

PEGESANAN BERITA PALSU: PENDEKATAN PEMBELAJARAN MENDALAM

MUHAMMAD QASHFULLAH BIN ABDUL GAFFAR, ZALINDA OTHMAN

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

ABSTRAK

Dalam pendekatan menggunakan pembelajaran mesin, projek ini bertujuan untuk mengesan dan membezakan berita palsu daripada berita sebenar dengan lebih cekap. Dalam era media digital dan sosial, penyebaran berita palsu adalah ancaman serius kepada kepercayaan dan ketelusan maklumat. Projek ini menggunakan analisis teks dan ciri artikel berita untuk mengenal pasti tanda berita palsu, termasuk unsur sensasi, kandungan yang meragukan dan tajuk berita yang menarik perhatian. Hasil projek ini akan menyediakan alat yang lebih baik dalam memerangi penyebaran berita palsu dan menjana kesedaran tentang masalah ini dalam dunia digital yang semakin kompleks. Ini adalah langkah penting dalam memastikan ketelusan dan kepercayaan terhadap maklumat yang kami terima. Projek ini melibatkan pengumpulan set data yang merangkumi pelbagai sumber berita serta latihan model pembelajaran mesin yang boleh membezakan ciri-ciri berita palsu dan sebenar. Model ini akan melibatkan teknik seperti penggunaan pengelasan seperti Naive Bayes atau Rangkaian Neural Konvolusi (RNK) yang lebih kompleks. Selain itu, ciri seperti asal sumber berita dan kajian trend boleh disertakan untuk meningkatkan prestasi model. Projek ini juga akan merangkumi antara muka pengguna (UI) menggunakan aplikasi web Flask untuk menyediakan platform yang lebih mesra pengguna. Aplikasi web ini akan membolehkan pengguna memasukkan teks berita yang mereka ingin analisis dengan mudah dan menerima keputusan pengesanan dengan cepat dan cekap. Rangka kerja Flask akan membolehkan pembangunan antara muka web yang responsif dan interaktif, meningkatkan pengalaman pengguna dan kebolehcapaian. Penilaian prestasi projek ini akan menggunakan metrik seperti ketepatan, ketepatan, ingatan semula dan skor F1. Kejayaan projek ini bukan sahaja membantu dalam memerangi penyebaran berita palsu tetapi juga boleh memberi sumbangan besar kepada masyarakat dalam mengekalkan ketelusan, integriti maklumat, dan mendidik orang ramai mengenali tanda-tanda berita palsu. Projek ini

memainkan peranan penting dalam membantu masyarakat membuat keputusan yang lebih termaklum dan sedar dalam dunia yang semakin digital ini. Dengan keselamatan dan privasi sebagai keutamaan, adalah penting untuk memastikan bahawa data dan maklumat pengguna yang digunakan dalam projek ini dilindungi dengan selamat dan mematuhi undang-undang dan peraturan yang berkaitan. Secara keseluruhannya, projek "Pengesanan Berita Palsu: Pendekatan Pembelajaran Mendalam" berpotensi untuk menjana kesedaran, ketelusan dan kepercayaan dalam dunia yang semakin kompleks yang penuh dengan lebih maklumat digital.

PENGENALAN

Dalam era digital ini, maklumat telah menjadi satu komoditi yang sangat berharga dalam kehidupan seharian kita. Namun, dengan ledakan internet dan media sosial, maklumat juga mudah dimanipulasi, membolehkan penularan berita palsu atau kononnya "hoax" yang berpotensi merosakkan integriti maklumat itu sendiri. Berita palsu boleh memberi impak yang meluas, daripada mempengaruhi pilihan raya politik kepada mewujudkan kekeliruan dalam masyarakat.

Penyiaran berita palsu merupakan cabaran utama dalam dunia media dan komunikasi hari ini, dan untuk mengatasi cabaran ini, pendekatan yang canggih dan cekap diperlukan. Salah satu alat yang boleh digunakan untuk mengatasi masalah ini ialah pendekatan pembelajaran mendalam, yang telah menjadi popular dalam pelbagai bidang, termasuk pemrosesan bahasa semula jadi dan penglihatan komputer.

Projek ini bertujuan untuk meneroka peranan pembelajaran mendalam dalam mengenal pasti dan memerangi berita palsu. Kita akan melihat bagaimana teknik pembelajaran mendalam, seperti rangkaian saraf tiruan dan model bahasa, boleh digunakan untuk analisis berita dan mengesan corak yang menunjukkan berita palsu. Selain itu, kami juga akan meneroka kaedah sedia ada dalam literatur berkaitan dan mengenal pasti jurang pengetahuan yang perlu diisi untuk membangunkan pendekatan yang lebih baik.

Projek ini disusun dalam beberapa bab yang merangkumi tinjauan literatur, metodologi kajian, pelaksanaan model, penilaian, dan kesimpulan. Kami percaya bahawa pemahaman yang lebih mendalam tentang cara berita palsu tersebar dan cara mengesannya dengan lebih ketepatan adalah langkah penting dalam melindungi integriti maklumat dan media yang kami gunakan setiap hari. Melalui penyelidikan ini, kami berharap dapat memberi sumbangan yang berharga kepada usaha menentang berita palsu dan memajukan pemahaman kami tentang penggunaan pembelajaran mendalam dalam bidang ini. Dengan kerjasama dan komitmen bersama, kita boleh mencapai dunia maklumat yang lebih dipercayai dan jujur.

Copyright@FTSM
UKM

METODOLOGI KAJIAN

Metodologi yang digunakan dalam pembangunan projek ini ialah Agile yang menggunakan lelaran pembangunan dan pengujian berterusan sepanjang kitaran hayat pembangunan perisian projek. Metodologi ini dipilih kerana ia merangkumi proses dinamik yang membenarkan perubahan walaupun sudah jauh dalam kitaran hayat pembangunan. Metodologi ini adalah amat sesuai dengan projek ini yang memerlukan pendekatan pembangunan permainan yang lebih fleksibel dan berorientasikan kepada pengguna. Dengan menggunakan metodologi Agile untuk projek ini, produk akhir yang berkualiti boleh dihasilkan.

1. Fasa Keperluan dan Rancangan

Dalam fasa Keperluan dan Reka Bentuk untuk Sistem Pengesanan Berita Palsu, adalah penting untuk mewujudkan keperluan teras sistem dan menggariskan fungsi keseluruhannya. Ini memerlukan memastikan keupayaan sistem untuk membezakan dengan tepat antara artikel berita palsu dan sah dengan ketepatan tinggi, pengesanan pantas selepas penerbitan, skalabiliti untuk memproses jumlah data yang besar, dan kebolehsuaian kepada taktik maklumat salah yang berkembang. Reka bentuk ini merangkumi sumber data daripada pelbagai platform, pemprosesan teks, pemodelan pembelajaran mesin, pengekstrakan ciri, pemarkahan kredibiliti dan mekanisme pengambilan tindakan. Sumber data hendaklah termasuk teks, imej dan video daripada pelbagai platform, manakala pelabelan data dengan tepat adalah penting untuk melatih model pembelajaran mesin. Pengekstrakan ciri melibatkan analisis aspek seperti sentimen, gaya penulisan, kredibiliti sumber dan maklumat kontekstual. Setelah dilatih, model memberikan skor kredibiliti kepada artikel berita, membolehkan sistem mengambil tindakan seperti membenderakan kandungan yang mencurigakan, mengeluarkan amaran atau melaporkan kepada pihak berkuasa. Perancangan yang teliti ini membentuk asas untuk membangunkan Sistem Pengesanan Berita Palsu yang berkesan.

2. Fasa Membangun

Dalam Fasa Pembangunan, tumpuan utama adalah membina komponen teras Sistem Pengesanan Berita Palsu. Ini melibatkan pembangunan komponen perisian untuk pengumpulan data, pemrosesan teks dan algoritma pembelajaran mesin. Perisian pengumpulan data direka untuk mengumpulkan maklumat daripada pelbagai sumber, termasuk artikel berita dalam talian, platform media sosial dan platform perkongsian kandungan. Data ini penting untuk melatih dan menguji model pembelajaran mesin. Alat pemrosesan teks dibangunkan untuk membersihkan, menghuraikan dan menyediakan data teks daripada artikel berita, menggunakan teknik pemrosesan bahasa semula jadi (NLP) apabila perlu. Inti fasa ini terletak pada melatih model pembelajaran mesin menggunakan set data yang sesuai yang termasuk contoh berlabel artikel berita palsu dan sah. Tambahan pula, jika sistem ditujukan untuk pengguna akhir, antara muka pengguna (UI) dibina untuk memudahkan interaksi dengan sistem. UI mungkin termasuk ciri seperti input teks untuk semakan ketulenan, memaparkan skor kredibiliti dan menyediakan pemberitahuan atau tindakan yang diperlukan untuk pengguna. Kerjasama yang berkesan dalam kalangan pembangun perisian, pakar NLP dan pakar domain adalah penting semasa fasa ini untuk memastikan kebolehpercayaan sistem, pemrosesan data yang tepat dan persembahan hasil yang mesra pengguna

3. Fasa Pengujian

Dalam Fasa Pengujian, objektif utama adalah untuk menilai dengan teliti prestasi Sistem Pengesanan Berita Palsu untuk memastikan ia berfungsi dengan berkesan. Data ujian yang pelbagai digunakan untuk menilai ketepatan sistem dalam membezakan antara artikel berita palsu dan sah. Metrik prestasi seperti ketepatan, ketepatan, kepekaan (ingat semula), dan skor F1 digunakan untuk mengukur keupayaan sistem untuk mengklasifikasikan item berita dengan betul. Melalui fasa ini, prestasi sistem dinilai secara menyeluruh, dan sebarang kelemahan atau bidang untuk penambahbaikan dikenal pasti. Setelah prestasi sistem memenuhi metrik yang ditetapkan dan dianggap memuaskan, ia boleh dianggap bersedia untuk penggunaan yang lebih luas dan penyepaduan ke dalam persekitaran yang berkaitan untuk memerangi ancaman berita palsu.

4. Fasa Implementasi

Dalam Fasa Pelaksanaan, selepas ujian yang berjaya, Sistem Pengesanan Berita Palsu disepadukan ke dalam penggunaan dunia sebenar untuk memberikan nilai kepada pengguna.

Ini melibatkan pemasangan perisian, konfigurasi pelayan dan persediaan rangkaian untuk menjalankan sistem dalam persekitaran pengeluaran. Memastikan akses pengguna yang betul, yang mungkin termasuk latihan, penciptaan akaun dan penugasan kebenaran, adalah penting untuk memudahkan interaksi pengguna dengan sistem dengan berkesan. Latihan pengguna mungkin diperlukan untuk memahami penggunaan sistem, navigasi UI dan tafsiran output sistem. Penyelenggaraan dan pemantauan berterusan adalah penting untuk menjamin prestasi sistem yang konsisten, merangkumi pemantauan pelayan, kemas kini perisian dan menangani sebarang isu yang mungkin timbul. Penilaian prestasi berterusan memastikan sistem kekal berkesan dalam mengenal pasti berita palsu dan menyesuaikan diri dengan ancaman yang muncul, dengan pilihan untuk mengemas kini model pembelajaran mesin secara berkala untuk prestasi yang lebih baik. Fasa Pelaksanaan menandakan peralihan kepada penggunaan praktikal, di mana sistem secara aktif memerangi berita palsu dan menyampaikan faedah dunia sebenar.

5. Fasa Semakan

Dalam Fasa Semakan, Sistem Pengesanan Berita Palsu sentiasa dipantau untuk sebarang isu atau penambahbaikan yang mungkin diperlukan dalam operasi berterusannya. Ini termasuk pemantauan prestasi proaktif untuk mengenal pasti sebarang penyelewengan dalam ketepatan, responsif atau kadar ralat, yang mesti ditangani dengan segera. Pengumpulan data baharu adalah penting untuk mencerminkan trend dan taktik terkini dalam penyebaran berita palsu, membolehkan latihan semula atau pengemaskinian model pembelajaran mesin untuk meningkatkan ketepatan sistem dalam mengesan berita palsu. Selain itu, penyelenggaraan dan penambahbaikan sistem yang berterusan dijalankan untuk menangani sebarang cabaran operasi yang mungkin timbul. Fasa ini juga menekankan keselamatan dan kebolehpercayaan untuk memastikan sistem kekal teguh terhadap potensi ancaman sambil mengekalkan ketersediaan perkhidmatan yang konsisten. Analisis data tetap dijalankan untuk mendapatkan cerapan tentang arah aliran dan corak yang muncul dalam maklumat salah. Fasa Semakan adalah penting untuk keberkesanan jangka panjang sistem dalam menentang berita palsu dengan sentiasa menyesuaikan diri dengan cabaran yang berkembang.

KEPUTUSAN DAN PERBINCANGAN

Penilaian Prestasi Model

Selepas menjalankan ujian ke atas pelbagai model pembelajaran mesin untuk pengesanan berita penipuan, hasil berikut diperhatikan:

Naive Bayes mencapai ketepatan 93.1%, menunjukkan keupayaannya untuk mengklasifikasikan sebahagian besar artikel berita dengan betul. Ketepatan adalah 92.9%, menunjukkan bahawa peratusan tinggi artikel yang dilabel sebagai penipuan adalah sememangnya penipuan. Penarikan balik berjumlah 93.8%, mencerminkan keupayaannya untuk mengenal pasti sebahagian besar artikel penipuan sebenar. Skor F1 sebanyak 93.3% mewakili prestasi seimbang antara ketepatan dan ingat semula. Matriks kekeliruan memaparkan 3998 negatif benar, 333 positif palsu, 287 negatif palsu dan 4383 positif benar.

Mesin Vektor Sokongan (SVM) menunjukkan prestasi unggul dengan ketepatan 99.3%. Ketepatannya adalah 99.4%, menunjukkan kebolehpercayaan yang sangat baik dalam mengenal pasti berita penipuan dengan betul. Penarikan balik adalah 99.2%, menekankan keberkesannya dalam menangkap hampir semua artikel penipuan. Skor F1 ialah 99.3%, mengesahkan keteguhan model itu. Matriks kekeliruan termasuk 4301 negatif benar, 30 positif palsu, 37 negatif palsu dan 4633 positif benar.

Random Forest mencapai ketepatan yang luar biasa pada 99.7%, mempamerkan kecekapannya yang tinggi dalam membezakan antara artikel penipuan dan tidak penipuan. Kedua-dua ketepatan dan ingatan adalah sama pada 99.7%, menunjukkan konsistensi model dalam mengenal pasti dan menangkap berita palsu. Skor F1 juga mencapai 99.7%. Matriks kekeliruan merekodkan 4316 negatif benar, 15 positif palsu, 15 negatif palsu dan 4655 positif benar.

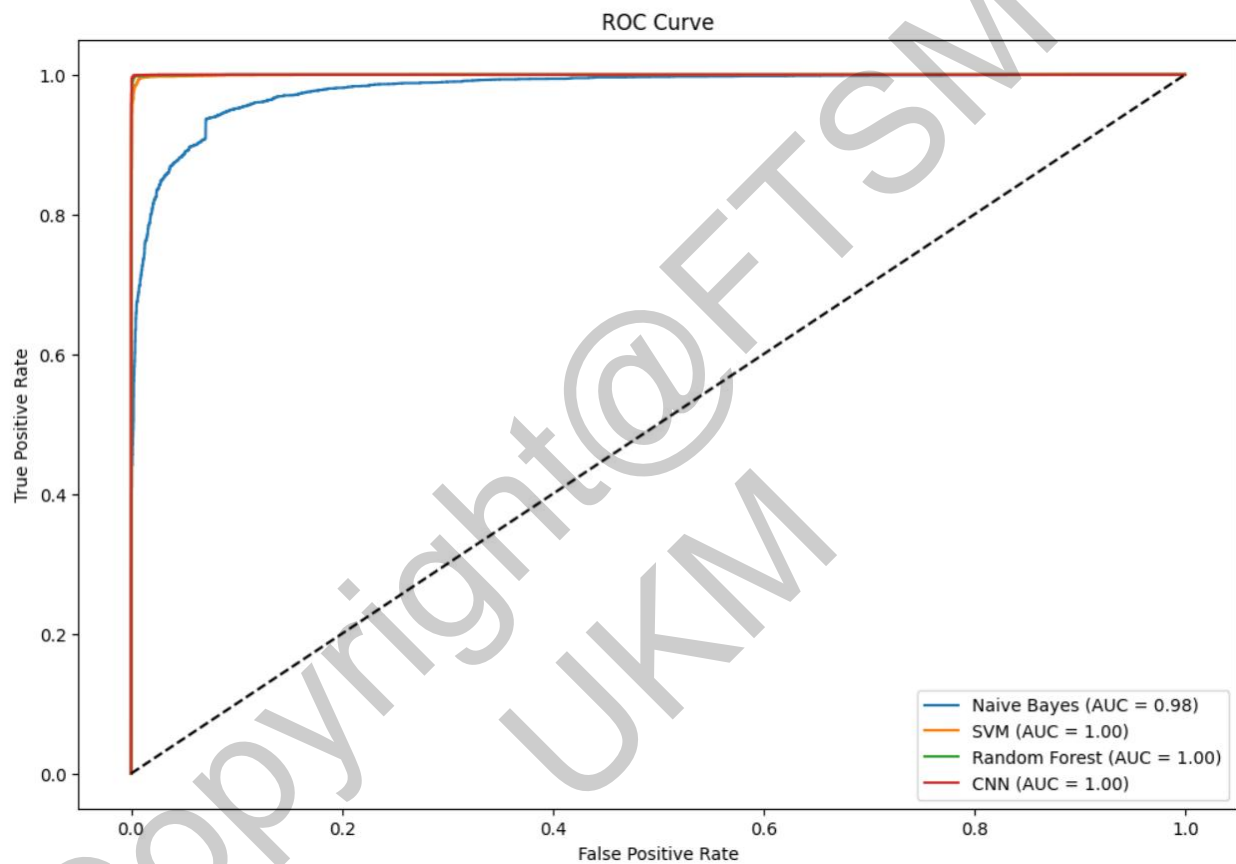
Rangkaian Neural Konvolusi (RNK) muncul sebagai pemain terbaik dengan ketepatan 99.8%. Ia mencapai ketepatan 99.8% dan penarikan balik sebanyak 99.9%, menyerlahkan keupayaan luar biasanya untuk mengklasifikasikan dan mengesan berita penipuan dengan betul. Skor F1 ialah 99.8%, mengukuhkan prestasi cemerlangnya. Matriks kekeliruan mendedahkan 4321 negatif benar, 10 positif palsu, 7 negatif palsu dan 4663 positif benar.

Penilaian Prestasi Model Berdasarkan Visualisasi

Lengkung ROC

Imej membentangkan lengkung Ciri Operasi Penerima (ROC), iaitu plot grafik yang menggambarkan keupayaan diagnostik pengelasan binari. Dalam kes ini, lengkung ROC untuk empat model pembelajaran mesin Naive Bayes, Support Vector Machine (SVM), Random Forest dan optimum. Rangkaian Neural Konvolusi (RNK) ditunjukkan. Berikut ialah penerangan terperinci tentang keputusan:

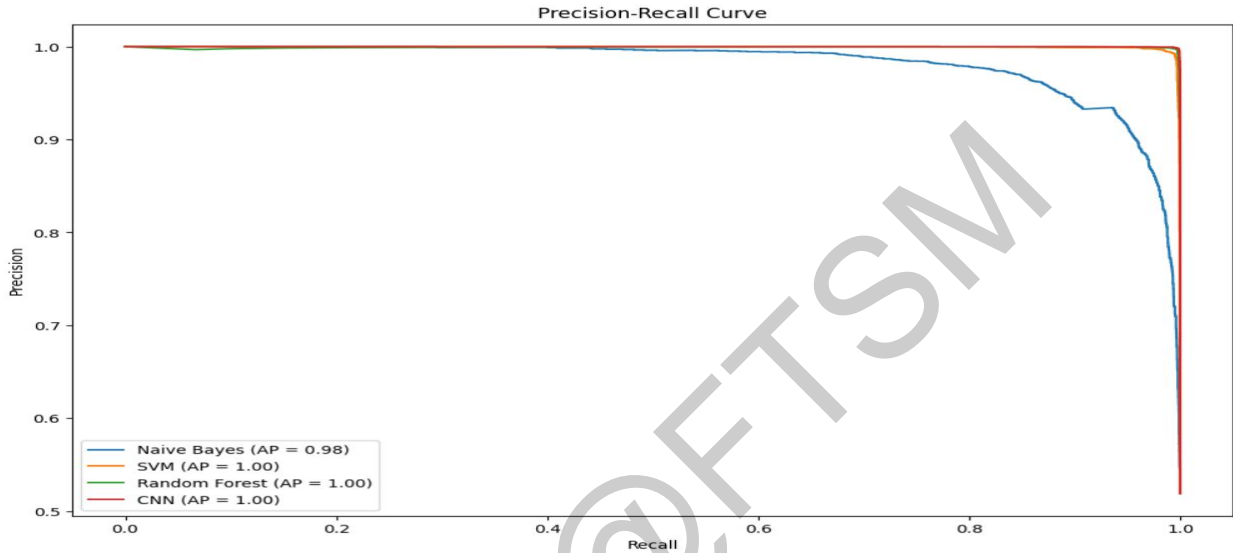
Analisis Lengkung ROC



Lengkung ROC dan nilai AUC menunjukkan bahawa walaupun semua model menunjukkan prestasi yang kukuh, model SVM, Random Forest dan RNK mempamerkan keupayaan pengelasan yang sempurna seperti yang dibuktikan oleh skor AUC mereka 1.00. Model Naive Bayes, dengan AUC 0.98, masih menunjukkan prestasi yang mengagumkan tetapi kurang berkesan berbanding tiga model lain. Secara keseluruhan, analisis ROC menyokong penemuan daripada metrik penilaian sebelumnya, menyerlahkan prestasi luar biasa SVM, Random Forest dan RNK dalam pengesanan berita penipuan.

Lengkung Ingatan Ketepatan

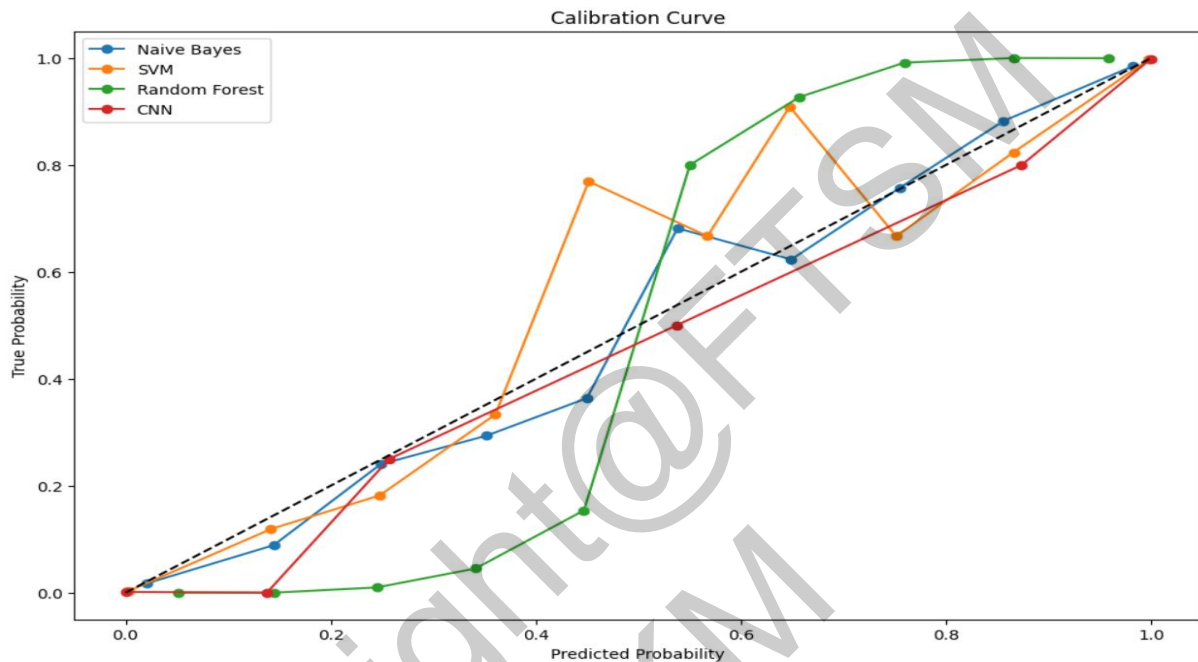
Imej itu membentangkan Lengkung Ingatan Ketepatan untuk empat model pembelajaran mesin Naive Bayes, Support Vector Machine (SVM), Random Forest dan Rangkaian Neural Konvolusi (RNK). Berikut ialah huraian terperinci tentang keputusan:



Lengkung Ingatan Ketepatan dan nilai AP menunjukkan bahawa walaupun semua model menunjukkan prestasi yang kukuh, model SVM, Random Forest dan RNK mempamerkan keupayaan pengelasan yang sempurna seperti yang dibuktikan oleh skor AP mereka 1.00. Model Naive Bayes, dengan AP 0.98, masih menunjukkan prestasi yang mengagumkan tetapi kurang berkesan berbanding tiga model lain. Secara keseluruhan, analisis Ingatan Ketepatan menyokong penemuan daripada matrik penilaian sebelumnya dan keluk ROC, menyerlahkan prestasi luar biasa SVM, Random Forest dan RNK dalam pengesanan berita penipuan.

Lengkung Kalibrasi

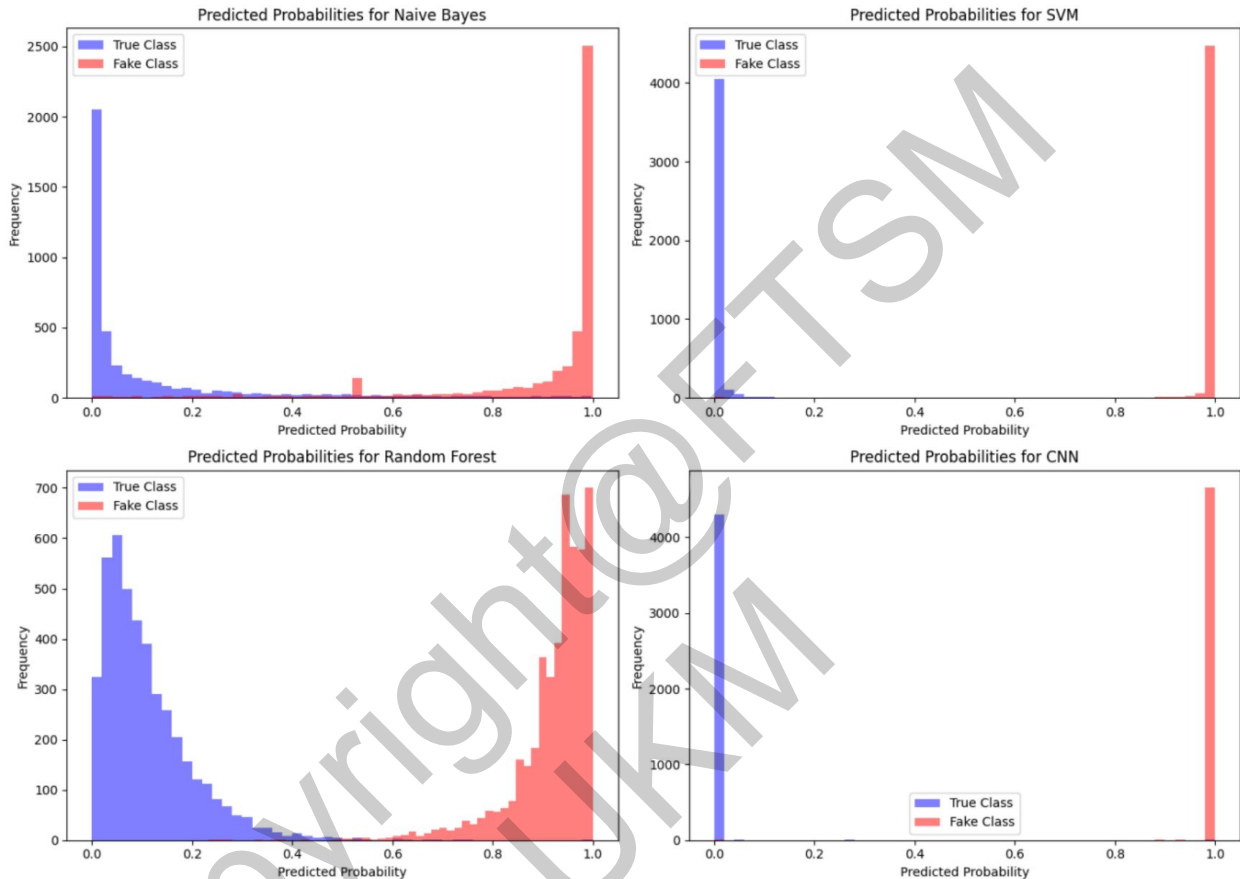
Imej itu membentangkan Lengkung Kalibrasi untuk empat model pembelajaran mesin Naive Bayes, Mesin Vektor Sokongan (SVM), Hutan Rawak dan Rangkaian Neural Konvolusi (CNN). Lengkung penentukuran, juga dikenali sebagai gambar rajah kebolehppercayaan, digunakan untuk menilai ketepatan kebarangkalian yang diramalkan.



Model SVM dan RNK mempunyai kebarangkalian yang ditentukan dengan lebih baik, dengan lengkungnya lebih dekat dengan garis pepenjuru, menunjukkan anggaran kebarangkalian yang boleh dipercayai. Naive Bayes mempunyai kebarangkalian yang ditentukan dengan munasabah tetapi dengan beberapa penyelewengan, mencadangkan ruang untuk penambahbaikan dalam anggaran kebarangkalian. Random Forest menunjukkan sisihan paling ketara daripada garis pepenjuru, menunjukkan penentukuran yang lemah dan anggaran kebarangkalian yang kurang dipercayai.

Kebarangkalian Ramalan

Imej ini menunjukkan kebarangkalian ramalan untuk empat model pembelajaran mesin yang berbeza (Naive Bayes, SVM, Random Forest dan CNN) dalam tugas pengelasan binari untuk membezakan antara "Kelas Benar" dan "Kelas Palsu".



Secara keseluruhan, SVM dan RNK menunjukkan ramalan yang paling yakin (mungkin terlalu yakin), manakala Naive Bayes dan Random Forest menunjukkan lebih ketidakpastian dalam ramalan mereka. Ini mungkin menunjukkan bahawa SVM dan CNN mungkin terlalu sesuai dengan data latihan, manakala Naive Bayes dan Random Forest mungkin memberikan anggaran kebarangkalian yang lebih benuansa.

Perbincangan Keputusan

Hasilnya menggariskan keteguhan SVM, Random Forest dan CNN dalam pengesanan berita penipuan. SVM dan Random Forest berkesan dalam memanfaatkan teknik pembelajaran ensemble untuk mencapai ketepatan tinggi dan klasifikasi yang mantap. Keupayaan CNN untuk menangkap corak rumit dalam data teks menyumbang kepada metrik prestasi unggulnya, menjadikannya sangat sesuai untuk tugas yang memerlukan pemahaman semantik yang mendalam.

Cadangan Penambahbaikan

Berdasarkan penemuan dan cabaran yang dikenal pasti sepanjang kajian mengenai pengesanan berita palsu menggunakan model pembelajaran mesin, beberapa cadangan utama boleh dibuat untuk penyelidikan masa depan. Ini termasuk pelaksanaan algoritma penyesuaian yang mampu berkembang pesat sebagai tindak balas kepada taktik penipuan baharu, peningkatan analisis kontekstual untuk meningkatkan ketepatan dengan memahami nuansa linguistik dan isyarat budaya yang kompleks, pengurangan bias data dengan menyusun set data latihan yang lebih seimbang dan mewakili, penyepaduan kaedah ensemble atau seni bina pembelajaran mendalam lanjutan untuk meningkatkan prestasi model, dan ujian berterusan dan penentukuran model yang digunakan untuk memastikan ramalan yang boleh dipercayai dan tepat. Penambahbaikan ini adalah penting untuk memajukan teknologi pengesanan berita palsu dan mengukuhkan pertahanan kami terhadap penyebaran maklumat yang mengelirukan dalam era digital.

KESIMPULAN

Secara keseluruhannya, kajian pengesanan berita palsu menggunakan model pembelajaran mesin ini telah berjaya dijalankan dengan data yang dianalisis dan diperolehi. Objektif dan keperluan yang ditetapkan pada awal kajian telah tercapai. Walaupun terdapat beberapa halangan, mereka diatasi dengan menggunakan pelbagai kaedah. Diharapkan kajian ini dapat dijadikan sebagai rujukan untuk kajian yang akan datang.

Kekuatan Sistem

Kekuatan sistem pengesanan berita palsu ini termasuk keupayaannya untuk menyesuaikan diri dan berkembang sebagai tindak balas kepada taktik baharu yang digunakan oleh mereka yang menyebarkan maklumat palsu. Sistem ini juga mendapat manfaat daripada set data latihan yang mantap dan pelbagai, yang meningkatkan ketepatan dan kebolehpercayaannya. Selain itu, penyepaduan seni bina pembelajaran mendalam termaju dan kaedah ensemble telah meningkatkan prestasi sistem dengan ketara.

Kelemahan Sistem

Satu batasan sistem ini ialah pergantungannya pada sambungan internet yang berterusan, yang boleh menjadi kekangan walaupun model itu bertujuan untuk diakses secara meluas. Semasa pembangunan, cabaran seperti kekurangan sumber pembelajaran yang komprehensif dan bimbingan pakar menimbulkan halangan yang ketara.

PENGHARGAAN

Penulis kajian ini ingin ucapkan setinggi-tinggi penghargaan dan jutaan terima kasih kepada Prof. Madya DR. Zalinda Binti Othman, 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

1. Jamal Abdul Nasir, Osama Subhani Khan, & Iraklis Varlamis. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. Data Science Institute, National University of Ireland Galway, Ireland; Department of Computer Science, International Islamic University Islamabad, Pakistan; Department of Informatics & Telematics, Harokopio University of Athens, Greece
2. Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake News Detection Using Machine Learning Ensemble Methods. Department of Computer Science and Information Technology, University of Engineering and Technology, Peshawar, Pakistan and Department of Mathematics and Computer Science, Karlstad University, Karlstad, Sweden.
3. Sharma, U., Saran, S., & Patil, S. M. (2021). Fake News Detection using Machine Learning Algorithms. International Journal of Engineering Research & Technology (IJERT), Special Issue. Mumbai, India
4. Gandhi, R. (2018). Support Vector Machine + Introduction to Machine Learning Algorithms. Towards Data Science. Retrieved July 11, 2024, from <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>
5. Saha, S. (2018). A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. Towards Data Science., from <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
6. Simplilearn. (n.d.). How to Create a Fake News Detection System. Simplilearn. , from <https://www.simplilearn.com/tutorials/machine-learning-tutorial/how-to-create-a-fake-news-detection-system>

7. Analytics Vidhya. (2023, February). Tackling Fake News with Machine Learning. Analytics Vidhya, from <https://www.analyticsvidhya.com/blog/2023/02/tackling-fake-news-with-machine-learning/>
8. AltexSoft. (n.d.). Language Models: How Do They Work? AltexSoft., from <https://www.altexsoft.com/blog/language-models-gpt/>
9. GeeksforGeeks. (n.d.). How to plot ROC curve in Python? GeeksforGeeks. Retrieved July 11, 2024, from <https://www.geeksforgeeks.org/how-to-plot-roc-curve-in-python/>
10. Machine Learning Mastery. (n.d.). ROC Curves and Precision-Recall Curves for Classification in Python. Machine Learning Mastery. from <https://machinelearningmastery.com/roc-curves-and-precision-recall-curves-for-classification-in-python/>
11. GeeksforGeeks. (n.d.). Calibration Curves. GeeksforGeeks. from <https://www.geeksforgeeks.org/calibration-curves/>
12. Real Python. (n.d.). Python Histograms: A Visual Introduction. Real Python. from <https://realpython.com/python-histograms/>

Muhammad Qashfullah bin Abdul Gaffar (A192597)

Prof. Madya DR. Zalinda Binti Othman

Fakulti Teknologi & Sains Maklumat

Universiti Kebangsaan Malaysia