

## **PENGESAHAN BERITA PALSU MELALUI MODEL PEMBELAJARAN MESIN**

Kai Xuan Koh<sup>1</sup>

Wandeep Kaur A/P Ratan Singh<sup>2</sup>

<sup>1,2</sup>*Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia, 43600 UKM Bangi,  
Selangor Darul Ehsan, Malaysia*

### **Abstrak**

Dalam era maklumat yang terlalu sarat di mana maklumat salah boleh tersebar dengan cepat yang menjadikan berita palsu menjadi kebimbangan serius dalam masyarakat. Pengesanan berita palsu telah menjadi satu bidang kajian yang penting. Penyelidikan ini berobjektif untuk menerangkan kaedah bagi mengesan berita palsu menggunakan teknik pembelajaran mesin. Teknik yang dicadangkan menggunakan julat luas ciri teks dan konteks seperti TF-IDF, N-gram, CBOW. Ciri ini kemudiannya digunakan untuk melatih pengelas pembelajaran mesin, seperti Regresi Logistik, Naïve Bayes, Mesin Vektor Sokongan (SVM), Hutan Rawak, dan Perceptron Berbilang Lapisan (MLP) untuk membezakan antara berita sebenar dan palsu. Percubaan pada set data penanda aras berskala besar yang berbeza membandingkan prestasi pengelas pada ketepatan, ketepatan (precision), nilai ingatan semula, skor F1 dan nilai AUC. Dapatan utama projek ini adalah ciri CBOW dan ciri gabungan antara CBOW dan N-gram (CBOW x N-gram) menunjukkan keputusan yang baik untuk semua set data yang telah diuji berbanding dengan ciri lain. Selain itu, setiap model pembelajaran mesin mencapai prestasi yang berbeza pada setiap set data penanda aras yang telah digunakan disebabkan setiap set data memberi fokus kepada isu topik yang berbeza. Oleh itu, pelbagai eksperimen akan dijalankan pada set data penanda aras dunia sebenar untuk menunjukkan ciri tekstual

dan model pembelajaran mesin yang dicadangkan dengan menunjukkan prestasi yang ketara untuk pengesanan berita palsu.

**Kata kunci:** [Pengesanan berita palsu, Klasifikasi teks, Pembelajaran mesin, CBOW, N-grams ]

### Pengenalan

Disebabkan perkembangan teknologi dan perkongsian maklumat, perkataan “berita palsu” telah banyak didengar dalam kehidupan harian kita (Tandoc, Jenkins & Craft 2018). Berita palsu ditakrifkan sebagai maklumat mengelirukan atau cerita yang seolah-olah berita dan tersebar dengan pantas di seluruh internet menggunakan pelbagai platform media sosial seperti Facebook, Twitter, Instagram, dan TikTok (Shu, Wang & Liu 2019). Berita palsu dikongsi secara meluas melalui platform media sosial terutamanya Facebook dan Twitter adalah tempat pembiakan berita palsu (Allcott & Gentzkow 2017). Berita palsu disifatkan sebagai satu maklumat yang mengandungi maklumat sensitif dan bertujuan untuk memperdaya pembaca. Motif utama penyebaran berita palsu adalah untuk mengelirukan pembaca, merosakkan reputasi seseorang dan mendapat keuntungan daripada sensasi (Thota 2018). Walau bagaimanapun, berita tepat adalah sukar untuk dikategorikan kerana terdapat kedua-dua maklumat yang boleh dipercayai dan tidak boleh dipercayai di internet. Media sosial adalah berasaskan kepada internet dan membolehkan pengguna berkongsi kandungan seperti maklumat peribadi, dokumen, filem dan imej secara berlangsung (Srikanth 2019). Pengguna berinteraksi dengan media sosial menggunakan perisian atau aplikasi berasaskan web pada komputer, tablet atau telefon pintar. Platform media sosial membolehkan pengguna meluahkan perasaan mereka dengan cara yang tidak ditapis dan secara suka hati mereka (Darr & Doss 2022). Menurut penyelidikan mengenai kelajuan penyebaran berita palsu, *tweet* yang mengandungi maklumat yang tidak benar sampai

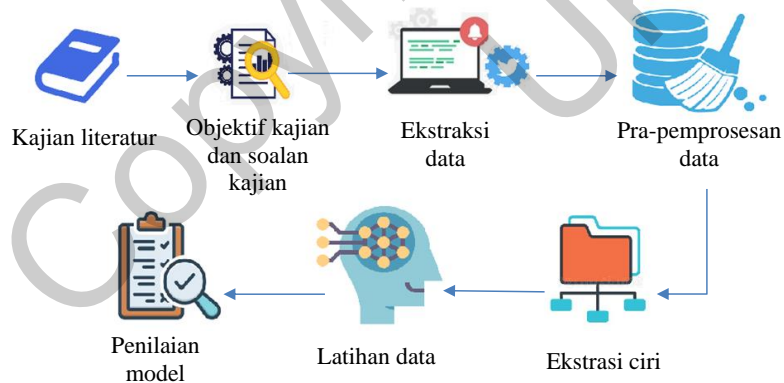
kepada individu di Twitter adalah enam kali lebih cepat daripada fakta (Thota 2018). Berita palsu telah menjadi isu kritikal yang berlaku di seluruh dunia kerana hal ini mempunyai pengaruh besar yang akan memudaratkan masyarakat dan isu ini telah mendapat perhatian dalam kalangan cendekiawan, kerajaan dan orang ramai (Egelhofer & Lecheler 2019). Sesiapa sahaja yang mempunyai akses kepada internet boleh menyebarkan maklumat palsu dengan mudah dan kebebasan untuk menyiarkan atau berkongsi maklumat di platform media sosial membolehkan penyebaran maklumat palsu yang boleh mengelirukan orang ramai (Buchanan 2020). Penyebar berita palsu tanpa identity akan kekal menyembunyi di sebalik internet dan memperoleh keuntungan perniagaan daripada setiap klik di tapak mereka sebaik sahaja pengguna mengkliknya. Dalam projek ini, model pengesanan berita palsu yang dicadangkan akan menggunakan pendekatan pembelajaran mesin dengan pelbagai algoritma modelan pembelajaran mesin. Pelbagai kajian telah dilakukan untuk mengesan berita palsu menggunakan pendekatan pembelajaran mesin secara mendalam bagi mengatasi isu berita palsu. Oleh itu, model pengesanan berita palsu perlu dilatih dengan baik dengan ciri tekstual yang dicadangkan agar dapat mengesan maklumat palsu yang tersebar di internet dengan lebih tepat.

Objektif projek ini adalah untuk mengenal pasti ciri tekstual bagi pengesanan berita palsu yang dapat diekstrakan dan pelaksanaan ciri tekstual dalam model pembelajaran mesin yang digunakan. Objektif projek yang kedua adalah untuk mencadangkan model pembelajaran mesin dengan menggunakan ciri-ciri yang dikenal pasti untuk mengesan berita palsu. Dengan ini, model pembelajaran mesin yang sesuai untuk pengesanan berita palsu dapat dikenal pasti dan mengenal pasti cara untuk meningkatkan prestasi model pembelajaran mesin tersebut.

Skop projek ini ialah penyelesaian yang dicadangkan ini tidak mempunyai keupayaan untuk menganalisis data dari sumber media sosial yang lain. Hal ini disebabkan semua data yang diekstrak untuk dilatih dan diuji adalah dari set data penanda aras yang telah dikumpulkan sahaja, namun data lain dari sumber yang berbeza tidak dapat dianalisis. Selain itu, berita palsu

yang secara tidak langsung tersebar seperti khabar angin tidak dapat dianalisis kerana tidak disiarkan di media sosial ataupun di platform lain. Skop dan kekangan kedua adalah berkaitan dengan bahasa data yang diekstrakkan. Hal ini kerana data yang dikumpul hanya dalam Bahasa Inggeris, bukannya bahasa lain seperti Bahasa Melayu, Bahasa Cina atau Bahasa Tamil yang merupakan lingua franca yang dituturkan oleh kebanyakan rakyat Malaysia. Ketiga, kekangan dalam projek ini adalah data hanya memberi focus kepada berita di Amerika Syarikat sahaja. Walau bagaimanapun, terdapat banyak data berita palsu yang tersebar melalui internet di seluruh dunia dalam setiap saat. Oleh itu, data selain daripada set data penanda aras yang telah dikumpulkan akan diabaikan dan tidak dianalisis walaupun data tersebut berada di bawah topik yang disasarkan.

Projek ini perlu dilaksanakan kerana dapat menyemak kesahihan berita dengan menggunakan model pembelajaran mesin agar dapat mengesan berita yang tersebar dengan lebih tepat dalam masa yang tersingkat. Hasil projek ini dapat menyumbang kepada ke penggunaan ciri tekstual dan model pembelajaran mesin yang paling sesuai bagi sistem pengesanan berita palsu pada masa hadapan.



Rajah 1: Aliran proses metodologi model yang dicadangkan

Metodologi projek ini adalah ditunjukkan seperti di rajah di atas. Proses kajian literatur

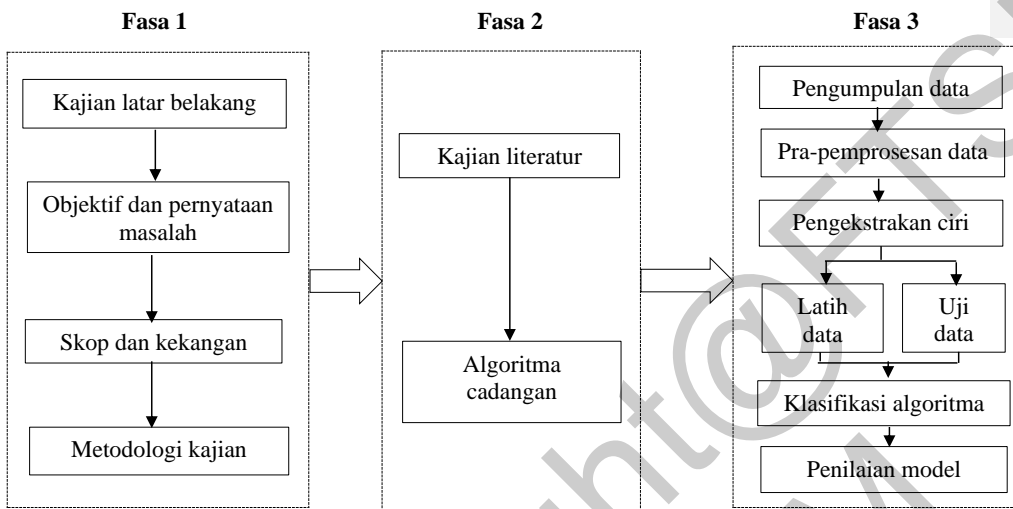
adalah fasa di mana domain projek ini perlu difahami dan topik telah disahkan. Objektif kajian dan soalan kajian bertujuan untuk membangunkan pelan dan objektif penyelidikan yang jelas. Dalam peringkat pengekstrakan data, set data penanda aras yang akan digunakan dalam projek ini akan dikumpulkan. Peringkat pra-pemprosesan data termasuk pembersihan data dan proses pra-pemprosesan di mana semua data yang dikumpulkan perlu dibersihkan dan mengeluarkan bunyi supaya data mentah dapat difahami dengan mudah oleh mesin. Dalam peringkat pengekstrakan ciri, ciri-ciri penting yang akan membantu mengesan berita palsu akan diekstrak. Peringkat latihan adalah satu proses di mana model pembelajaran mesin yang dicadangkan akan dibina dan latihan model yang dicadangkan dengan set data yang berbeza akan dibuat untuk mendapatkan hasil ketepatan yang lebih tinggi. Dalam proses penilaian, model yang dicadangkan akan dinilai dan diuji dengan set data dunia sebenar untuk membuktikan bahawa model yang dicadangkan lebih baik dalam mengklasifikasikan berita palsu.

Sorotan sastera projek ini dapat memberikan gambaran keseluruhan tentang kajian yang telah dijalankan. Kaedah yang dicadangkan oleh (Faustini & Covões 2020) untuk pengesanan berita palsu dalam pelbagai bahasa seperti Germanic, Latin dan Slavic dijalankan dengan 5 set data iaitu FakeBrCorpus, TwitterBR, Fakenewsdata1, Fake\_or\_real\_news dan btvlifestyle. Setiap set data dijalankan dengan penilaian berbeza menggunakan ciri tersuai iaitu *Word2Vec*, *DCDistance* dan *Bag Of Words*. Kemudian, setiap set data dilatih pada algoritma klasifikasi yang berbeza iaitu *KNN*, *Random Forest Breiman*, *Gaussian Naive Bayes* dan *Support Vector Machine (SVM)*. Model yang dicadangkan oleh (Reis et al. 2019) mencadangkan untuk memperkenalkan set ciri baharu untuk pengesanan berita palsu yang lebih baik. Sebanyak 141 ciri teks telah dinilai dan dikumpulkan ke dalam set daripada 2282 artikel berita yang berkaitan dengan Pilihan Raya Amerika Syarikat. Modelan telah dijalankan dengan model seperti *k-Nearest Neighbors (KNN)*, *Naive Bayes (NB)*, *Random Forests (RF)*, *Support Vector Machine dengan kernel RBF (SVM)*, dan *XGBoost (XGB)*. Berdasarkan kaedah yang dicadangkan oleh

(Choudhary & Arora 2021), model linguistik dicadangkan untuk menentukan kandungan yang akan menjana ciri terdorong bahasa yang boleh mengekstrak ciri tatabahasa, sentimen dan kebolehbacaan. Berdasarkan kaedah yang dicadangkan oleh (Giachanou, Rosso & Crestani 2019), beliau mencadangkan bahawa isyarat emosi harus dipertimbangkan untuk model pengesanan berita palsu. Set data *Politifact-1* dan *Politifact-2* telah digunakan untuk menjalankan ujian dengan model cadangan *LSTM* yang menggabungkan isyarat emosi dengan ketepatan 60.8% dan 62.8% masing-masing untuk setiap set data. (Bhutani et al. 2019) mencadangkan untuk memasukkan sentimen sebagai ciri untuk meningkatkan ketepatan pengesanan berita palsu. Penyelesaian yang dicadangkan itu menggabungkan persamaan kosinus, *TF-IDF* dan skor analisis sentimen untuk membantu dalam mengklasifikasikan berita palsu menggunakan model *Random Forest* dan *M-Naïve Bayes*. (Ajao, Bhowmik & Zargari 2018) mencadangkan bahawa pengecaman automatik ciri-ciri dalam catatan *Twitter* tanpa mengetahui topik perbincangan menggunakan model pembelajaran mendalam hibrid aplikasi *CNN* dan rangkaian berulang (*RNN*) *LSTM*. Model yang dicadangkan boleh mengklasifikasikan berita palsu tanpa pengetahuan awal tentang topik pada ketepatan 82%. Satu lagi kaedah cadangan oleh (Del Vicario et al. 2019) mencadangkan ciri model tambahan yang merangkumi satu siri ciri berkenaan dengan tingkah laku pengguna di media sosial. Kemudian, projek itu dijalankan dengan pengumpulan set data daripada Facebook dalam bahasa Itali yang merangkumi 300K berita rasmi dan 50K berita palsu menggunakan *KNN*, *Regressi Linear*, *Linear regression*, *Logistic Regression* dan algoritma *Decision Tree*.

Dalam bahagian pendahuluan, latar belakang, objektif, skop, justifikasi dan kepentingan, metodologi dan sorotan sastera kajian ini akan dibincangkan.

### Metodologi Kajian



Rajah 2: Rangka kerja operasi projek yang dicadangkan

Rajah 2 di atas menunjukkan rangka kerja operasi yang digunakan dalam projek pengesanan berita palsu ini. Dalam fasa 1, kajian latar belakang mengenai topik yang dipilih sedang dibincangkan seperti berita palsu, analisis emosi dan pembelajaran mesin. Pengenalan kepada topik yang difokuskan juga dijelaskan bahawa bagaimana isu ini memberi kesan kepada masyarakat. Objektif, pernyataan masalah, metodologi, skop, dan kekangan telah dibincangkan untuk pemahaman yang jelas mengenai projek ini. Dalam fasa 2, kajian literatur dengan model sedia ada sedang dibentangkan dan algoritma yang dicadangkan juga telah diperkenalkan. Dalam fasa 3, bahagian praktikal projek ini akan dilaksanakan untuk menguji prestasi model yang dicadangkan. Pertama sekali, sejumlah lima set data penanda aras iaitu set data FakeNewsNet, set data ISOT, set data Pengesanan Berita Palsu Kaggle (FakeNewsDetection Kaggle), set data LIAR dan set data FNC daripada kajian terdahulu akan digunakan untuk menguji dengan

algoritma yang dicadangkan. Pra-pemprosesan data akan dilakukan untuk mengalih keluar data yang tidak dibersihkan dengan perkataan henti, nilai hilang dan pendua. Data pra-diproses akan divisualisasikan dalam carta untuk memberikan pemahaman yang jelas tentang maklumat yang dikumpul. Kemudian, ciri-ciri akan diekstrak dan menjalani pemilihan ciri seperti TF-IDF, N-gram dan pembenaman perkataan CBOV akan dilaksanakan kemudian memisahkan data kepada data latihan dan ujian masing-masing. Data latihan akan digunakan untuk melatih model kemudian menjalani penilaian model untuk prestasi klasifikasi yang lebih baik. Akhir sekali, model akan disahkan dengan data ujian untuk penilaian untuk melihat prestasinya untuk mengesan berita palsu dengan tepat.

Set data yang digunakan dalam projek ini ialah set data penanda aras yang dikumpulkan daripada kajian terdahulu dan akan diuji berjalan di Google Colab. Di bawah ialah jadual yang menunjukkan pecahan setiap set data sebelum dan selepas pra-pemprosesan.

Set data penanda aras		Mentah	Selepas pembersihan data
FakeNewsNet dataset <sup>1</sup> (Kaliyar, Goswami & Narang 2021a)	Politifact	1056	1056
	Gossipfake	22140	22140
ISOT dataset <sup>2</sup> (Shaikh & Patil 2020)		44898	39119
Fake News Detection dataset <sup>3</sup> (P et al. 2021)		4009	3522
LIAR dataset <sup>4</sup> (Nasir, Khan & Varlamis 2021)		11505	9117
FNC dataset <sup>5</sup> (Umer et al. 2020)		49972	25413

Jadual 1: Gambaran keseluruhan set data penanda aras

Data-data tersebut akan dianalisis dengan kaedah pemilihan ciri bagi membangunkan model klasifikasi yang lebih baik. Kaedah pemilihan ciri yang akan digunakan ialah Kekerapan Jangka-kekerapan dokumen songsang (TF-IDF), N-grams, dan Beg Perkataan Yang Berterusan (CBOV).

Pengukur dan alat ukur yang digunakan dalam projek ini adalah penilaian metrik, ketepatan, precision, recall, skor F1, matriks kekeliruan, dan lengkung AUC-ROC yang mampu mengukur

<sup>1</sup> FakeNewsNet dataset: <https://github.com/KaiDMML/FakeNewsNet>

<sup>2</sup> ISOT dataset: <https://onlineacademiccommunity.uvic.ca/isot/2022/11/27/fake-news-detection-datasets/>

<sup>3</sup> Fake News Detection dataset: <https://www.kaggle.com/datasets/jruvika/fake-news-detection>

<sup>4</sup> LIAR dataset: <https://www.kaggle.com/datasets/khandalayan/liar-preprocessed-dataset>

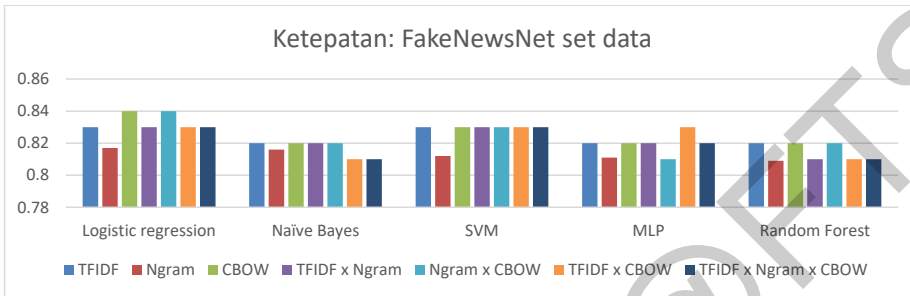
<sup>5</sup> FNC dataset: <https://github.com/FakeNewsChallenge/fnc-1>



kualiti prestasi model pembelajaran mesin dengan memberi kesan kepada bagaimana prestasi dan keberkesanan model dijejaki dan dibandingkan.

**Keputusan dan Perbincangan**

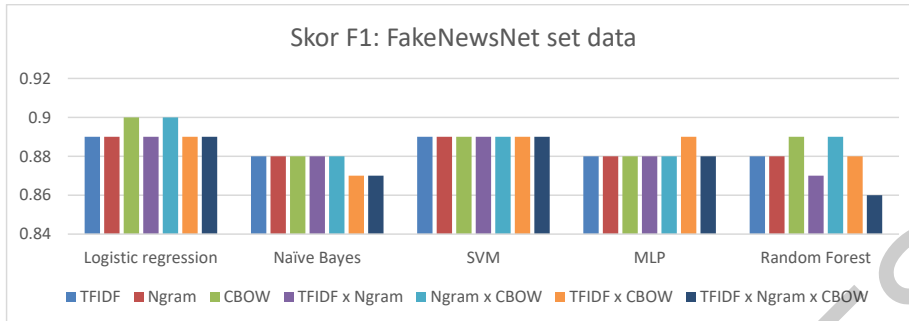
Keputusan set data: FakeNewsNet



Rajah 3: Visualisasi ketepatan pada set data FakeNewsNet

	TFIDF	N-gram	CBOW	TFIDF x N-gram	N-gram x CBOW	TFIDF x CBOW	TFIDF x N-gram x CBOW
Logistic regression	0.83	2 n-gram: 0.817 3 n-gram: 0.775	0.84	0.83	0.84	0.83	0.83
Naïve Bayes	0.82	2 n-gram: 0.816 3 n-gram: 0.775	0.82	0.82	0.82	0.81	0.81
SVM	0.83	2 n-gram: 0.812 3 n-gram: 0.776	0.83	0.83	0.83	0.83	0.83
MLP	0.82	2 n-gram: 0.811 3 n-gram: 0.773	0.82	0.82	0.81	0.83	0.82
Random Forest	0.82	2 n-gram: 0.809 3 n-gram: 0.773	0.82	0.81	0.82	0.81	0.81

Jadual 2: Ketepatan ciri yang berbeza pada set data FakeNewsNet



Rajah 4 Visualisasi skor F1 pada set data FakeNewsNet

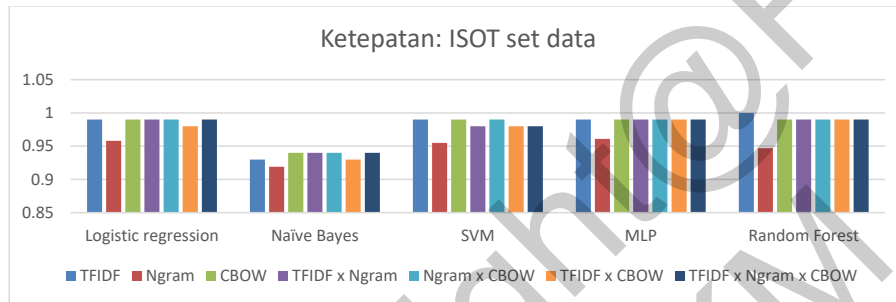
	TFIDF	N-gram	CBOW	TFIDF x N-gram	N-gram x CBOW	TFIDF x CBOW	TFIDF x N-gram x CBOW
Logistic regression	0.89	2 n-gram: 0.89 3 n-gram: 0.89	0.90	0.89	0.90	0.89	0.89
Naïve Bayes	0.88	2 n-gram: 0.88 3 n-gram: 0.88	0.88	0.88	0.88	0.87	0.87
SVM	0.89	2 n-gram: 0.89 3 n-gram: 0.89	0.89	0.89	0.89	0.89	0.89
MLP	0.88	2 n-gram: 0.88 3 n-gram: 0.88	0.88	0.88	0.88	0.89	0.88
Random Forest	0.88	2 n-gram: 0.88 3 n-gram: 0.88	0.89	0.87	0.89	0.88	0.88

Jadual 3 Skor F1 bagi ciri berbeza pada set data FakeNewsNet

Dalam set data FakeNewsNet, ciri CBOW dan N-gram dan CBOW (N-gram x CBOW) dengan model regresi logistik menunjukkan hasil yang lebih tepat berbanding ciri dan model lain. Dengan hanya TFIDF sebagai ciri, kedua-dua regresi logistik dan model SVM mengatasi model lain dengan ketepatan 0.83 masing-masing. Ciri N-gram menunjukkan prestasi yang ketara dalam bigram dan bukannya trigram dalam semua model di mana model regresi logistik mempunyai ketepatan yang lebih tinggi iaitu 0.817 seperti ditunjukkan di atas. Manakala gabungan TFIDF dan N-gram (TFIDF x N-gram), kedua-dua regresi logistik dan SVM mempunyai ketepatan yang paling tinggi antara model lain seperti model Naïve Bayes, MLP dan

Random Forest. Untuk ciri gabungan TFIDF dan CBOW (TFIDF x CBOW), model seperti regresi logistik, SVM dan MLP masing-masing mempunyai ketepatan 0.83 manakala untuk gabungan semua ciri (TFIDF x N-gram x CBOW) tidak menunjukkan banyak peningkatan dalam ketepatan manakala regresi logistik dan model SVM menunjukkan ketepatan 0.83 masing-masing. Regresi Logistik dan model SVM mencapai skor F1 tertinggi yang menunjukkan 0.89 ramalan positif telah dikenal pasti dengan betul. Oleh itu, Regresi Logistik dan model SVM berprestasi terbaik berbanding model pembelajaran mesin lain dalam set data ini.

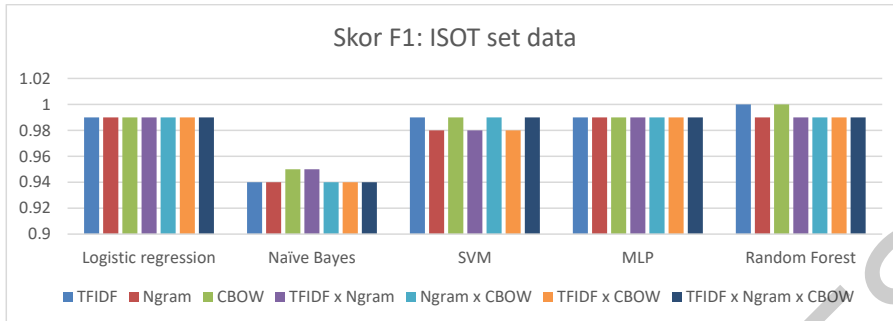
Keputusan set data: ISOT



Rajah 5: Visualisasi ketepatan pada set data ISOT

	TFIDF	N-gram	CBOW	TFIDF x N-gram	N-gram x CBOW	TFIDF x CBOW	TFIDF x N-gram x CBOW
Logistic regression	0.99	2 n-gram: 0.958 3 n-gram: 0.889	0.99	0.99	0.99	0.98	0.99
Naive Bayes	0.93	2 n-gram: 0.919 3 n-gram: 0.812	0.94	0.94	0.94	0.93	0.94
SVM	0.99	2 n-gram: 0.955 3 n-gram: 0.889	0.99	0.98	0.99	0.98	0.98
MLP	0.99	2 n-gram: 0.961 3 n-gram: 0.889	0.99	0.99	0.99	0.99	0.99
Random Forest	1.00	2 n-gram: 0.947 3 n-gram: 0.881	0.99	0.99	0.99	0.99	0.99

Jadual 4: Ketepatan ciri yang berbeza pada set data ISOT



Rajah 6: Visualisasi skor F1 pada set data ISOT

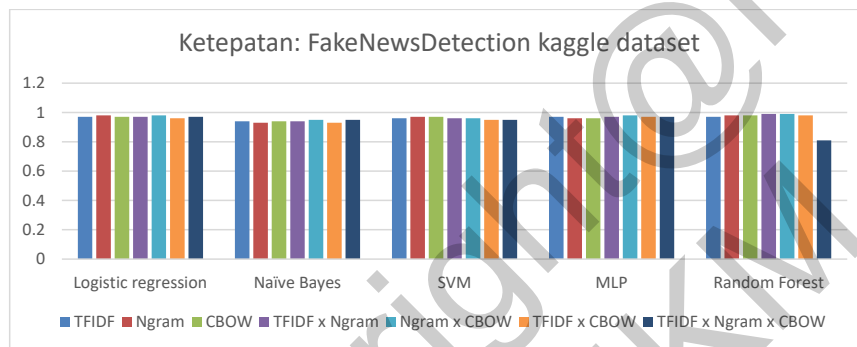
	TFIDF	N-gram	CBOW	TFIDF x N-gram	N-gram x CBOW	TFIDF x CBOW	TFIDF x N-gram x CBOW
Logistic regression	0.99	2 n-gram: 0.99 3 n-gram: 0.99	0.99	0.99	0.99	0.99	0.99
Naive Bayes	0.94	2 n-gram: 0.94 3 n-gram: 0.94	0.95	0.95	0.94	0.94	0.94
SVM	0.99	2 n-gram: 0.98 3 n-gram: 0.98	0.99	0.98	0.99	0.98	0.99
MLP	0.99	2 n-gram: 0.99 3 n-gram: 0.99	0.99	0.99	0.99	0.99	0.99
Random Forest	1.00	2 n-gram: 0.99 3 n-gram: 0.99	1.00	0.99	0.99	0.99	0.99

Jadual 5: Skor F1 bagi ciri berbeza pada set data ISOT

Dalam set data ISOT, ciri menunjukkan prestasi yang baik dalam semua model kecuali model Naive Bayes. Untuk ciri TFIDF, model Random Forest mencapai ketepatan 1.00 manakala regresi logistik, model SVM dan MLP menghasilkan ketepatan 0.99 namun model Naive Bayes hanya menunjukkan ketepatan 0.93. Bagi N-gram, MLP menunjukkan prestasi yang lebih baik dengan ketepatan 0.961 berbanding model lain. Ciri CBOW menunjukkan regresi logistik, SVM, MLP dan Random Forest menghasilkan ketepatan 0.99 manakala model Naive Bayes hanya menunjukkan ketepatan 0.94. Untuk ciri gabungan (TFIDF x N-gram) dan (N-gram x CBOW) model logistik, SVM, MLP dan Random Forest melakukan ketepatan 0.99 manakala Naive

Bayes hanya melakukan ketepatan 0.94. Manakala untuk semua ciri gabungan (TFIDF x N-gram x CBOW) regresi logistik, model MLP dan Random Forest mencapai ketepatan 0.99, Naïve Bayes mencapai ketepatan 0.94 dan SVM mencapai ketepatan 0.98. Model Regresi Logistik, MLP dan Hutan Rawak mencapai skor F1 tertinggi yang menunjukkan 0.99 ramalan positif telah dikenal pasti dengan betul. Oleh itu, model Logistik Regresi, MLP dan Hutan Rawak berprestasi terbaik berbanding model pembelajaran mesin lain dalam set data ini.

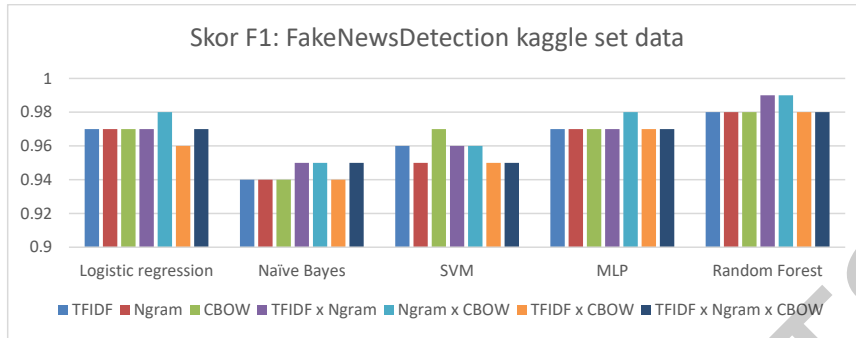
Keputusan set data: FakeNewsDetection Kaggle



Rajah 7: Visualisasi ketepatan pada set data FakeNewsDetection Kaggle

	TFIDF	N-gram	CBOW	TFIDF x N-gram	N-gram x CBOW	TFIDF x CBOW	TFIDF x N-gram x CBOW
Logistic regression	0.97	2 n-gram: 0.98 3 n-gram: 0.92	0.97	0.97	0.98	0.96	0.97
Naïve Bayes	0.94	2 n-gram: 0.93 3 n-gram: 0.79	0.94	0.94	0.95	0.93	0.95
SVM	0.96	2 n-gram: 0.97 3 n-gram: 0.92	0.97	0.96	0.96	0.95	0.95
MLP	0.97	2 n-gram: 0.96 3 n-gram: 0.91	0.96	0.97	0.98	0.97	0.97
Random Forest	0.97	2 n-gram: 0.98 3 n-gram: 0.92	0.98	0.99	0.99	0.98	0.98

Jadual 6: Ketepatan ciri yang berbeza pada set data FakeNewsDetection Kaggle



Rajah 8: Visualisasi skor F1 pada set data FakeNewsDetection Kaggle

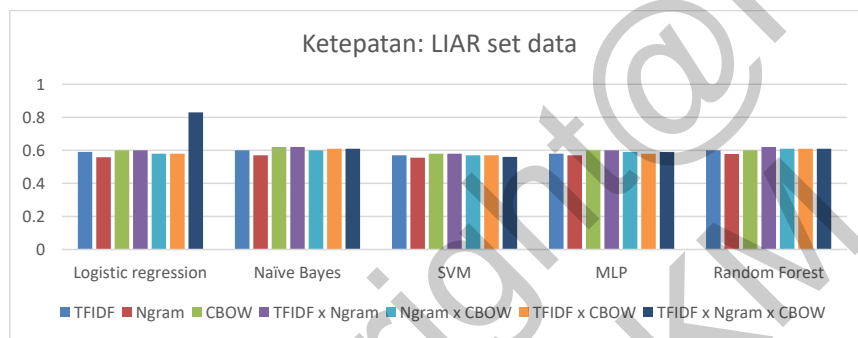
	TFIDF	N-gram	CBOW	TFIDF x N-gram	N-gram x CBOW	TFIDF x CBOW	TFIDF x N-gram x CBOW
Logistic regression	0.97	2 n-gram: 0.97 3 n-gram: 0.97	0.97	0.97	0.98	0.96	0.97
Naïve Bayes	0.94	2 n-gram: 0.94 3 n-gram: 0.94	0.94	0.95	0.95	0.94	0.95
SVM	0.96	2 n-gram: 0.95 3 n-gram: 0.95	0.97	0.96	0.96	0.95	0.95
MLP	0.97	2 n-gram: 0.97 3 n-gram: 0.97	0.97	0.97	0.98	0.97	0.97
Random Forest	0.98	2 n-gram: 0.98 3 n-gram: 0.98	0.98	0.99	0.99	0.98	0.98

Jadual 7: Skor F1 bagi ciri berbeza pada set data FakeNewsDetection Kaggle

Dalam set data FakeNewsDetection dengan ciri TFIDF, regresi logistik, model MLP dan Random Forest mempunyai prestasi yang lebih baik dengan ketepatan 0.97 berbanding model Naïve Bayes dan SVM dengan model 0.94 dan 0.96 masing-masing. Bagi ciri N-gram, Regresi Logistik dan Hutan Rawak menunjukkan keputusan ketepatan tertinggi pada bigram dengan 0.98 manakala model Naïve Bayes, SVM dan MLP menunjukkan ketepatan masing-masing 0.93, 0.97 dan 0.96. Untuk CBOW, model Hutan Rawak mempunyai hasil ketepatan yang lebih tinggi, iaitu 0.98 dan 0.97, 0.94, 0.97 dan 0.96 untuk Model Logistik, Naïve Bayes, SVM dan MLP masing-masing. Ciri gabungan (N-gram x CBOW) menunjukkan peningkatan dalam hasil menggunakan model seperti Naïve Bayes, MLP dan Hutan Rawak dengan ketepatan masing-masing 0.95, 0.98

dan 0.99. Walaupun dengan semua ciri gabungan (TFIDF x N-gram x CBOW), model Naïve Bayes menunjukkan peningkatan dengan ketepatan 0.95 manakala model Hutan Rawak mencapai ketepatan tertinggi 0.98. Oleh itu, model Hutan Rawak mengatasi model pembelajaran mesin yang lain dalam set data yang diberikan ini pada ciri berbeza yang digunakan. Model MLP dan Hutan Rawak mencapai skor F1 tertinggi yang menunjukkan 0.99 ramalan positif telah dikenal pasti dengan betul. Oleh itu, model Hutan Rawak berprestasi terbaik berbanding dengan model pembelajaran mesin lain dalam set data ini.

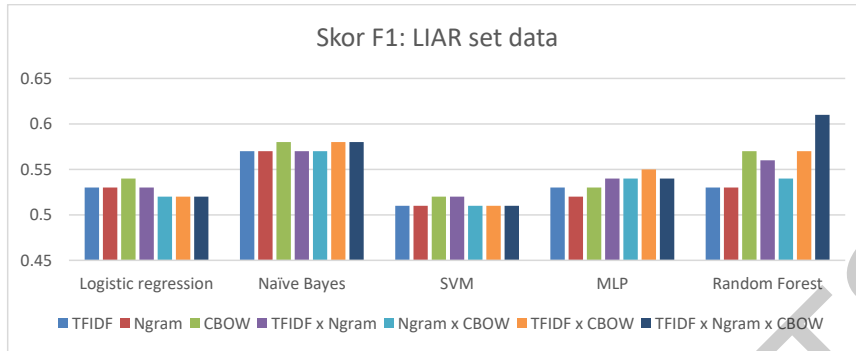
Keputusan set data: LIAR



Rajah 9: Visualisasi ketepatan pada set data LIAR

	TFIDF	N-gram	CBOW	TFIDF x N-gram	N-gram x CBOW	TFIDF x CBOW	TFIDF x N-gram x CBOW
Logistic regression	0.59	2 n-gram: 0.558 3 n-gram: 0.550	0.60	0.60	0.58	0.58	0.58
Naïve Bayes	0.60	2 n-gram: 0.570 3 n-gram: 0.551	0.62	0.61	0.60	0.61	0.61
SVM	0.57	2 n-gram: 0.556 3 n-gram: 0.543	0.58	0.58	0.57	0.57	0.56
MLP	0.58	2 n-gram: 0.570 3 n-gram: 0.541	0.60	0.60	0.59	0.58	0.59
Random Forest	0.60	2 n-gram: 0.578 3 n-gram: 0.544	0.62	0.60	0.61	0.61	0.61

Jadual 8: Ketepatan ciri yang berbeza pada set data ISOT



Rajah 10: Visualisasi skor F1 pada set data ISOT

	TFIDF	N-gram	CBOW	TFIDF x Ngram	Ngram x CBOW	TFIDF x CBOW	TFIDF x Ngram x CBOW
Logistic regression	0.53	2 n-gram: 0.53 3 n-gram: 0.53	0.54	0.53	0.52	0.52	0.52
Naïve Bayes	0.57	2 n-gram: 0.57 3 n-gram: 0.57	0.58	0.57	0.57	0.58	0.58
SVM	0.51	2 n-gram: 0.51 3 n-gram: 0.51	0.52	0.52	0.51	0.51	0.51
MLP	0.53	2 n-gram: 0.52 3 n-gram: 0.52	0.53	0.54	0.54	0.55	0.54
Random Forest	0.53	2 n-gram: 0.53 3 n-gram: 0.54	0.57	0.56	0.54	0.57	0.55

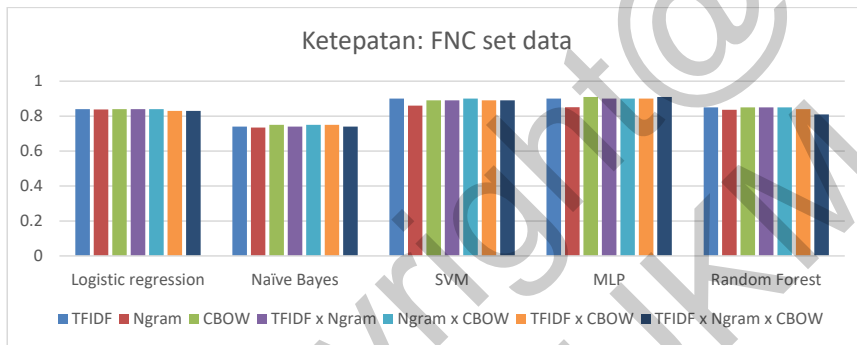
Jadual 9: Skor F1 bagi ciri berbeza pada set data LIAR

Dalam set data LIAR, ciri CBOW pada model Naïve Bayes dan Hutan Rawak mengatasi ciri lain pada model berbeza dengan ketepatan 0.62 setiap satu. Bagi ciri TFIDF, ciri Naïve Bayes dan Hutan Rawak menunjukkan hasil ketepatan yang paling tinggi iaitu 0.60 manakala bagi ciri N-gram model Hutan Rawak mempunyai ketepatan 0.578 iaitu yang paling tinggi berbanding dengan model lain. Untuk ciri gabungan (TFIDF x N-gram), (N-gram x CBOW), (TFIDF x CBOW) dan (TFIDF x N-gram x CBOW) menunjukkan hasil ketepatan 0.61 setiap satu pada kedua-dua model Naïve Bayes dan Hutan Rawak. Oleh itu, model Naïve Bayes dan Hutan Rawak menunjukkan prestasi yang lebih baik dalam set data ini. Model Naïve Bayes mencapai skor F1 tertinggi yang



menunjukkan 0.58 ramalan positif telah dikenal pasti dengan betul. Oleh itu, model Naïve Bayes melakukan yang terbaik berbanding dengan model pembelajaran mesin lain dalam set data ini.

Keputusan set data: FNC

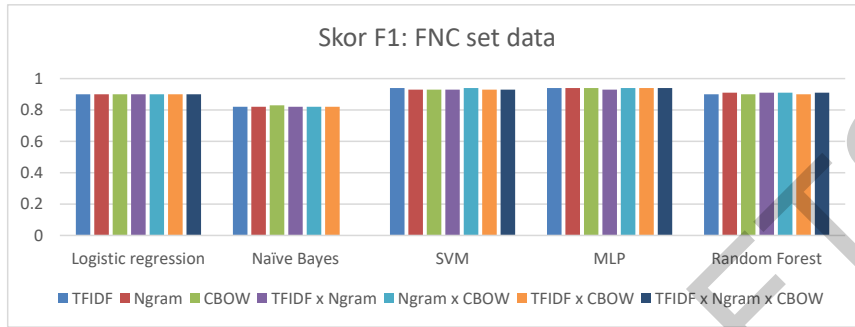


Rajah 11: Visualisasi ketepatan pada set data FNC

	TFIDF	N-gram	CBOW	TFIDF x N-gram	N-gram x CBOW	TFIDF x CBOW	TFIDF x N-gram x CBOW
Logistic regression	0.84	2 n-gram: 0.838 3 n-gram: 0.859	0.84	0.84	0.84	0.83	0.83
Naïve Bayes	0.74	2 n-gram: 0.735 3 n-gram: 0.793	0.75	0.74	0.75	0.75	0.74
SVM	0.90	2 n-gram: 0.860 3 n-gram: 0.864	0.89	0.89	0.90	0.89	0.89
MLP	0.90	2 n-gram: 0.851 3 n-gram: 0.846	0.91	0.90	0.90	0.90	0.91

Random Forest	0.85	2 n-gram: 0.836 3 n-gram: 0.846	0.85	0.85	0.85	0.84	0.85
---------------	------	------------------------------------	------	------	------	------	------

Jadual 10: Ketepatan ciri yang berbeza pada set data FNC



Rajah 12: Visualisasi skor F1 pada set data FNC

	TFIDF	N-gram	CBOW	TFIDF x Ngram	Ngram x CBOW	TFIDF x CBOW	TFIDF x Ngram x CBOW
Logistic regression	0.90	2 n-gram: 0.90 3 n-gram: 0.90	0.90	0.90	0.90	0.90	0.90
Naïve Bayes	0.82	2 n-gram: 0.82 3 n-gram: 0.82	0.83	0.82	0.82	0.82	0.81
SVM	0.94	2 n-gram: 0.93 3 n-gram: 0.93	0.93	0.93	0.94	0.93	0.93
MLP	0.94	2 n-gram: 0.94 3 n-gram: 0.94	0.94	0.93	0.94	0.94	0.94
Random Forest	0.90	2 n-gram: 0.91 3 n-gram: 0.91	0.90	0.91	0.91	0.90	0.91

Jadual 11: Skor F1 bagi ciri berbeza pada set data FNC

Dalam set data FNC, model MLP mencapai hasil ketepatan tertinggi sebanyak 0.91 dengan kedua-dua CBOW dan ciri gabungan (TFIDF x N-gram x CBOW). Dengan ciri TFIDF, model SVM dan MLP menunjukkan ketepatan tertinggi masing-masing 0.90. Manakala dalam ciri Ngram, model SVM menunjukkan 0.86 yang tertinggi di kalangan yang lain dan model MLP menunjukkan ketepatan 0.91 dalam ciri CBOW berbanding dengan model lain. Dengan ciri gabungan (TFIDF x NGRAM), (N-gram x CBOW), (TFIDF x CBOW) dan (TFIDF x N-gram x CBOW), model MLP mencapai ketepatan tertinggi berbanding model lain dengan 0.90, 0.90,

0.90 dan 0.91 ketepatan masing-masing. Oleh itu, dalam set data ini model MLP mengatasi model pembelajaran mesin lain yang diuji menggunakan ciri yang berbeza. Model SVM dan MLP mencapai skor F1 tertinggi yang menunjukkan 0.94 ramalan positif telah dikenal pasti dengan betul. Oleh itu, model SVM dan MLP berprestasi terbaik berbanding model pembelajaran mesin lain dalam set data ini.

Keputusan keseluruhan dan perbincangan

Set data	Model	Ketepatan tertinggi dicapai
FakeNewsNet set data	Logistic Regression & SVM	0.83
ISOT set data	MLP	0.99
FakeNewsDetection Kaggle set data	Random Forest	0.99
LIAR set data	Naïve Bayes & Random Forest	0.62
FNC set data	MLP	0.91

Jadual 12: Model terbaik untuk setiap set data berdasarkan ketepatan

Selepas beberapa eksperimen telah dijalankan, beberapa kesimpulan dapat dinyatakan seperti di bawah. Setiap set data memberi fokus secara khusus pada topik yang berbeza justeru keputusan ketepatan juga menunjukkan prestasi yang berbeza menggunakan pelbagai jenis model pembelajaran mesin. Untuk set data FakeNewsNet, Regresi Logistik dan model SVM mencapai ketepatan tertinggi 0.83. Manakala untuk dataset ISOT dan FakeNewsDetection Kaggle, ketepatan tertinggi yang dicapai ialah 0.99 oleh model MLP dan Hutan Rawak. Walau bagaimanapun, dalam dataset LIAR kedua-dua Naïve Bayes dan Hutan Rawak mencapai ketepatan 0.62, tetapi ia adalah yang paling rendah antara semua set data. Akhir sekali dalam set

data FNC, model MLP menunjukkan ketepatan tertinggi yang dicapai sebanyak 0.91. Setiap model pembelajaran mesin yang digunakan dalam eksperimen mengatasi prestasi dalam set data yang berbeza kerana setiap pendedaran data set data adalah berbeza walaupun ia berada dalam domain yang sama. Oleh itu, setiap model pembelajaran mesin adalah bebas daripada setiap set data.

Set data	Ciri terbaik
FakeNewsNet set data	CBOW, N-gram x CBOW
ISOT set data	TFIDF, CBOW
FakeNewsDetection Kaggle set data	N-gram x CBOW
LIAR set data	CBOW, TFIDF x N-gram
FNC set data	TFIDF, N-gram x CBOW

Jadual 13: Ciri terbaik pada setiap set data

Dalam eksperimen di atas, ciri sedang diuji secara berasingan dan digabungkan untuk menguji prestasi ciri yang diekstrak dan model pembelajaran mesin. Seperti yang diperhatikan, CBOW ialah ciri yang baik untuk semua set data penanda aras secara bebas atau digabungkan dengan ciri lain. Ciri CBOW menangkap makna semantik perkataan berdasarkan konteks perkataan dan mempelajari penyusunan perkataan yang merupakan perwakilan vektor padat perkataan dengan mempertimbangkan perkataan jiran untuk meramalkan perkataan berikut seterusnya. Selain itu, gabungan CBOW dan N-gram (CBOW x N-gram) menghasilkan ciri yang baik kerana kedua-dua ciri menggunakan kekerapan perkataan dalam data latihan untuk memahami kemungkinan urutan perkataan atau kejadian perkataan tertentu, dengan CBOW menganggarkan

kebarangkalian perkataan berdasarkan kekerapan kejadian bersama perkataan dan model N-gram bergantung pada kekerapan urutan panjang N tertentu.

Selain itu, ciri N-gram menunjukkan bahawa bigram ( $n=2$ ) berprestasi lebih baik daripada trigram ( $n=3$ ) dalam semua model pembelajaran mesin. Hal ini disebabkan oleh isu sparsity data di mana kekerapan beberapa trigram mungkin terlalu rendah untuk menganggarkan kebarangkalian yang boleh dipercayai. Bigrams juga dipercayai bahawa mempunyai kebolehan menangkap yang lebih baik untuk kejadian bersama yang lebih biasa dengan julat konteks yang lebih luas. Set data yang diperolehi kesemuanya adalah dalam bahasa Inggeris, bigram berprestasi lebih baik daripada trigram kerana bahasa itu sering menunjukkan kebergantungan yang lebih kuat antara perkataan bersebelahan dan bukannya perkataan yang berturutan.

Set data ISOT telah menunjukkan prestasi tertinggi berbanding set data penanda aras yang lain. Walaupun ini adalah hasil yang baik, tetapi ia mungkin membawa kepada masalah overfitting. Set data ISOT boleh mencapai prestasi terbaiknya adalah kerana set data ini adalah set data seimbang yang dijelaskan oleh pakar manusia yang terdiri daripada artikel sebenar yang dikumpul daripada Reuters.com manakala artikel palsu yang dikumpul daripada PolitiFact dan Wikipedia memfokuskan pada artikel antara 2016 dan 2017 oleh penyelidikan dari University of Victoria dan University of British Columbia pada 2018. Walau bagaimanapun, set data LIAR yang lain menunjukkan ketepatan set data yang lebih rendah daripada set data LIAR yang lain. Hal ini disebabkan data yang tidak mencukupi di mana data adalah hanya satu ayat pendek dan data kurang terlatih.

Cadangan masa hadapan bagi projek ini adalah menggunakan kaedah ciri Word2Vec kerana ciri ini menunjukkan prestasi yang agak tinggi dalam melaksanakan tugas teks klasifikasi, analisis sentiment, pengecaman entity bernama dan terjemahan mesin. Ciri ini juga dapat menggambarkan perkataan sebagai vector berterusan yang dapat menangkap lebih banyak maklumat tentang maknanya berbanding dengan perwakilan jarang tradisional.

### **Kesimpulan**

Kesimpulan yang dapat dihasilkan untuk projek ini adalah pertama sekali setiap model pembelajaran mesin bertindak secara berbeza dalam setiap set data penanda aras yang telah digunakan. Hal ini disebabkan set data yang berbeza mempunyai pengedaran data yang berbeza dan model pembelajaran mesin tersebut yang dilatih pada satu set data tidak digeneralisasikan dengan baik kepada set data yang lain secara ketara. Selain itu, kesimpulan kedua adalah ciri tekstual yang telah digunakan iaitu Beg Perkataan Yang Berterusan (CBOW) menunjukkan prestasi yang terbaik dalam semua set data yang telah diuji. Hal ini kerana ciri CBOW dapat memerangkap makna semantik perkataan berdasarkan konteks perkataan yang dapat memberi pertimbangan perkataan sekelilingnya untuk meramalkan perkataan yang seterusnya. Kedua-kedua objektif kajian ini telah dicapai iaitu mengenalpasti ciri tekstual dan model pembelajaran mesin yang sesuai untuk pengesanan berita palsu. Impak hasil projek ini adalah dapat mengekalkan integriti informasi yang tersebar di internet. Hal ini juga dapat memulihkan keyakinan dan kepercayaan dalam saluran berita yang sah dan penyedia informasi maklumat yang boleh dipercayai.

Cadangan penambahbaikan masa hadapan boleh dibuat dengan melaksanakan set data masa nyata semasa yang mana akan terdapat lebih banyak latihan oleh model pembelajaran mesin

menggunakan ciri yang dicadangkan untuk meningkatkan prestasi pengesanan berita palsu. Selain itu, data yang dikumpul boleh dalam pelbagai bahasa pertuturan biasa atau bahasa campuran agar lebih sesuai dengan budaya berbilang kaum di Malaysia dengan topik berita popular semasa Malaysia. Selain itu, penalaan hiperparameter dengan mencari satu set nilai hiperparameter optimum ataupun menggunakan model rangkaian neural untuk memaksimumkan prestasi model pembelajaran mesin dengan ralat yang lebih sedikit dihasilkan.

### **Penghargaan**

Terlebih dahulu, saya ingin mengucapkan jutaan terima kasih kepada penyelia saya Dr. Wandeep Kaur A/P Ratan Singh yang banyak memberi bimbingan dan tunjuk ajar sepanjang kajian ilmiah saya. Beliau telah banyak memberikan dedikasi dan kesabaran kepada saya sepanjang tempoh saya menyiapkan projek ini. Jutaan terima kasih kepada para pensyarah dari FTSM yang telah banyak memberi didikan dan ilmu pengetahuan sepanjang pengajian saya di Universiti Kebangsaan Malaysia. Selain itu, saya ingin mengambil peluang ini kepada ibu bapa saya yang telah banyak memberikan sokongan dan bantuan yang menjadi sumber inspirasi sepanjang projek tahun akhir ini dijalankan. Tidak lupa juga kepada rakan-rakan yang sentiasa memberi bantuan dan pandangan kepada saya apabila saya memerlukannya. Akhir sekali, saya ingin mengucapkan terima kasih juga kepada pihak-pihak yang terlibat secara langsung dan tidak langsung sepanjang pelaksanaan kajian ilmiah tahun akhir saya ini.

## RUJUKAN

- Allcott, H. & Gentzkow, M. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31(2): 211–236.
- Ajao, O., Bhowmik, D. & Zargari, S. 2018. Fake news identification on Twitter with hybrid CNN and RNN models. *ACM International Conference Proceeding Series*: 226–230.
- Bhutani, B., Rastogi, N., Sehgal, P. & Purwar, A. 2019. Fake News Detection Using Sentiment Analysis. 2019 12th International Conference on Contemporary Computing, IC3 2019: 1–5.
- Buchanan, T. 2020. Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. *PLoS ONE*, hlm.
- Choudhary, A. & Arora, A. 2021. Linguistic feature based learning model for fake news detection and classification. *Expert Systems with Applications* 169: 114171.
- Darr, C.R. & Doss, E.F. 2022. The Fake One is the Real One: Finstas, Authenticity, and Context Collapse in Teen Friend Groups. *Journal of Computer-Mediated Communication* 27(4).
- Del Vicario, M., Quattrociocchi, W., Scala, A. & Zollo, F. 2019. Polarization and fake news: Early warning of potential misinformation targets. *ACM Transactions on the Web* 13(2).
- Domenico, G. Di, Sit, J., Ishizaka, A. & Nunan, D. 2021. Fake news, social media and marketing: A systematic review. *Journal of Business Research* 124(December 2020): 329–341.
- Egelhofer, J.L. & Lecheler, S. 2019. Fake news as a two-dimensional phenomenon: a framework and research agenda. *Annals of the International Communication Association* 43(2): 97–116.
- Faustini, P.H.A. & Covões, T.F. 2020. Fake news detection in multiple platforms and languages. *Expert Systems with Applications* 158: 113503.
- Giachanou, A., Rosso, P. & Crestani, F. 2019. Leveraging emotional signals for credibility detection. *SIGIR 2019 - Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*: 877–880.



- Kaliyar, R.K., Goswami, A. & Narang, P. 2021a. DeepFake: improving fake news detection using tensor decomposition-based deep neural network. *Journal of Supercomputing* 77(2): 1015–1037.
- Nasir, J.A., Khan, O.S. & Varlamis, I. 2021. Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights* 1(1).
- P, D., Chakraborty, T., Long, C. & G, S.K. 2021. Deep Learning for Fake News Detection. *Information Retrieval Series* 42: 71–100.
- Reis, J.C.S., Correia, A., Murai, F., Veloso, A., Benevenuto, F. & Cambria, E. 2019. Supervised Learning for Fake News Detection. *IEEE Intelligent Systems* 34(2): 76–81.
- Shaikh, J. & Patil, R. 2020. Fake news detection using machine learning. *Proceedings - 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security, iSSSC 2020* 2020.
- Shu, K., Wang, S. & Liu, H. 2019. Beyond news contents: The role of social context for fake news detection. *WSDM 2019 - Proceedings of the 12th ACM International Conference on Web Search and Data Mining*(i): 312–320.
- Srikanth, J. 2019. SCRUTINY OF SOCIAL MEDIA CONTENT AND MEDIA XI(I): 1–14.
- Tandoc, E.C., Jenkins, J. & Craft, S. 2018. Fake News as a Critical Incident in Journalism. *Journalism Practice* 2786.
- Thota, A. 2018. SMU Data Science Review Fake News Detection : A Deep Learning Approach Fake News Detection : A Deep Learning Approach 1(3).
- Umer, M., Imtiaz, Z., Ullah, S., Mehmood, A., Choi, G.S. & On, B.W. 2020. Fake news stance detection using deep learning architecture (CNN-LSTM). *IEEE Access* 8: 156695–156706.

PTA-FTSM-2023-A179338

Koh Kai Xuan (A179338)  
Dr. Wandeep Kaur a/p Ratan Singh  
Fakulti Teknologi & Sains Maklumat,  
Universiti Kebangsaan Malaysia

**Commented [U1]:** Nama dan No. Matriks Pelajar, serta Nama  
Penyelia

Copyright@FTSM  
UKM