi perkataan yang sering dijumpai dalam tulisan emosi gembira seperti “*love*” (cinta), “*good*” (baik), dan “*thank*” (terima kasih). Hal ini menunjukkan tema kasih sayang dan positif. Persamaan antara istilah yang terdapat dalam tulisan sentimen positif, tulisan bukan bunuh diri, dan tulisan emosi gembira membuktikan sentimen positif merendahkan kemungkinan idea bunuh diri.

Sebaliknya, kekerapan istilah “*people*” (orang) dan “*time*” (masa) adalah lebih tinggi dalam tulisan dengan sentimen negatif berbanding dengan tulisan dengan sentimen positif. Menurut Zhang et al. (2021), istilah yang berkaitan dengan “*orang*” kerap muncul dalam nota bunuh diri. Hal ini mencadangkan terdapat hubungan antara tulisan bersentimen negatif dan idea bunuh diri.



Rajah 9 Awan perkataan bagi tulisan dengan sentimen positif

al. (2022) iaitu 90.3%. Kelemahan ini mungkin disebabkan oleh perbendaharaan kata (*vocabulary*) model Word2Vec Google News yang telah dihadkan kepada 500,000 patah perkataan daripada 3 juta patah perkataan untuk mengatasi masalah kekurangan RAM. Selain itu, set data yang digunakan untuk melatih dan menguji adalah kecil ($N = 3840$). Hal ini boleh menjejaskan keupayaan model klasifikasi untuk membuat generalisasi terhadap input.

Model klasifikasi yang menggunakan skor intensiti emosi sahaja tidak menunjukkan prestasi yang baik, dengan kejituan 63.28%. Satu kemungkinan yang menyebabkan prestasi rendah ialah matriks skor intensiti emosi yang jarang. Hal ini disebabkan oleh kebanyakan istilah Inggeris mempunyai intensiti emosi sifar kerana istilah tersebut tidak dikaitkan dengan sesuatu emosi. Selain itu, saiz perbendaharaan kata NRC-EIL mungkin tidak mencukupi, kerana NRC-EIL hanya mengandungi intensiti emosi untuk 14,671 patah perkataan sahaja. Sebagai perbandingan, perbendaharaan kata Word2Vec yang digunakan untuk melatih model ini adalah sebanyak 500,000 patah perkataan dengan 300 dimensi untuk pembedaan. Skor intensiti emosi mungkin tidak memberikan maklumat yang mencukupi kepada model klasifikasi untuk mengklasifikasikan idea bunuh diri dengan tepat daripada tulisan media sosial.

Cadangan Penambahbaikan

Selepas menjalankan kajian yang menyeluruh, cadangan untuk menambahbaik permainan serius ini pada masa hadapan adalah dengan memanfaatkan *Bidirectional Encoder Representations from Transformers* (BERT) bagi tugas pengesanan idea bunuh diri. Hal ini demikian kerana BERT ialah model perwakilan bahasa yang canggih. BERT telah menunjukkan prestasi yang luar biasa dalam berbagai-bagai tugas mengenai pemrosesan bahasa tabii (Devlin et al. 2018). Walau bagaimanapun, BERT tidak dapat digunakan dalam projek ini atas kekurangan kelengkapan komputer yang berspesifikasi tinggi. Sekiranya kerja masa hadapan mampu memanfaatkan BERT dengan menggunakan kelengkapan berspesifikasi tinggi, kerja masa hadapan berpotensi untuk mencapai prestasi yang lebih tinggi dalam tugas pengesanan idea bunuh diri.

Selain itu, kaedah mengekstrak skor intensiti emosi yang digunakan dalam projek ini boleh diperbaiki dengan menggunakan model yang dibuat khusus untuk tugas pengesanan emosi. Memandangkan emosi manusia sangat rumit, model khusus ini boleh menjalankan penilaian intensiti emosi dengan lebih tepat daripada tulisan media sosial.

Tambahan pula, selain daripada kandungan teks, penyelidikan masa depan boleh menambahkan kandungan media sosial yang lain seperti imej, mesej suara, dan video. Kebanyakan platform media sosial membolehkan penggunanya untuk memuat naik kandungan multimedia. Penambahan kandungan multimedia dapat memberikan lebih banyak maklumat untuk dianalisis. Kandungan multimedia mungkin mengandungi informasi yang relevan untuk mengenal pasti keadaan emosi seseorang pengguna. Analisis gabungan teks dengan kandungan multimedia boleh meningkatkan lagi tahap pemahaman emosi pengguna dan meningkatkan prestasi model pengesanan idea bunuh diri.

KESIMPULAN

Secara keseluruhannya, model pembelajaran mendalam untuk mengesan idea bunuh diri ini telah berjaya dibangunkan dengan menggunakan set data eRisk. Objektif yang telah ditetapkan pada permulaan projek ini telah berjaya dicapai. Walaupun terdapat beberapa halangan, ia berjaya diatasi menggunakan pelbagai cara. Diharapkan model pengesanan idea bunuh diri ini dijadikan titik kajian untuk kajian lain pada masa hadapan.

Kekuatan Sistem

Kekuatan projek ini adalah pada penemuan perbezaan intensiti emosi antara tulisan bunuh diri dengan tulisan bukan bunuh diri. Tulisan bunuh diri mempunyai skor intensiti yang lebih tinggi bagi emosi negatif seperti marah, jijik, takut, dan sedih. Sebaliknya, tulisan bukan bunuh diri menunjukkan emosi positif seperti gembira.

Kelemahan Sistem

Kelemahan utama projek ini adalah pada kekurangan kuasa skor intensiti emosi pada prestasi model klasifikasi LSTM-Attention-CNN. Hasil dapatan projek ini menunjukkan bahawa penggabungan skor intensiti emosi dengan pembenaman Word2Vec tidak meningkatkan prestasi model dalam pengesanan idea bunuh diri. Pembenaman Word2Vec sudah mencukupi untuk menangkap nuansa dalam penulisan media sosial. Walaupun analisis emosi adalah salah satu aspek penting dalam proses memahami idea bunuh diri, penggabungannya ke dalam model klasifikasi tidak memberikan peningkatan seperti yang dijangkakan.

Selain itu, satu lagi kelemahan projek ini adalah pada kebolehtafsiran model pembelajaran mendalam (deep learning). Walaupun model mengeluarkan output tentang sama ada sesuatu siaran (post) mengandungi kandungan idea bunuh diri atau tidak, model itu tidak memberikan sebarang maklumat tentang alasan di sebalik ramalannya. Proses untuk memahami apakah ciri dan corak yang model bergantung padanya semasa membuat ramalan adalah sangat mencabar.

PENGHARGAAN

Dengan sukacitanya, penulis kajian ini ingin ucapkan setinggi-tinggi penghargaan dan jutaan terima kasih kepada Dr. Wandeep Kaur a/p Ratan Singh, penyelia penulis kajian ini yang telah memberi tunjuk ajar serta bimbingan untuk menyiapkan projek ini dengan jayanya.

Penulis kajian ini juga ingin mengucapkan terima kasih kepada semua pihak yang membantu secara langsung mahupun tidak langsung dalam menyempurnakan projek ini. Segala bantuan yang telah dihulurkan amatlah dihargai kerana tanpa bantuan mereka, projek ini tidak dapat dilaksanakan dengan baik. Semoga Tuhan merahmati dan memberikan balasan yang terbaik.

RUJUKAN

- Acuña Caicedo, R. W., Gómez Soriano, J. M. & Melgar Sasieta, H. A. 2020. Assessment of supervised classifiers for the task of detecting messages with suicidal ideation. *Heliyon* 6(8): e04412.
- Ait Skourt, B., El Hassani, A. & Majda, A. 2022. Mixed-pooling-dropout for convolutional neural network regularization. *Journal of King Saud University - Computer and Information Sciences* 34(8, Part A): 4756-4762.
- Ananthakrishnan, G., Jayaraman, A. K., Trueman, T. E., Mitra, S., Abinesh, A. K. & Murugappan, A. 2022. Suicidal Intention Detection in Tweets Using BERT-Based Transformers. *3rd IEEE 2022 International Conference on Computing, Communication, and Intelligent Systems, ICCIS 2022*, hlm. 322-327.
- Arendt, F., Scherr, S. & Romer, D. 2019. Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults. *New Media & Society* 21(11-12): 2422-2442.
- Ashok, K. J., Trueman, T. E. & Abinesh, A. K. 2021. Suicidal risk identification in social media. *Procedia CIRP*, hlm. 368-373.
- Bird, S., Klein, E. & Loper, E. 2009. *Natural Language Processing with Python*.
- Birjali, M., Kasri, M. & Beni-Hssane, A. 2021. A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems* 226(107134).
- Chadha, A., Gupta, A. & Kumar, Y. 2022. Suicidal Ideation Detection on Social Media: A Machine Learning Approach. *2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*. Anjuran
- Chai, C. P. 2023. Comparison of text preprocessing methods. *Natural Language Engineering* 29(3): 509-553.
- Chan, S. Y. & Chng, C. K. 2022. Analysis of Risk Factors Affecting Suicidal Ideation among Public University Students in Malaysia using Analytic Hierarchy Process. *ASM Science Journal* 17(
- Chatterjee, M., Kumar, P., Samanta, P. & Sarkar, D. 2022. Suicide ideation detection from online social media: A multi-modal feature based technique. *International Journal of Information Management Data Insights* 2(2): 100103.
- Chatterjee, M., Samanta, P., Kumar, P. & Sarkar, D. 2022. Suicide Ideation Detection using Multiple Feature Analysis from Twitter Data. *2022 IEEE Delhi Section Conference, DELCON 2022*, hlm.
- Clef Initiative. 2017. eRisk 2017: Early Risk Prediction on the Internet: Experimental Foundations. <https://early.irlab.org/2017/index.html> [June 16, 2024].
- Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*
- Galassi, A., Lippi, M. & Torroni, P. 2021. Attention in Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems* 32(10): 4291-4308.
- Ghosal, S. & Jain, A. 2023. Depression and Suicide Risk Detection on Social Media using fastText Embedding and XGBoost Classifier. *Procedia Computer Science* 218(1631-1639).
- Glenn, J. J., Nobles, A. L., Barnes, L. E. & Teachman, B. A. 2020. Can Text Messages Identify Suicide Risk in Real Time? A Within-Subjects Pilot Examination of Temporally Sensitive Markers of Suicide Risk. *Clinical Psychological Science* 8(4): 704-722.
- Grant, J. B., Batterham, P. J., Mccallum, S. M., Werner-Seidler, A. & Cleave, A. L. 2023. Specific anxiety and depression symptoms are risk factors for the onset of suicidal ideation and suicide attempts in youth. *Journal of Affective Disorders* 327(299-305).

- Haque, R., Islam, N., Islam, M. & Ahsan, M. M. 2022. A Comparative Analysis on Suicidal Ideation Detection Using NLP, Machine, and Deep Learning. *Technologies* 10(3):
- Harnedy, L. E., Dreier, M. J. & Selby, E. A. 2023. Assessing positive emotion dampening and suicide risk presence and severity among college students. *Personality and Individual Differences* 214(
- Hutto, C. J. & Gilbert, E. 2015. *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*.
- Ibrahim, N., Che Din, N., Ahmad, M., Amit, N., Ghazali, S. E., Wahab, S., Abdul Kadir, N. B. Y., Halim, F. W. & A. Halim, M. R. T. 2019. The role of social support and spiritual wellbeing in predicting suicidal ideation among marginalized adolescents in Malaysia. *BMC Public Health* 19(S4):
- Ji, S. 2022. Towards Intention Understanding in Suicidal Risk Assessment with Natural Language Processing. *Findings of the Association for Computational Linguistics: EMNLP 2022*, hlm. 4057-4067.
- Kancharapu, R. & A Ayyagari, S. N. 2023. A comparative study on word embedding techniques for suicide prediction on COVID-19 tweets using deep learning models. *International Journal of Information Technology* 15(6): 3293-3306.
- Kim, S.-W. & Gil, J.-M. 2019. Research paper classification systems based on TF-IDF and LDA schemes. *Human-centric Computing and Information Sciences* 9(1): 30.
- Komati, N. 2021. Suicide and Depression Detection. Kaggle.
- Kumar, L., Tummalapalli, S., Rathi, S. C., Murthy, L. B., Krishna, A. & Misra, S. 2023. Machine learning with word embedding for detecting web-services anti-patterns. *Journal of Computer Languages* 75(101207).
- Lew, B., Kölves, K., Lester, D., Chen, W. S., Ibrahim, N. B., Khamal, N. R. B., Mustapha, F., Chan, C. M. H., Ibrahim, N., Siau, C. S. & Chan, L. F. 2021. Looking Into Recent Suicide Rates and Trends in Malaysia: A Comparative Analysis. *Front Psychiatry* 12(770252).
- Liu, Q., Teng, C. C., Sun, I., Muñoz, R. F., Garza, M., Liu, N. H., Barakat, S. & Leykin, Y. 2024. Suicide attempts in the absence of depression: Differences between broad cultural groups. *Journal of Affective Disorders* 356(722-727).
- Liu, X., Huang, J., Yu, N. X., Li, Q. & Zhu, T. 2020. Mediation Effect of Suicide-Related Social Media Use Behaviors on the Association Between Suicidal Ideation and Suicide Attempt: Cross-Sectional Questionnaire Study. *Journal of Medical Internet Research* 22(4): e14940.
- Losada, D., Crestani, F. & Parapar, J. 2017. *eRISK 2017: CLEF Lab on Early Risk Prediction on the Internet: Experimental Foundations*.
- Low, D. M., Rumker, L., Talkar, T., Torous, J., Cecchi, G. & Ghosh, S. S. 2020. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on reddit during COVID-19: Observational study. *Journal of Medical Internet Research* 22(10):
- Macrynika, N., Auad, E., Menjivar, J. & Miranda, R. 2021. Does social media use confer suicide risk? A systematic review of the evidence. *Computers in Human Behavior Reports* 3(100094).
- Malik, M. S. I., Nazarova, A., Jamjoom, M. M. & Ignatov, D. I. 2023. Multilingual hope speech detection: A Robust framework using transfer learning of fine-tuning RoBERTa model. *Journal of King Saud University - Computer and Information Sciences* 35(8): 101736.
- Mohammad, S. 2017. Word Affect Intensities.
- Mohd Fadhli, S. A., Liew Suet Yan, J., Ab Halim, A. S., Ab Razak, A. & Ab Rahman, A. 2022. Finding the Link between Cyberbullying and Suicidal Behaviour among Adolescents in Peninsular Malaysia. *Healthcare* 10(5): 856.

- Parrott, S., Britt, B. C., Hayes, J. L. & Albright, D. L. 2020. Social Media and Suicide: A Validation of Terms to Help Identify Suicide-related Social Media Posts. *Journal of Evidence-Based Social Work (United States)* 17(5): 624-634.
- Pertubuhan Kesehatan Sedunia. 2023. Suicide. <https://www.who.int/news-room/fact-sheets/detail/suicide> [1 November 2023].
- Pham, T. V., Fetter, A. K., Wiglesworth, A., Rey, L. F., Prairie Chicken, M. L., Azarani, M., Riegelman, A. & Gone, J. P. 2022. Suicide interventions for American Indian and Alaska Native populations: A systematic review of prevention strategies, logics, and rationales. *SSM - Mental Health* 2(100139).
- Pusat Kawalan Dan Pencegahan Penyakit. 2023. Facts About Suicide. <https://www.cdc.gov/suicide/facts/index.html> [20 November 2023].
- Rabani, S. T., Ud Din Khanday, A. M., Khan, Q. R., Hajam, U. A., Imran, A. S. & Kastrati, Z. 2023. Detecting suicidality on social media: Machine learning at rescue. *Egyptian Informatics Journal* 24(2): 291-302.
- Raudsepp, L. & Kais, K. 2019. Longitudinal associations between problematic social media use and depressive symptoms in adolescent girls. *Preventive Medicine Reports* 15(100925).
- Renjith, S., Abraham, A., Jyothi, S. B., Chandran, L. & Thomson, J. 2022. An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms. *Journal of King Saud University - Computer and Information Sciences* 34(10): 9564-9575.
- Rozek, D. C., Tyler, H., Fina, B. A., Baker, S. N., Moring, J. C., Smith, N. B., Baker, J. C., Bryan, A. O., Bryan, C. J. & Dondanville, K. A. 2023. Suicide Intervention Practices: What Is Being Used by Mental Health Clinicians and Mental Health Allies? *Archives of Suicide Research* 27(3): 1034-1046.
- Sarsam, S. M., Al-Samarraie, H., Alzahrani, A. I., Alnumay, W. & Smith, A. P. 2021. A lexicon-based approach to detecting suicide-related messages on Twitter. *Biomedical Signal Processing and Control* 65(102355).
- Sedgwick, R., Epstein, S., Dutta, R. & Ougrin, D. 2019. Social media, internet use and suicide attempts in adolescents. *Current Opinion in Psychiatry* 32(6): 534-541.
- Shah, N. 2022. Open source Emoticons and Emoji detection library: emot. PyPI.
- Soumya, K. & Garg, V. K. 2022. Named Entity Emotion Intensity Tagging for Suicidal Ideation Detection From Social Media Texts During MT. *International Journal of Intelligent Engineering and Systems* 15(6): 224-236.
- Steinert, S. & Dennis, M. J. 2022. Emotions and Digital Well-Being: on Social Media's Emotional Affordances. *Philosophy & Technology* 35(2): 36.
- Testoni, I., Piol, S. & De Leo, D. 2021. Suicide Prevention: University Students' Narratives on Their Reasons for Living and for Dying. *International Journal of Environmental Research and Public Health* 18(15): 8029.
- Uban, A.-S., Chulvi, B. & Rosso, P. 2021. An emotion and cognitive based analysis of mental health disorders from social media data. *Future Generation Computer Systems* 124(480-494).
- Valero-Carreras, D., Alcaraz, J. & Landete, M. 2023. Comparing two SVM models through different metrics based on the confusion matrix. *Computers & Operations Research* 152(106131).
- Wan Ismail, W. S., Sim, S. T., Tan, K. A., Bahar, N., Ibrahim, N., Mahadevan, R., Nik Jaafar, N. R., Baharudin, A. & Abdul Aziz, M. 2020. The relations of internet and smartphone addictions to depression, anxiety, stress, and suicidality among public university students in Klang Valley, Malaysia. *Perspectives in Psychiatric Care* 56(4): 949-955.

- Wang, N., Luo, F., Shivtare, Y., Badal, V., Subbalakshmi, K. P., Chandramouli, R. & Lee, E. 2021. Learning Models for Suicide Prediction from Social Media Posts. *Computational Linguistics and Clinical Psychology: Improving Access, CLPsych 2021 - Proceedings of the 7th Workshop, in conjunction with NAACL 2021*, hlm. 87-92.
- Zhang, T., Schoene, A. M. & Ananiadou, S. 2021. Automatic identification of suicide notes with a transformer-based deep learning model. *Internet Interventions* 25(100422).
- Zhang, Y.-Y., Lei, Y.-T., Song, Y., Lu, R.-R., Duan, J.-L. & Prochaska, J. J. 2019. Gender differences in suicidal ideation and health-risk behaviors among high school students in Beijing, China. *Journal of Global Health* 9(1):
- Zirikly, A., Resnik, P., Uzuner, Ö. & Hollingshead, K. 2019. CLPsych 2019 Shared Task: Predicting the Degree of Suicide Risk in Reddit Posts. *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology*

Stuart Heng Fu Yu (A189660)

Dr. Wandeep Kaur a/p Ratan Singh

Fakulti Teknologi & Sains Maklumat

Universiti Kebangsaan Malaysia