

**ANALISIS KINERJA *FEED FORWARD NEURAL NETWORK*
DAN *SUPPORT VECTOR MACHINE* DALAM
MENGKLASIFIKASIKAN BERITA PALSU DI TWITTER/X**



Oleh:

Agus Suparman

NIM: 20206051008

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

**PROGRAM STUDI INFORMATIKA
PROGRAM MAGISTER FAKULTAS SAINS DAN
TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA**

2024

PERNYATAAN KEASLIAN

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Agus Suparman
NIM : 20206051008
Jenjang : Magister
Program Studi : Informatika

Menyatakan bahwa naskah tesis ini secara keseluruhan adalah hasil penelitian/karya saya sendiri, kecuali pada bagian-bagian yang dirujuk sumbernya.

Yogyakarta, 8 Agustus 2024
menyatakan,



Agus Suparman
NIM: 20206051008

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PERNYATAAN BEBAS PLAGIASI

PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan di bawah ini:

Nama : Agus Suparman
NIM : 20206051008
Jenjang : Magister
Program Studi : Informatika

Menyatakan bahwa naskah tesis ini secara keseluruhan benar-benar bebas dari plagiasi. Jika di kemudian hari terbukti melakukan plagiasi, maka saya siap ditindak sesuai ketentuan hukum yang berlaku.

Yogyakarta, 8 Agustus 2024

menyatakan,



Agus Suparman
NIM: 20206051008

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PENGESAHAN TUGAS AKHIR



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-1498/Un.02/DST/PP.00.9/08/2024

Tugas Akhir dengan judul : ANALISIS KINERJA FEED FORWARD NEURAL NETWORK DAN SUPPORT VECTOR MACHINE DALAM MENGLASIFIKASIKAN BERITA PALSU DI TWITTER/X

yang dipersiapkan dan disusun oleh:

Nama : AGUS SUPARMAN, S.Kom.
Nomor Induk Mahasiswa : 20206051008
Telah diujikan pada : Senin, 19 Agustus 2024
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Dr. Agung Fatwanto, S.Si., M.Kom.
SIGNED

Valid ID: 66c443df54946



Penguji I

Ir. Muhammad Taufiq Nuruzzaman, S.T.
M.Eng., Ph.D.
SIGNED

Valid ID: 66c3e5bc3fd52



Penguji II

Dr. Ir. Bambang Sugiantoro, M.T., IPU.,
ASEAN Eng.
SIGNED

Valid ID: 66c315fed2fd



Yogyakarta, 19 Agustus 2024
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 66c4779f45ecf

NOTA DINAS PEMBIMBING

NOTA DINAS PEMBIMBING

Kepada Yth.,
Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga
Yogyakarta

Assalamu 'alaikum wr. wb.

Setelah melakukan bimbingan, arahan, dan koreksi terhadap penulisan tesis yang berjudul:

ANALISIS KINERJA *FEED FORWARD NEURAL NETWORK* DAN *SUPPORT VECTOR MACHINE* DALAM MENGLASIFIKASIKAN BERITA PALSU DI TWITTER/X

Yang ditulis oleh:

Nama : Agus Suparman
NIM : 20206051008
Jenjang : Magister
Program Studi : Informatika

Saya berpendapat bahwa tesis tersebut sudah dapat diajukan kepada Magister Informatika UIN Sunan Kalijaga untuk diujikan dalam rangka memperoleh gelar Magister Informatika.

Wassalamu'alaikum wr. wb.

Yogyakarta, 8 Agustus 2024
Pembimbing



Dr. Agung Fatwanto, S.Si., M.Kom.
NIP. 19770103 200501 1 003

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

ABSTRAK

Media sosial telah berkembang sebagai media penyebaran informasi (termasuk berita palsu) secara massal. Berita palsu di media sosial seperti Twitter/X sering mendapatkan *retweet* yang cepat dan tersebar luas, bahkan terkadang menjadi viral. Oleh karena itu, diperlukan mekanisme yang dapat mengklasifikasi berita palsu di media sosial secara otomatis. Penelitian ini bertujuan untuk membandingkan kinerja dua model pembelajaran mesin, yaitu *Feed Forward Neural Network* (FFNN) dan *Support Vector Machine* (SVM) untuk mengklasifikasi berita palsu secara otomatis. Selain itu, penelitian ini juga bermaksud membandingkan pengaruh proporsi data uji terhadap data latih serta pengaruh keseimbangan proporsi kelas dalam *dataset* terhadap kinerja klasifikasi. Penelitian ini menggunakan pendekatan kuantitatif eksperimental. Proses eksperimen dilakukan terhadap model FFNN dan SVM dengan cara melatih dan menguji kedua model menggunakan *dataset* metadata postingan di Twitter/X yang meliputi jumlah pengikut dari pengguna, jumlah kutip *tweet*, jumlah *retweet*, dan jumlah *like*. Data yang digunakan untuk eksperimen sebanyak 2.104 yang diambil melalui Astramaya oleh DroneEmprit dan 1.781 data tambahan yang diambil dari situs TurnBackHoax. Evaluasi atas hasil klasifikasi kedua model diukur menggunakan *Confusion Matrix* untuk menghitung *accuracy*, *precision*, *recall*, dan *f1-score*. Hasil penelitian menunjukkan model SVM secara umum memiliki tingkat *accuracy*, *recall*, dan *f1-score* yang lebih tinggi dibandingkan dengan model FFNN (yang hanya unggul dalam tingkat *precision*). Kedua model secara umum mengalami penurunan kinerja seiring dengan peningkatan proporsi data uji (kecuali pada proporsi data uji 40% yang lebih baik

dibandingkan saat menggunakan proporsi data uji 20%). Kedua model menunjukkan kinerja klasifikasi yang lebih baik saat dilatih dengan *dataset* yang memiliki proporsi kelas yang seimbang, dibandingkan dengan *dataset* yang memiliki proporsi kelas tidak seimbang atau yang diseimbangkan menggunakan teknik *Oversampling* dengan metode SMOTE. Model SVM menghasilkan tingkat klasifikasi yang lebih baik dibandingkan dengan model FFNN. Kinerja kedua model juga dipengaruhi oleh proporsi data latih dan data uji. Secara umum, semakin tinggi proporsi data latih, semakin baik kinerja kedua model. Selain itu, keseimbangan proporsi kelas dalam *dataset* yang digunakan untuk pelatihan dan pengujian juga mempengaruhi kinerja klasifikasi model. Pada umumnya, kinerja kedua model lebih baik saat dilatih menggunakan *dataset* dengan proporsi kelas yang seimbang secara alami.

Kata kunci: Klasifikasi Berita Palsu, *Feed Forward Neural Network*, *Support Vector Machine*, Twitter/X, *Machine Learning*.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

ABSTRACT

Social media has evolved as a medium for the mass dissemination of information (including fake news). Fake news on social media platforms like Twitter/X often gets *retweeted* quickly and spreads widely, sometimes even going viral. Therefore, a mechanism is needed to automatically classify fake news on social media. This study aims to compare the performance of two machine learning models, namely *Feed Forward Neural Network* (FFNN) and *Support Vector Machine* (SVM), in automatically classifying fake news. Additionally, this research intends to compare the impact of the test data proportion relative to the training data, as well as the effect of class balance in the *dataset* on classification performance. This study employs an experimental quantitative approach. The experiment process involved training and testing the FFNN and SVM models using a *dataset* of Twitter/X post metadata, including the number of followers, number of quote *tweets*, number of *retweets*, and number of *likes*. The data used for the experiment totaled 2,104 samples collected through Astramaya by DroneEmprit, with an additional 1,781 samples obtained from the TurnBackHoax site. The classification results of both models were evaluated using a *Confusion Matrix* to calculate *accuracy*, *precision*, *recall*, and *F1-score*. The findings show that the SVM model generally achieved higher *accuracy*, *recall*, and *F1-score* compared to the FFNN model, which only excelled in *precision*. Both models generally experienced a decline in performance with an increasing proportion of test data (except for a 40% test data proportion, which performed better than a 20% test data proportion). Both models demonstrated better classification performance when trained on a *dataset* with balanced class

proportions, as opposed to an imbalanced *dataset* or one balanced using the SMOTE *Oversampling* technique. The SVM model produced better classification results than the FFNN model. The performance of both models was also influenced by the proportion of training and test data. In general, the higher the proportion of training data, the better the performance of both models. Moreover, the balance of class proportions in the *dataset* used for training and testing also impacted the models' classification performance. Overall, both models performed better when trained on a *dataset* with naturally balanced class proportions.

Keywords: *Fake News Classification, Feed Forward Neural Network, Support Vector Machine, Twitter/X, Machine Learning.*



KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT, atas berkat dan rahmat yang diberikan-Nya sehingga penulis dapat menyelesaikan penelitian ini. Tentunya selama penulisan dan penelitian yang dilakukan banyak pihak yang membantu dan memberikan dukungan, oleh karenanya penulis ingin mengucapkan terimakasih kepada:

1. Bapak Sudirman & Ibu Warsini kedua orang tua yang telah mendukung & memberikan do'a terbaik hingga saat sekarang ini.
2. Bapak Dr. Bambang Sugiantoro, M.T. selaku Kepala Program Studi Magister Informatika.
3. Ibu Maria Ulfah Siregar, S.Kom., MIT., Ph.D. selaku Dosen Pembimbing Akademik Program Studi Magister Informatika.
4. Bapak Dr. Agung Fatwanto, S.Si., M.Kom. selaku Dosen pembimbing yang dengan ilmu, kesabaran dan do'a beliau selama penyusunan tesis.
5. Bapak Ir. Muhammad Taufiq Nuruzzaman, S.T. M.Eng., Ph.D. selaku dosen penguji 1.
6. Bapak Dr. Ir. Bambang Sugiantoro, M.T., IPU., ASEAN Eng. selaku dosen penguji 2.
7. Bapak / Ibu Dosen Magister Informatika yang telah memberikan ilmu dan pengalaman terbaiknya dalam proses pengajaran.
8. Teman-teman Program Studi Magister Informatika angkatan 2020.

Penulis menyadari karya tulis ini masih jauh dari kata sempurna, oleh sebab itu kritik serta saran senantiasa penulis harapkan demi

perbaikan karya ilmiah ini. Akhir kata, semoga penelitian ini dapat bermanfaat bagi pembaca.



DAFTAR ISI

PERNYATAAN KEASLIAN	ii
PERNYATAAN BEBAS PLAGIASI.....	iii
PENGESAHAN TUGAS AKHIR	iv
NOTA DINAS PEMBIMBING	v
ABSTRAK.....	vi
ABSTRACT	viii
KATA PENGANTAR	x
DAFTAR ISI	xii
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR	xv
DAFTAR SINGKATAN	xvii
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah	1
B. Rumusan Masalah	4
C. Batasan Masalah.....	4
D. Tujuan dan Manfaat.....	5
BAB II KAJIAN PUSTAKA DAN LANDASAN TEORI.....	7
A. Kajian Pustaka	7
B. Landasan Teori	9
1. Berita Palsu.....	9

2.	<i>Feed Forward Neural Network</i>	11
3.	<i>Support Vector Machine</i>	34
4.	<i>Confusion Matrix</i>	40
BAB III METODE PENELITIAN.....		43
A.	Lokasi dan Waktu Penelitian	43
B.	Alat dan Bahan Penelitian	43
1.	Komputer Personal	43
2.	Google Colab	44
C.	Tahapan Penelitian	44
1.	Studi Literatur	45
2.	Pengumpulan Data	45
3.	Prapemrosesan Data	45
4.	Eksperimen	47
5.	Evaluasi Model	51
D.	Metode Analisis Data	52
BAB IV HASIL DAN PEMBAHASAN.....		53
BAB V PENUTUP.....		65
A.	Kesimpulan	65
B.	Saran	66
DAFTAR PUSTAKA		67
DAFTAR RIWAYAT HIDUP		76

DAFTAR TABEL

Tabel 2.1 Fungsi <i>Kernel</i> untuk <i>Support Vector Machine</i> Berbasis <i>Kernel</i>	38
Tabel 2.2 <i>Confusion Matrix</i>	40
Tabel 3.1 Deskripsi Atribut Identifikasi Berita Palsu.....	46
Tabel 3.2 Parameter yang Digunakan dalam Model FFNN	48
Tabel 3.3 <i>Hyperparameter</i> yang Digunakan dalam Model FFNN.....	48
Tabel 3.4 Strategi yang Digunakan dalam Model FFNN	49
Tabel 3.5 Parameter yang Digunakan dalam Model SVM.....	50
Tabel 3.6 <i>Hyperparameter</i> yang Digunakan dalam Model SVM	50
Tabel 3.7 Strategi yang Digunakan dalam Model SVM	50
Tabel 3.8 Proporsi Data yang Digunakan untuk Menguji Kinerja Model Setelah Dilatih	51
Tabel 4.1 Evaluasi Kinerja Model pada Data yang Tidak Seimbang..	53
Tabel 4.2 Evaluasi Kinerja Model Menggunakan Teknik <i>Oversampling</i> (SMOTE)	54
Tabel 4.3 Evaluasi Kinerja Model pada Data yang Seimbang	54

DAFTAR GAMBAR

Gambar 2.1 Arsitektur Dasar Jaringan Saraf.....	12
Gambar 2.2 Contoh <i>Feed Forward Neural Network</i>	18
Gambar 2.3 Ilustrasi Menghitung Nilai Keluaran di Lapisan Tersembunyi pada <i>Feed Forward Neural Network</i>	19
Gambar 2.4 Ilustrasi Menghitung Nilai Keluaran di Lapisan Output pada <i>Feed Forward Neural Network</i>	20
Gambar 2.5 Ilustrasi Menghitung Nilai Kesalahan pada <i>Feed Forward Neural Network</i>	23
Gambar 2.6 Ilustrasi Pembaruan Bobot di Antara Lapisan Output dan Lapisan Tersembunyi pada <i>Feed Forward Neural Network</i>	25
Gambar 2.7 Ilustrasi Mencari Nilai Kesalahan di Lapisan Tersembunyi pada <i>Feed Forward Neural Network</i>	27
Gambar 2.8 Ilustrasi Menghitung Nilai Kesalahan di Lapisan Tersembunyi pada <i>Feed Forward Neural Network</i>	28
Gambar 2.9 Ilustrasi Memperbarui Bobot di Antara Lapisan Tersembunyi dan Lapisan Input pada <i>Feed Forward Neural Network</i>	30
Gambar 2.10 Ilustrasi Memperbarui Bias di Seluruh Lapisan pada <i>Feed Forward Neural Network</i>	31
Gambar 2.11 <i>Support Vector Machine</i> Linear untuk Kasus yang Dapat Dipisahkan	35
Gambar 2.12 <i>Support Vector Machine</i> untuk Kasus yang Tidak Dapat Dipisahkan Secara Linear	37
Gambar 2.13 Ilustrasi Pemetaan Ruang Input ke Ruang Fitur dalam Kasus yang Tidak Dapat Dipisahkan Secara Linear	39

Gambar 3.1 Tahapan Penelitian.....	44
Gambar 4.1 Grafik Perbandingan Nilai Rata-rata Kinerja Model <i>Feed Forward Neural Network</i> dan <i>Support Vector Machine</i>	57
Gambar 4.2 Grafik Perbandingan Kinerja Model <i>Feed Forward Neural Network</i> dan <i>Support Vector Machine</i> dengan Berbagai ukuran Set Pengujian	59
Gambar 4.3 Grafik Perbandingan Kinerja Model <i>Feed Forward Neural Network</i> dan <i>Support Vector Machine</i> terhadap Perbedaan Distribusi Kelas dalam Dataset	60



DAFTAR SINGKATAN

FFNN	: <i>Feed Forward Neural Network</i>
SVM	: <i>Support Vector Machine</i>
RBF	: <i>Radial Basis Function</i>
SMOTE	: <i>Synthetic Minority Over-sampling Technique</i>
Acc	: <i>Accuracy</i>
Prec	: <i>Precision</i>
Rec	: <i>Recall</i>
F1	: <i>F1-Score</i>



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Kemajuan teknologi digital memainkan peran penting dalam perkembangan interaksi sosial. Salah satu perkembangan tersebut adalah munculnya media sosial sebagai platform komunikasi yang mampu menghubungkan banyak orang. Media sosial memberikan fasilitas yang memungkinkan pengguna untuk melakukan berbagai aktivitas sosial dengan mudah, termasuk berbagi informasi secara instan.

Namun, kemudahan berbagi informasi yang ditawarkan oleh internet dan media sosial juga membuka pintu bagi penyebaran berita palsu. Platform digital memiliki pengaruh besar dalam menyebarkan informasi yang sering kali salah dan tidak disadari oleh penggunanya.

Layanan media sosial memberikan kewenangan bagi setiap pengguna untuk memiliki kemampuan dalam membuat berita tanpa perlu ada yang memastikan informasi itu benar, sehingga membuat semua orang dapat menulis berita palsu (Greifeneder dkk., 2020). Kurangnya kesadaran banyak orang yang cenderung tidak memeriksa kebenaran berita sebelum membagikannya dengan teman, yang kemudian dapat menjangkau banyak pengguna lainnya dan memperburuk masalah ini (Buchanan, 2020). Ketika orang-orang mempercayai informasi palsu sebagai kebenaran, mereka condong menyebarkannya kembali, sehingga memperbesar penyebaran berita palsu tersebut.

Platform media sosial kini menjadi kanal utama untuk pembuatan dan pertukaran informasi serta berita. Meskipun media sosial memberikan ruang untuk penyebaran berita, platform ini juga menjadi tempat berkembangnya informasi yang tidak akurat. Salah satu media sosial yang terkenal dalam menyebarkan berita palsu adalah Twitter/X (Assiroj dkk., 2018). Penelitian telah menunjukkan bahwa berita kebohongan di Twitter/X biasanya di-*retweet* oleh lebih banyak orang dan lebih cepat daripada berita yang sebenarnya (Vosoughi dkk., 2018). Sebuah studi terhadap 2,8 juta kali berita yang dibagikan di Twitter/X menunjukkan bahwa 59% dari berita tersebut dibagikan tanpa dibaca terlebih dahulu oleh penggunanya (Gabelkov dkk., 2016).

Penelitian lain mengungkapkan bahwa indikator popularitas di media sosial, seperti jumlah suka, bagikan, atau komentar yang didapatkan dalam sebuah postingan dapat menarik lebih banyak perhatian dan meningkatkan kemungkinan untuk mendapatkan lebih banyak interaksi lagi (Tandoc dkk., 2017). Oleh karena itu, dalam penelitian ini, data analitik postingan di Twitter/X, termasuk jumlah pengikut pengirim *tweet* serta tingkat interaksi pada postingan seperti jumlah kutip *tweet*, jumlah *retweet*, dan jumlah *like*, akan digunakan sebagai fitur dalam *dataset* untuk mengklasifikasikan berita palsu. Fitur-fitur tersebut digunakan sebagai input dalam model pembelajaran mesin untuk mendeteksi berita palsu secara otomatis, yang diharapkan dapat meningkatkan keberhasilan dalam mendeteksi berita palsu.

Penelitian ini membandingkan kinerja antara salah satu algoritma klasifikasi berbasis jaringan saraf dan pembelajaran

mesin tradisional. Algoritma jaringan saraf yang dipilih adalah *Feed Forward Neural Network* (FFNN) karena memiliki waktu komputasi lebih cepat dibandingkan dengan *Recurrent Neural Network* (RNN) dan penerapannya yang lebih mudah daripada *Convolutional Neural Network* (CNN) (Ali dkk., 2015; Kareem dkk., 2021; Pang dkk., 2020). Sementara itu, sebagai perwakilan algoritma pembelajaran mesin tradisional, *Support Vector Machine* (SVM) dipilih karena kinerjanya yang lebih unggul dibandingkan dengan algoritma lain seperti *K-Nearest Neighbor*, *Genetic Algorithm*, *Decision Tree*, *Naive Bayes*, dan *Random Forest* (Bansal dkk., 2022; Biswas dkk., 2022; Narayan, 2021; Yu dkk., 2017).

Namun, sering ditemui masalah keterbatasan *dataset* dalam model pembelajaran mesin (Kumar dkk., 2021). Oleh karena itu, penelitian ini juga membandingkan pengaruh keseimbangan proporsi kelas dalam *dataset* terhadap kinerja klasifikasi, serta pengaruh proporsi data uji terhadap data latih. Namun, penting untuk dicatat bahwa dalam penelitian ini, isi atau konten twit tidak digunakan sebagai bahan untuk melakukan proses klasifikasi. Pendekatan ini berlandaskan pada asumsi bahwa pola penyebaran berita palsu dipengaruhi oleh interaksi pengguna dan ketenaran pihak yang memposting konten berita, yang sering kali berkaitan dengan kecepatan dan jangkauan penyebarannya yang lebih mudah terlihat oleh para pengguna Twitter/X. Berdasarkan asumsi ini, data analitik dari Twitter/X yang menunjukkan popularitas pihak yang memposting konten berita dapat menggambarkan tingkat kepercayaan pengguna lain terhadap konten. Hal ini berpotensi

menjadi indikator penting dalam mengidentifikasi berita yang dicurigai palsu. Dibandingkan dengan menggunakan teks sebagai data input yang memerlukan proses untuk mengubahnya menjadi data numerik, menggunakan data analitik postingan di Twitter/X memungkinkan model bekerja dengan lebih cepat dan efisien.

Evaluasi dan perbandingan kinerja kedua model dilakukan dengan menggunakan *Confusion Matrix*, yang merupakan metode untuk mengukur dan merangkum kinerja algoritma klasifikasi. Alat ini memungkinkan penghitungan berbagai *performance metrics* seperti, *accuracy*, *precision*, *recall*, dan, *f1-score*. Dengan menganalisis semua matriks ini, memungkinkan peneliti untuk menentukan model yang terbaik dalam mengklasifikasikan berita palsu berdasarkan data analitik yang diperoleh dari Twitter/X.

B. Rumusan Masalah

Dari latar belakang masalah yang telah dipaparkan sebelumnya, penulis membuat dua rumusan masalah, yaitu:

1. Bagaimana kemampuan model *Feed Forward Neural Network* dan *Support Vector Machine* dalam mengklasifikasikan berita palsu menggunakan data analitik di Twitter/X?
2. Bagaimana proporsi data uji terhadap data latih mempengaruhi kinerja model *Feed Forward Neural Network* dan *Support Vector Machine*?
3. Bagaimana pengaruh proporsi kelas dalam *dataset* terhadap kinerja klasifikasi model *Feed Forward Neural Network* dan *Support Vector Machine*?

C. Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya membandingkan kinerja dari dua jenis metode algoritma *machine learning*, yaitu *Feed Forward Neural Network* dan *Support Vector Machine*, dalam klasifikasi berita palsu.
2. Penelitian ini hanya menggunakan data analitik di Twitter/X, termasuk jumlah pengikut dan interaksi dalam postingan seperti, jumlah kutip *tweet*, jumlah *retweet*, dan jumlah *like*. Isi atau konten twit tidak digunakan sebagai bagian dari proses klasifikasi berita palsu.
3. Evaluasi dan perbandingan kinerja dari kedua model dilakukan dengan menggunakan *Confusion Matrix*. Pengukuran kinerja model terbatas pada matriks seperti *accuracy*, *precision*, *recall*, dan *f1-score*.

D. Tujuan dan Manfaat

Tujuan utama dari penelitian ini adalah untuk melakukan analisis komparatif terhadap kinerja dua model *machine learning*, yaitu *Feed Forward Neural Network* dan *Support Vector Machine*, dalam tugas klasifikasi berita palsu di platform media sosial Twitter/X. Secara lebih spesifik, penelitian ini bertujuan sebagai berikut:

1. Menentukan model yang paling baik dalam mengklasifikasikan berita palsu berdasarkan data analitik yang diperoleh dari Twitter/X, termasuk jumlah pengikut dan interaksi dalam postingan seperti jumlah kutip *tweet*, jumlah *retweet*, dan jumlah *like*.
2. Mengkaji pengaruh proporsi data uji terhadap data latih terhadap kinerja model *Feed Forward Neural Network* dan *Support Vector Machine*.

3. Menganalisis pengaruh proporsi kelas dalam *dataset* terhadap kinerja klasifikasi model *Feed Forward Neural Network* dan *Support Vector Machine*.

Secara teoretis, penelitian ini diharapkan dapat memberikan manfaat berupa:

1. Penelitian ini diharapkan dapat memberikan sumbangan ilmiah mengenai kinerja model klasifikasi *Feed Forward Neural Network* dan *Support Vector Machine* dalam klasifikasi berita palsu.
2. Penelitian ini diharapkan dapat digunakan sebagai landasan atau dasar bagi penelitian selanjutnya dalam mengembangkan model klasifikasi berita palsu yang lebih baik.

Secara praktis, penelitian ini diharapkan dapat memberikan manfaat berupa:

1. Menambah pengalaman bagi penulis dalam melakukan penelitian.
2. Hasil penelitian ini diharapkan dapat meningkatkan kesadaran dan pemahaman pengguna media sosial terhadap pentingnya memeriksa informasi atau berita terlebih dahulu sebelum membagikannya.

BAB V

PENUTUP

A. Kesimpulan

Berdasarkan hasil analisis yang telah dilakukan untuk mengkaji kemampuan model *Feed Forward Neural Network* dan *Support Vector Machine* dalam mengklasifikasikan berita palsu berdasarkan data analitik di Twitter/X, terdapat beberapa kesimpulan yang dapat diambil. Pertama, model *Support Vector Machine* menunjukkan kinerja yang lebih unggul dalam hal *accuracy*, *recall*, dan *f1-score* dibandingkan dengan model *Feed Forward Neural Network* (FFNN).

Temuan kedua dari penelitian ini menunjukkan bahwa proporsi data uji mempengaruhi kinerja model SVM dan *Feed Forward*. Kedua model mengalami penurunan *accuracy* seiring dengan penambahan proporsi data uji. Namun, ketika proporsi data uji sebesar 40%, kedua model menghasilkan tingkat *accuracy* yang tertinggi.

Temuan ketiga dari penelitian ini menunjukkan bahwa penggunaan *dataset* dengan proporsi kelas yang seimbang meningkatkan kinerja klasifikasi kedua model. Selain itu penggunaan *dataset* pelatihan yang memiliki proporsi jenis kelas yang seimbang secara alami juga meningkatkan kinerja klasifikasi kedua model dibandingkan penggunaan *dataset* pelatihan yang diseimbangkan menggunakan teknik *Oversampling* dengan metode SMOTE.

B. Saran

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, beberapa saran yang dapat diberikan untuk penelitian selanjutnya antara lain:

1. Peneliti berikutnya disarankan untuk mengkaji penggunaan model atau algoritma *machine learning* lain yang mungkin memberikan kinerja yang lebih baik.
2. Penelitian mendatang dapat mempertimbangkan fitur lain dari data Twitter/X yang belum digunakan dalam penelitian ini.
3. Pengembangan arsitektur model *Feed Forward Neural Network* yang lebih kompleks dapat dilakukan dengan menambahkan jumlah lapisan tersembunyi atau *neuron*, serta menyesuaikan *hyperparameter* atau parameter lainnya untuk meningkatkan kinerja klasifikasi. Diperlukan eksperimen berulang dan evaluasi terhadap berbagai arsitektur jaringan untuk menemukan konfigurasi yang optimal.
4. Penelitian lebih lanjut diperlukan untuk mengeksplorasi metode lain selain menggunakan teknik *Oversampling* dengan metode SMOTE dalam menangani ketidakseimbangan kelas dalam *dataset*.
5. Disarankan untuk melakukan eksperimen dengan berbagai proporsi data uji dan latih yang belum digunakan dalam penelitian ini.

DAFTAR PUSTAKA

- Abbas, A. K., Al-haideri, N. A., & Bashikh, A. A. (2019). Implementing artificial neural networks and support vector machines to predict lost circulation. *Egyptian Journal of Petroleum*, 28(4), 339–347. <https://doi.org/https://doi.org/10.1016/j.ejpe.2019.06.006>
- Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., & Notarnicola, C. (2015). Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. *Remote Sensing*, 7. <https://doi.org/10.3390/rs71215841>
- Asniar, Maulidevi, N. U., & Surendro, K. (2022). SMOTE-LOF for noise identification in imbalanced data classification. *Journal of King Saud University - Computer and Information Sciences*, 34(6, Part B), 3413–3423. <https://doi.org/https://doi.org/10.1016/j.jksuci.2021.01.014>
- Assiroj, P., Meyliana, Hidayanto, A. N., Prabowo, H., & Warnars, H. L. H. S. (2018). Hoax News Detection on Social Media: A Survey. 2018 Indonesian Association for Pattern Recognition International Conference (INAPR), 186–191. <https://doi.org/10.1109/INAPR.2018.8627053>
- Bansal, M., Goyal, A., & Choudhary, A. (2022). A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning. *Decision Analytics Journal*, 3, 100071. <https://doi.org/https://doi.org/10.1016/j.dajour.2022.100071>
- Bhatia, T., Manaskasemsak, B., & Rungsawang, A. (2023). Detecting Fake News Sources on Twitter Using Deep Neural Network. 2023 11th International Conference on Information and Education

- Technology (ICIET), 508–512.
<https://doi.org/10.1109/ICIET56899.2023.10111446>
- Biswas, N., Uddin, K. M. M., Rikta, S. T., & Dey, S. K. (2022). A comparative analysis of machine learning classifiers for stroke prediction: A predictive analytics approach. *Healthcare Analytics*, 2, 100116.
<https://doi.org/https://doi.org/10.1016/j.health.2022.100116>
- Borah, P., & Gupta, Dr. D. (2017). Review: Support Vector Machines in Pattern Recognition. *International Journal of Engineering and Technology*, 9, 43–48.
<https://doi.org/10.21817/ijet/2017/v9i3/170903S008>
- Bowyer, K. W., Chawla, N. V., Hall, L. O., & Kegelmeyer, W. P. (2011). SMOTE: Synthetic Minority Over-sampling Technique. *CoRR*, abs/1106.1813. <http://arxiv.org/abs/1106.1813>
- 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*, 15(10), 1–33. <https://doi.org/10.1371/journal.pone.0239666>
- Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, 189–215.
<https://doi.org/https://doi.org/10.1016/j.neucom.2019.10.118>
- Collins. (t.t.). Definition of fake news. Diambil 1 Februari 2024, dari <https://www.collinsdictionary.com/dictionary/english/fake-news>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>

- da Silva, I. N., Spatti, D. H., Flauzino, R. A., Liboni, L. H. B., & Alves, S. F. dos R. (2018). *Artificial Neural Networks: A Practical Course* (1st ed.). Springer Publishing Company, Incorporated.
- Elreedy, D., & Atiya, A. F. (2019). A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance. *Information Sciences*, 505, 32–64. <https://doi.org/https://doi.org/10.1016/j.ins.2019.07.070>
- Elreedy, D., Atiya, A. F., & Kamalov, F. (2024). A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Machine Learning*, 113(7), 4903–4923. <https://doi.org/10.1007/s10994-022-06296-4>
- Fauzi, A., Setiawan, E., & Baizal, A. (2019). Hoax News Detection on Twitter using Term Frequency Inverse Document Frequency and Support Vector Machine Method. *Journal of Physics: Conference Series*, 1192, 12025. <https://doi.org/10.1088/1742-6596/1192/1/012025>
- Gabiolkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). Social Clicks: What and Who Gets Read on Twitter? *Proceedings of the 2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science*. <https://api.semanticscholar.org/CorpusID:7532016>
- García Cabello, J. (2022). *Mathematical Neural Networks*. *Axioms*, 11(2). <https://doi.org/10.3390/axioms11020080>
- Ghosh, S., Dasgupta, A., & Swetapadma, A. (2019). A Study on Support Vector Machine based Linear and Non-Linear Pattern Classification. *2019 International Conference on Intelligent*

- Sustainable Systems (ICISS), 24–28.
<https://doi.org/10.1109/ISS1.2019.8908018>
- Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. Dalam Y. W. Teh & M. Titterton (Ed.), *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* (Vol. 9, hlm. 249–256). PMLR. <https://proceedings.mlr.press/v9/glorot10a.html>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for Multi-Class Classification: an Overview.
- Greifeneder, R., Jaffé, M., Newman, E., & Schwarz, N. (2020). What is New and True about Fake News? (hlm. 1–8). <https://doi.org/10.4324/9780429295379-1>
- Greifeneder, R., Müller, P., Stahlberg, D., van den Bos, K., & Bless, H. (2011). Guiding Trustful Behavior: The Role of Accessible Content and Accessibility Experiences. *Journal of Behavioral Decision Making*, 24, 498–514. <https://doi.org/10.1002/bdm.705>
- Guan, H., Zhao, L., Dong, X., & Chen, C. (2023). Extended natural neighborhood for SMOTE and its variants in imbalanced classification. *Engineering Applications of Artificial Intelligence*, 124, 106570. <https://doi.org/https://doi.org/10.1016/j.engappai.2023.106570>
- Kareem, S. W., Hamad, Z. J., & Askar, S. K. (2021). An evaluation of CNN and ANN in prediction weather forecasting: A review. *Sustainable Engineering and Innovation*. <https://api.semanticscholar.org/CorpusID:238996288>

- Kencana, C., Setiawan, E., & Kurniawan, I. (2020). Hoax Detection System on Twitter using Feed-Forward and Back-Propagation Neural Networks Classification Method. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4, 655–663. <https://doi.org/10.29207/resti.v4i4.2038>
- Kumar, P., Bhatnagar, R., Gaur, K., & Bhatnagar, A. (2021). Classification of Imbalanced Data: Review of Methods and Applications. *IOP Conference Series: Materials Science and Engineering*, 1099, 12077. <https://doi.org/10.1088/1757-899X/1099/1/012077>
- Li, F., Wang, W., Xu, J., Yi, J., & Wang, Q. (2019). Comparative study on vulnerability assessment for urban buried gas pipeline network based on SVM and ANN methods. *Process Safety and Environmental Protection*, 122, 23–32. <https://doi.org/https://doi.org/10.1016/j.psep.2018.11.014>
- Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, 21(6), 335–346. <https://doi.org/10.1038/s41583-020-0277-3>
- Lyons, B., & Reifler, J. (2020). How bad is the fake news problem? The effect of baseline information in public perceptions.
- Meng, D., & Li, Y. (2022). An imbalanced learning method by combining SMOTE with Center Offset Factor. *Applied Soft Computing*, 120, 108618. <https://doi.org/https://doi.org/10.1016/j.asoc.2022.108618>
- Narayan, Y. (2021). Comparative analysis of SVM and Naive Bayes classifier for the SEMG signal classification. *Materials Today:*

- Proceedings, 37, 3241–3245.
<https://doi.org/https://doi.org/10.1016/j.matpr.2020.09.093>
- Niu, X., Yang, C., Wang, H., & Wang, Y. (2017). Investigation of ANN and SVM based on limited samples for performance and emissions prediction of a CRDI-assisted marine diesel engine. *Applied Thermal Engineering*, 111, 1353–1364.
<https://doi.org/https://doi.org/10.1016/j.applthermaleng.2016.10.042>
- Ojha, V. K., Abraham, A., & Snášel, V. (2017). Metaheuristic design of feedforward neural networks: A review of two decades of research. *Engineering Applications of Artificial Intelligence*, 60, 97–116.
<https://doi.org/https://doi.org/10.1016/j.engappai.2017.01.013>
- Otchere, D. A., Arbi Ganat, T. O., Gholami, R., & Ridha, S. (2021). Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *Journal of Petroleum Science and Engineering*, 200, 108182.
<https://doi.org/https://doi.org/10.1016/j.petrol.2020.108182>
- Pang, Z., Niu, F., & O'Neill, Z. (2020). Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons. *Renewable Energy*, 156, 279–289.
<https://doi.org/https://doi.org/10.1016/j.renene.2020.04.042>
- Patil, A., Framewala, A., & Kazi, F. (2020). Explainability of SMOTE Based Oversampling for Imbalanced Dataset Problems. 2020 3rd International Conference on Information and Computer Technologies (ICICT), 41–45.
<https://doi.org/10.1109/ICICT50521.2020.00015>

- Santoso, H. A., Rachmawanto, E. H., & Hidayati, U. (2020). Fake Twitter Account Classification of Fake News Spreading Using Naïve Bayes. *Scientific Journal of Informatics*, 7(2), 228–237. <https://doi.org/10.15294/sji.v7i2.25747>
- Sekeroglu, B., Kirsal Ever, Y., Dimililer, K., & Al-Turjman, F. (2022). Comparative Evaluation and Comprehensive Analysis of Machine Learning Models for Regression Problems. *Data Intelligence*, 4, 620–652. https://doi.org/10.1162/dint_a_00155
- Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97, 105524. <https://doi.org/https://doi.org/10.1016/j.asoc.2019.105524>
- Singhania, S., Fernandez, N., & Rao, S. (2017, Juli). 3HAN: A Deep Neural Network for Fake News Detection. https://doi.org/10.1007/978-3-319-70096-0_59
- Smith, K. (2020, Januari 2). 60 Incredible and Interesting Twitter Stats and Statistics. <https://www.brandwatch.com/blog/twitter-stats-and-statistics/>.
- Tandoc, E., Lim, Z., & Ling, R. (2017). Defining “Fake News”: A typology of scholarly definitions. *Digital Journalism*, 6, 1–17. <https://doi.org/10.1080/21670811.2017.1360143>
- Thölke, P., Mantilla-Ramos, Y.-J., Abdelhedi, H., Maschke, C., Dehgan, A., Harel, Y., Kemtur, A., Mekki Berrada, L., Sahraoui, M., Young, T., Bellemare Pépin, A., El Khantour, C., Landry, M., Pascarella, A., Hadid, V., Combrisson, E., O’Byrne, J., & Jerbi, K. (2023). Class imbalance should not throw you off balance: Choosing the right classifiers and performance metrics for brain

- decoding with imbalanced data. *NeuroImage*, 277, 120253.
<https://doi.org/https://doi.org/10.1016/j.neuroimage.2023.120253>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151.
<https://doi.org/10.1126/science.aap9559>
- Wardle, C., & Derakhshan, H. (2017). INFORMATION DISORDER :
 Toward an interdisciplinary framework for research and policy making
 Information Disorder Toward an interdisciplinary framework for research and policymaking.
- Waspada, I., Bahtiar, N., Wirawan, P. W., & Awan, B. D. A. (2020). Performance Analysis of Isolation Forest Algorithm in Fraud Detection of Credit Card Transactions. *Khazanah Informatika Jurnal Ilmu Komputer dan Informatika*, 6(2).
<https://doi.org/10.23917/khif.v6i2.10520>
- Wei Qiong AND Dunbrack, J. R. L. (2013). The Role of Balanced Training and Testing Data Sets for Binary Classifiers in Bioinformatics. *PLOS ONE*, 8(7), 1–12.
<https://doi.org/10.1371/journal.pone.0067863>
- Wirasati, I., Rustam, Z., Aurelia, J. E., Hartini, S., & Saragih, G. S. (2021). Comparison some of kernel functions with support vector machines classifier for thalassemia dataset. *IAES International Journal of Artificial Intelligence*, 10(2), 430–437.
<https://doi.org/10.11591/IJAI.V10.I2.PP430-437>
- Xu, Y., & Goodacre, R. (2018). On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. *Journal of Analysis and*

Testing, 2(3), 249–262. <https://doi.org/10.1007/s41664-018-0068-2>

Yu, P.-S., Yang, T.-C., Chen, S.-Y., Kuo, C.-M., & Tseng, H.-W. (2017). Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. *Journal of Hydrology*, 552, 92–104. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2017.06.020>

