PENGESANAN JENAYAH SIBER DALAM KALANGAN REMAJA MENGGUNAKAN PEMBELAJARAN MENDALAM

[1]Intan Humaira Binti Mohammad Fyaizul, [2]Rohizah Abd Rahman

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

Abstrak

Jenayah siber merujuk kepada sebarang aktiviti jenayah yang dilakukan melalui teknologi digital seperti peranti pintar, dan internet. Terdapat pelbagai jenis jenayah siber seperti penggodaman, penipuan dalam talian, kecurian identiti, dan penyebaran perisian hasad. Perkara ini memberi kesan dari segi kewangan, reputasi, dan ancaman keselamatan terhadap mangsa. Penglibatan ini memberi kesan buruk kepada pembangunan sahsiah, emosi, dan intelektual remaja, sekaligus mengancam potensi mereka. Objektif utama kajian ini bertujuan untuk mengesan dan menganalisis jenayah siber yang melibatkan remaja menggunakan algoritma pembelajaran mendalam. Metodologi kajian yang akan digunakan adalah berpandukan model Cross Industry Standard Process for Data Mining (CRISP-DM). Pembangunan model pengesanan ini akan merangkumi enam fasa utama iaitu pemahaman bisnes, pemahaman data, penyediaan data, pembangunan model, penilaian dan penyebaran. Kajian ini akan dijalankan menggunakan bahasa pengaturcaraan Python. Terdapat rangkuman empat langkah utama: (1) menjalankan proses pembersihan dan analisis eksplorasi data; (2) pemilihan model pembelajaran mendalam yang sesuai; (3) membangunkan model menggunakan pembelajaran mendalam; dan (4) membina antara muka untuk menganalisis prestasi model dan teknik menggunakan StreamLit. Projek ini menggunakan teknik algoritma pembelajaran mendalam seperti Recurrent Neural Networks (RNN), Long Short- Term Memory (LSTM) dan Bidirectional LSTM (BiLSTM) yang bertujuan untuk pembangunan model dan pemprosesan bahasa semula jadi bagi membantu dalam pengesanan jenayah siber dalam kalangan remaja. Hasil daripada kajian ini menunjukkan bahawa model LSTM mencatatkan model yang terbaik berdasarkan nilai ketetapan dan *F1-Score* yang tinggi berbanding model yang lain iaitu dengan nilai 0.68% dan 0.28%. Oleh sebab itu, model LSTM telah dipilih untuk digunakan dalam proses pengesanan jenayah siber melalui antara muka pengguna *Streamlit* bagi tujuan pengujian data. Kesimpulannya, projek ini diharapkan dapat memberikan sumbangan yang bermakna dalam membangunkan sistem yang mampu mengesan jenayah siber dengan lebih berkesan. Tambahan, model yang dibangunkan diharap boleh digunapakai untuk algoritma pembelajaran mendalam terkini yang dipertingkatkan pada masa akan datang.

Kata kunci: Jenayah siber, Pembelajaran Mesin, Pembelajaran Mendalam

Copyright@FTSM UKM

Abstract

Cybercrime is any criminal activity through digital technology, such as smart devices and the internet. There are different types of cybercrime, such as hacking, online fraud, identity theft, and the spread of malware. Cybercrime has an impact in terms of financial, reputational, and security threats to the victim. This involvement adversely affects the adolescent's personality, emotional, and intellectual development, thus threatening their potential. This study aims to detect and analyze cybercrime involving adolescents using deep learning algorithms. The research methodology is based on the Cross-Industry Standard Process for Data Mining (CRISP-DM) model. The development of this tracking model will include six primary phases: business understanding, data understanding, data preparation, model development, evaluation, and dissemination. This study will be conducted using Python. There is a summary of four main steps: (1) conducting a data cleansing and exploratory analysis process; (2) selection of an appropriate Deep Learning model; (3) developing a model using deep learning; and (4) building an interface to analyze the performance of models and techniques using StreamLit. The project uses deep learning algorithmic techniques such as Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Bidirectional LSTM (BiLSTM) aimed at the development of models and natural language processing to assist in the detection of cybercrime among adolescents. The findings of this study show that the LSTM model achieved the best performance based on its higher accuracy and F1-Score compared to the other models, with values of 0.68 and 0.28 respectively. Therefore, the LSTM model was chosen for use in the cybercrime detection process via the Streamlit user interface for data testing purposes. In conclusion, this project is expected to make a meaningful contribution toward developing a system capable of detecting cybercrime more effectively. Additionally, it is hoped that the developed model can be adapted for use with improved deep learning algorithms in the future.

Keywords: Cybercrime, Machine Learning, Deep Learning

Copyright@FTSM UKM

1.0 PENGENALAN

Jenayah siber dalam kalangan remaja di Malaysia semakin membimbangkan seiring dengan perkembangan teknologi dan penggunaan internet yang meluas. Proses globalisasi dan perkembangan tamadun moden telah membawa kepada peralihan masyarakat industri kepada masyarakat maklumat. Dunia siber kini menjadi sebahagian penting dalam kehidupan manusia moden. Ia berperanan besar dan telah mengubah banyak aspek kehidupan. Namun begitu, selain membawa kesan positif, dunia siber juga boleh memberi impak negatif kepada masyarakat,

khususnya golongan remaja yang mudah terdedah kepada aktiviti jenayah (Malek & Kamil 2010).

Selain itu, daya fikir yang turut berkembang menyebabkan lahirnya sebuah pengetahuan yang meluas. Namun, tidak semua orang mampu memanfaatkan pengetahuan tersebut dengan bijak, sehingga perkara ini boleh mendatangkan kerugian kepada banyak pihak. Antara isu utama yang berlaku termasuk buli siber, penipuan dalam talian, penggodaman akaun peribadi, eksploitasi seksual, dan ketagihan permainan dalam talian. Buli siber melibatkan penghinaan, ancaman, atau penyebaran maklumat palsu yang mencemarkan maruah seseorang melalui media sosial atau aplikasi mesej. Hal ini boleh memberi kesan psikologi seperti tekanan emosi, kemurungan, dan ketakutan sosial. Selain itu, penipuan dalam talian sering melibatkan remaja yang menjadi mangsa skim palsu, seperti pembelian barangan yang tidak wujud atau penipuan kewangan melalui laman e-dagang. Penggodaman akaun pula berlaku apabila penggodam mencuri data peribadi atau menggunakan akaun media sosial mangsa untuk menyebarkan kandungan yang tidak wajar. Dalam kes yang lebih serius, terdapat remaja yang menjadi mangsa eksploitasi seksual, di mana gambar atau video sensitif digunakan untuk ugutan atau disebarkan tanpa kebenaran. Tambahan pula, ketagihan permainan dalam talian juga mendorong remaja melakukan penipuan seperti bersubahat melakukan jenayah penipuan dalam talian yang akhirnya membawa kepada kesan sosial pada awal usia remaja (Othman 2025).

Dalam menangani isu ini, pembelajaran mendalam menawarkan penyelesaian yang berpotensi besar membantu dalam pengecaman masalah ini. Pembelajaran mendalam, yang merupakan cabang kecerdasan buatan (Artificial Intelligence - AI), menggunakan Rangkaian Neural Berlapis untuk menganalisis data yang kompleks dan mengenal pasti pola tertentu. Sebagai contoh, sistem keselamatan yang dikuasakan oleh pembelajaran mendalam boleh mengesan aktiviti mencurigakan dalam akaun pengguna, seperti log masuk dari lokasi luar biasa atau pola penggunaan yang menyerupai aktiviti penipuan. Dalam pencegahan penipuan dalam talian, model seperti Recurrent Neural Networks (RNN) juga dapat mengenal pasti transaksi yang tidak normal berdasarkan corak sejarah pengguna, sekali gus membantu melindungi remaja daripada menjadi mangsa. Selain itu, ia juga boleh digunakan untuk memantau interaksi permainan dalam talian bagi mengesan tanda-tanda ketagihan atau tingkah laku yang tidak wajar.

Copyright@FTSM UKM

Oleh itu, laporan teknikal ini disusun kepada beberapa bahagian utama iaitu kajian literatur, metodologi kajian, hasil kajian, kesimpulan, penghargaan, dan rujukan. Selepas itu, bahagian kajian literatur mengumpulkan analisis daripada penyelidikan terdahulu yang berkaitan dengan topik pengesanan jenayah siber. Metodologi kajian menerangkan kaedah serta pendekatan yang digunakan sepanjang pelaksanaan kajian ini. Hasil kajian pula mempersembahkan penemuan yang diperoleh. Kesimpulan merumuskan keseluruhan kajian dengan memberikan ringkasan keputusan serta implikasinya. Bahagian penghargaan menyatakan ucapan terima kasih kepada individu, kumpulan, atau organisasi yang telah memberikan bantuan dan sokongan sepanjang kajian dijalankan. Akhir sekali, rujukan menyenaraikan semua sumber dan bahan yang dirujuk dalam projek ini.

2.0 KAJIAN LITERATUR

Kajian-kajian terdahulu menunjukkan bahawa penggunaan teknik pembelajaran mesin dan pembelajaran mendalam memainkan peranan penting dalam pengesanan jenayah siber dengan keberkesanan yang tinggi. Kajian oleh Al-Khater et al. (2020) menyediakan pandangan menyeluruh terhadap pelbagai teknik seperti Hidden Markov Model, Bayesian Network, dan Outlier Detection Algorithm. Mereka turut menguji teknik pembelajaran mesin seperti Naive Bayes, KNN, dan K-Means serta teknik pembelajaran mendalam seperti CNN, LSTM, BiLSTM dan FeedForward Neural Network. Dapatan menunjukkan bahawa model pembelajaran mendalam mampu mencapai ketepatan melebihi 90%, berbanding model lain seperti Decision Tree dan Instance-Based Learner yang hanya mencatat sekitar 78.5%. Seterusnya, Roy et al. (2022) menjalankan kajian ke atas pengesanan pancingan melalui URL menggunakan model LSTM, BiLSTM dan GRU. BiLSTM mencatatkan ketepatan tertinggi iaitu 99.0%, membuktikan kemampuannya dalam menangani isu vanishing gradient serta mengenal pasti URL yang mencurigakan secara lebih tepat. Kajian ini turut menilai prestasi menggunakan metrik seperti precision, recall, F1-score dan confusion matrix.

Dalam kajian Alharbi et al. (2024), LSTM digunakan untuk mengesan akaun Instagram palsu dengan dua set data yang menunjukkan ketepatan sebanyak 97.42% dan 94.21%. Penilaian lanjut menunjukkan bahawa penggunaan optimizer seperti Adam menghasilkan prestasi terbaik berbanding pengoptimum lain seperti SGD, RMSprop dan Adagrad. Keputusan ini memperlihatkan bagaimana pemilihan pengoptimum yang tepat dapat meningkatkan

Copyright@FTSM UKM

keberkesanan model mengikut jenis data. Kajian oleh Amer et al. (2022) pula menilai pelbagai algoritma NLP ke atas 60,001 data tweet jenayah siber. Model BiLSTM konsisten mencatat ketepatan tertinggi (hingga 97.20%), diikuti oleh LSTM dan CNN, manakala model seperti SVM dan Count Vectorization memberikan keputusan yang kurang memuaskan. KNN dan Random Forest menunjukkan prestasi terbaik dalam pembelajaran mesin apabila digabungkan dengan TF-IDF. Akhir sekali, kajian oleh Biodoumoye et al. (2023) memfokuskan kepada pengesanan jenayah siber dengan menggunakan model seperti DistilBERT, LSTM dan transformer BERT. Model DistilBERT muncul sebagai model paling berkesan dengan ketepatan 98.73%, disusuli oleh BERT dan LSTM. Proses pemodelan melibatkan pembersihan metadata laman web defaced serta pengoptimuman menggunakan TensorFlow Lite dan strategi pemilihan hyperparameter. Kajian ini juga menekankan kepentingan teknik pemprosesan data seperti pengendalian nilai hilang dan pengurangan ciri bagi meningkatkan prestasi sistem.

Secara keseluruhan, kesemua kajian menunjukkan bahawa model pembelajaran mendalam terutamanya BiLSTM, LSTM, GRU dan DistilBERT amat sesuai untuk pengesanan dan pemprofilan jenayah siber. Pemilihan model, teknik pra-pemprosesan data serta pengoptimum yang bersesuaian adalah faktor utama dalam menentukan ketepatan dan keberkesanan sistem pengesanan. Maka, pendekatan berasaskan deep learning dilihat sangat relevan dan berpotensi besar dalam menangani ancaman jenayah siber masa kini. Jadual 1.0 menunjukkan perbandingan kajian pengesanan menggunakan pembelajaran mendalam dan pembelajara mesin.

Copyright@FTSM UKM

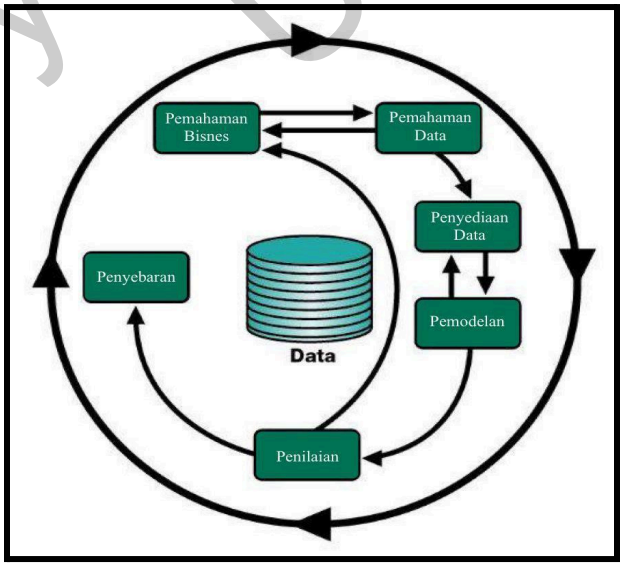
Jadual 1.0 Perbandingan kajian pengesanan menggunakan pembelajaran mendalam dan pembelajaran mesin

Artikel	Algoritma	Rumusan	Ketetapan (%)
Al-Khater et al. 2020	Multilayer Feed Foward	Fokus kepada penggunaan teknik pelbagai dalam pengesanan	95.00
	ANN		99.93
	SVM		98.00
			97.80
Roy et al. 2022	LSTM	Kajian pengesanan pancingan melalui URL	96.90
	Bi-LSTM		99.00
	GRU		97.50
Alharbi et al. 2024	LSTM	Pengesanan akaun palsu dan perbandingan penggunaan optimizer	99.00
	K-Means		97.98
	LR		96.20

Amer et al. 2022	CNN (Word2Vec)	Pengesanan menggunakan pelbagai algoritma NLP keatas data tweet jenayah siber	94.16
	LSTM (Word2Vec)		96.00
	Bi-LSTM (Word2Vec)		97.01
	GRU (Word2Vec)		93.90
	CNN (FastText)		95.00
	LSTM (FastText)		96.14
	Bi-LSTM (FastText)		97.20
	GRU (FastText)		95.19
Biodoumoye et al. 2023	DistilBert	Pengesanan jenayah siber melibatkan pembersihan metadata laman web defaced	98.73
	LSTM		89.63
	BERT		92.57

3.0 METADOLOGI

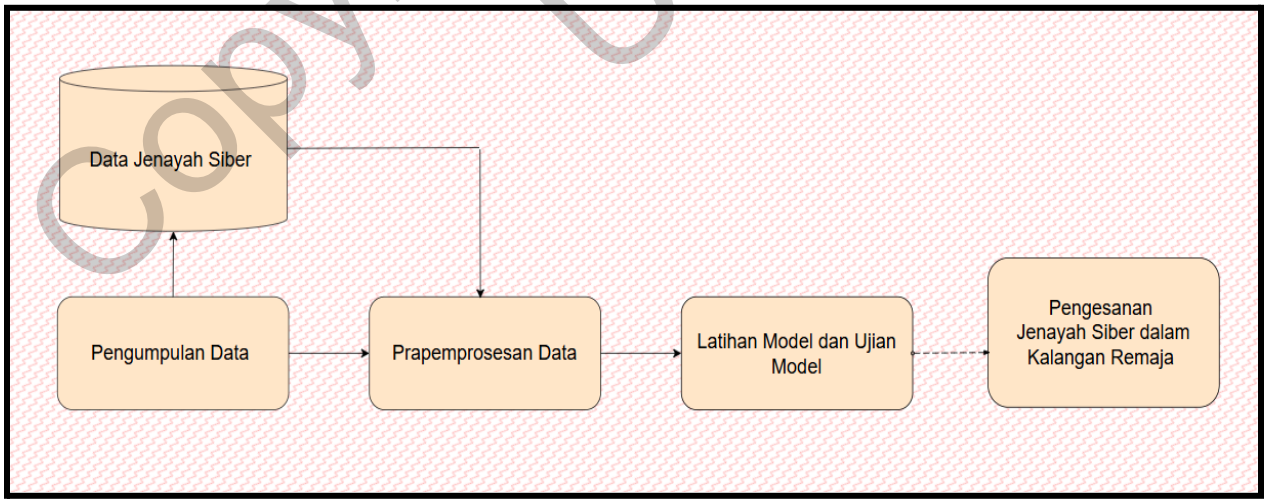
Projek ini akan dijalankan berdasarkan metodologi yang dikenali sebagai model Cross Industry Standard Process for Data Mining (CRISP-DM). Model ini merupakan satu kitaran hayat dan rujukan utama untuk pelbagai projek perlombongan data yang melibatkan 6 fasa utama iaitu pemahaman bisnes, pemahaman data, penyediaan data, permodelan, penilaian, dan penyebaran. Selain itu, model ini tidak berakhir pada fasa penyebaran tetapi akan melalui kitaran semula untuk proses penambahbaikan projek itu sendiri (Ghaedi 2018). Rajah 1.0 menunjukkan gambaran visual seni bina model CRISP-DM yang terlibat dalam projek ini.



Rajah 1.0 Seni Bina Model CRISP-DM
(Sumber: Ghaedi 2018)

Copyright@FTSM UKM

Proses pembangunan projek ini merangkumi enam fasa utama berdasarkan pendekatan standard dalam perlombongan data. Fasa pertama, Pemahaman Bisnes, bertujuan memahami objektif projek dari perspektif jenayah siber dalam kalangan remaja dan merangka pelan awal. Seterusnya, Fasa Pemahaman Data melibatkan eksplorasi data daripada sumber Kaggle bagi mengenal pasti struktur dan isu kualiti data. Set data yang dipilih mengandungi 7400 baris pemerhatian terdiri daripada 11 lajur ciri. Fasa Penyediaan Data pula memfokuskan pada pembersihan dan pemformatan data seperti mengisi nilai hilang dan menormalisasi atribut untuk memastikan kesesuaian bagi pemodelan. Dalam Fasa Pemodelan, pelbagai algoritma digunakan termasuk RNN, LSTM, BiLSTM serta model pembelajaran mesin seperti Logistic Regression, Random Forest, KNN dan SVM untuk perbandingan prestasi. Fasa Penilaian kemudiannya dilakukan bagi menilai ketepatan model dan melakukan penambahbaikan jika perlu menggunakan metrik penilaian seperti *accuracy, precision, recall* dan *F1-Score*. Akhir sekali, Fasa Penyebaran memfokuskan kepada penyampaian hasil melalui antara muka interaktif seperti Streamlit bagi memudahkan pengguna menganalisis dan memahami hasil kajian ini. Rajah 2.0 menunjukkan reka bentuk algoritma bagi teknik pengesanan jenayah siber dan rajah 3.0 menunjukkan reka bentuk senibina proses pembangunan pengesanan jenayah siber.



Rajah 2.0 Reka bentuk algoritma bagi teknik pengesanan jenayah siber

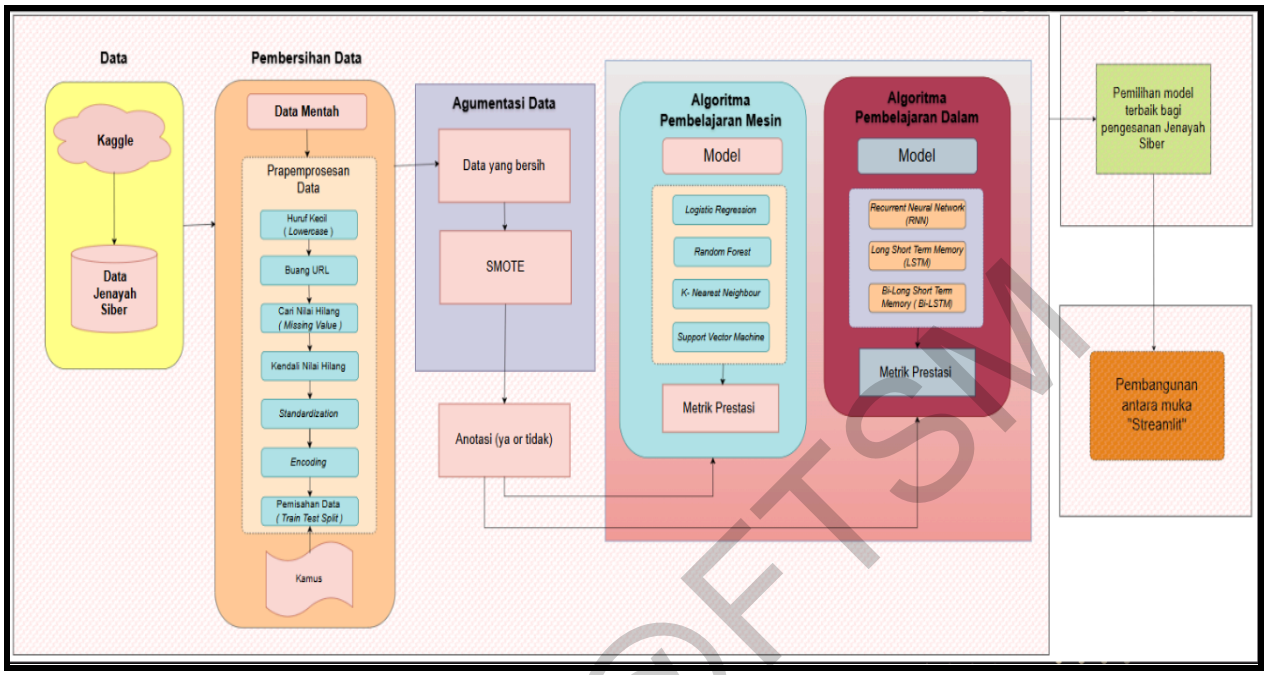
Projek ini dimulakan dengan pengumpulan data daripada dua sumber utama, iaitu dataset "Cybercrime Forensic Dataset" dari Kaggle dan data kaji selidik kesedaran jenayah siber dalam

Copyright@FTSM UKM

kalangan remaja. Dataset tersebut mengandungi atribut seperti jenis anomali, percubaan log masuk, tindakan, dan jenis aktiviti yang membantu dalam mengenal pasti dan menganalisis pola serangan siber. Gabungan data teknikal dan data kaji selidik memberikan gambaran menyeluruh tentang ancaman dan tahap kesedaran remaja terhadap jenayah siber. Seterusnya, proses prapemprosesan data melibatkan pembersihan, penyeragaman, pengekodan, serta penskalaan ciri-ciri penting untuk memastikan data berada dalam format yang sesuai untuk pembelajaran mesin dan mendalam. Proses ini juga merangkumi pengurusan nilai hilang, penyingkiran ciri tidak relevan dan pembahagian data kepada set latihan dan ujian (80:20). Label encoding digunakan untuk menyediakan sasaran pengelasan yang konsisten, manakala *standardization* membantu mempercepatkan proses pembelajaran dan meningkatkan prestasi model.

Bagi latihan dan ujian model, tiga algoritma pembelajaran mendalam utama digunakan iaitu RNN, LSTM dan BiLSTM. RNN sesuai untuk data urutan tetapi terhad oleh isu *vanishing gradient*, manakala LSTM direka untuk mengekalkan maklumat jangka panjang melalui penggunaan sel memori dan gerbang kawalan. BiLSTM pula menawarkan kelebihan melihat urutan dari dua arah bagi menangkap konteks yang lebih luas. Selain itu, beberapa model pembelajaran mesin seperti Logistic Regression, Random Forest, KNN dan SVM turut diuji sebagai perbandingan prestasi awal, terutamanya untuk menilai kesan saiz data yang terhad. Model-model ini dipilih berdasarkan kelebihan masing-masing dalam mengendalikan data klasifikasi dan keupayaan penyesuaian terhadap pelbagai bentuk data. Penilaian akhir dilakukan menggunakan metrik seperti ketepatan (*accuracy*) dan skor F1, bagi memastikan prestasi yang stabil dan kebolehpercayaan model dalam mengesan jenayah siber. Pendekatan gabungan ini membolehkan pembangunan sistem pengesanan yang mantap, tepat dan berpotensi digunakan dalam aplikasi sebenar.

Copyright@FTSM UKM



Rajah 3.0 Reka bentuk senibina proses pembangunan pengesanan jenayah siber

Reka bentuk seni bina adalah satu rajah menyeluruh yang menggambarkan proses-proses utama dalam pembangunan projek ini. Untuk projek ini, reka bentuk seni bina secara umum dimulakan dengan proses pengumpulan data dari platform seperti Kaggle dan diakhiri dengan pembangunan model untuk kajian pengesanan jenayah siber berdasarkan analisis data.

4.0 HASIL

4.1 Pembangunan Proses Pengimbangan Kelas Tidak Seimbang

Data latihan yang digunakan dalam kajian ini menunjukkan ketidakseimbangan kelas yang ketara, di mana hanya sekitar 19.35% mewakili kelas minoriti (*Suspicious*). Ketidakseimbangan ini boleh menjejaskan prestasi model, kerana ia cenderung mengabaikan pola daripada kelas minoriti. Untuk mengatasi masalah ini, teknik Synthetic Minority Oversampling Technique (SMOTE) telah digunakan bagi menjana sampel sintetik kelas minoriti tanpa mengurangkan data kelas majoriti.

Teknik SMOTE dipilih kerana ia mampu memperkayakan kelas minoriti dengan lebih realistik berbanding kaedah lain seperti *undersampling* atau *oversampling* rawak. Langkah ini penting untuk memastikan model pembelajaran mendalam dapat belajar secara adil dan berkesan dalam mengesan aktiviti mencurigakan dalam konteks pengesanan jenayah siber.

Copyright@FTSM UKM

Oleh itu, data latihan yang telah diimbang menggunakan teknik pengimbangan data telah tersedia untuk dijadikan input kepada algoritma Pembelajaran Mesin dan Pembelajaran Mendalam. Subtopik 4.2 dan 4.3 akan menunjukkan proses pembangunan model Pembelajaran Mesin dan pembelajaran mendalam serta ujian prestasi ke atas setiap satu daripada model-model itu. Penilaian ke atas prestasi model dijalankan dengan menggunakan metrik penilaian yang bersesuaian untuk memastikan keberkesanan dan ketepatan klasifikasi jenayah siber yang dijalankan. Prestasi model adalah berdasarkan metrik seperti ketepatan (*accuracy*), *recall*, *precision*, dan *F1-score*.

4.2 Pembangunan dan Perbandingan Model Pembelajaran Mesin

Objektif utama pembangunan model Pembelajaran Mesin ini adalah untuk membandingkan kesesuaiannya serta ketetapannya daripada Pembelajaran Mendalam. Proses ini akan melibatkan empat model Pembelajaran Mesin iaitu Logistic Regression, Random Forest, K-Nearest Neighbour dan Support Vector Machine yang diimport daripada pustaka *scikit-learn*. Selain itu, laporan klasifikasi dan matriks kekeliruan (*confusion matrix*) turut digunakan bagi memperoleh pemahaman yang lebih mendalam mengenai prestasi model dalam mengesan serangan serta aktiviti yang normal.

Secara keseluruhan, model Random Forest menunjukkan prestasi paling seimbang dengan ketepatan sebanyak 62.47%, *recall* yang agak baik untuk kelas positif (50%), dan skor ROC-AUC tertinggi antara keempat-empat model iaitu 57.12%, menjadikannya pilihan terbaik dalam kumpulan ini. *Logistic Regression* pula mencatatkan recall kedua tertinggi untuk kelas positif (54%) tetapi dengan *precision* yang rendah (19%) serta skor ROC-AUC 56.21%, menunjukkan model ini mengesan banyak kes positif tetapi dengan banyak juga kes palsu positif.

Seterusnya, Support Vector Machine (SVM) mencatatkan *accuracy* 59.14%, *recall* kelas positif 44%, dan ROC-AUC 55.50% walaupun *recall* agak sederhana, *precision* (18%) dan F1-score (26%) masih rendah, menunjukkan banyak kes positif berjaya dikesan tetapi dengan kadar kesalahan yang tinggi. Akhir sekali, K-Nearest Neighbors (KNN) pula menunjukkan ketepatan keseluruhan paling tinggi (82.90%) tetapi ini agak mengelirukan kerana *recall* untuk kelas positif sangat rendah (1%), bermakna model ini hampir langsung gagal mengesan kes

Copyright@FTSM UKM

positif dan berat sebelah kepada kelas majoriti. Skor ROC-AUC KNN juga rendah (47.83%), menunjukkan kemampuan pemisahan kelas yang lemah.

Kesimpulannya, walaupun Random Forest menunjukkan sebagai model dengan prestasi paling seimbang di antara keempat-empat model, kesemuanya masih mempunyai kelemahan dari segi *precision* dan *F1-score* untuk kelas positif. Oleh itu, penambahbaikan seperti penukaran parameter, pengimbangan data, atau penggunaan teknik *ensemble* mungkin diperlukan untuk meningkatkan prestasi model secara keseluruhan. Walau bagaimanapun, pengujian ke atas model-model pembelajaran mesin ini dijalankan semata-mata untuk menilai prestasi awal sebelum dibandingkan dengan model pembelajaran mendalam (*deep learning*). Ini kerana set data yang digunakan dalam kajian ini berskala kecil dan berkemungkinan mempunyai kekurangan dari segi kualiti dan kuantiti. Jadual 2.0 menunjukkan keseluruhan perbandingan prestasi daripada kesemua model pembelajaran mesin.

Jadual 2.0 Perbandingan prestasi model pembelajaran mesin

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Random Forest	0.62	0.22	0.50	0.30	0.57
Support Vectort Machine	0.59	0.18	0.44	0.26	0.55
K-Nearest Neighbours	0.83	0.17	0.03	0.01	0.48
Logistic Regression	0.56	0.19	0.54	0.28	0.56

4.3 Pembangunan Model Pembelajaran Mendalam

Pembelajaran mendalam ialah satu cabang dalam pembelajaran mesin yang menggunakan rangkaian neural tiruan dengan banyak lapisan untuk memproses dan menganalisis data secara automatik. Ia mampu mengenal pasti corak dan ciri kompleks dalam data seperti teks atau imej, dan sering digunakan dalam aplikasi seperti pengecaman suara, analisis bahasa semula jadi, dan pengesanan jenayah siber. Objektif utama pembangunan model pembelajaran mendalam ini adalah untuk menilai keberkesanan dan kesesuaiannya dalam menganalisis data berjumlah kecil. Tiga model utama yang digunakan dalam kajian ini ialah Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) dan Bidirectional LSTM (BiLSTM). Pemilihan model-model ini adalah kerana fokus utama projek adalah untuk mengkaji prestasi pembelajaran mendalam

Copyright@FTSM UKM

dan diambil daripada kajian lepas namun diaplikasikan ke atas set data yang kecil. Walaupun hasil prestasi model yang dibina tidak begitu memuaskan, projek ini memberikan penemuan penting bahawa saiz set data sangat memainkan peranan yang sangat signifikan dalam menentukan keberkesanan model pembelajaran mendalam.

Model pembelajaran mendalam yang digunakan dalam kajian ini iaitu RNN, LSTM dan BiLSTM telah menunjukkan prestasi berbeza dalam tugas pengesanan jenayah siber. Model RNN mencatatkan keputusan paling rendah dengan accuracy 62.7%, precision 0.197, dan f1-score hanya 0.2697. Walaupun recall agak sederhana (0.426), ROC-AUC model ini hanya 0.5586, menghampiri nilai rawak (0.50), menandakan keupayaan pembezaan kelas yang lemah. LSTM pula menunjukkan peningkatan dari segi accuracy (67.7%) dan f1-score (0.2766), dengan nilai ROC-AUC sebanyak 0.5638, menjadikannya lebih stabil dan seimbang berbanding RNN. BiLSTM mencatatkan accuracy tertinggi iaitu 69.4%, tetapi f1-score sedikit menurun kepada 0.2628, dengan ROC-AUC 0.5625 yang hampir setara dengan LSTM. Walaupun BiLSTM dilihat terbaik dari segi ketepatan keseluruhan, nilai precision dan recall yang masih rendah menunjukkan bahawa ketiga-tiga model menghadapi kesukaran mengesan kelas minoriti "*Suspicious*" dengan tepat. Keputusan ini mencerminkan cabaran biasa dalam pengesanan jenayah siber yang menggunakan data tidak seimbang, di mana model cenderung lebih baik mengenal pasti kelas majoriti iaitu "*Normal*".

Dalam konteks data yang terhad dan tidak seimbang seperti dalam kajian ini, pemilihan metrik penilaian yang sesuai amat penting. Ketepatan (*accuracy*) sahaja tidak mencukupi kerana model boleh mencatat nilai tinggi dengan hanya mengklasifikasikan kelas majoriti. Oleh itu, metrik seperti f1-score lebih relevan kerana ia mengambil kira keseimbangan antara *precision* dan *recall*, terutamanya dalam situasi di mana kesalahan mengenal pasti aktiviti mencurigakan boleh membawa implikasi serius. Di samping itu, ROC-AUC digunakan untuk menilai keupayaan pembezaan model terhadap dua kelas pada pelbagai ambang, namun ia turut terjejas apabila distribusi kelas tidak seimbang. Berdasarkan semua metrik ini, walaupun BiLSTM menunjukkan accuracy tertinggi, LSTM mencatat nilai f1-score dan ROC-AUC yang sedikit lebih baik, menjadikannya pilihan paling sesuai dalam mengenal pasti jenayah siber dalam kalangan remaja, terutamanya apabila matlamat utama adalah untuk mengesan kelas minoriti dengan lebih berkesan. Penambahbaikan lanjutan seperti pelarasan ambang klasifikasi, pemilihan

Copyright@FTSM UKM

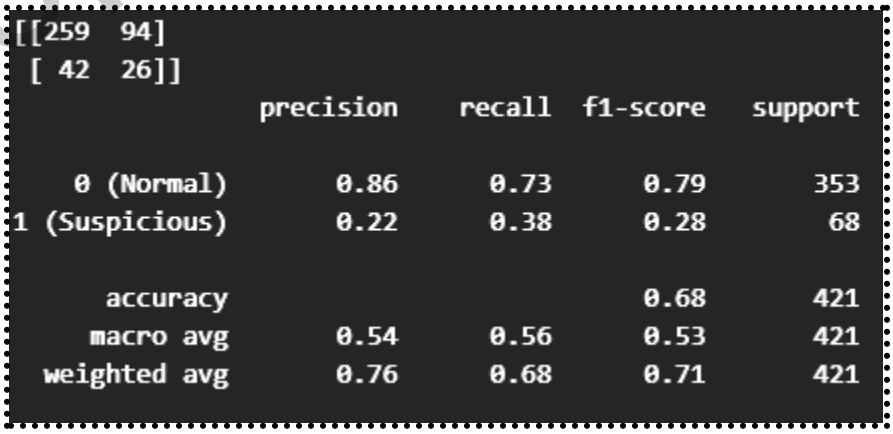
ciri yang lebih signifikan, dan teknik pengimbangan data tambahan amat disarankan bagi meningkatkan prestasi model pada masa akan datang. Oleh itu, jadual 3.0 menunjukkan keseluruhan perbandingan prestasi daripada kesemua model pembelajaran mendalam.

Jadual 3.0 Perbandingan prestasi model pembelajaran mendalam

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Recurrent Neural Network	0.63	0.20	0.43	0.27	0.55
Long Short-Term Memory	0.68	0.22	0.38	0.28	0.56
Bi-Long Short-Term Memory	0.69	0.21	0.34	0.26	0.56

4.4 Pemilihan Model Terbaik

Long Short-Term Memory (LSTM) merupakan varian seterusnya daripada RNN yang direka untuk mengatasi masalah "*vanishing gradient*" dan dapat mengingati ketergantungan jangka panjang dalam data urutan. Ia menggunakan struktur pintu (*gates*) yang kompleks untuk memutuskan maklumat mana yang disimpan atau dilupakan, menjadikannya lebih berkesan dalam mengenal pola yang kompleks dalam data bersiri. Dalam konteks pengesanan jenayah siber, LSTM digunakan untuk mengenal pasti corak tingkah laku yang mencurigakan dengan lebih tepat berbanding model asas. Oleh itu, Rajah 4.0 menunjukkan prestasi metrik bagi model Long Short -Term Memory (LSTM) dan Jadual 4.0 menunjukkan prestasi bagi model Long Short-Term Memory (LSTM).



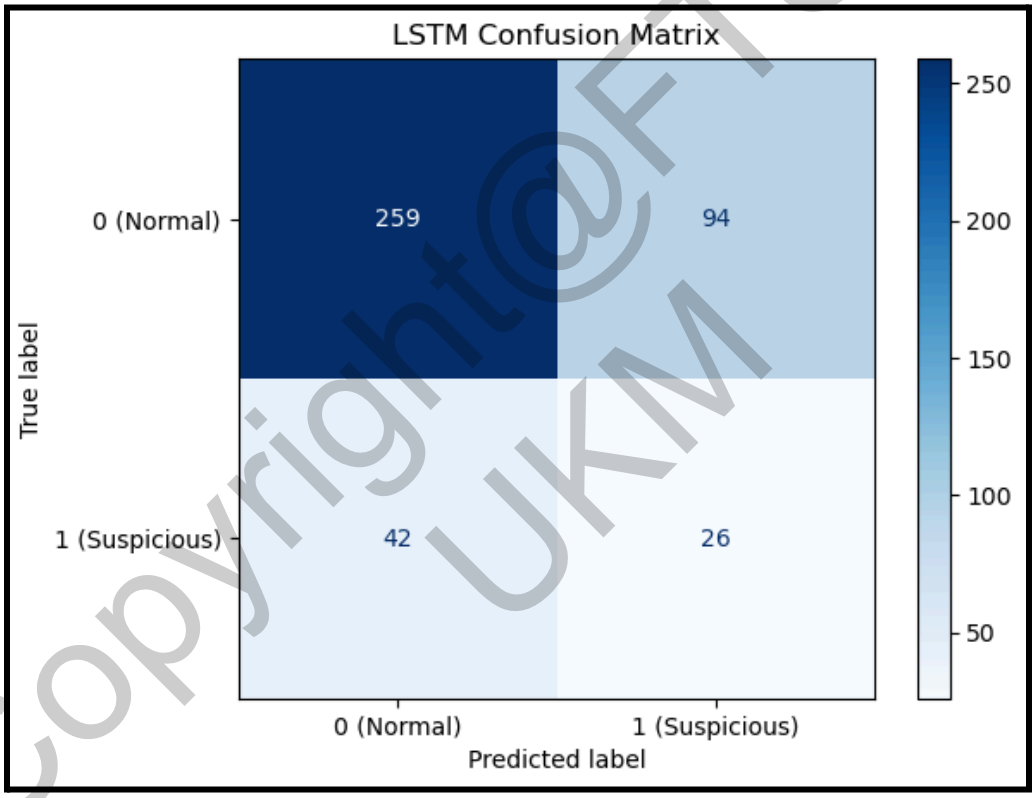
```
[[259  94]
 [ 42  26]]
                precision    recall  f1-score   support

   0 (Normal)       0.86      0.73      0.79       353
1 (Suspicious)      0.22      0.38      0.28        68

     accuracy                           0.68       421
    macro avg       0.54      0.56      0.53       421
 weighted avg       0.76      0.68      0.71       421
```

Rajah 4.0 Prestasi metrik bagi model Long Short-Term Memory (LSTM).

Copyright@FTSM UKM

Jadual 4.0 Prestasi Long Short-Term Memory (LSTM)

Metrik Penilaian	Nilai
Accuracy	0.68
Precision	0.22
Recall	0.38
F1-Score	0.28
ROC-AUC	0.56

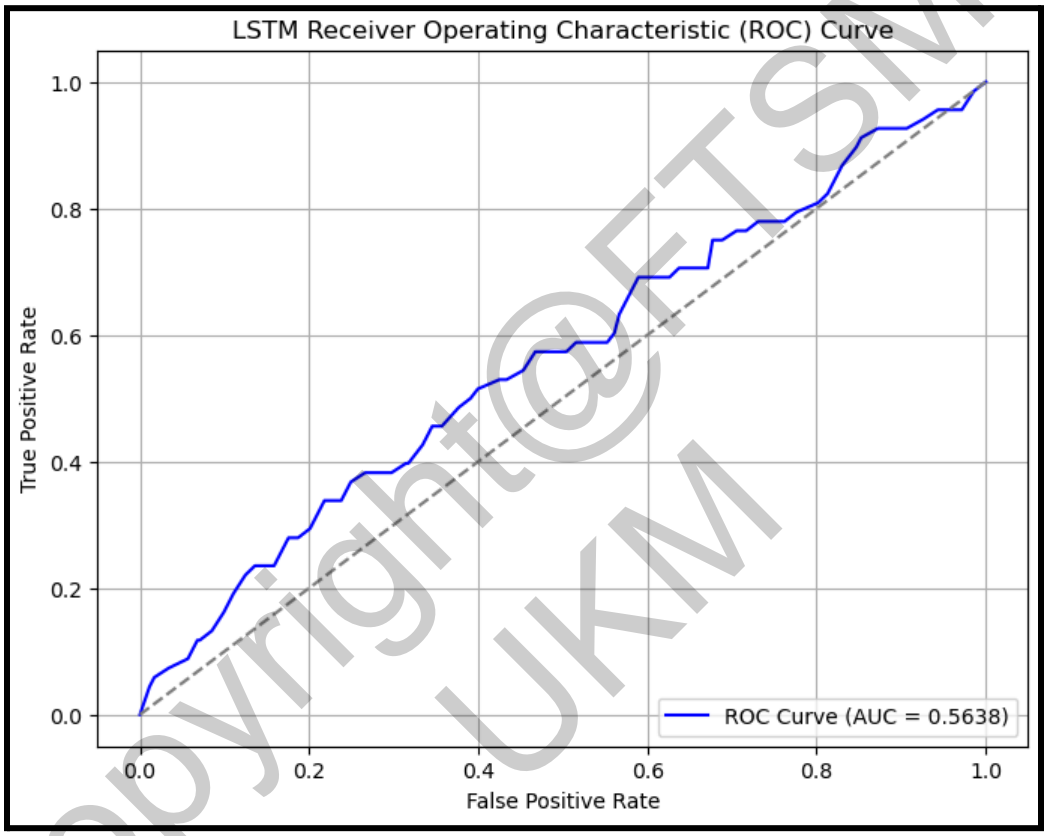


Rajah 5.0 Matriks kekeliruan bagi model Long Short-Term Memory (LSTM).

Berdasarkan Rajah 5.0 keputusan Matrik kekeliruan (*Confusion Matrix)* bagi model LSTM dengan *threshold* 0.4, sebanyak 259 daripada 353 sampel kelas '*Normal*' berjaya dikelaskan dengan betul, manakala 94 sampel Normal tersalah dikelaskan sebagai '*Suspicious'*. Bagi kelas *Suspicious*, 26 daripada 68 sampel dikesan dengan betul, sementara 42 sampel *Suspicious* tersalah dikelaskan sebagai Normal. Walaupun terdapat peningkatan dalam jumlah pengesanan kelas *Suspicious* berbanding ramalan rawak, jumlah salah klasifikasi bagi kelas

Copyright@FTSM UKM

tersebut masih tinggi. Ini mencadangkan bahawa walaupun model LSTM mampu mengenal pasti sebahagian pola tingkah laku mencurigakan, ia masih menghadapi cabaran dalam membezakan sepenuhnya kedua-dua kelas dengan tepat, terutama dalam mengurangkan kes tersalah klasifikasi *Suspicious* sebagai Normal. Seterusnya, Rajah 6.0 menunjukkan lengkung AUC-ROC (*Receiver Operating Characteristic*) bagi model Long Short-Term Memory (LSTM).



Rajah 6.0 Lengkung AUCROC bagi model Long Short-Term Memory (LSTM).

Model LSTM menunjukkan nilai kawasan di bawah lengkung ROC (AUC-ROC) sebanyak 0.5638. Nilai ini hanya sedikit melebihi nilai rawak 0.50, yang menandakan bahawa keupayaan model untuk membezakan antara kelas *Normal* dan *Suspicious* masih rendah. Walaupun terdapat peningkatan kecil berbanding tebakan rawak, ia belum cukup untuk dianggap sebagai pembezaan yang kukuh. Nilai AUC-ROC sekitar 0.56 ini menunjukkan bahawa model masih kerap mengelirukan kedua-dua kelas dalam ramalan berasaskan kebarangkalian.

4.5 Pembangunan Streamlit

Copyright@FTSM UKM

Streamlit ialah library Python sumber terbuka yang membolehkan pembinaan antara muka yang interaktif dengan cepat tanpa memerlukan pengetahuan mendalam dalam HTML, CSS, atau JavaScript. Dengan Streamlit, pembangun boleh menumpukan perhatian kepada logik aplikasi dan analisis data, sementara elemen antara muka pengguna dapat dibangunkan terus menggunakan kod Python.

Antara ciri utama *Streamlit* termasuk keupayaan untuk memaparkan teks, gambar, graf, dan elemen interaktif seperti butang, '*sliders'*, '*checkboxes'*, dan '*select boxes'* yang membolehkan pengguna berinteraksi secara langsung dengan aplikasi. Tujuan utama pembangunan aplikasi menggunakan *Streamlit* dalam kajian ini adalah untuk menyediakan platform sumber terbuka (*open source*) yang membolehkan pengguna memasukkan input seperti maklumat cubaan log masuk (*Login_Attempts)* dan tindakan (*Actions*) untuk meramal kemungkinan berlakunya jenayah siber yang dianalisis oleh model LSTM. Antara muka yang direka membolehkan pengguna memasukkan data dengan mudah dan melihat keputusan ramalan secara langsung tanpa perlu memuat semula halaman. Ciri interaktif *Streamlit* membolehkan sistem memaparkan hasil klasifikasi serta visualisasi prestasi model dengan cara yang mudah difahami oleh pengguna. Oleh itu, rajah 7.0 seterusnya menunjukkan kod bagi *user prediction* dalam *Streamlit* dan fail komponen penting dimuat naik.

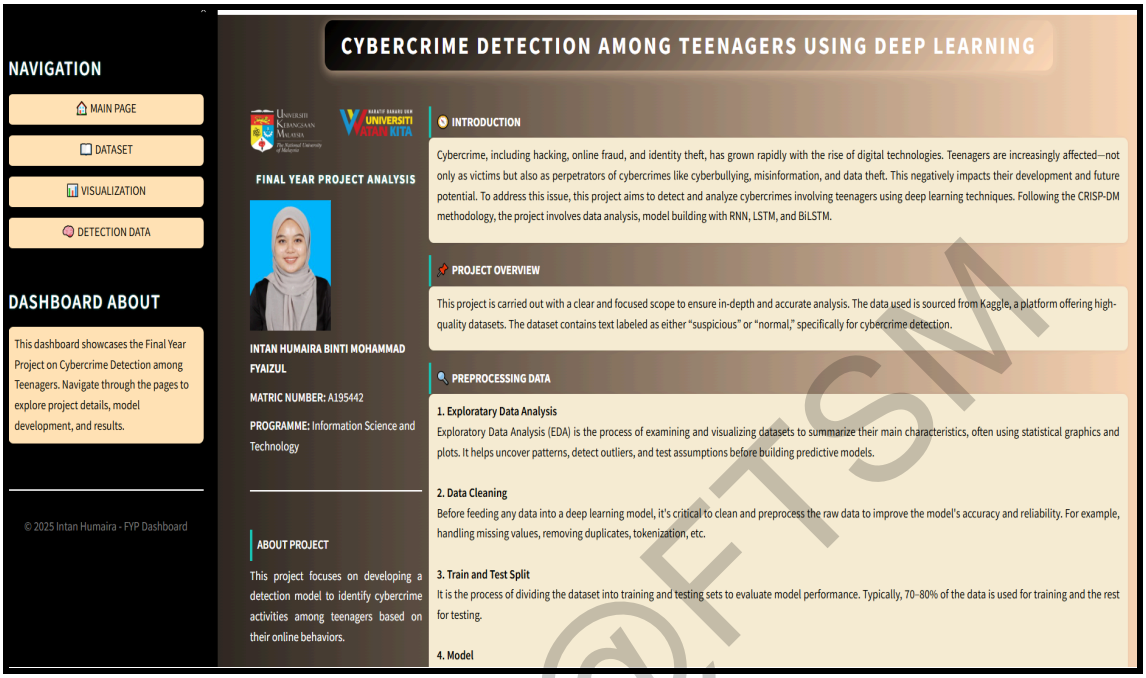
Copyright@FTSM UKM

```
# ==============================
# Part 2: Live User Prediction
# ==============================
st.header("🔮 Predict New User Activity")
st.caption("Use the form below to classify a new activity as **Normal** or **Suspicious**.")

try:
    model = load_model('lstm_model.h5')
    with open('scaler.pkl', 'rb') as f:
        scaler = pickle.load(f)
    with open('label_encoder.pkl', 'rb') as f:
        label_encoder = pickle.load(f)
except Exception as e:
    st.error(f"❌ Error loading model or encoders: {e}")
    st.stop()

with st.form("prediction_form"):
    st.subheader("📝 Input Features")
    st.caption("Provide the user activity details below.")

    # Slider for login attempts
    login_attempts = st.slider(
        '🔢 Number of Login Attempts',
        min_value=1,
        max_value=10,
        value=1,
        help="How many times the user tried to log in."
    )
```

Rajah 7.0 Kod bagi *User Prediction* dan fail komponen penting

Blok kod ini membina bahagian ramalan aktiviti pengguna secara langsung (*live prediction*) dalam aplikasi *Streamlit*, yang membolehkan pengguna memasukkan data baharu dan mengklasifikasikannya sebagai '*Normal'* atau *'Suspicious'* berdasarkan model yang telah dilatih. Kod kemudian menunjukkan model LSTM (lstm_model.h5) serta dua komponen penting: *scaler* untuk menormalkan input dan '*label_encoder'* untuk menukarkan kod label. Sekiranya berlaku ralat semasa pemuatan, aplikasi akan menunjukkan mesej ralat dan menghentikan proses. Dalam borang *prediction_form*, pengguna diminta untuk memberikan input ciri aktiviti pengguna, dan antara input yang diminta ialah bilangan percubaan log masuk (*login attempts*), yang dikumpul melalui *slider* dari nilai 1 hingga 10. Komponen ini merupakan permulaan kepada satu sistem pengesanan siber interaktif yang menggunakan model pembelajaran mendalam untuk menilai risiko aktiviti pengguna berdasarkan input semasa. Akhir sekali, Rajah 8.0 dan Rajah 9.0 menunjukkan dua contoh halaman reka bentuk antara muka menggunakan Streamlit.

Copyright@FTSM UKM



Rajah 8.0 Halaman reka bentuk antara muka menggunakan Streamlit

Rajah 9.0 Halaman reka bentuk antara muka *(detection page)*

Hak Cipta@FTSM
Universiti Kebangsaan Malaysia

5.0 KESIMPULAN

Secara keseluruhan, projek ini telah berjaya membangunkan satu sistem pengesanan jenayah siber dalam kalangan remaja dengan pendekatan yang menyeluruh, merangkumi aspek pembinaan model, pemprosesan data, serta pembangunan antara muka pengguna menggunakan Streamlit. Kekuatan utama projek ini terletak pada eksperimen yang melibatkan variasi model pembelajaran mendalam seperti Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) dan Bidirectional LSTM (BiLSTM), yang menunjukkan kesungguhan dalam mencari pendekatan paling sesuai. Walaupun prestasi model masih belum mencapai tahap terbaik, usaha perbandingan ini telah menghasilkan satu garis dasar (*baseline*) penting yang boleh dijadikan rujukan oleh penyelidik lain dalam kajian masa hadapan. Ini membuka ruang kepada pemahaman yang lebih mendalam terhadap kelebihan dan kelemahan setiap model, khususnya dalam konteks analisis tingkah laku digital yang kompleks berkait rapat dengan jenayah siber.

Projek ini turut menyerlahkan kepentingan kualiti data dalam pembangunan model pembelajaran mendalam. Isu ketidakseimbangan kelas dalam set data asal, yang menunjukkan hanya 19.35% data tergolong dalam kategori mencurigakan, merupakan satu cabaran besar. Bagi menangani isu ini, teknik Synthetic Minority Oversampling Technique (SMOTE) digunakan untuk menyeimbangkan taburan kelas, membolehkan model dilatih dengan lebih adil dan mengurangkan bias terhadap kelas majoriti. Walau bagaimanapun, walaupun SMOTE membantu, ketidakseimbangan mungkin masih berlaku dalam subset tertentu, yang memberi kesan kepada hasil akhir model. Tambahan pula, saiz data yang kecil selepas proses pembersihan, serta jumlah atribut yang terhad, menjejaskan lagi kemampuan model untuk mengesan corak yang rumit. Ini membuktikan bahawa kejayaan model bukan semata-mata bergantung kepada algoritma yang digunakan, tetapi juga sangat bergantung kepada kekayaan dan kesesuaian data latihan.

Selain itu, antara muka pengguna yang dibangunkan menggunakan Streamlit berjaya memberikan platform interaktif untuk ujian dan penggunaan sistem. Namun, pembangunan antaramuka turut mencabar dari segi pemprosesan input pengguna yang kadang kala tidak selaras dengan cara data ujian diproses, mengakibatkan ketidaktepatan ramalan. Masalah ini menekankan keperluan terhadap prapemprosesan yang konsisten serta kawalan input yang lebih

Copyright@FTSM UKM

teliti melalui validasi, penskalaan, dan pengekodan automatik. Walaupun begitu, usaha membina sistem yang boleh digunakan secara langsung oleh pengguna menunjukkan kelebihan projek dari sudut aplikasi sebenar, menjadikannya bukan sahaja sebagai eksperimen saintifik tetapi juga sebagai prototaip yang boleh dikembangkan lagi.

Cadangan penambahbaikan yang dikemukakan dalam projek ini juga cukup jelas dan realistik, antaranya termasuk meningkatkan kualiti set data melalui penggabungan sumber data baharu, menggunakan teknik imputasi lanjutan untuk menangani nilai hilang, serta memperluaskan pemilihan ciri melalui teknik *feature engineering*. Selain itu, eksplorasi model baharu seperti CNN dan transformer, serta pelarasan *hyperparameter* yang lebih menyeluruh, juga boleh membuka peluang kepada peningkatan prestasi model. Dari segi pembangunan sistem, penambahan ciri interaktif seperti paparan kebarangkalian ramalan, visualisasi penjelasan model (contohnya SHAP atau LIME), serta panduan penggunaan boleh meningkatkan kebolehgunaan dan kebolehfahaman sistem.

Akhir sekali, projek ini telah menunjukkan asas yang kukuh dalam usaha membangunkan sistem pengesanan jenayah siber berasaskan pembelajaran mendalam. Meskipun berdepan dengan beberapa cabaran dari aspek data dan pelaksanaan teknikal, pendekatan yang digunakan tetap relevan dan berpotensi tinggi untuk diperluaskan pada masa akan datang. Ia bukan sahaja membuktikan keupayaan model pembelajaran mendalam dalam menganalisis tingkah laku digital yang mencurigakan, tetapi juga menyerlahkan keperluan terhadap pengurusan data yang berkualiti, pembangunan sistem yang berorientasikan pengguna, dan pelaksanaan strategi pembelajaran mesin yang holistik. Dengan pelaksanaan cadangan penambahbaikan secara berperingkat, projek ini boleh ditingkatkan menjadi sistem pengesanan jenayah siber yang lebih efisien, tepat dan berdaya guna, sekali gus menyumbang kepada usaha melindungi remaja daripada ancaman dalam dunia digital yang semakin kompleks.

Copyright@FTSM UKM

6.0 PENGHARGAAN

Alhamdulillah, bersyukur ke hadrat Allah S.W.T, Yang Maha Pengasih lagi Maha Penyayang kerana dengan izin dan berkat-Nya telah memberikan saya kesihatan yang baik, masa yang cukup dan kematangan fikiran untuk saya menyiapkan laporan projek tahun akhir ini. Saya dengan rendah hati menadah tangan tanda kesyukuran, terharu kerana telah menyiapkan projek tahun akhir saya bagi memenuhi sebahagian daripada syarat memperolehi Ijazah Sarjana Muda Sains Komputer dengan Kepujian dalam tempoh masa yang ditetapkan. Saya sekali lagi berasa sangat bersyukur kerana telah diberikan kekuatan untuk saya mengatasi segala masalah dan kekangan yang timbul sepanjang projek ini dijalankan. Saya sedar bahawa dugaan yang hadir ini telah membentuk peribadi saya untuk sentiasa peka dan disiplin dalam melakukan sesuatu kerja.
Seterusnya, saya ingin merakamkan setinggi-tinggi penghargaan kepada penyelia saya, Ts. Rohizah Binti Abd. Rahman, atas bimbingan dan dorongan sepanjang saya menyiapkan usulan projek ini tanpa lelah dan jemu. Beliau telah banyak memberikan tunjuk ajar serta bantuan, teguran dan nasihat yang sangat berguna kepada saya. 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. Tidak lupa juga kepada Fakulti Teknologi dan Sains Maklumat dan Universiti Kebangsaan Malaysia atas kemudahan dan sumber yang disediakan kerana telah banyak membantu saya dalam menyiapkan projek ini serta pensyarah-pensyarah khususnya kepada mereka yang pernah menabur ilmu pengetahuan kepada saya sepanjang pengajian.

Akhir sekali, dengan kesempatan ini saya menghulurkan setinggi-tinggi ucapan penghargaan kepada beberapa pihak yang lain sama ada terlibat secara langsung mahupun tidak langsung dalam menjayakan penyiapan projek ini. Tidak lupa juga kepada semua ahli keluarga saya khususnya ibunda tercinta, Puan Junaizah Jumahat serta sahabat seperjuangan terutamanya Nur ‘Eiza Athira dan Iliana Hanin yang telah memberi sokongan dan semangat yang tinggi serta sentiasa berdoa terhadap kejayaan saya di universiti. Saya juga ingin memohon maaf sekiranya terdapat kesalahan sepanjang perlaksanaan projek akhir tahun ini. Akhirnya, tidak lupa untuk berterima kasih kepada diri sendiri kerana sudah bersusah payah mengharungi detik-detik akhir untuk menjadi seorang pelajar sarjana muda sains komputer dengan jayanya. Sekian, terima kasih.

Copyright@FTSM UKM

7.0 RUJUKAN

Al-Khater, Wadha Abdullah, et al. “Comprehensive Review of Cybercrime Detection Techniques.” *IEEE Access*, vol. 8, no. 1, 2020, pp. 1–1.

Alharbi, N., Alkalifah, B., Alqarawi, G., & Rassam, M. A. (2024). Countering social media cybercrime using deep learning: Instagram Fake accounts detection. Future Internet, 16(10), 367.

Amer, A., Siddiqui, T., & Athamena, B. (2022). Detecting Cybercrime: an evaluation of machine learning and deep learning using natural language processing techniques on the social network. Research Square (Research Square).

Bheemaiah, K., Esposito, M., & Tse, T. (2017). What is machine learning? The Conversation. https://theconversation.com/what-is-machine-learning-76759

Biodoumoye George Bokolo, et al. Deep Learning Assisted Cyber Criminal Profiling. 7 July 2023,*2*, 226–231.

Ghaedi, M. (2018, March 7). CRISP-DM methodology. https://www.linkedin.com/pulse/crisp-dm-methodology-mani-ghaedi/

Graves, A., & Graves, A. (2012). Long short-term memory. Supervised sequence labelling with recurrent neural networks, 37-45.

Malek, M. D. A., & Mohamed Kamil, I. S. (2010). Jenayah dan masalah sosial di kalangan remaja: Cabaran dan realiti dunia siber.

Roy, S. S., Awad, A. I., Amare, L. A., Erkihun, M. T., & Anas, M. (2022). Multimodel phishing URL detection using LSTM, bidirectional LSTM, and GRU models. Future Internet, 14(11), 340.

Tom Fawcett. (2016). Learning from imbalanced classes. Guiding Tech Media.

Tranung, K. (2024, August 27). Penguatkuasaan Akta Keselamatan Siber 2024 (Akta 854). Laman Web MKN. https://www.mkn.gov.my/web/ms/2024/08/26/penguatkuasaan-akta-keselamatan-siber-2024-akta-854/

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

Copyright@FTSM UKM

https://www.who.int/europe/news/item/27-03-2024-one-in-six-school-aged-children-experiences-cyberbullying--finds-new-who-europe-study

Xie, D., Zhang, L., & Bai, L. (2017). Deep learning in visual computing and signal processing. Applied Computational Intelligence and Soft Computing, 2017(1), 1320780.

Zoldi, Nor. Lailatul. A. (2019). Sistem Pengesanan Emosi pada masa nyata (SPEMN) [Thesis]. Fakulti Seni, Komputeran dan Industri Kreatif, Universiti Pendidikan Sultan Idris. https://ir.upsi.edu.my/files/docs/2020/4966_1594262970.pdf

Intan Humaira Binti Mohammad Fyaizul (A195442)
Ts. Rohizah binti Abd Rahman
Fakulti Teknologi & Sains Maklumat
Universiti Kebangsaan Malaysia

Copyright@FTSM
UKM