PENGESANAN BULI SIBER MENGGUNAKAN PENDEKATAN PEMBELAJARAN MENDALAM

Nurul Aleeya binti Adnan & Salwani Abdullah

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

Abstrak

Dalam meniti arus kemodenan ini, isu buli siber semakin meningkat seiring dengan kemajuan teknologi di Malaysia. Buli siber merupakan satu bentuk jenayah yang menggunakan teknologi digital untuk mengganggu dan menakut-nakutkan seseorang. Kegiatan buli ini juga berlaku melalui pelbagai platform dalam talian terutamanya Twitter. Jika dibiarkan, buli siber ini ia akan memberi pelbagai impak yang negatif kepada mangsa seperti kemurungan, masalah mental dan boleh membawa maut. Kaedah yang sedia ada mengambil masa yang lama dan kurang efisien untuk mengenal pasti buli tersebut. Oleh itu, Kajian ini telah dibina untuk membangunkan satu model yang menggunakan pendekatan pembelajaran mendalam untuk mengenal pasti buli siber. Model yang digunakan ialah Memori Jangka Pendek Panjang (LSTM), Memori Jangka Pendek Panjang Dua Hala (BiLSTM) dan Unit Ulangan Berpintu (GRU). Dataset yang mengandungi teks berkaitan buli siber dipraproses sebelum dilatih dengan setiap model. Prestasi model dinilai menggunakan metrik seperti ketepatan, kepersisan, dapatan semula dan skor F1. Hasil kajian menunjukkan bahawa model GRU memberikan prestasi terbaik dengan skor F1 tertinggi, diikuti oleh BiLSTM dan LSTM. Sistem ini kemudiannya diintegrasikan ke dalam aplikasi web untuk memudahkan pengguna mengesan unsur buli siber secara automatik dan efisien. Secara keseluruhan, kajian ini menyumbang kepada usaha mencegah buli siber melalui teknologi yang bersifat proaktif dan berasaskan kecerdasan buatan.

Kata kunci: Buli Siber, Pembelajaran Mendalam, GRU

Pengenalan

Buli siber merupakan satu bentuk jenayah yang menggunakan teknologi digital untuk mengganggu dan menakut-nakutkan seseorang hingga menimbulkan perasaan malu atau terhina kepada individu tersebut. Buli siber ini boleh berlaku secara langsung atau tidak langsung. Ia berlaku melalui platform-platform seperti Facebook, Twitter, Instagram dan Tiktok. Pihak Kerajaan telah menubuhkan satu Akta Keselamatan Siber 2024 (Akta 854) yang bertujuan meningkatkan tahap keselamatan siber di Malaysia (Majlis Keselamatan Negara, 2024). Menurut satu kajian statistik oleh World Health Organization (WHO), ianya telah membuktikan bahawa 1 daripada 6 kanak-kanak sekolah telah terdedah dan menjadi mangsa buli siber. Punca utama mengapa buli siber ini berlaku belum dikenal pasti tetapi ada yang mengatakan bahawa individu yang melakukan buli siber berani melakukannya kerana mereka rasa lebih berkuasa berlindung disebalik paparan untuk mereka mendapat kepentingan tertentu. Buli siber lebih berbahaya berbanding buli secara fizikal kerana ianya boleh dilakukan secara tidak diketahui. Jika buli siber ini terus berlaku, ia akan memberi impak yang negatif kepada mangsa seperti gangguan mental, kemurungan dan jika lebih teruk akan membawa kepada maut.

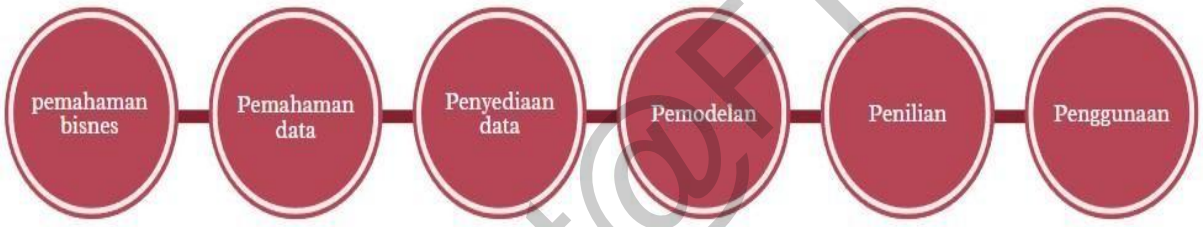
Untuk mengatasi masalah yang diajukan dalam projek ini, pendekatan pembelajaran mendalam akan digunakan untuk menganalisis corak komunikasi dalam talian dan mengenal pasti tingkah laku yang menunjukkan buli siber. Dengan membangunkan model pengesanan yang tepat, ia dapat membantu sekolah dan institusi pendidikan dalam mengenal pasti dan mengatasi buli siber lebih awal, serta memberi sokongan yang diperlukan kepada mangsa. Data-data seperti komen di aplikasi twitter digunakan sebagai sumber untuk mengesan buli siber. Masalah buli siber semakin membimbangkan dan memberi kesan jangka panjang terhadap kesihatan mental mangsa. Memandangkan komunikasi digital berlaku dalam jumlah besar dan sukar dipantau secara manual, kaedah berasaskan kecerdasan buatan sangat diperlukan untuk mengesan dan mencegah insiden buli secara proaktif. Sistem yang dibangunkan juga boleh digunakan sebagai rujukan untuk penambahbaikan sistem sokongan di institusi pendidikan dan agensi berkaitan.

Sebelum melatih data, prapemprosesan data perlu dijalankan seperti pembersihan data dan penaskalaan data. Platfrom yang akan digunakan adalah *Google Colab*. Setelah proses prapemprosesan data serta proses data latih dan uji, proses ini diteruskan lagi dengan menggunakan model pembelajaran mendalam seperti Memori Jangka Pendek Panjang (LSTM), Memori Jangka Pendek Panjang Dua Hala (BiLSTM) dan Unit Ulangan Berpintu (GRU). Hasil daripada permodelan tersebut akan divisualisasikan dalam bentuk graf dan model pembelajaran mendalam akan dibandingkan untuk mencari model terbaik. Objektif kajian ini adalah untuk membangunkan model pembelajaran mendalam menggunakan data penghantaran teks di platform twitter untuk mengesan buli siber. Bagi mencapai matlamat tersebut, beberapa objektif telah disenaraikan. Objektif kajian ini adalah untuk melaksanakan pra-pemprosesan teks untuk memastikan data sesuai digunakan bagi latihan model, Membangunkan model pembelajaran mendalam seperti LSTM, BiLSTM dan GRU untuk mengenal pasti sama ada teks tersebut mengandungi unsur buli siber atau tidak dan menilai peratusan ketepatan model melalui data latih dan uji untuk mengukur keberkesanan prestasi model pembelajaran mendalam.

Set data yang akan digunakan ialah data yang diperoleh dari platfrorm Kaggle. https://www.kaggle.com/datasets/andrewmvd/cyberbullying-classification. Data ini merupakan pengumpulan dari platform twitter yang berkaitan dengan buli siber. Data tersebut mempunyai teks dan label seperti bukan buli siber, manakala untuk buli siber, ia mempunyai lima kelas iaitu umur, agama, etnik, jantina dan lain-lain buli. Model pembelajaran mendalam yang akan dibangunkan ialah LSTM, BiLSTM dan GRU. Metodologi CRISP-DM telah dipilih kerana fleksibiliti dan keupayaannya dalam pemprosesan data yang kompleks. Ia merangkumi enam fasa iaitu pemahaman perniagaan, pemahaman data, penyediaan data, pemodelan, penilaian dan penggunaan. Ketiga-tiga model akan dibangunkan, diuji dan dibandingkan untuk mengenal pasti model paling berkesan.

Metodologi Kajian

Metod yang digunakan dalam projek ini ialah *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Metod ini telah diterbitkan pada 1999. Keboleh upayaan metodologi ini yang fleksibel dan berguna telah menjadi metod ini terkenal untuk perlombongan data dan analisis data. Terdapat 6 fasa di dalam CRISP-DM iaitu fasa pemahaman perniagaan, fasa pemahaman data, fasa penyediaan data, fasa pemodelan, fasa penilaian serta fasa penggunaan. Fasa- fasa di model ini juga fleksibel dan tidak perlu mengikut turutan kerana ia boleh mengundur untuk melangkau ke fasa yang lain serta boleh mengulang aktiviti pada fasa yang sama.



Rajah 1 Metod CRISP-DM. Sumber: (Hotz 2022)

Langkah pertama dalam CRISP-DM ialah fasa pemahaman terhadap perniagaan. Fasa ini membantu memastikan projek perlombongan data difokuskan untuk memenuhi keperluan dan objektif perniagaan atau organisasi. Pada fasa ini, matlamat utama adalah untuk memahami isu yang dihadapi dan bagaimana penyelesaian automatik dapat membantu. Dalam kes ini, masalah yang ingin ditangani adalah peningkatan buli siber di media sosial, terutamanya di Twitter, yang sering mengandungi ucapan kebencian yang disasarkan kepada individu, kumpulan, atau entiti tertentu. Selain itu, fasa pemahaman data ialah langkah kedua CRISP-DM, dan ia tertumpu pada penerokaan dan menganalisis data yang dikumpul untuk mendapatkan pemahaman yang lebih baik tentang kandungan, struktur dan kualitinya. Fasa ini melibatkan pemahaman terhadap data yang akan digunakan dalam sistem. Dalam kes ini, data yang digunakan adalah teks yang diambil dari Twitter. Data ini mungkin mengandungi teks yang tidak berstruktur, termasuk pelbagai jenis perkataan, simbol dan perkataan yang tidak relevan atau *noise* seperti pautan, tanda baca, atau nama pengguna. Oleh itu, data perlu dibersihkan dan diproses dengan teknik seperti tokenisasi untuk memudahkan analisis selanjutnya.

Fasa penyediaan data merupakan tahap kritikal dalam pembinaan sistem pembelajaran mendalam, khususnya dalam pengesanan buli siber berdasarkan teks. Tujuannya adalah untuk menyediakan set data yang bersih, berstruktur, dan sesuai digunakan oleh model yang telah dipilih. Berdasarkan set data cyberbullying_tweets.csv, beberapa proses pembersihan dan transformasi data telah dilakukan seperti penerokaan struktur data, pembersihan teks, pembuangan kata pendua, normalisasi, serta tokenisasi untuk penyediaan input kepada model. Proses pembersihan dilakukan untuk membuang data pendua, menghapuskan mesej kosong, serta membersihkan teks menggunakan regex, pembuangan emoji, dan pembuangan kata henti. Langkah-langkah yang dilaksanakan ialah menukar semua huruf kepada huruf kecil, membuang tanda nama, URL, simbol khas dan tanda baca, membuang perkataan berhenti, menggunakan stemming untuk normalisasi kata asas, membuang mesej yang kurang daripada 4 perkataan dan membuang kata pendua.

Seterusnya, *WordCloud* digunakan untuk memvisualkan perkataan paling kerap muncul bagi setiap kelas *bully_type. Countplot* pula memaparkan pengedaran bilangan tweet mengikut panjang ayat selepas pembersihan. Visualisasi ini membantu dalam mengenal pasti konteks perkataan dominan serta panjang purata mesej bagi setiap kelas. Akhir sekali, teks yang telah dibersihkan akan ditukar kepada urutan berangka melalui proses tokenisasi menggunakan *Tokenizer* dari *Keras*. Kemudian, urutan ini akan dipendekkan atau dipanjangkan menggunakan fungsi *pad_sequences* bagi memastikan semua input mempunyai panjang yang konsisten sebelum dimasukkan ke dalam model. Semua langkah ini digabungkan membentuk pipeline prapemprosesan yang lengkap dan sesuai untuk tugasan pengesanan buli siber dalam platform media sosial. Data yang akan digunakan dibahagikan kepada data latih dan data uji dengan nisbah 70:30.

Fasa yang seterusnya ialah pemodelan. Fasa pemodelan dalam penyelidikan ini melibatkan pembangunan dan ujian model pembelajaran mendalam untuk mengenal pasti corak dan hubungan dalam data. Fasa pemodelan akan menggunakan pelbagai model untuk mencari satu model yang

paling bagus dan berkesan dalam pengesanan buli siber ini. Proses ini bermula dengan menggunakan set data latihan untuk melatih model, diikuti dengan mengujinya untuk menilai keberkesanannya. Model-model yang akan digunakan dalam penyelidikan ini ialah Memori Jangka Pendek Panjang (LSTM), Memori Jangka Pendek Panjang Dua Hala (BiLSTM) dan Unit Ulangan Berpintu (GRU).

Rangkaian Neural Berulang (RNN), yang kita kenali sebagai kelas rangkaian neural buatan dengan sambungan antara nod, mempunyai beberapa kekangan disebabkan oleh jumlah lapisan rangkaian yang berlebihan. Ini menjadikan ia suatu tugas yang mencabar, namun satu kajian terkini menunjukkan bahawa rangkaian Memori Jangka Pendek-Panjang (LSTM) merupakan penyelesaian yang efektif bagi cabaran ini kerana struktur berangkainya yang menyerupai modul rangkaian neural pelbagai dalam RNN (Arshad, 2019). LSTM terdiri daripada pelbagai pintu kawalan, termasuk pintu input, pintu *output*, dan pintu lupa yang digunakan dalam model LSTM. Pintu-pintu ini digunakan untuk menentukan bagaimana maklumat diterima dan ditolak sepanjang rangkaian. Model yang digunakan dalam kajian ini ialah model LSTM yang dibina menggunakan perpustakaan *Keras Sequential API*. Model ini hanya mengimbas data dalam satu arah secara berurutan. Struktur model yang telah dibina terdiri daripada Lapisan *Embedding* yang memetakan setiap perkataan kepada vektor berdimensi 100, berdasarkan 10,000 perkataan teratas daripada tokenisasi. Dua lapisan LSTM berturutan, masing-masing dengan 64 dan 10 neuron. Lapisan ini digunakan untuk menangkap kebergantungan jangka panjang dalam data teks. Beberapa lapisan *Dense* selepas LSTM, 256, 128, 64 dan 32 neuron dengan fungsi pengaktifan ReLU. Satu lapisan *Dropout* dengan kadar 0.3 digunakan untuk mengurangkan masalah *overfitting*. Akhir sekali, lapisan *output Dense* dengan 5 neuron dan fungsi pengaktifan softmax digunakan untuk melakukan klasifikasi pelbagai kelas bagi lima jenis buli siber.

Seterusnya, model BiLSTM mengekalkan dua laluan input berasingan dan mengendalikan keadaan input ke hadapan yang dkalkan oleh dua LSTM yang berbeza. LSTM pertama berfungsi sebagai

urutan biasa yang memproses dari awal perenggan, manakala LSTM kedua memproses input dalam urutan yang bertentangan. Konsep utama rangkaian dua hala ini adalah untuk mengumpul maklumat daripada konteks sekeliling input iaitu sebelum dan selepas bagi memberikan pemahaman yang lebih menyeluruh terhadap data. Kaedah dua hala ini biasanya memperoleh maklumat dengan lebih cepat berbanding pendekatan sehala, namun keberkesanannya masih bergantung kepada jenis tugasan. Model BiLSTM yang digunakan dalam kajian ini dibina menggunakan rangka kerja *Keras Sequential API* dan mengandungi beberapa lapisan penting yang dirangka bagi menangkap maklumat urutan secara dua hala dengan lebih efektif. Struktur model terdiri daripada dua lapisan BiLSTM yang mempunyai 64 neuron dan mengembalikan urutan penuh bagi membolehkan pemprosesan berlapis, kemudian 10 neuron untuk menyempurnakan pemhaman konteks dari dua arah. Kemudian, Tiga lapisan padat ditambah iaitu 256 neuron dengan fungsi pengaktifan ReLU, 128 neuron, 64 neuron dan 32 neuron. Seterusnya, Tiga lapisan *dropout* digunakan di antara lapisan-lapisan utama dengan kadar 0.3, dan 0.4 secara berturutan. Akhir sekali, lapisan output menggunakan fungsi pengaktifan *softmax* dengan 5 neuron.

GRU telah digunakan sebagai varian terkini bagi RNN yang direka untuk menangani masalah ingatan jangka pendek, serupa seperti LSTM. Namun, berbeza dengan LSTM yang mempunyai *cell state*, GRU hanya menggunakan keadaan tersembunyi untuk membawa maklumat sepanjang urutan. Di mana zt mewakili pintu kemas kini, yang berfungsi sama seperti gabungan pintu lupa dan pintu input dalam LSTM. Ia bertanggungjawab untuk menentukan maklumat mana yang perlu disimpan atau dibuang. rt ialah pintu reset, yang menentukan berapa banyak maklumat daripada masa lalu yang perlu dilupakan. σ adalah fungsi sigmoid yang menghasilkan nilai antara 0 hingga 1. w, U, dan b masing-masing ialah parameter matriks dan vektor. ht ialah vektor output pada masa t, dan xt ialah vektor input pada masa t. Disebabkan GRU mempunyai struktur yang lebih ringkas berbanding LSTM, ia membolehkan proses latihan model berjalan dengan lebih cepat dan efisien, tanpa

mengorbankan keupayaan untuk memahami konteks urutan dalam data teks seperti dalam tugas pengesanan buli siber.

Selain itu, Fasa Penilaian dibahagikan kepada tiga tugas utama iaitu menilai hasil, mengkaji semula proses dan menentukan langkah seterusnya. Ia menilai sejauh mana model memenuhi objektif perniagaan dan menguji model pada aplikasi ujian jika masa dan kos mencukupi. Selain itu, ia juga melakukan kajian semula yang lebih mendalam terhadap keseluruhan proses perlombongan data untuk mengenal pasti faktor atau tugas penting yang mungkin telah terlepas, mengenal pasti masalah jaminan kualiti dan merumuskan kajian semula proses serta aktiviti yang terlepas atau perlu diulang. Bagi menentukan Langkah seterusnya, Ia menilai cara untuk meneruskan projek. Dalam bahagian ini, ianya penting untuk menyenaraikan tindakan selanjutnya bersama alasan bagi setiap pilihan serta cara untuk melangkah kehadapan. Dalam konteks penilaian model atau pengesanan buli siber, istilah seperti *True Positive (TP), False Positive (FP), True Negative (TN), dan False Negative (FN)* adalah penting untuk menilai prestasi model dalam klasifikasi. Ini merujuk kepada bagaimana model mengklasifikasikan data dan mengukur ketepatan keputusan yang diambil berdasarkan data yang diuji. Istilah ini digunakan dalam matriks kekeliruan, yang membantu memahami bagaimana model membuat keputusan dan memberi gambaran tentang kesalahan yang mungkin berlaku.

Akhir sekali, fasa Penggunaan adalah langkah terakhir dalam proses CRISP-DM. Dalam konteks kajian pengesanan buli siber ini, fasa ini merujuk kepada penggunaan model pengesanan buli yang telah dibangunkan dan diuji untuk mengklasifikasikan mesej atau komunikasi dalam talian. Walau bagaimanapun, dalam kajian ini, model pengesanan buli siber yang dibangunkan tidak digunakan sepenuhnya dalam fasa penggunaan kerana kajian ini merupakan kajian awal yang bertujuan untuk menilai keberkesanan model dalam mengenal pasti buli siber dengan lebih tepat. Di masa hadapan, penambahbaikan pada model ini diharapkan dapat diterapkan dalam platform sosial untuk membantu dalam usaha pencegahan dan pengurusan buli siber dengan lebih efektif.

Keputusan dan Perbincangan

Bahagian ini membentangkan hasil penilaian matriks untuk model klasifikasi, membandingkan keputusan penilaian matriks model LSTM, BiLSTM dan GRU yang dipaparkan dalam bentuk jadual dan rajah.

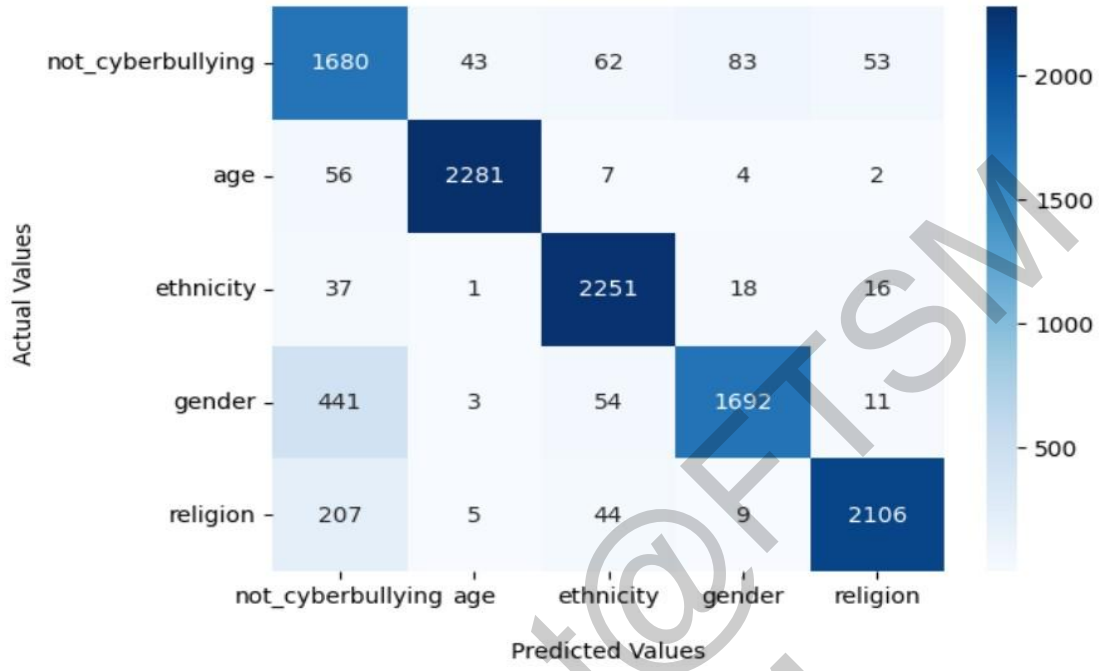
Penilaian Model Memori Jangka Pendek Panjang (LSTM)

Jadual 1 Keputusan penilaian matriks model LSTM

Label	**Penilaian Matriks**			
	Kepersisan	Dapatan Semula	Skor F1	Sokongan
not_cyberbullying	0.69	0.87	0.77	1921
age	0.98	0.97	0.97	2350
ethnicity	0.93	0.97	0.95	2323
gender	0.94	0.77	0.84	2201
religion	0.96	0.89	0.92	2371
		Kejituan: 0.90		

Jadual 1 menunjukkan keputusan penilaian matriks bagi model LSTM dalam pengesanan teks buli siber. Model ini telah mencapai kejituan keseluruhan sebanyak 0.90, menunjukkan bahawa 90% daripada data ujian telah berjaya diklasifikasikan dengan betul. Dari segi analisis matriks individu, prestasi tertinggi dicatatkan bagi label *age* dan *ethnicity* dengan skor F1 masing-masing sebanyak 0.97 dan 0.95, menunjukkan kebolehan model dalam mengenalpasti kedua-dua kategori dengan tepat dan konsisten. Walaubagaimanapun, label *not_cyberbullying* mencatatkan kepersisan yang rendah iaitu 0.69, walaupun nilai dapatan semulanya adalah tinggi. Ini menunjukkan bahawa model cenderung memberikan terlalu banyak klasifikasi positif untuk label ini, menyebabkan peningkatan dalam kes positif palsu. Begitu juga, label *gender* mempunyai skor F1 yang lebih rendah berbanding kategori lain, iaitu 0.84, disebabkan oleh nilai dapatan semula yang lebih rendah, kemungkinan model menghadapi kesukaran untuk mengesan kes buli siber berkaitan jantina secara menyeluruh. Secara keseluruhannya, model LSTM menunjukkan prestasi yang memuaskan dan stabil dalam kebanyakan

kategori, namun masih terdapat ruang untuk penambahbaikan terutama dalam meningkatkan kepersisan klasifikasi bagi label *not_cyberbullying* dan *gender*.



Rajah 2 Matriks Kekeliruan model LSTM

Rajah di atas menunjukkan matriks kekeliruan bagi model LSTM dalam pengelasan teks kepada lima kategori buli siber dan satu kategori bukan buli siber. Matriks ini memberikan gambaran prestasi model dalam membezakan antara label sebenar dan label yang diramalkan. Model menunjukkan ketepatan yang tinggi dalam mengklasifikasikan kategori *age*, dengan 2,281 daripada 2,350 sampel berjaya diklasifikasikan dengan betul. Begitu juga dengan label *ethnicity* dan *religion*, masingmasing mempunyai 2,251 dan 2,106 ramalan betul, mencerminkan keberkesanan model dalam mengesan kategori yang berkaitan. Walau bagaimanapun, kelemahan ketara dapat dilihat dalam klasifikasi label *gender* dan *not_cyberbullying*. Bagi kategori *gender*, hanya 1,692 sampel daripada 2,201 diklasifikasikan dengan betul, manakala sejumlah besar iaitu 441 sampel telah disalahklasifikasikan sebagai *not_cyberbullying*. Fenomena ini menjelaskan nilai dapatan semula yang rendah dalam matriks penilaian bagi label *gender*, yang juga mencerminkan kemungkinan kekeliruan model dalam membezakan antara teks neutral dan teks berunsur buli siber kelas *gender*. Begitu juga, bagi label

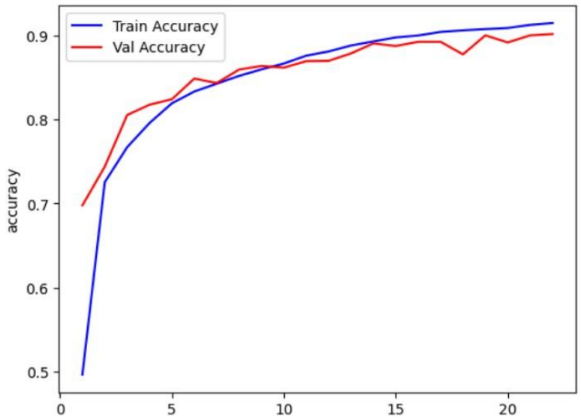
not_cyberbullying, hanya 1,680 daripada 1,921 sampel diklasifikasikan dengan tepat, manakala sejumlah yang ketara telah diklasifikasikan secara salah kepada kategori lain seperti *gender* (83 kes) dan *ethnicity* (62 kes). Ini menunjukkan bahawa model masih mengalami kesukaran dalam membezakan antara kandungan yang benar-benar bukan buli siber dan kandungan buli siber yang tersirat.

Jadual 2 Keputusan penilaian matriks model LSTM

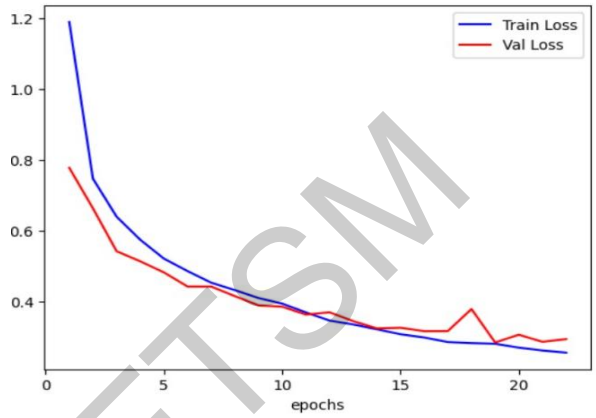
Epoch	Kehilangan Latihan	Kejituan Latihan	Kehilangan Ujian	Kejituan Ujian
1	0.2869	0.9037	0.3778	0.8772
2	0.2826	0.9076	0.2833	0.8998
3	0.2665	0.9081	0.3052	0.8914
4	0.2670	0.9081	0.2852	0.8998
5	0.2488	0.9166	0.2928	0.9014

Jadual 4.5 menunjukkan perubahan nilai kehilangan dan kejituan bagi proses latihan dan ujian model LSTM sepanjang lima epoch. Jumlah epoch yang dijalankan untuk LSTM ini ialah 22, yang dianalisis ialah epoch 5 terakhir, ianya berhenti kerana fungsi *early stop*. Pada epoch pertama, nilai kehilangan latihan adalah sebanyak 0.2869 dengan kejituan 90.37%, manakala kehilangan ujian adalah 0.3778 dengan kejituan 87.72%. Sepanjang epoch kedua hingga kelima, model menunjukkan penambahbaikan yang stabil dalam kedua-dua matriks. Pada epoch kedua, kehilangan ujian menurun dengan ketara kepada 0.2833 manakala kejituan ujian meningkat kepada 89.98%, mencerminkan keupayaan model dalam belajar dan menyesuaikan diri dengan data. Walaupun terdapat sedikit naik turun kecil pada epoch ketiga dan keempat, nilai-nilai tersebut kekal konsisten dan tidak menunjukkan tanda-tanda *overfitting* yang ketara. Pada epoch kelima, model mencatatkan prestasi terbaik dengan kehilangan latihan paling rendah iaitu 0.2488 dan kejituan latihan tertinggi iaitu 91.66%. Dalam masa yang sama, kehilangan ujian mencatatkan nilai 0.2928 dengan kejituan 90.14%, menunjukkan bahawa

model mampu mengekalkan prestasi baik pada data yang tidak pernah dilihat semasa latihan. Ini menggambarkan bahawa model LSTM berjaya melakukan generalisasi dengan berkesan ke atas data ujian serta menunjukkan kestabilan prestasi setelah beberapa epoch.



Rajah 3 Graf Kejituan Latihan dan Ujian LSTM



Rajah 4 Graf Kehilangan Latihan dan Ujian LSTM

Rajah 3 dan Rajah 4 menunjukkan corak kejituan dan kehilangan model LSTM sepanjang 22 epoch semasa proses latihan dan penilaian. Dalam Rajah 3, garis biru mewakili ketepatan data latihan manakala garis merah mewakili ketepatan data penilaian. Model menunjukkan peningkatan yang konsisten dalam kejituan sepanjang proses latihan. Ketepatan penilaian juga meningkat secara progresif pada peringkat awal dan mengekalkan kestabilan sekitar 90% selepas epoch ke-10. Perkara ini menunjukkan bahawa model tidak menghadapi isu *underfitting* atau *overfitting* yang ketara, malah berjaya mempelajari corak data dengan baik. Rajah 4 pula menggambarkan nilai kehilangan bagi kedua-dua set latihan dan penilaian. Kehilangan bagi data latihan menunjukkan penurunan mendadak dari awal hingga sekitar epoch ke-15, menandakan model sedang belajar dengan efektif. Kehilangan data penilaian juga menurun selari dengan kehilangan latihan, namun terdapat sedikit peningkatan kecil dalam beberapa epoch terakhir, yang boleh menunjukkan kemungkinan model mula mengalami *overfitting* yang ringan. Walau bagaimanapun, secara keseluruhan, penurunan nilai kehilangan yang stabil dan sejajar antara latihan dan penilaian menunjukkan prestasi yang baik bagi model LSTM.

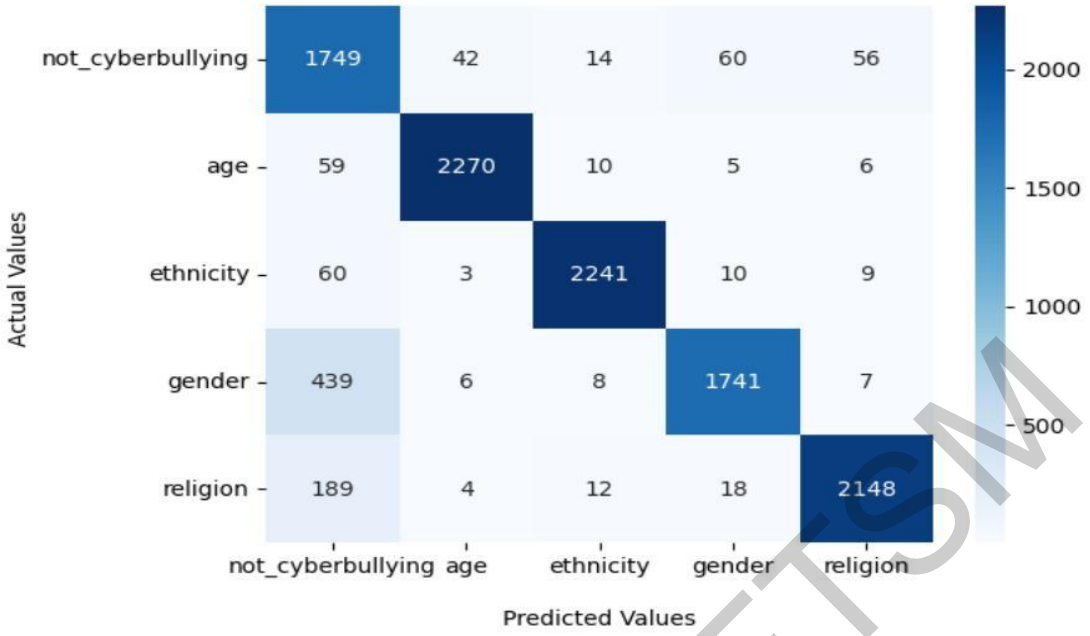
Penilaian Model Memori Jangka Pendek Panjang Dua Hala (BiLSTM)

Jadual 3 menunjukkan keputusan penilaian matriks bagi model BiLSTM, termasuk kepersisan, dapatan semula, skor F1 dan juga sokongan keputusan klasifikasi model bagi lima jenis kategori.

Jadual 3 Keputusan penilaian matriks model BiLSTM

Label	**Penilaian Matriks**			
	Kepersisan	Dapatan Semula	Skor F1	Sokongan
not_cyberbullying	0.70	0.91	0.79	1921
age	0.98	0.97	0.97	2350
ethnicity	0.98	0.96	0.97	2323
gender	0.95	0.79	0.86	2201
religion	0.96	0.91	0.93	2371
Kejituan: 0.91				

Model BiLSTM ini menunjukkan bahawa dalam kategori *age, ethnicity, dan religion*, prestasi yang tinggi telah dicapai iaitu dengan nilai F1 melebihi 0.93 bagi setiap kelas. Ini menunjukkan bahawa model ini sangat berkesan dalam mengesan teks yang berkaitan dengan bentuk buli siber berdasarkan umur, etnik dan agama. Bagi kategori *not_cyberbullying* mencatat kepersisan paling rendah iaitu 0.70, walaupun nilai dapatan semula agak tinggi iaitu sebanyak 0.91. Ini menunjukkan bahawa model kerap mengelirukan teks biasa sebagai buli siber, namun masih mampu mengenalpasti kebanyakan teks yang bukan buli siber. Manakala Prestasi pada kategori *gender* pula agak sederhana, dengan skor F1 sebanyak 0.86, disebabkan dapatan semula yang lebih rendah iaitu 0.79. Ini mungkin menandakan bahawa model masih sukar untuk mengesan semua kes buli siber berkaitan jantina. Secara kejituan keseluruhan, Model BiLSTM mencapai kejituan keseluruhan sebanyak 91%, menunjukkan prestasi umum yang baik dalam mengklasifikasikan teks kepada lima kelas dengan betul.



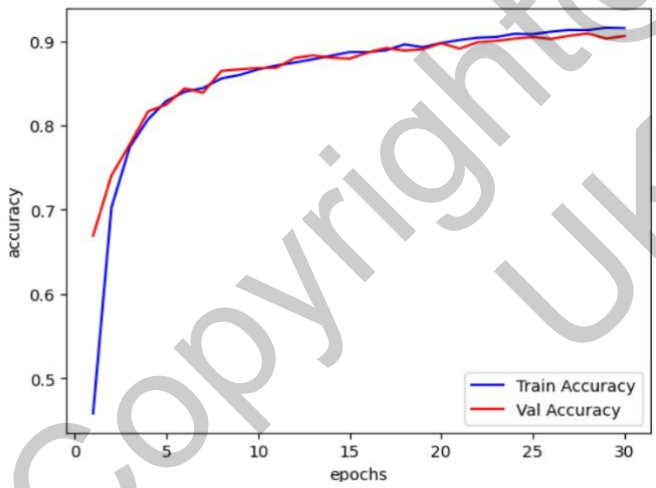
Rajah 5 Matriks Kekeliruan model BiLSTM

Matriks kekeliruan bagi model BiLSTM menunjukkan keupayaan model dalam mengklasifikasikan lima jenis buli siber. Secara keseluruhan, model menunjukkan prestasi tinggi dalam mengenal pasti buli berdasarkan umur (2,270 betul), etnik (2,241 betul), dan agama (2,148 betul). Namun, terdapat kelemahan ketara dalam pengesanan buli berdasarkan jantina, di mana sejumlah besar data (439 teks) tersalah diklasifikasikan sebagai bukan buli siber. Kesilapan turut berlaku bagi kelas *not_cyberbullying,* dengan sejumlah 172 teks diramal sebagai kelas buli lain. Ini menunjukkan bahawa walaupun model mempunyai kejituan keseluruhan yang baik, ia masih perlu dipertingkatkan dalam mengenal pasti ayat yang lebih neutral terutamanya berkaitan jantina.

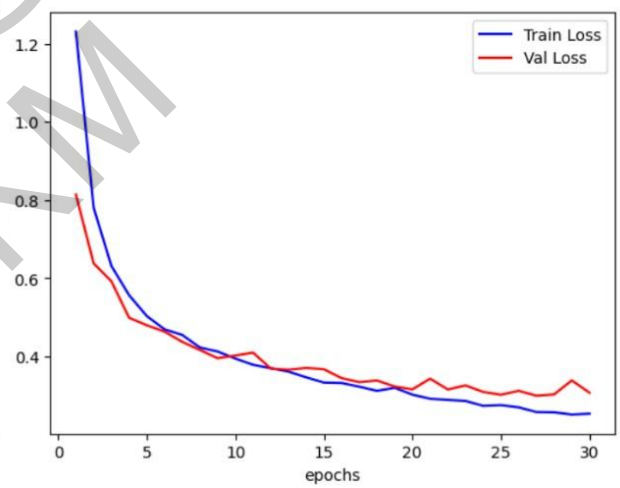
Jadual 4 Kehilangan dan Kejituan Latihan dan Ujian Model BiLSTM

Epoch	Kehilangan Latihan	Kejituan Latihan	Kehilangan Ujian	Kejituan Ujian
1	0.2724	0.9079	0.3118	0.9025
2	0.2617	0.9110	0.2993	0.9064
3	0.2570	0.9115	0.3026	0.9091
4	0.2540	0.9137	0.3384	0.9029
5	0.2466	0.9167	0.3069	0.9060

Jadual 4 menunjukkan prestasi model BiLSTM sepanjang epoch untuk diuji. Jumlah epoch bagi ujian BiLSTM ini sebenarnya tiga puluh dengan menggunakan fungsi *early stop* untuk mengelakkan *overfit*. Pada jadual di atas, hanya epoch ke 25 – 30 yang di ambil untuk analisis. lima epoch terakhir dari segi kehilangan dan kejituan bagi set latihan dan set ujian. Pada epoch pertama, model mencatatkan kehilangan latihan sebanyak 0.2724 dengan kejituan 90.79%, manakala kehilangan ujian adalah 0.3118 dengan kejituan 90.25%. Sepanjang epoch kedua hingga kelima, kehilangan latihan secara umum semakin berkurangan, menunjukkan bahawa model semakin mempelajari dengan baik daripada data latihan. Kejituan latihan juga meningkat secara konsisten, mencapai 91.67% pada epoch kelima. Bagi set ujian pula, prestasi model adalah stabil dengan kejituan sekitar 90% hingga 91%, menandakan bahawa model tidak mengalami *overfitting* yang ketara. Walaupun terdapat sedikit peningkatan nilai kehilangan ujian pada epoch keempat, kejituan masih kekal tinggi.



Rajah 6 Graf Kejituan Latihan dan Uji BiLSTM



Rajah 7 Graf Kehilangan Latihan dan Uji BiLSTM

Rajah prestasi model BiLSTM semasa latihan menunjukkan peningkatan ketara dalam ketepatan sepanjang 30 epoch. Ketepatan latihan bermula sekitar 46% dan meningkat dengan konsisten melepasi 90%, menunjukkan bahawa model berjaya mempelajari corak dalam data dengan baik. Pada masa yang sama, nilai kehilangan latihan menurun secara signifikan daripada lebih 1.2 kepada sekitar 0.25, yang menunjukkan bahawa ralat ramalan model semakin berkurang. Corak ini menandakan proses pembelajaran yang stabil tanpa isu *overfitting* pada data latihan. Rajah prestasi pada set ujian

juga menunjukkan kenaikan positif yang konsisten dengan data latihan. Ketepatan meningkat dengan baik dan berjaya mencecah lebih 90%, menunjukkan bahawa model mampu mengadaptasi dengan baik terhadap data baharu. Nilai kehilangan ujian pula menurun daripada sekitar 0.8 ke bawah 0.4, namun mula menunjukkan sedikit peningkatan menjelang akhir epoch. Ini mungkin menandakan awal tanda-tanda overfitting, tetapi masih dalam julat terkawal. Secara keseluruhan, model BiLSTM menunjukkan prestasi yang kukuh dalam kedua-dua fasa latihan dan ujian.

Penilaian Model Unit Ulangan Berpintu (GRU)

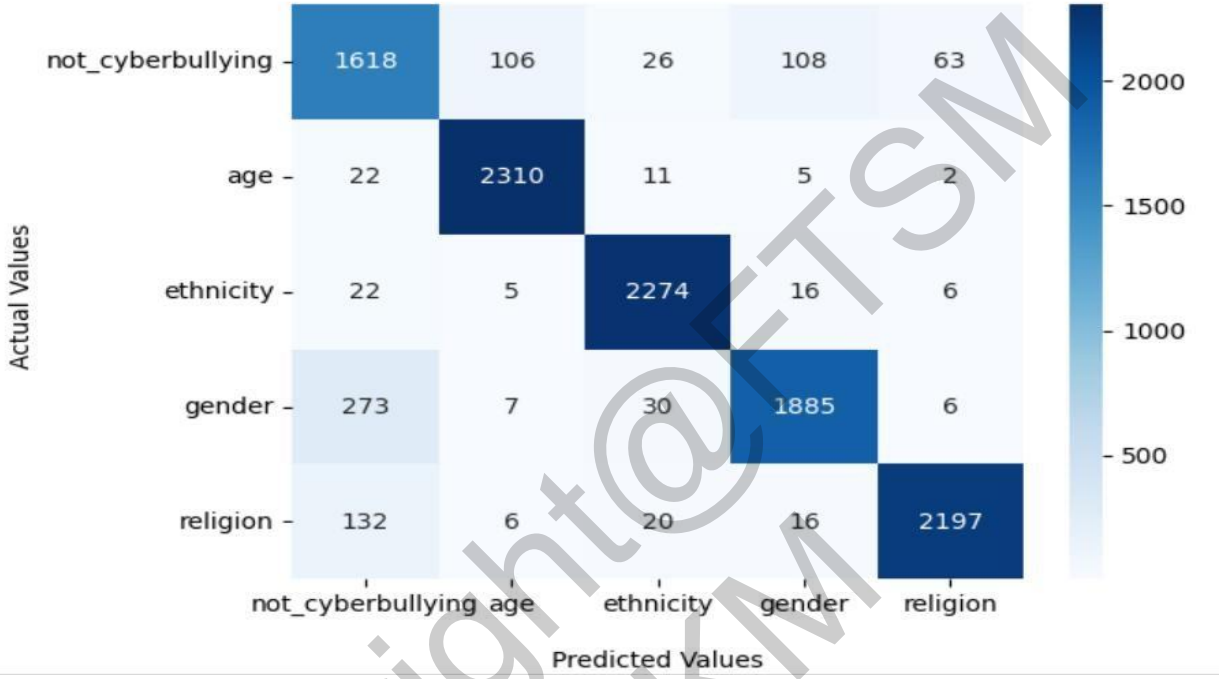
Jadual 4.8 menunjukkan keputusan penilaian matriks bagi model GRU, termasuk kepersisan, dapatan semula, skor F1 dan juga sokongan keputusan klasifikasi model.

Jadual 5 Keputusan penilaian matriks model GRU

Label	**Penilaian Matriks**			
	Kepersisan	Dapatan Semula	Skor F1	Sokongan
not_cyberbullying	0.78	0.84	0.81	1921
age	0.95	0.98	0.97	2350
ethnicity	0.96	0.98	0.97	2323
gender	0.93	0.86	0.89	2201
religion	0.97	0.93	0.95	2371
Kejituan: 0.92				

Model GRU ini telah menunjukkan prestasi yang paling baik antara model lain dalam klasifikasi teks berkaitan jenis buli siber. Berdasarkan Jadual 4.8, model mencatatkan kejituan keseluruhan sebanyak 92%, menandakan kebolehan yang sangat baik dalam membezakan antara lima kategori. Secara terperinci, model memperoleh skor F1 tertinggi sebanyak 0.97 bagi kategori *age* dan *ethnicity*, menunjukkan keseimbangan yang kukuh antara kepekaan dan kejituan dalam mengenal pasti jenis buli berdasarkan umur dan etnik. Kategori *religion* turut menunjukkan prestasi cemerlang dengan skor F1 sebanyak 0.95. Walaupun kategori *not_cyberbullying* menunjukkan prestasi yang lebih rendah berbanding kategori lain (0.81), kepekaannya yang tinggi menandakan model masih berjaya

mengenal pasti kebanyakan teks yang bukan berbentuk buli. Secara keseluruhan, matriks penilaian ini membuktikan bahawa model GRU berupaya untuk menangani tugas pengelasan berbilang kelas dengan ketepatan dan keseimbangan yang baik, menjadikannya sesuai untuk digunakan dalam sistem pengesanan buli siber secara automatik.



Rajah 8 Matriks Kekeliruan model GRU

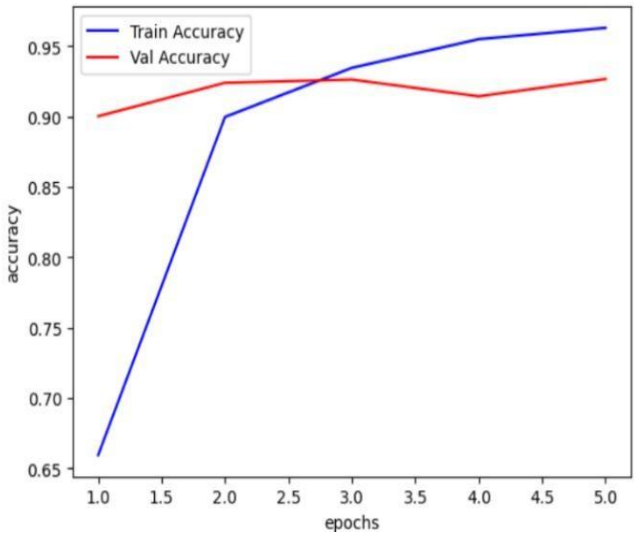
Gambar rajah di atas menunjukkan matriks kekeliruan bagi model GRU, yang menggambarkan prestasi klasifikasi model terhadap lima kategori utama iaitu *not_cyberbullying*, *age*, *ethnicity*, *gender*, dan *religion*. Secara keseluruhan, model telah menunjukkan prestasi klasifikasi yang sangat baik dengan bilangan ramalan tepat yang tinggi bagi setiap kelas. Sebagai contoh, sebanyak 2310 sampel kategori *age* telah dikelaskan dengan betul, begitu juga dengan 2274 bagi *ethnicity*, dan 2197 bagi *religion*, menunjukkan bahawa model sangat berkesan dalam mengenal pasti jenis buli siber berdasarkan umur, etnik, dan agama. Namun begitu, terdapat kekeliruan yang ketara dalam kategori *gender* dan *not_cyberbullying*, di mana sebanyak 273 teks dari *gender* telah tersalah diklasifikasikan sebagai *not_cyberbullying*, dan 108 teks dari *not_cyberbullying* telah diramalkan sebagai *gender*. Situasi ini mungkin berlaku disebabkan oleh persamaan kandungan bahasa dalam kategori tersebut.

Walaupun terdapat beberapa kesilapan klasifikasi silang, secara umumnya matriks kekeliruan ini mengesahkan bahawa model GRU mampu menjalankan pengelasan pelbagai kelas dengan tahap ketepatan yang tinggi, serta menunjukkan keupayaan mengenal pasti corak yang kompleks dalam data teks buli siber.

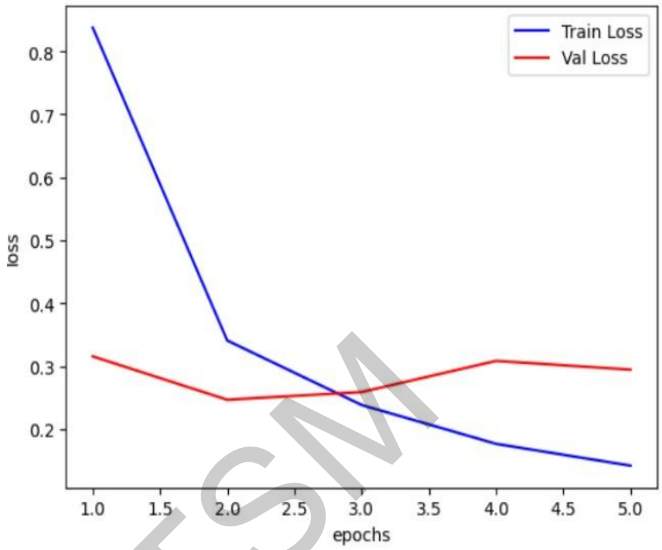
Jadual 6 Kehilangan dan Kejituan Latihan dan Ujian Model GRU

Epoch	Kehilangan Latihan	Kejituan Latihan	Kehilangan Ujian	Kejituan Ujian
1	1.1696	0.4910	0.3155	0.9002
2	0.3645	0.8902	0.2466	0.9240
3	0.2405	0.9345	0.2588	0.9263
4	0.1696	0.9576	0.3083	0.9144
5	0.1405	0.9649	0.2945	0.9267

Jadual 4.9 menunjukkan nilai kehilangan dan kejituan bagi set latihan dan ujian untuk model GRU sepanjang lima epoch latihan. Pada permulaan Latihan, nilai kehilangan latihan adalah tinggi iaitu 1.1696 dengan kejituan hanya 49.10%, menandakan model masih dalam proses pembelajaran awal. Namun, berlaku peningkatan ketara selepas epoch ke-2 apabila kehilangan menurun kepada 0.3645 dan kejituan meningkat kepada 89.02%. Progres ini berterusan sehingga epoch ke-5, di mana model mencapai kehilangan serendah 0.1405 dan kejituan setinggi 96.49% dalam set latihan, menunjukkan bahawa model telah belajar dengan baik daripada data. Bagi set ujian pula, kejituan kekal stabil dan tinggi sepanjang lima epoch, bermula dari 90.02% (epoch 1) sehingga mencapai 92.67% (epoch 5), walaupun terdapat sedikit turun naik pada nilai kehilangan. Secara keseluruhannya, jadual ini membuktikan bahawa model GRU bukan sahaja mampu belajar corak dalam data dengan baik, tetapi juga menunjukkan keupayaan penyesuaian yang baik terhadap data ujian, tanpa tandatanda *overfitting* yang ketara dalam tempoh latihan tersebut.



Rajah 9 Graf Kejituan Latihan dan Ujian GRU



Rajah 10 Graf Kehilangan Latihan dan Ujian GRU

Rajah ini memaparkan perbandingan prestasi kejituan model GRU bagi set data latihan dan set data pengesahan sepanjang lima epoch latihan. Garis berwarna biru mewakili kejituan pada set latihan, manakala garis merah menunjukkan kejituan pada set pengesahan. Prestasi model meningkat dengan ketara dari epoch pertama hingga epoch ketiga, dengan kejituan latihan mencapai lebih 95% dan kejituan pengesahan kekal sekitar 92%. Kestabilan garis pengesahan yang tidak menunjukkan penurunan mendadak menandakan model tidak mengalami isu *overfitting* yang serius, malah berjaya mencapai tahap generalisasi yang baik. Pada rajah kedua yang menggambarkan perubahan nilai kehilangan sepanjang proses latihan bagi model GRU. Garis biru mewakili kehilangan pada set latihan, manakala garis merah mewakili kehilangan pada set pengesahan. Dari epoch pertama ke epoch ketiga, kedua-dua nilai kehilangan menunjukkan penurunan yang konsisten, mencerminkan penyesuaian model yang baik terhadap data. Namun, selepas epoch ketiga, kehilangan pengesahan menunjukkan sedikit peningkatan walaupun kehilangan latihan terus menurun. Ini berkemungkinan menunjukkan permulaan kepada *overfitting*, di mana model menjadi terlalu spesifik terhadap data latihan. Justeru, fungsi seperti *early stopping* adalah penting untuk mengelakkan penurunan prestasi ke atas data baru. Penilaian visual ini menunjukkan bahawa model GRU berjaya mencapai prestasi yang sangat baik dalam masa yang singkat, namun pemantauan berterusan diperlukan bagi mengelakkan *overfitting* sekiranya latihan diteruskan.

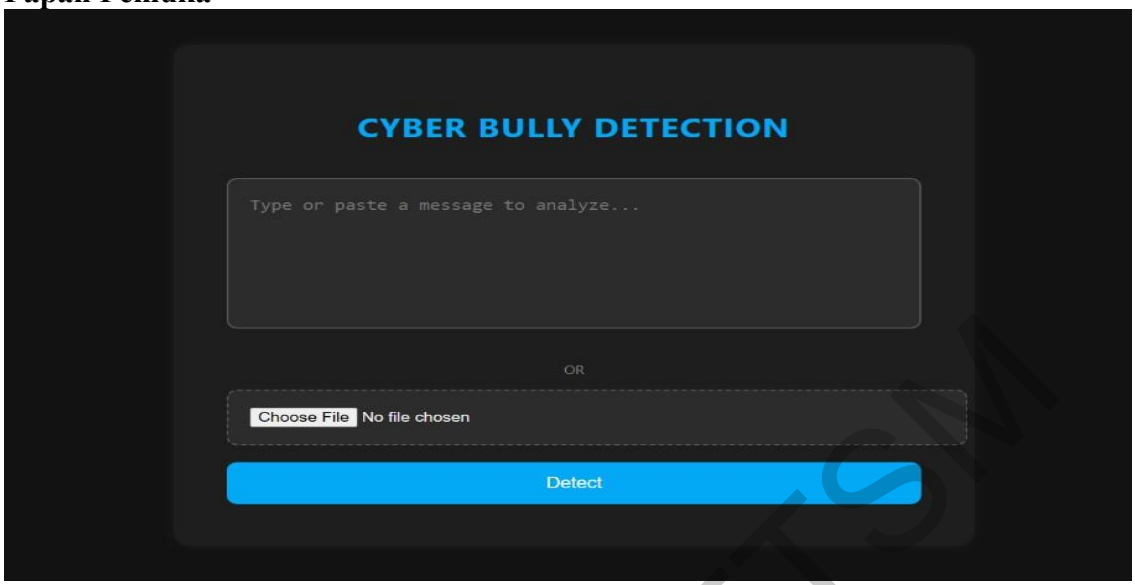
Perbandingan Model secara Keseluruhan

Jadual 7 Perbandingan keputusan model klasifikasi

Model	Penilaian Matriks				Ranking
	Kepersisan	Dapatan Semula	Kejituan	Skor F1	
GRU	0.92	0.92	0.92	0.92	1
BiLSTM	0.91	0.91	0.91	0.90	2
LSTM	0.90	0.90	0.90	0.89	3

Secara keseluruhan, perbandingan antara ketiga-tiga model pembelajaran mendalam yang digunakan, iaitu LSTM, BiLSTM dan GRU menunjukkan bahawa model GRU memberikan prestasi terbaik dalam tugas pengesanan buli siber. Berdasarkan penilaian matriks yang melibatkan kepersisan, dapatan semula, kejituan dan skor F1, GRU mencatatkan kejituan tertinggi iaitu 0.92, mengatasi BiLSTM (0.91) dan LSTM (0.90). GRU juga menunjukkan prestasi paling seimbang dalam mengklasifikasikan kelima-lima label, terutamanya dalam kelas *not_cyberbullying* dan *gender* yang sering menjadi cabaran kepada model lain. Manakala BiLSTM pula hampir menyamai prestasi GRU dalam label *age* dan *ethnicity* tetapi sedikit menurun dari segi dapatan semula untuk kelas *gender*. Model LSTM menunjukkan prestasi yang stabil, namun kurang cemerlang jika dibandingkan dengan GRU dan BiLSTM. Dari segi kecekapan, GRU juga lebih ringan berbanding LSTM dan BiLSTM, menjadikannya pilihan yang lebih sesuai untuk digunakan dalam sistem sebenar seperti aplikasi web atau mudah alih. Oleh itu, model GRU dipilih sebagai model terbaik dalam projek ini kerana ia memberikan ketepatan klasifikasi tertinggi, prestasi yang seimbang berbanding semua kelas, dan kecekapan dari aspek pengiraan serta pelaksanaan.

Papan Pemuka



Rajah 11 papan pemuka sistem pengesanan buli siber

Rajah 12 Papan pemuka apabila contoh teks buli siber jenis jantina dimasukkan

Rajah 13 Papan pemuka apabila contoh teks bukan buli siber dimasukkan

Kesimpulan

Secara keseluruhannya, projek ini membuktikan bahawa pendekatan pembelajaran mendalam berpotensi besar dalam membantu mengesan buli siber secara automatik dan berkesan. Dengan menggunakan model seperti LSTM, BiLSTM dan GRU, sistem mampu mengenal pasti teks yang mengandungi unsur buli berdasarkan pola bahasa yang dianalisis. Hasil kajian menunjukkan bahawa model GRU memberikan prestasi terbaik dari segi ketepatan dan skor F1, sekali gus memperkukuhkan keberkesanan pendekatan ini dalam aplikasi sebenar. Sistem yang dibangunkan telah berjaya diintegrasikan ke dalam aplikasi web bagi membolehkan pengguna mengesan unsur buli siber dengan lebih mudah dan cepat. Kajian ini bukan sahaja menyumbang kepada penyelesaian masalah sosial yang semakin teruk, malah menunjukkan bagaimana teknologi kecerdasan buatan boleh dimanfaatkan untuk memberi amaran awal, sokongan kepada mangsa, dan mencegah berlakunya insiden yang lebih serius.

Namun begitu, kekangan seperti ketidakseimbangan kelas data, kesukaran mengenal pasti mesej sarkastik, keperluan sumber pemprosesan tinggi dan kekangan sumber data daripada satu platform masih menjadi cabaran dalam meningkatkan kebolehgunaan sistem secara menyeluruh. Untuk masa depan, penggunaan model NLP moden seperti BERT atau RoBERTa boleh dipertimbangkan bagi pemahaman konteks lebih baik. Isu ketidakseimbangan data boleh diatasi dengan teknik seperti SMOTE. Selain itu, meluaskan aplikasi kepada pelbagai bahasa dan platform akan menjadikan sistem ini lebih inklusif dan bersifat global. Dengan itu, projek ini diharap dapat menjadi asas kepada pembangunan sistem pengesanan buli siber yang lebih proaktif dan komprehensif pada masa akan datan

Penghargaan

Segala puji dan syukur bagi Allah, Yang Maha Pengasih lagi Maha Penyayang. Saya dengan rendah hati menadah tangan tanda kesyukuran, terharu kerana telah menyiapkan projek tahun akhir saya dengan bimbingan dan rahmat yang dikurniakan kepada saya. Saya ingin merakamkan setinggi-tinggi penghargaan kepada penyelia saya, Prof. Dr. Salwani Abdullah, atas bimbingan, sokongan dan dorongan sepanjang saya menyiapkan usulan projek ini sebagai penyelia. Kepakaran dan pandangan beliau yang tidak ternilai telah memainkan peranan penting dalam membentuk hala tuju dan hasil usulan projek saya. Saya amat berterima kasih atas masa dan usaha yang beliau telah korbankan untuk saya dan kerja saya. Saya tidak akan dapat menyiapkan usulan projek ini tanpa bimbingan beliau. Terima kasih, Prof. Dr. Salwani.

Saya juga ingin mengucapkan terima kasih juga kepada para pensyarah dari FTSM yang telah menaburkan ilmu pengetahuan kepada saya sepanjang pengajian saya di Universiti Kebangsaan Malaysia (UKM). Selain itu, Saya ingin merakamkan ribuan terima kasih kepada ahli keluarga saya kerana sentiasa memberi sokongan dan semangat serta sentiasa berdoa terhadap kejayaan saya di universiti. Akhir sekali, terima kasih kepada semua yang terlibat secara langsung atau tidak langsung dalam menghasilkan projek tahun akhir ini. Saya juga ingin memohon maaf sekiranya terdapat kesilapan sepanjang perlaksanaan projek akhir tahun ini.

Sekian, terima kasih.

RUJUKAN

Altayeva, A., Sultan, D., Abdrakhmanov, R., Tolep, A., & Toktarova, A. (2024). Hybrid LSTM–CNN model for detecting cyberbullying in social media text. *CEUR Workshop Proceedings*.

Arshad, S. (2019, September 21). Sentiment analysis / text classification using RNN (bi-LSTM). *Medium*.

Daineko, J., Moghimi, M., & Bernard, J. (2022). Cyberbullying detection solutions based on deep learning architectures. *Journal of Intelligent & Fuzzy Systems, 43*(3), 3073–3087.

Daraghmi, D., Al-Qurashi, A., & Ghaleb, B. (2024). An integrated CNN–BiLSTM– GRU model for Arabic cyberbullying detection. *Neural Computing and Applications*.

DataScience-PM. (n.d.). CRISP-DM: The complete data science methodology. *DataSciencePM*.

FutureLearn. (2022, October 25). Updates, insights, and news from FutureLearn | Online learning for you. *FutureLearn*.

Hasan, M. T., Hossain, M. A. E., Mukta, M. S. H., Akter, A., Ahmed, M., & Islam, S. (2023). A review on deep-learning-based cyberbullying detection. *Future Internet, 15*(5), 179.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-term Memory. *Neural Computation, 9*, 1735–1780.

Hotz, N. (2024, December 9). What is CRISP-DM? *Data Science PM*.

Kumar, R., & Bhat, A. (2022). A study of machine learning-based models for detection, control, and mitigation of cyberbullying in online social media. *International Journal of Information Security, 21*(6), 1409–1431.

Luna, Z. (2022, January 5). CRISP-DM Phase 5: Evaluation phase. *Analytics Vidhya Medium*.

Maity, K., Bhattacharya, S., Saha, S., & Seera, M. (2023). A deep learning framework for Malay hate speech detection using BiLSTM and RNN. *Neural Computing and Applications*.

Majlis Keselamatan Negara. (2024, August 26). Penguatkuasaan Akta Keselamatan Siber 2024 (Akta 854). *MKN*.

Mukhopadhyay, D., Mishra, K., Mishra, K., & Tiwari, L. (2021). Cyber bullying detection based on Twitter dataset. In *Machine Learning for Predictive Analysis* (pp. 87–94). Springer Singapore.

Ojha, M., Patil, N. M., & Joshi, M. (2024, November 19). Cyberbullying detection and prevention using machine learning. *Grenze International Journal of Engineering and Technology*.

Paul, S., Saha, S., & Hasanuzzaman, M. (2020). Identification of cyberbullying: A deep learning based multimodal approach. *Multimedia Tools and Applications*.

Sanner, M. F. (1999). Python: A programming language for software integration and development. *Journal of Molecular Graphics and Modelling, 17*(1), 57–61.

Schmidt, T. (2019). Recurrent neural networks: A comprehensive review of architectures, variants, and applications. *MDPI*.

Sharma, H. K., Kshitiz, K., & Shailendra. (2018). NLP and machine learning techniques for detecting insulting comments on social networking platforms. In *2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)*.

Taye, M. M. (2023). Theoretical understanding of convolutional neural networks: Concepts, architectures, applications, future directions. *Computers, 11*(3), 52.

Tranung, K., & Tranung, K. (2024, August 27). Penguatkuasaan Akta Keselamatan Siber 2024 (Akta 854). *Laman Web MKN*.

Unnava, S., & Parasana, S. R. (2024). A study of cyberbullying detection and classification techniques: A machine learning approach. *Engineering, Technology & Applied Science Research, 14*(4), 15607–15613.

Vallejo, W., Díaz-Uribe, C., & Fajardo, C. (2022). Google Colab and virtual simulations: Practical e-learning tools to support the teaching of thermodynamics and to introduce coding to students. *ACS Omega, 7*(8), 7421– 7429.

Vietnamese Authors. (2019). Hate speech detection on Vietnamese social media text using Bi-GRU-LSTM-CNN. *arXiv preprint*.

Wang, J., Fu, K., & Lu, C.-T. (2020). SOSNet: A graph convolutional network approach to fine-grained cyberbullying detection. In *2020 IEEE International Conference on Big Data (Big Data)*.

World Health Organization. (2024, March 27). One in six school-aged children experiences cyberbullying, finds new WHO Europe study. *WHO Europe*.

Zhang, S., Li, M., & Yan, C. (2022). The empirical analysis of Bitcoin price prediction based on deep learning integration method. *Computational Intelligence and Neuroscience, 2022*, Article ID 1265837.

Nurul Aleeya binti Adnan (A195935)
Prof Dr. Salwani Abdullah
Fakulti Teknologi & Sains Maklumat,
Universiti Kebangsaan Malaysia