

PENGESANAN BERITA PALSU MENGGUNAKAN PENDEKATAN PEMBELAJARAN MENDALAM CNN

NUR HUSNINA BINTI FAIZAL

DR SAIDAH BINTI SAAD

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

ABSTRAK

Penyebaran maklumat palsu di laman media sosial seperti Instagram, TikTok, dan Twitter telah menimbulkan masalah serius yang memberi kesan buruk kepada masyarakat. Penyelidikan ini bertujuan untuk menangani isu ini dengan membangunkan dan menguji sistem pengesanan berita palsu berasaskan pembelajaran mesin. Sistem ini menggunakan model Convolutional Neural Network (CNN) dan Local Interpretable Model-agnostic Explanations (LIME) untuk menganalisis dan menilai artikel berita. Masalah yang ditangani oleh penyelidikan ini ialah penyebaran berita palsu yang pesat dan meluas, yang menjejaskan kepercayaan orang ramai dan memutarbelitkan wacana awam. Untuk menyelesaikan masalah ini, sistem teguh yang bukan sahaja mengesan berita palsu tetapi juga memberikan ketelusan dalam proses membuat keputusannya, dicadangkan. Strategi pembangunan melibatkan penggunaan CNN untuk kecekapannya dalam mengendalikan data teks dan LIME untuk meningkatkan kebolehtafsiran. Model ini dilatih pada set data daripada Kaggle, yang termasuk artikel berita benar dan palsu. Google Colab digunakan sebagai platform utama untuk latihan model dan analisis data kerana sumber pengiraan dan ciri kolaboratifnya. Selain itu, sistem berasaskan web dibangunkan menggunakan Flask untuk menyediakan antara muka mesra pengguna, memudahkan pengguna memasukkan artikel berita dan menerima hasil analisis. Hasil daripada projek ini ialah sistem pengesanan berita palsu yang sangat tepat dan boleh ditafsir. Model CNN, disepadukan dengan LIME, memastikan pengguna dapat memahami sebab artikel berita tertentu diklasifikasikan sebagai palsu atau benar. Sistem akhir menunjukkan metrik prestasi yang kukuh, dengan ketepatan 97.92%, mempamerkan keberkesanannya sebagai alat dalam memerangi penyebaran berita palsu dalam era digital. Sistem ini bukan sahaja membantu dalam mengesan maklumat palsu tetapi juga meningkatkan kepercayaan pengguna melalui penjelasan yang telus.

Kata kunci: CNN, LIME, maklumat palsu

Pengenalan

Penyebaran berita palsu yang semakin meluas melalui platform media sosial seperti Twitter, Instagram dan TikTok telah menjadi satu kebimbangan terhadap individu dan masyarakat. Penyebaran berita palsu menjadi lebih tidak terkawal dalam era kini di mana maklumat disebar

dan diterima secara pantas, seterusnya mempengaruhi pendapat awam, proses membuat keputusan dan kestabilan sosial. Berita palsu, yang ditakrifkan sebagai penyebaran maklumat palsu secara sengaja, merupakan topik yang rumit dan telah menarik perhatian ahli akademik, penyelidik dan penggubal dasar. Penyebaran berita palsu dapat memberikan pelbagai impak, termasuk rasa kurang percaya terhadap media, penyelewengan perbincangan awam dan mungkin menggalakkan ketidakstabilan sosial. Oleh itu, mengatasi cabaran pengesanan berita palsu adalah penting untuk mengurangkan impak tersebut. Untuk membezakan antara berita yang benar dan berita palsu, teknologi moden seperti algoritma pembelajaran mesin, perlu diaplikasikan.

Projek ini mengkaji pembangunan sistem pengesanan berita palsu dengan menggunakan pendekatan pembelajaran mesin seperti Convolutional Neural Network (CNN) dan teknik interpretasi model seperti *Local Interpretable Model-agnostic Explanations* (LIME), dengan matlamat untuk meningkatkan ketelusan dan kebolehtafsiran dalam pengesanan berita palsu dan model dibangunkan menggunakan TensorFlow dan Keras, dan diuji dalam aplikasi web menggunakan Flask. Selain daripada membanteras salah komunikasi dalam era digital, penyelidikan ini juga bertujuan untuk menjelaskan kesan pelbagai algoritma dan pelbagai kaedah pengekstrakan ciri terhadap prestasi sistem yang ingin dibangunkan. Kepentingan topik ini ditekankan oleh pertumbuhan penyelidikan yang menekankan keperluan alat yang kukuh dalam membanteras wabak berita palsu. Dalam konteks lanskap digital yang terus mengembang, terdapat alasan yang kukuh untuk penyelesaian inovatif yang bukan hanya mengesan berita palsu dengan berkesan tetapi juga memberikan wawasan tentang proses membuat keputusan, yang akan memperkuat kepercayaan dalam sistem automatik yang direka untuk mengatasi masalah penyebaran berita palsu.

METODOLOGI KAJIAN

Bagi projek ini, model proses yang sesuai adalah model Pembangunan Berperingkat. Pembangunan berperingkat ialah strategi pembangunan perisian yang melibatkan pembinaan perisian dalam bentuk fasa. Setiap peringkat menambah baik sistem dan memperkenalkan ciri tambahan. Ia berbeza daripada model 'Waterfall,' 'Linear' dan 'One-Way,' di mana setiap peringkat harus diselesaikan sebelum meneruskan ke peringkat seterusnya. Setiap peringkat dalam model Pembangunan Inkremental mungkin termasuk banyak kitaran pembangunan pendek seperti ujian, peningkatan dan pengoptimuman.

Fasa Penyelidikan Awal

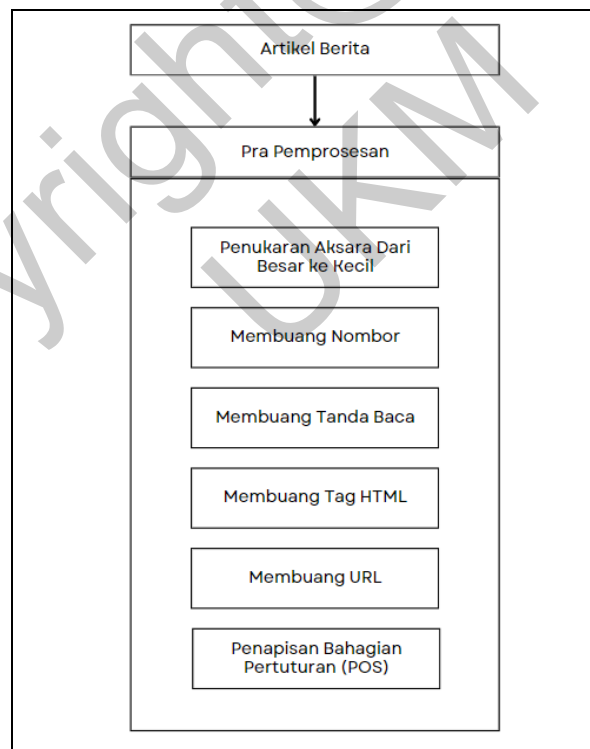
Dalam fasa penyelidikan awal, projek bermula dengan mengumpul keperluan pengguna dan sistem. Ini melibatkan pemahaman masalah yang dihadapi oleh pengguna, mentakrifkan objektif projek, dan mengenal pasti keperluan fungsian dan bukan fungsian yang mesti dipenuhi oleh sistem. Menjalankan kajian literatur yang komprehensif membantu dalam memahami pendekatan sedia ada untuk pengesanan berita palsu dan mengenal pasti jurang yang perlu ditangani. Perancangan projek terperinci berikut, menggariskan garis masa, fasa pembangunan, dan sumber yang diperlukan. Fasa ini menetapkan asas untuk projek dengan menyediakan peta jalan yang jelas dan pemahaman tentang matlamat dan cabaran projek.

Fasa Pengumpulan Data

Fasa pengumpulan data melibatkan pengumpulan set data yang berkaitan daripada pelbagai sumber. Untuk projek ini, set data yang mengandungi berita palsu dan sebenar diperoleh daripada repositori yang tersedia secara umum. Memastikan kualiti dan kaitan data adalah penting untuk kejayaan model. Data yang dikumpul kemudiannya disimpan dalam format berstruktur, seperti fail CSV, untuk memudahkan akses dan manipulasi yang mudah semasa fasa berikutnya. Organisasi dan penyimpanan data yang betul adalah penting untuk pemprosesan dan analisis data yang cekap.

Fasa Pra Pemprosesan Data

Pra pemprosesan data ialah fasa kritikal di mana data mentah dibersihkan dan disediakan untuk analisis. Langkah ini termasuk mengalih keluar bendasing dan unsur yang tidak berkaitan seperti URL, tag HTML, tanda baca dan nombor. Membersihkan data memastikan bahawa hanya teks yang bermakna disimpan untuk analisis. Data teks kemudiannya dinormalisasi dengan menukar semua aksara kepada huruf kecil dan mengalih keluar ruang putih tambahan bagi mengekalkan konsistensi merentas set data. Selain itu, penapisan Bahagian Pertuturan (POS) digunakan untuk mengekalkan hanya perkataan yang berkaitan, seperti kata nama, kata kerja dan kata sifat, yang penting untuk analisis.



Rajah 1 Proses Pra Pemprosesan

Fasa Pengekstrakan Ciri

Dalam fasa pengekstrakan ciri, teks yang dibersihkan dan dinormalkan ditukar menjadi urutan token menggunakan Tokenizer. Tokenisasi memecahkan teks kepada perkataan atau token individu yang boleh diproses selanjutnya oleh model. *Padding* digunakan pada jujukan token

ini untuk memastikan semua jujukan mempunyai panjang yang sama, yang diperlukan untuk memasukkan data ke dalam model rangkaian saraf. Langkah ini mengubah data teks kepada format yang sesuai untuk algoritma pembelajaran mesin.

Fasa Latihan Model

Semasa fasa latihan model, set data dibahagikan kepada set latihan dan ujian. Set latihan digunakan untuk melatih model, manakala set ujian digunakan untuk menilai prestasinya. Model Rangkaian Neural Konvolusi (CNN) dilaksanakan dan dilatih menggunakan data pra-proses dalam Google Colab. Proses latihan melibatkan memasukkan data ke dalam model, melaraskan berat dan mengoptimimumkan model untuk meningkatkan ketepatannya dalam mengesan berita palsu. Setelah dilatih, model dan tokenizer disimpan untuk kegunaan masa hadapan, memastikan model terlatih boleh digunakan dan digunakan tanpa latihan semula.

Fasa Penilaian Model

Fasa penilaian model melibatkan penilaian prestasi model terlatih menggunakan metrik seperti ketepatan dan laporan klasifikasi. Metrik ini memberikan pandangan tentang keberkesanan model dalam membezakan antara berita palsu dan sebenar. Berdasarkan keputusan penilaian, penambahbaikan dan pengoptimuman dilaksanakan untuk meningkatkan prestasi model. Penilaian dan pengoptimuman berterusan adalah penting untuk mengekalkan tahap ketepatan dan kebolehpercayaan yang tinggi.

Fasa Penerapan Model

Dalam fasa penggunaan model, model terlatih digunakan untuk mengesan berita palsu dalam artikel berita baharu. Ini melibatkan penggunaan model pada data dunia sebenar dan menjana ramalan. Selain itu, teknik tafsiran model seperti LIME digunakan untuk menerangkan keputusan model kepada pengguna, memberikan ketelusan dan membantu pengguna memahami sebab artikel berita tertentu diklasifikasikan sebagai palsu atau nyata.

Fasa Pembentangan Keputusan

Fasa terakhir melibatkan pembentangan hasil sistem pengesanan berita palsu. Ini termasuk menyediakan perwakilan visual hasil pengesanan dan menjana laporan terperinci tentang analisis. Hasilnya dipaparkan dalam cara yang mesra pengguna, memastikan pengguna dapat memahami dan mentafsir penemuan dengan mudah. Pembentangan hasil yang berkesan adalah penting untuk membina kepercayaan pengguna dan memastikan aplikasi praktikal sistem dalam senario dunia sebenar.

Reka bentuk algoritma Sistem Pengesanan Berita Palsu direka dengan teliti untuk menangani cabaran rumit dalam mengenal pasti maklumat salah dalam artikel berita. Dengan menggunakan algoritma pembelajaran mesin lanjutan seperti CNN dan teknik tafsiran seperti LIME, sistem ini memastikan analisis yang mantap dan komprehensif. Setiap fasa, dari penyelidikan awal hingga pembentangan akhir, disusun dengan teliti untuk meningkatkan ketepatan dan kebolehpercayaan sistem. Melalui langkah terperinci yang merangkumi pengumpulan data, pra pemprosesan, pengekstrakan ciri, latihan model, penilaian, penggunaan dan pembentangan hasil, projek ini bukan sahaja bertujuan untuk mengesan berita palsu dengan

berkesan tetapi juga memberikan hasil yang telus dan boleh ditafsir kepada pengguna. Pendekatan berstruktur ini adalah penting dalam membangunkan alat yang boleh dipercayai untuk memerangi penyebaran berita palsu dalam era digital hari ini.

KEPUTUSAN DAN PERBINCANGAN

Projek Pengesanan Berita Palsu yang efektif untuk penggunaan pihak ketiga telah berjaya dibangunkan dan semua dokumentasinya telah dilengkapkan. Semasa proses pembangunan, projek ini dibangunkan menggunakan Flask dengan bahasa pengaturcaraannya yang dipanggil Python. Teknologi pembelajaran mesin yang digunakan termasuk TensorFlow dan Keras. Model dilatih dan diuji menggunakan Google Colab untuk memanfaatkan sumber komputasi awan. Selain itu, Pandas, NumPy, NLTK, SpaCy, dan scikit-learn digunakan untuk pelbagai tugas pemprosesan data dan pembelajaran mesin. Antara muka pengguna dibangunkan dengan HTML, CSS, dan JavaScript untuk memastikan pengalaman pengguna yang mesra dan berkesan.

Antara Muka Memuat Naik Dokumen

Antara muka ini membolehkan pengguna memasukkan teks berita atau memuat naik fail berita untuk mengesan kesahihannya. Antara muka mempunyai kotak teks yang besar untuk menampal teks berita dan butang "*Choose File*" untuk memuat naik fail. Setelah input diberikan, pengguna boleh mengklik butang "*Detect*" untuk menganalisis berita untuk kesahihannya. Antara muka adalah ringkas dan mesra pengguna, direka untuk memudahkan input mudah data berita untuk analisis.



Rajah 2 Antara Muka Memuat Naik Dokumen

Antara Muka Input Teks

Selepas memilih fail atau menampal teks, antara muka menunjukkan teks berita yang dimasukkan dalam kotak input. Pengguna boleh menyemak teks dan memastikan ia betul sebelum meneruskan. Setelah berpuas hati, mereka mengklik butang "*Detect News*" untuk

memulakan analisis.



Rajah 3 Antara Muka Input Teks

Antara Muka Pengesanan dalam Proses

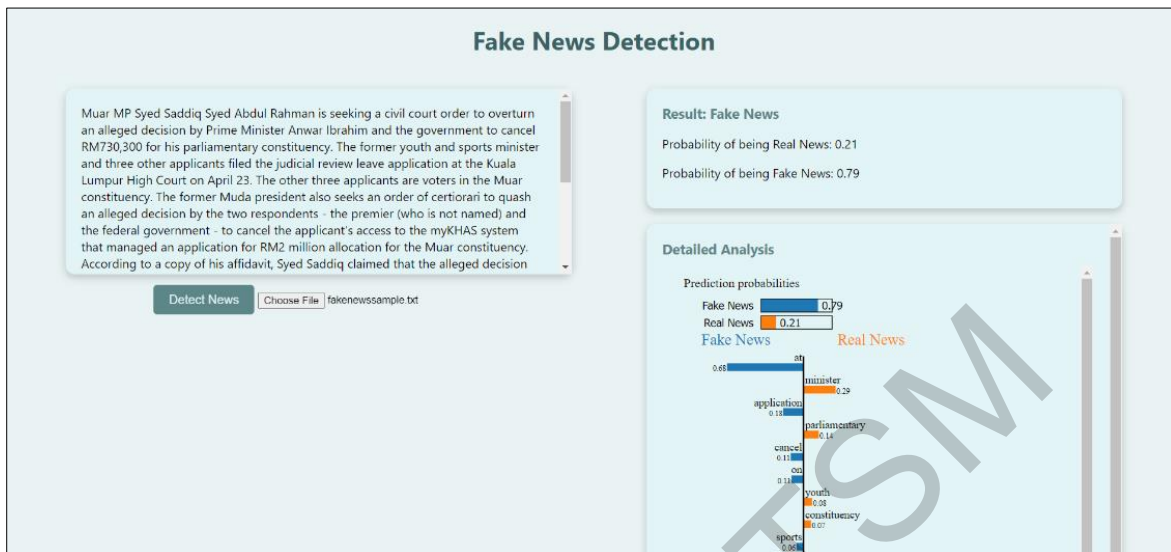
Antara muka ini menunjukkan bahawa proses pengesanan berita sedang dijalankan. Butang "Detect News" bertukar menjadi "Detecting...", memaklumkan pengguna bahawa sistem sedang menganalisis input. Ini memberikan maklum balas kepada pengguna bahawa permintaan mereka sedang diproses.



Rajah 4 Antara Muka Pengesanan dalam Proses

Antara Muka Output Hasil

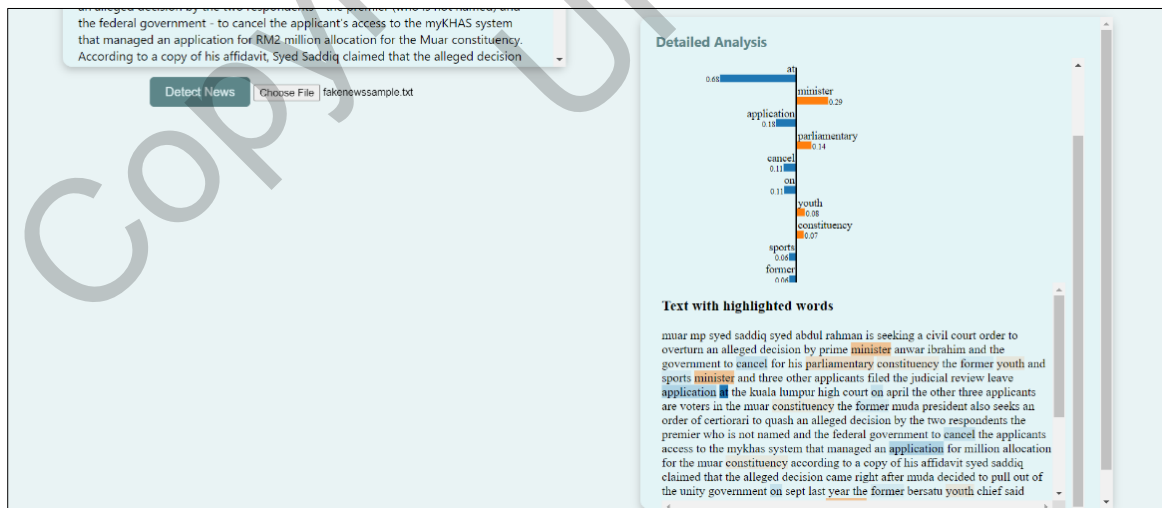
Setelah selesai analisis, antara muka ini memaparkan keputusan pada skrin. Ia termasuk kebarangkalian berita itu benar atau palsu. Sebagai contoh, ia menunjukkan kebarangkalian 0.79 untuk berita palsu dan 0.21 untuk berita sebenar. Hasilnya dibentangkan dengan jelas untuk membantu pengguna memahami kesahihan berita dengan cepat.



Rajah 5 Antara Muka Output Hasil

Antara Muka Analisis Terperinci

Antara muka ini memberikan penjelasan yang lebih mendalam tentang keputusan. Ia termasuk perwakilan grafik kebarangkalian ramalan dan pecahan faktor penyumbang. Contohnya, kata kunci dan frasa yang mempengaruhi klasifikasi berita palsu diserlahkan, membantu pengguna memahami sebab berita itu diklasifikasikan sebagai palsu. Penjelasan kapur dalam kotak analisis terperinci menawarkan penjelasan yang komprehensif, menjadikan proses pengesanan telus dan bermaklumat.



Rajah 6 Antara Muka Analisis Terperinci

Pengujian Kebolehgunaan

Bahagian ini melaporkan hasil ujian yang dijalankan untuk sistem Pengesanan Berita Palsu. Ujian ini menilai pelbagai aspek sistem, termasuk ketepatan, precision, recall, skor-F1, dan prestasi keseluruhannya.

Model terlatih dengan ketepatan 97.92% pada set data ujian, menunjukkan model mampu mengenal pasti artikel berita palsu dan sebenar dengan tingkat kebolehpercayaan yang tinggi. Precision model untuk mengesan berita palsu adalah 0.98, dan untuk berita sebenar adalah 0.99. Recall untuk berita palsu adalah 0.99, dan untuk berita sebenar adalah 0.98. Model ini mencapai skor-F1 0.98 untuk kedua-dua kelas berita palsu dan sebenar, menunjukkan keseimbangan yang baik antara precision dan recall.

	precision	recall	f1-score	support
0	0.98	0.98	0.98	4733
1	0.98	0.98	0.98	4247
accuracy			0.98	8980
macro avg	0.98	0.98	0.98	8980
weighted avg	0.98	0.98	0.98	8980
Accuracy: 0.9791759465478842				

Rajah 7 Laporan Klasifikasi

Cadangan Penambahbaikan

Untuk meningkatkan lagi pencapaian sistem, beberapa penambahbaikan boleh dipertimbangkan. Mengemas kini dan mengembangkan set data secara berterusan untuk menyertakan artikel berita yang lebih pelbagai dan terkini boleh meningkatkan ketepatan dan keteguhan model. Melaksanakan teknik latihan yang lebih cekap dan menggunakan sumber pengiraan yang lebih berkuasa boleh mengurangkan masa latihan dan meningkatkan prestasi sistem. Meneroka teknik lanjutan untuk menyepadukan LIME atau kaedah kebolehtafsiran lain boleh meningkatkan ketelusan dan kebolehgunaan sistem. Memperluas sistem untuk mengesan berita palsu dalam pelbagai bahasa boleh meluaskan kebolehgunaan dan kegunaannya. Selain itu, menggabungkan gelung maklum balas pengguna untuk memperjelas dan menambah baik model secara berterusan berdasarkan penggunaan dan maklum balas dunia sebenar boleh meningkatkan keberkesanan sistem dengan ketara.

KESIMPULAN

Secara keseluruhannya, sistem ini telah berjaya dibangunkan dengan menggunakan data yang telah dikaji dan diperolehi. Objektif kajian dan keperluan yang telah ditetapkan sebelum ini telah berjaya dicapai. Walaupun terdapat beberapa halangan, ia berjaya diatasi menggunakan pelbagai cara.

Kekuatan Sistem

Projek ini mempamerkan beberapa kekuatan. Penggunaan CNN dan teknik pra pemrosesan yang komprehensif menghasilkan sistem pengesanan yang sangat tepat dan boleh dipercayai. Menggabungkan LIME memberikan penjelasan yang jelas untuk klasifikasi, meningkatkan

kepercayaan dan pemahaman pengguna. Pembangunan aplikasi web berasaskan Flask memastikan pengguna boleh berinteraksi dengan sistem dengan mudah, menjadikannya boleh diakses dan praktikal untuk kegunaan dunia sebenar. Selain itu, reka bentuk modular dan penggunaan sumber berasaskan awan seperti Google Colab membenarkan pengendalian yang cekap bagi set data yang besar dan kebolehskalaan sistem.

Kelemahan Sistem

Walau bagaimanapun, projek itu juga menghadapi batasan tertentu. Proses latihan model boleh memakan masa, terutamanya dengan set data yang besar, yang mungkin menjejaskan kecekapan sistem. Walaupun LIME menyediakan kebolehtafsiran yang berharga, menyepadukannya dengan lancar ke dalam aplikasi web menimbulkan cabaran, menunjukkan ruang untuk pengoptimuman selanjutnya. Prestasi sistem sangat bergantung pada kualiti dan kepelbagaian set data latihan, yang mungkin mengehadkan keberkesannya dalam mengesan jenis berita palsu yang lebih baharu atau kurang biasa.

PENGHARGAAN

Segala pujian bagi Allah atas kurniaan ilhamNya di sepanjang berlangsungnya perancangan dan pembangunan bagi projek ini.

Terima kasih yang tidak terhingga kepada penyelia projek tahun akhir saya, Dr. Saidah binti Saad atas dorongan dan panduan bermakna yang telah diberikan selama dua semester kajian ini dilakukan, dan juga kepada para pensyarah dan pihak Fakulti Teknologi & Sains Maklumat yang telah banyak memberikan tunjuk ajar melalui sesi taklimat projek tahun akhir mingguan serta bahan-bahan rujukan berkaitan.

Terima kasih yang tidak terkata kepada ibu bapa, keluarga, dan rakan-rakan seperjuangan atas doa, sokongan dan tunjuk ajar yang telah diberikan kepada saya. Akhir bicara, terima kasih dan tahniah kepada diri saya sendiri kerana telah berjaya melengkapkan projek tahun akhir untuk tahun ini dengan penuh dedikasi walau pelbagai kekangan dan halangan yang muncul sepanjang kajian ini dijalankan.

RUJUKAN

- Ahmed, A. A. A., Aljabouh, A., Donepudi, P. K., & Choi, M. S. (2021). Detecting fake news using machine learning: A systematic literature review. *arXiv preprint arXiv:2102.04458*.
- Ain, Q. T., Ali, M., Riaz, A., Noreen, A., Kamran, M., Hayat, B., & Rehman, A. (2017). Sentiment analysis using deep learning techniques: a review. *International Journal of Advanced Computer Science and Applications*, 8(6).
- Al-Fedaghi, S. (2021). UML sequence diagram: an alternative model. *arXiv preprint arXiv:2105.15152*.
- Berrar, D. (2018). Bayes' theorem and naive Bayes classifier. *Encyclopedia of bioinformatics and computational biology: ABC of bioinformatics*, 403, 412.

- Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, 2(01), 20-28.
- Choudhary, M., Jha, S., Saxena, D., & Singh, A. K. (2021, May). A review of fake news detection methods using machine learning. In *2021 2nd International Conference for Emerging Technology (INCET)* (pp. 1-5). IEEE.
- Gupta, S., & Meel, P. (2021). Fake news detection using passive-aggressive classifier. In *Inventive Communication and Computational Technologies: Proceedings of ICICCT 2020* (pp. 155-164). Springer Singapore.
- Jain, A., & Kasbe, A. (2018, February). Fake news detection. In *2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)* (pp. 1-5). IEEE.
- Khalid, S., Khalil, T., & Nasreen, S. (2014, August). A survey of feature selection and feature extraction techniques in machine learning. In *2014 science and information conference* (pp. 372-378). IEEE.
- Khanam, Z., Alwasel, B. N., Sirafi, H., & Rashid, M. (2021, March). Fake news detection using machine learning approaches. In *IOP conference series: materials science and engineering* (Vol. 1099, No. 1, p. 012040). IOP Publishing.
- Kong, S. H., Tan, L. M., Gan, K. H., & Samsudin, N. H. (2020, April). Fake news detection using deep learning. In *2020 IEEE 10th symposium on computer applications & industrial electronics (ISCAIE)* (pp. 102-107). IEEE.
- Leon, F., Floria, S. A., & Bădică, C. (2017, July). Evaluating the effect of voting methods on ensemble-based classification. In *2017 IEEE international conference on INnovations in intelligent Systems and applications (INISTA)* (pp. 1-6). IEEE.
- Nagashri, K., & Sangeetha, J. (2021). Fake news detection using passive-aggressive classifier and other machine learning algorithms. In *Advances in Computing and Network Communications: Proceedings of CoCoNet 2020, Volume 2* (pp. 221-233). Springer Singapore.
- Nordin, W. A. F. B., Alfred, R., Yee, C. P., Tanalol, S. H., Loudin, R. V., & Iswandono, Z. (2022, August). Malay Fake News Classification Using a Deep Learning Approach. In *International Conference on Computational Science and Technology* (pp. 17-32). Singapore: Springer Nature Singapore.
- Oshikawa, R., Qian, J., & Wang, W. Y. (2018). A survey on natural language processing for fake news detection. *arXiv preprint arXiv:1811.00770*.
- Rolland, C., & Achour, C. B. (1998). Guiding the construction of textual use case specifications. *Data & Knowledge Engineering*, 25(1-2), 125-160.
- Shaikh, J., & Patil, R. (2020, December). Fake news detection using machine learning. In *2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC)* (pp.1-5). IEEE.
- Shukla, S., Maheshwari, A., & Johri, P. (2021, December). Comparative analysis of ml algorithms & stream lit web application. In *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)* (pp. 175-180). IEEE.
- Singh, A. K., & Shashi, M. (2019). Vectorization of text documents for identifying unifiable news articles. *International Journal of Advanced Computer Science and Applications*,

10(7).

Zakiah Hanum Hj Hashim, & Nor Samsiah Sani (2022). Gabungan Teknik Pengekstrakan Fitur TF-IDF Dengan Pembelajaran Ensembel Voting untuk Pengesanan Berita Palsu. PS-FTSM-2022-008.

Nur Husnina binti Faizal (A188456)

Dr. Saidah Saad

Fakulti Teknologi & Sains Maklumat

Universiti Kebangsaan Malaysia

Copyright@FTSM
UKM