IMPLEMENTASI WFRAUD ALERT SEBAGAI PREDIKSI PESAN PENIPUAN WHATSAPP MENGGUNAKAN MODEL NAIVE BAYES

SKRIPSI

Diajukan Untuk Memenuhi Syarat Mendapatkan Gelar Sarjana Strata Satu (S1) Pada

Program Studi Teknik Elektro Fakultas Teknik Universitas Islam Nusantara

Oleh :

Julian

41037002211004

PROGRAM STUDI TEKNIK ELEKTRO

FAKULTAS TEKNIK

UNIVERSITAS ISLAM NUSANTARA

2025

LEMBAR KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini:

Nama: Julian

NIM: 41037002211004

Program Studi: Teknik Elektro

Menyatakan bahwa Skripsi yang berjudul:

WFRAUD ALERT SEBAGAI PREDIKSI PESAN PENIPUAN WHATSAPP MENGGUNAKAN MODEL NAIVE BAYES

Dibuat dengan sebenar-benarnya dari penelitian, pemikiran, dan pemaparan hasil saya sendiri, untuk melengkapi sebagai pernyataan menjadi Sarjana (S1) pada jurusan Teknik Elektro Fakultas Teknik Universitas Islam Nusantara Bandung, sejauh yang saya ketahui bukan merupakan tiruan atau duplikasi dari buku Skripsi yang sudah dipublikasikan dan atau pernah dipakai untuk mendapatkan jenjang Sarjana (S1) di lingkungan Teknik Elektro Fakultas Teknik Universitas Islam Nusantara Bandung maupun perguruan-perguruan tinggi atau instansi manapun kecuali bagian yang sumber informasi dicantumkan sebagaimana mestinya.

Bandung, 27 Desember 2024

Yang membuat pernyataan,

JULIAN

NIM 41037002211004

LEMBAR PENGESAHAN

WFRAUD ALERT SEBAGAI PREDIKSI PESAN PENIPUAN WHATSAPP MENGGUNAKAN MODEL NAIVE BAYES

Disusun dan diajukan oleh :

JULIAN

41037002211004

Disetujui dan disahkan pada sidang skripsi

pada tanggal :

Bandung, 3 Februari 2025

Pembimbing 1	Pembimbing 2
Ganis Sanhaji, S.Si., M.Sc.	**Dr. Iksal Rachman, M.T.**

Mengetahui :

Dekan Fakultas Teknik	Ketua Prodi Teknik Elektro
Dr. Ricky Yoseptry, S.T., M.M.Pd.	**Ganis Sanhaji, S.Si., M.Sc.**

LEMBAR PENGESAHAN

REVISIAN SKRIPSI

WFRAUD ALERT SEBAGAI PREDIKSI PESAN PENIPUAN WHATSAPP MENGGUNAKAN MODEL NAIVE BAYES

Telah Direvisi oleh :

JULIAN

41037002211004

Bandung, 3 Februari 2025

Mengesahkan,

Penguji 1	Penguji 2
Dr. Tedjo Darmanto, M.T.	**Galih, S.T., M.Kom.**

Ketua Sidang

Dr.Ricky Yoseptry, M.M.Pd.

BIODATA PENULIS

`

Nama	: Julian
Tempat, Tanggal Lahir	: Sumedang, 26 Agustus 2002
Telepon	: 0857-9733-9179
Email	: julianfikoma@gmail.com
Riwayat Pendidikan	: SDN Giri Jaya
	SMPN 1 Tomo
	SMAN Tomo

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT, Tuhan Yang Maha Kuasa, atas limpahan rahmat, kasih sayang, serta hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "WFraud Alert sebagai Prediksi Pesan Penipuan WhatsApp Menggunakan Model Naive Bayes". Skripsi ini merupakan salah satu syarat untuk menyelesaikan pendidikan Strata Satu (S1) di Universitas Islam Nusantara, Fakultas Teknik, Program Studi Teknik Elektro.

Dalam perjalanan panjang penyusunan skripsi ini, penulis tidak akan sampai pada titik ini tanpa dukungan, doa, dan bantuan dari berbagai pihak. Dengan hati yang penuh rasa syukur dan hormat, penulis menyampaikan rasa terima kasih yang mendalam kepada:

1. Allah SWT, yang senantiasa memberikan kekuatan, kesehatan, dan petunjuk di setiap langkah penulis. Segala sesuatu adalah karena kehendak-Nya, dan penulis memohon semoga karya ini menjadi bermanfaat.
2. Ibu Tini Sumartini, Abah Cece Kurniawan, Bapak Uho, Emak Iin, dan Bapak Johanes, sebagai orang tua yang selalu menjadi sumber kekuatan terbesar. Terima kasih atas segala doa yang tak pernah terputus, dukungan moral, motivasi, serta bantuan material yang telah diberikan. Juga kepada kakak tercinta Jimmy dan Carolline, serta adik-adik tercinta Jordan, Junior, Jennifer, dan Jevelline, yang selalu memberikan semangat dan cinta yang tak ternilai. Tak lupa, ucapan terima kasih kepada seluruh keluarga besar penulis yang telah menjadi pilar kekuatan selama ini.
3. Bapak dan Ibu Herlan Syah, yang telah menjadi orang tua kedua bagi saya, sekaligus sumber inspirasi dan teladan dalam perjuangan akademik serta kehidupan.
4. Bapak Ganis Sanhaji, S.Si., M.Sc., selaku Ketua Program Studi Teknik Elektro sekaligus dosen pembimbing skripsi, atas bimbingan, arahan, dan kesabaran dalam membantu penulis menyelesaikan skripsi ini.
5. Dr. Ricky Yoseptry, S.T., M.M.Pd., selaku Dekan Fakultas Teknik, yang telah memberikan dukungan dan motivasi kepada penulis selama masa perkuliahan.

6. Seluruh dosen Fakultas Teknik, yang telah dengan tulus memberikan ilmu dan pengalaman berharga yang menjadi bekal bagi penulis.
7. Seluruh staff Tata Usaha Fakultas Teknik, yang telah membantu penulis dalam menyelesaikan berbagai urusan administrasi kemahasiswaan dengan penuh kesabaran.
8. Prof. Dr. Endang Komara, M.Si., selaku Rektor Universitas Islam Nusantara, yang telah memberikan kesempatan kepada penulis untuk menimba ilmu di universitas tercinta ini.
9. Rekan seperjuangan skripsi: Herlan Syah, Decky Putra Kurnia, Anisa Febrianti, Vito Dwi Nur Hidayat, dan Hidayat, yang telah menjadi teman diskusi, penyemangat, dan partner perjuangan selama menyelesaikan skripsi ini.
10. Seluruh teman-teman Teknik Elektro angkatan 2021, yang telah memberikan warna, dukungan, dan kebersamaan selama masa perkuliahan.
11. Semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah membantu, mendukung, dan mendoakan penulis selama ini.

Akhirnya, penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, saran dan kritik yang membangun sangat diharapkan demi perbaikan di masa mendatang. Semoga karya ini dapat memberikan manfaat bagi siapa saja yang membacanya.

Bandung, 27 Desember 2024

Penulis

ABSTRAK

Kemajuan teknologi digital mempermudah komunikasi melalui platform seperti WhatsApp, namun juga meningkatkan ancaman kejahatan siber, terutama pesan penipuan. Penelitian ini mengembangkan WFraud Alert, aplikasi untuk memprediksi pesan penipuan WhatsApp menggunakan algoritma Naïve Bayes. Dataset terdiri dari 156 pesan yang dikategorikan sebagai pesan normal, pesan penipuan, dan promosi judi daring. Proses preprocessing meliputi case folding, normalisasi, penghapusan stopword, dan stemming untuk meningkatkan kinerja model. Algoritma diuji menggunakan metrik presisi, recall, dan F1-score, dengan presisi 88%, recall 90%, f1-score 87%. Hasil menunjukkan bahwa preprocessing meningkatkan akurasi model secara signifikan. Selain deteksi, aplikasi ini juga bertujuan meningkatkan literasi digital pengguna dan mengurangi risiko kejahatan siber. Penelitian ini berkontribusi pada pengembangan pemrosesan bahasa alami (NLP) dan keamanan siber di Indonesia dengan menerapkan metode klasifikasi teks untuk mengatasi penipuan digital.

ABSTRACT

The advancement of digital technology has simplified communication through platforms like WhatsApp but has also increased the threat of cybercrime, particularly phishing messages. This research developed WFraud Alert, an application designed to predict WhatsApp phishing messages using the Naïve Bayes algorithm. The dataset consists of 156 messages categorized as normal messages, phishing messages, and online gambling promotions. The preprocessing steps include case folding, normalization, stopword removal, and stemming to enhance model performance. The algorithm was evaluated using precision, recall, and F1-score metrics, achieving 88% precision, 90% recall, and 87% F1-score. The results indicate that preprocessing significantly improves the model's accuracy. Beyond detection, the application aims to enhance users' digital literacy and reduce the risk of cybercrime. This study contributes to the development of natural language processing (NLP) and cybersecurity in Indonesia by applying text classification methods to combat digital fraud.

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB 1 PENDAHULUAN

1.1. Latar Belakang

Kemajuan teknologi informasi dan komunikasi telah menciptakan perubahan besar dalam kehidupan manusia, terutama dalam cara berinteraksi dan berbagi informasi. Media sosial, seperti WhatsApp, telah menjadi bagian tak terpisahkan dari kehidupan sehari-hari, menyediakan sarana komunikasi yang cepat, mudah, dan efisien tanpa batasan geografis. WhatsApp kini tidak hanya digunakan untuk kebutuhan pribadi, tetapi juga untuk keperluan bisnis, pendidikan, hingga layanan masyarakat. Popularitas platform ini menunjukkan besarnya pengaruh teknologi digital dalam menghubungkan masyarakat modern.

Namun, di balik manfaat yang ditawarkan, kemajuan ini juga membuka peluang bagi munculnya kejahatan digital, seperti penipuan daring. Penipuan melalui pesan WhatsApp menjadi salah satu modus operandi yang sering digunakan oleh pelaku kejahatan siber. Modus ini sangat mengkhawatirkan karena pesan-pesan tersebut kerap kali terlihat seperti komunikasi resmi, sehingga sulit dibedakan oleh penerima. Pesan penipuan bisa berbentuk permintaan transfer uang dengan alasan darurat, tautan berbahaya yang mencuri data pribadi, atau bahkan tawaran hadiah palsu yang dirancang untuk menarik perhatian.

Data dari Kementerian Komunikasi dan Informatika (Kominfo) menunjukkan betapa seriusnya masalah ini. Sejak Agustus 2018 hingga Februari 2023, terdapat lebih dari 1.730 kasus penipuan daring yang dilaporkan, dengan total kerugian mencapai Rp 18,7 triliun. Kerugian yang signifikan ini menunjukkan dampak besar kejahatan digital terhadap masyarakat, terutama dari sisi ekonomi. Selain itu, laporan ini juga menggambarkan pentingnya solusi yang efektif untuk mencegah dan mengidentifikasi pesan-pesan berbahaya yang dapat menipu pengguna.

Penipuan daring tidak hanya merugikan secara finansial, tetapi juga menciptakan tekanan psikologis bagi korban. Modus operandi yang digunakan sering kali memanfaatkan celah psikologis korban, seperti rasa panik atau kepercayaan berlebih pada

informasi yang diberikan. Hal ini membuat masyarakat rentan terhadap manipulasi digital. Dengan ancaman yang terus berkembang ini, diperlukan langkah-langkah preventif berbasis teknologi untuk membantu masyarakat mengenali dan menghindari pesan-pesan yang mencurigakan.

Salah satu metode yang menjanjikan untuk menghadapi masalah ini adalah penerapan algoritma pembelajaran mesin (machine learning), khususnya Naïve Bayes. Algoritma ini dikenal efektif dalam menangani masalah klasifikasi teks, terutama dalam domain pengolahan bahasa alami (Natural Language Processing). Keunggulan utama Naïve Bayes terletak pada kesederhanaannya dalam perhitungan dan kemampuannya untuk memberikan hasil yang akurat meskipun dengan jumlah data yang terbatas. Oleh karena itu, algoritma ini menjadi pilihan yang relevan dalam pengembangan aplikasi deteksi penipuan seperti WFraud Alert.

Penelitian ini bertujuan untuk mengembangkan aplikasi WFraud Alert yang mampu mengklasifikasikan pesan WhatsApp menjadi tiga kategori utama: pesan normal, pesan penipuan, dan pesan promosi judi daring. Proses klasifikasi dilakukan melalui berbagai tahap pengolahan awal data (preprocessing), seperti case folding, normalisasi kata, penghapusan kata-kata umum (stopword removal), dan stemming. Setiap tahapan ini dirancang untuk meningkatkan akurasi algoritma dalam mengenali pola dan karakteristik dari setiap kategori pesan.

Dengan menggunakan algoritma Naïve Bayes, aplikasi ini diharapkan mampu mengidentifikasi pesan-pesan berbahaya dengan tingkat presisi yang tinggi. Tidak hanya itu, aplikasi ini juga dapat menjadi alat edukasi bagi masyarakat untuk memahami pola-pola penipuan yang umum terjadi di platform digital. Dalam jangka panjang, WFraud Alert diharapkan dapat memberikan kontribusi nyata dalam meningkatkan literasi digital masyarakat sekaligus melindungi mereka dari ancaman kejahatan siber.

Penipuan daring melalui pesan WhatsApp tidak hanya menjadi permasalahan di tingkat individu, tetapi juga memiliki dampak luas terhadap keamanan ekonomi nasional. Masyarakat yang menjadi korban sering kali kehilangan kepercayaan terhadap sistem

digital, yang pada akhirnya dapat menghambat adopsi teknologi secara lebih luas. Oleh karena itu, solusi yang diusulkan melalui aplikasi WFraud Alert tidak hanya memiliki dampak langsung dalam melindungi individu, tetapi juga memberikan kontribusi strategis dalam memperkuat ekosistem digital yang aman dan terpercaya di Indonesia.

Dalam konteks akademik, penelitian ini memberikan kontribusi pada pengembangan metode klasifikasi teks berbasis algoritma Naïve Bayes, khususnya untuk teks berbahasa Indonesia. Selain itu, penelitian ini juga menyoroti pentingnya pendekatan yang terintegrasi, di mana teknologi pembelajaran mesin digabungkan dengan metode pengolahan bahasa alami untuk menghasilkan solusi yang lebih efektif dan efisien.

Penelitian ini tidak hanya relevan bagi masyarakat luas, tetapi juga menjadi langkah awal bagi pengembangan teknologi keamanan digital di Indonesia. Dengan ancaman kejahatan siber yang terus berkembang, solusi seperti WFraud Alert menjadi semakin penting untuk menghadapi tantangan era digital. Selain itu, penelitian ini juga membuka peluang bagi pengembangan lebih lanjut di bidang keamanan siber, baik dari sisi teknologi maupun kebijakan.

Melalui penelitian ini, diharapkan masyarakat dapat lebih waspada terhadap ancaman penipuan daring sekaligus lebih percaya diri dalam memanfaatkan teknologi digital secara aman. Dengan demikian, aplikasi WFraud Alert tidak hanya berfungsi sebagai alat pendeteksi, tetapi juga menjadi bagian dari solusi jangka panjang untuk membangun lingkungan digital yang lebih baik.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, terdapat beberapa permasalahan utama yang menjadi fokus penelitian ini, yaitu:

1. Bagaimana mengidentifikasi pesan WhatsApp yang tergolong sebagai pesan penipuan, promosi judi online, atau pesan normal dengan tingkat akurasi yang tinggi?
2. Seberapa efektif algoritma Naïve Bayes dalam mengklasifikasikan pesan WhatsApp berdasarkan kategori yang telah ditentukan?

3. Bagaimana tahapan preprocessing data, seperti case folding, normalisasi kata, penghapusan stopword, dan stemming, dapat meningkatkan performa model dalam klasifikasi pesan WhatsApp?
4. Seberapa signifikan aplikasi WFraud Alert dalam membantu masyarakat mencegah dampak negatif dari pesan penipuan di platform WhatsApp?

1.2. Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan utama, yaitu:

1. Mengembangkan aplikasi WFraud Alert yang mampu mengidentifikasi dan mengklasifikasikan pesan WhatsApp ke dalam tiga kategori utama: pesan normal, pesan penipuan, dan promosi judi daring.
2. Menggunakan algoritma Naïve Bayes untuk melatih model klasifikasi pesan WhatsApp secara efektif dan efisien.
3. Mengimplementasikan tahapan preprocessing data, seperti case folding, normalisasi kata, stopword removal, dan stemming, untuk meningkatkan performa model dalam mengenali pola pesan.
4. Mengevaluasi performa aplikasi WFraud Alert berdasarkan metrik seperti presisi, recall, dan F1-score, guna memastikan tingkat akurasi yang tinggi dalam klasifikasi pesan.
5. Memberikan solusi berbasis teknologi yang dapat membantu masyarakat mengenali dan menghindari ancaman pesan penipuan di platform WhatsApp.

1.3. Manfaat Penelitian

Manfaat Teoritis:

1. Memberikan kontribusi ilmiah dalam pengembangan algoritma pembelajaran mesin, khususnya Naïve Bayes, untuk klasifikasi teks berbahasa Indonesia.
2. Menambah wawasan akademis mengenai teknik preprocessing data dalam pengolahan bahasa alami (Natural Language Processing).
3. Menyediakan referensi penelitian untuk studi lanjutan yang berkaitan dengan keamanan siber dan deteksi penipuan daring.

Manfaat Praktis:

1. Membantu pengguna WhatsApp dalam mengenali dan menghindari pesan penipuan, sehingga melindungi mereka dari kerugian finansial dan dampak psikologis.
2. Meningkatkan literasi digital masyarakat dengan memberikan pemahaman tentang pola-pola penipuan yang sering terjadi.
3. Mendukung upaya pemerintah dan institusi dalam menciptakan ekosistem digital yang aman dan terpercaya.

Manfaat Sosial:

1. Mengurangi tingkat kejahatan siber yang berkaitan dengan penipuan melalui platform komunikasi digital.
2. Membangun kesadaran masyarakat akan pentingnya menjaga privasi dan keamanan dalam berkomunikasi di era digital.

BAB 2 LANDASAN TEORI

2.1. Definisi Kejahatan Siber

Kejahatan siber, atau yang sering disebut cybercrime, mengacu pada berbagai tindakan kriminal yang memanfaatkan teknologi informasi sebagai alat atau sasaran. Aktivitas ilegal ini berkembang pesat seiring dengan penetrasi internet dalam kehidupan masyarakat modern. Menurut Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (UU ITE), kejahatan siber mencakup pelanggaran hukum yang dilakukan melalui perangkat teknologi informasi, baik untuk tujuan penipuan, pencurian, maupun penyebaran konten yang merugikan pihak lain.

Kejahatan siber memiliki karakteristik khusus yang membedakannya dari tindak kriminal konvensional. Salah satu ciri utamanya adalah pelaku dan korban sering kali tidak memiliki interaksi langsung. Kejahatan ini disebut sebagai "faceless crime" karena dijalankan tanpa tatap muka dan menggunakan identitas palsu, sehingga sulit dilacak. Teknologi yang digunakan dalam kejahatan siber seringkali bersifat lintas negara, menciptakan tantangan besar bagi aparat penegak hukum dalam menetapkan yurisdiksi dan mengidentifikasi pelaku.

Selain itu, kejahatan siber dapat dilakukan oleh individu maupun kelompok terorganisir. Dalam kasus tertentu, pelaku memiliki kemampuan teknis tinggi, seperti meretas sistem keamanan, mencuri data sensitif, atau menyebarkan perangkat lunak berbahaya. Dampak dari kejahatan siber ini tidak hanya pada aspek ekonomi, tetapi juga mencakup ancaman terhadap keamanan data pribadi dan stabilitas sistem teknologi yang digunakan secara luas.

2.2. Penipuan Daring sebagai Bagian dari Kejahatan Siber

Penipuan daring adalah salah satu jenis kejahatan siber yang paling sering terjadi. Istilah ini mengacu pada aktivitas manipulatif yang dilakukan secara digital dengan tujuan untuk menipu individu atau organisasi demi keuntungan pribadi pelaku. Penipuan ini biasanya melibatkan penyebaran informasi palsu, pengelabuan, atau eksploitasi kelemahan teknis dari korban.

Keberadaan teknologi komunikasi dan internet telah membuka peluang besar bagi pelaku penipuan daring untuk menjangkau korban tanpa batasan geografis. Modus operandi dalam penipuan ini melibatkan berbagai teknik, seperti penggunaan situs palsu, pengiriman email penipuan, hingga eksploitasi media sosial. Selain itu, penipuan daring juga memanfaatkan kelemahan dalam literasi digital masyarakat, terutama mereka yang kurang memahami risiko keamanan dalam transaksi elektronik.

2.3. Karakteristik Penipuan Daring

Penipuan daring memiliki beberapa karakteristik yang membedakannya dari bentuk kejahatan lainnya. Berikut adalah ciri-ciri utamanya:

a. Tidak Ada Interaksi Fisik

Seluruh proses penipuan dilakukan melalui media digital tanpa pertemuan langsung antara pelaku dan korban. Hal ini mempermudah pelaku untuk menyembunyikan identitas dan lokasi mereka, sehingga sulit dilacak oleh aparat penegak hukum.

b. Penyalahgunaan Identitas

Pelaku sering menggunakan identitas palsu untuk meyakinkan korban. Identitas ini bisa berupa nama samaran, profil media sosial palsu, atau bahkan kredensial dari institusi yang dihormati.

c. Bervariasinya Media yang Digunakan

Penipuan daring dilakukan melalui berbagai saluran, seperti situs web palsu, email, telepon, pesan instan, dan media sosial. Tiap media memiliki metode penipuan yang berbeda, disesuaikan dengan karakteristik korban yang menjadi target.

d. Bersifat Global

Sifat digital dari penipuan ini memungkinkan pelaku untuk menjangkau korban di berbagai lokasi tanpa batasan geografis. Hal ini menciptakan hambatan yurisdiksi dalam penyelidikan dan penuntutan.

2.4. Modus Dalam Penipuan Daring

Berdasarkan data penelitian dari Kepolisian Daerah Jawa Timur, terdapat berbagai modus operandi yang sering digunakan dalam penipuan daring, antara lain:

a. Penipuan Melalui Situs Web (Web Fraud)

Dalam modus ini, pelaku membuat situs web palsu yang dirancang untuk menyerupai toko daring atau layanan resmi. Situs tersebut menawarkan produk atau layanan dengan harga yang sangat menarik. Setelah korban melakukan pembayaran, barang atau layanan yang dijanjikan tidak pernah dikirimkan.

b. Email Penipuan (Email Fraud)

Modus ini melibatkan pengiriman email palsu yang seolah berasal dari institusi terpercaya, seperti bank atau perusahaan besar. Tujuannya adalah mencuri data pribadi korban, seperti informasi rekening atau kartu kredit.

c. Penipuan Melalui Telepon (Telephone Fraud)

Pelaku menghubungi korban dengan berpura-pura menjadi anggota keluarga atau pihak berwenang, seperti polisi atau petugas bea cukai. Mereka kemudian meminta korban untuk mentransfer sejumlah uang dengan alasan tertentu, seperti membayar denda atau menebus barang.

d. Pesan Singkat Palsu (SMS Fraud)

Dalam modus ini, pelaku mengirimkan pesan singkat massal kepada banyak orang. Pesan tersebut sering berisi pemberitahuan bahwa penerima memenangkan hadiah atau diminta mentransfer uang untuk tujuan tertentu.

e. Pencurian Data Kartu Kredit (Credit Card Fraud)

Pelaku mencuri informasi kartu kredit korban, termasuk nomor kartu, kode CVV, dan alamat penagihan, untuk melakukan transaksi ilegal.

2.5. Dampak Penipuan Daring

Penipuan daring memiliki dampak yang luas, baik secara ekonomi maupun psikologis. Kerugian ekonomi sering kali signifikan, terutama karena korban kehilangan uang yang sulit dikembalikan. Berdasarkan survei Consumer Security Risks 2016 yang dilakukan oleh Kaspersky Lab, 48% pengguna internet menjadi target penipuan daring, sementara 6% dari mereka kehilangan uang dengan rata-rata kerugian sebesar USD 283 atau sekitar Rp3,6 juta. Lebih dari itu, hanya 54% korban yang berhasil mendapatkan kembali uang mereka.

Dari sisi psikologis, korban penipuan daring seringkali merasa tertekan dan kehilangan kepercayaan pada sistem digital. Hal ini dapat berdampak jangka panjang terhadap adopsi teknologi, khususnya di kalangan masyarakat yang sebelumnya kurang memahami risiko digital.

2.6. Upaya Penanggulangan Penipuan Daring

Penanggulangan penipuan daring memerlukan pendekatan yang komprehensif. Selain penegakan hukum yang lebih kuat, diperlukan pula edukasi masyarakat tentang risiko dunia digital dan pentingnya menjaga keamanan data pribadi. Pemerintah dan institusi terkait juga perlu meningkatkan kemampuan deteksi dan penanganan kasus kejahatan siber. Dengan kolaborasi yang baik antara semua pihak, diharapkan tingkat kejahatan daring dapat diminimalkan secara signifikan.

2.7. WhatsApp sebagai Media Komunikasi Digital

2.7.1 Pengertian WhatsApp

WhatsApp adalah salah satu aplikasi pesan instan yang paling populer di dunia, dirancang untuk mendukung komunikasi digital yang cepat dan efisien melalui jaringan internet. Aplikasi ini memungkinkan pengguna untuk mengirim pesan teks, gambar, video, pesan suara, dokumen, dan melakukan panggilan suara maupun video secara gratis, tanpa menggunakan pulsa konvensional. Dalam kategori media sosial, WhatsApp dikelompokkan sebagai platform perpesanan instan yang menghubungkan individu maupun kelompok dalam komunikasi waktu nyata.

WhatsApp dianggap sebagai revolusi dalam dunia komunikasi digital. Dengan antarmuka yang sederhana, aplikasi ini mudah diakses oleh berbagai kalangan, termasuk mereka yang memiliki keterbatasan dalam teknologi. Selain untuk komunikasi pribadi, WhatsApp juga digunakan untuk keperluan bisnis dan pendidikan, menjadikannya alat yang multifungsi dalam kehidupan sehari-hari.

2.7.2 Sejarah WhatsApp

WhatsApp didirikan oleh Jan Koum dan Brian Acton pada tahun 2009. Sebelumnya, keduanya bekerja selama lebih dari 20 tahun di perusahaan teknologi ternama, Yahoo. Inspirasi untuk mengembangkan aplikasi ini muncul ketika Jan Koum menyadari potensi besar industri aplikasi yang sedang berkembang pesat melalui platform App Store yang baru dirilis.

Konsep awal WhatsApp cukup sederhana. Koum ingin menciptakan sebuah aplikasi yang dapat menampilkan status pengguna di sebelah nama mereka. Setelah berdiskusi dengan Brian Acton, mereka memutuskan untuk merealisasikan ide ini dengan melibatkan Alex Fishman, seorang teman mereka. Fishman kemudian memperkenalkan Koum kepada Igor Solomennikov, seorang pengembang iOS asal Rusia, untuk membantu dalam pengembangan teknis.

Pengembangan awal WhatsApp penuh tantangan. Versi pertama aplikasi ini mengalami berbagai masalah teknis, seperti konsumsi baterai yang tinggi dan seringnya aplikasi mengalami gangguan. Bahkan, beberapa teman Koum memberikan tanggapan negatif terhadap prototipe WhatsApp. Kondisi ini hampir membuat Koum menyerah, tetapi dorongan dan dukungan dari Brian Acton memotivasi Koum untuk melanjutkan proyek tersebut.

Pada tanggal 24 Februari 2009, WhatsApp Inc. resmi didirikan, dan aplikasi WhatsApp pertama kali diluncurkan untuk perangkat iOS. Seiring berjalannya waktu, aplikasi ini berkembang dengan fitur-fitur baru dan diperluas ke berbagai platform, termasuk Android. Pada tahun 2014, WhatsApp diakuisisi oleh Facebook (sekarang

Meta) dengan nilai transaksi sebesar $19 miliar, menjadikannya salah satu akuisisi teknologi terbesar dalam sejarah.

2.7.3 Keunggulan WhatsApp

WhatsApp memiliki sejumlah keunggulan yang menjadikannya platform komunikasi favorit di berbagai belahan dunia. Beberapa keunggulan utama WhatsApp meliputi:

a. Kemudahan Penggunaan

WhatsApp memiliki antarmuka yang sederhana dan mudah dipahami, bahkan oleh pengguna dengan tingkat literasi digital yang rendah. Proses pendaftaran yang hanya memerlukan nomor telepon juga mempermudah aksesibilitasnya.

b. Kompatibilitas Lintas Platform

Aplikasi ini mendukung berbagai perangkat, termasuk ponsel pintar berbasis Android dan iOS, serta versi desktop melalui WhatsApp Web, yang memungkinkan pengguna untuk tetap terhubung di berbagai perangkat.

c. Fitur Lengkap

WhatsApp menyediakan fitur-fitur komunikasi yang lengkap, seperti pesan teks, panggilan suara dan video, berbagi dokumen, lokasi, hingga fitur status yang mirip dengan media sosial lainnya.

d. Keamanan Data

Dengan fitur enkripsi ujung-ke-ujung (end-to-end encryption), WhatsApp memastikan bahwa pesan hanya dapat dibaca oleh pengirim dan penerima, sehingga memberikan rasa aman kepada penggunanya.

2.7.4 Risiko dan Bentuk Penipuan di WhatsApp

Sebagai salah satu aplikasi dengan jumlah pengguna terbesar di dunia, WhatsApp sering menjadi target kejahatan siber. Berdasarkan penelitian Wahyudin et al. (2024), beberapa bentuk penipuan yang sering terjadi di platform ini adalah:

a. Phishing

Phishing merupakan salah satu metode penipuan yang sering digunakan. Pelaku menyamar sebagai institusi resmi, seperti bank atau lembaga pemerintah, untuk mengelabui korban agar memberikan informasi pribadi mereka. Biasanya, pelaku mengirimkan tautan palsu yang mengarahkan korban ke situs web berbahaya untuk mencuri data sensitif.

b. Sniffing

Sniffing adalah metode pencurian data melalui jaringan internet. Pelaku memanfaatkan Wi-Fi publik yang tidak aman untuk mengakses informasi pengguna, termasuk kata sandi atau kode OTP.

c. Tautan Palsu

Penipu mengirimkan tautan palsu yang menjanjikan hadiah atau promo tertentu. Saat korban mengklik tautan tersebut, malware dapat terinstal secara otomatis, memungkinkan pelaku mencuri data atau mengontrol perangkat korban.

d. Permintaan Transfer Uang

Dalam modus ini, pelaku menyamar sebagai teman atau anggota keluarga korban dan meminta uang dengan alasan mendesak, seperti keadaan darurat.

2.7.5 Dampak Penipuan Daring melalui WhatsApp

Penipuan daring melalui WhatsApp memiliki dampak yang luas, mencakup aspek ekonomi, sosial, dan psikologis:

a. Dampak Ekonomi

Korban penipuan sering mengalami kerugian finansial yang signifikan. Menurut laporan Badan Pusat Statistik (BPS), kerugian akibat penipuan daring di Indonesia mencapai miliaran rupiah setiap tahunnya, dengan WhatsApp menjadi salah satu saluran utama.

b. Dampak Sosial

Penipuan dapat merusak hubungan sosial, terutama jika melibatkan identitas orang terdekat. Korban sering kehilangan kepercayaan pada orang lain dan menjadi lebih waspada, yang dapat memengaruhi interaksi sosial mereka.

c. Dampak Psikologis

Korban sering mengalami stres, kecemasan, dan trauma akibat penipuan. Kehilangan uang atau data pribadi dapat membuat korban merasa malu dan takut untuk kembali menggunakan teknologi digital.

Untuk mengurangi risiko ini, literasi digital menjadi sangat penting. Edukasi mengenai cara mengenali modus penipuan dan menjaga privasi data harus terus digencarkan. Dengan langkah ini, diharapkan masyarakat dapat lebih waspada dan terlindungi dari ancaman kejahatan siber.

2.8. Natural Language Processing (NLP)

2.8.1 Definisi Natural Language Processing (NLP)

Natural Language Processing (NLP) atau Pemrosesan Bahasa Alami adalah cabang dari kecerdasan buatan (Artificial Intelligence) yang berfokus pada pengolahan bahasa manusia sehingga dapat dimengerti oleh mesin. NLP bertujuan untuk menjembatani komunikasi antara manusia dan komputer melalui bahasa alami. Menurut Aries Muslim dan Robby Kurniawan (2023), NLP memainkan peran penting dalam memahami, menganalisis, dan memproses bahasa manusia, yang mencakup aspek sintaksis, semantik, hingga pragmatik untuk mendukung berbagai aplikasi, termasuk komunikasi dengan aktor virtual.

Dalam sistem berbasis NLP, komputer memanfaatkan aturan linguistik dan algoritma pembelajaran mesin untuk memproses bahasa. NLP digunakan untuk berbagai aplikasi, seperti analisis teks, deteksi anomali, penerjemahan otomatis, dan bahkan komunikasi dengan agen virtual dalam lingkungan interaktif. Teknologi ini memungkinkan komputer memahami makna di balik kata atau kalimat dalam konteks tertentu. Kemampuan ini sangat penting terutama dalam pengembangan aplikasi modern seperti chatbot, sistem rekomendasi, dan alat analisis data berbasis teks.

2.8.2 Komponen Utama NLP

Untuk mencapai tujuan dalam pemrosesan bahasa alami, beberapa komponen utama digunakan dalam sistem NLP:

a. Tokenisasi (Tokenization)

Proses memecah teks menjadi unit kecil seperti kata atau kalimat. Hal ini penting untuk memahami struktur teks secara lebih terperinci. Misalnya, dalam sebuah paragraf, tokenisasi memungkinkan sistem mengidentifikasi kata-kata individual yang relevan untuk analisis lebih lanjut.

b. Penghapusan Stop Word

Menghapus kata-kata yang tidak memiliki nilai informasi yang signifikan, seperti "dan", "atau", dan "yang". Langkah ini membantu mengurangi kompleksitas data tanpa menghilangkan makna inti dari teks yang dianalisis.

c. Stemming dan Lemmatization

Mengubah kata ke bentuk dasar untuk mempermudah proses analisis. Sebagai contoh, kata "berlari", "berlari-lari", dan "berlarian" diubah menjadi "lari". Teknik ini memastikan bahwa variasi bentuk kata tidak memengaruhi hasil analisis.

d. Named Entity Recognition (NER)

Identifikasi entitas penting dalam teks, seperti nama orang, lokasi, atau waktu. Dalam aplikasi praktis, NER digunakan untuk mengekstrak informasi spesifik yang relevan dari dokumen besar.

e. Part-of-Speech Tagging

Menentukan kategori gramatikal setiap kata dalam teks, seperti kata benda, kata kerja, atau kata sifat. Informasi ini membantu dalam memahami struktur sintaksis teks.

f. Parsing

Menganalisis struktur sintaksis kalimat untuk memahami hubungan antar elemen dalam teks. Parsing membantu sistem mengidentifikasi pola yang lebih kompleks dalam

data teks. Metode parsing seperti Tree-Adjoining Grammar (TAG) dan Tree-Furcating Grammar (TFG) telah digunakan dalam berbagai penelitian NLP.

2.8.3 NLP dalam Pengelolaan Data Teks

Salah satu aplikasi utama NLP adalah pengelolaan data teks secara efisien untuk mendukung analisis lebih lanjut. Pengelolaan ini mencakup transformasi teks mentah menjadi bentuk yang dapat dimengerti oleh algoritma komputer melalui tahapan seperti tokenisasi, stemming, dan penghapusan stop word. Pada tahap lanjutan, fitur-fitur seperti TF-IDF (Term Frequency-Inverse Document Frequency) digunakan untuk mengekstrak informasi penting yang relevan dari dokumen.

Pengelolaan data teks ini menjadi pondasi bagi berbagai aplikasi, seperti analisis sentimen, klasifikasi dokumen, dan deteksi topik. Contoh nyata penerapan ini adalah penggunaan NLP untuk mengekstrak data dari ulasan pelanggan atau percakapan dalam aplikasi pesan instan, yang dapat memberikan wawasan berharga untuk pengambilan keputusan. Dalam penelitian modern, NLP juga digunakan untuk menganalisis data dalam berbagai bahasa, memungkinkan penanganan konteks multibahasa secara efisien.

2.8.4 NLP dalam Deteksi Pesan Penipuan

Dalam konteks penelitian ini, NLP digunakan untuk mendeteksi dan memprediksi pesan penipuan pada platform WhatsApp. Proses ini melibatkan langkah-langkah berikut:

Preprocessing Data: Melakukan pembersihan teks, seperti menghapus simbol khusus, tanda baca, atau angka yang tidak relevan. Langkah ini juga mencakup tokenisasi dan stemming untuk menyederhanakan teks. Tahapan ini bertujuan mengurangi noise dalam data yang dapat mengganggu akurasi analisis.

a. Ekstraksi Fitur

Mengonversi teks menjadi representasi numerik menggunakan metode seperti bag-of-words atau Term Frequency-Inverse Document Frequency (TF-IDF). Representasi ini memberikan gambaran tentang frekuensi dan relevansi kata dalam dokumen.

b. Klasifikasi

Menggunakan model Naive Bayes, pesan diklasifikasikan berdasarkan probabilitasnya sebagai penipuan atau bukan. Naive Bayes dipilih karena kemampuannya dalam menangani data teks dengan efisiensi tinggi. Dalam penelitian ini, algoritma ini berperan penting untuk mengidentifikasi pola dalam pesan yang mengandung elemen penipuan.

2.8.5 Tantangan dan Peluang NLP

Meskipun NLP menawarkan banyak manfaat, ada beberapa tantangan yang perlu diatasi, terutama dalam konteks bahasa alami:

1. **Ambiguitas Bahasa**

Bahasa manusia sering kali ambigu, dengan makna yang bergantung pada konteks tertentu. Contoh ambiguitas dapat ditemukan pada kata yang memiliki arti ganda, seperti "bisa" yang dapat berarti kemampuan atau racun.

2. **Variasi Ekspresi**

Bahasa manusia kaya akan variasi, termasuk penggunaan slang, idiom, dan ekspresi regional, yang sulit dimodelkan secara komprehensif. Hal ini menjadi tantangan besar bagi sistem NLP yang dirancang untuk bekerja dalam lingkungan multikultural.

3. **Kebutuhan Data yang Besar**

Untuk melatih model NLP dengan akurasi tinggi, diperlukan data pelatihan dalam jumlah besar, yang tidak selalu mudah diperoleh. Selain itu, data yang digunakan harus mencakup berbagai skenario untuk meningkatkan kemampuan generalisasi model.

Namun, tantangan ini membuka peluang bagi pengembangan teknologi baru, seperti integrasi NLP dengan teknologi kecerdasan buatan lainnya, untuk meningkatkan kapabilitasnya. Dalam konteks deteksi pesan penipuan, NLP dapat dikombinasikan dengan algoritma pembelajaran mendalam untuk meningkatkan akurasi klasifikasi. Teknologi ini juga dapat diperluas untuk mendeteksi pola lain dalam data teks, seperti potensi ancaman keamanan atau tren sosial.

2.9. Algoritma Naïve Bayes

2.9.1. Definisi Algoritma Naïve Bayes

Algoritma Naïve Bayes adalah salah satu metode klasifikasi yang sederhana namun kuat, didasarkan pada penerapan Teorema Bayes. Algoritma ini sering digunakan dalam berbagai bidang, seperti analisis teks, deteksi penipuan, dan klasifikasi data. Algoritma ini mengasumsikan bahwa semua fitur dalam dataset saling independen, meskipun asumsi ini jarang sepenuhnya benar dalam praktiknya.

Dasar dari penggunaan Naïve Bayes dalam pemrograman adalah rumus Bayes, di mana peluang kejadian h sebagai D ditentukan oleh peluang D saat h terjadi, peluang h, dan peluang D. Persamaan yang digunakan dapat diungkapkan sebagai berikut :

$$P(h|D) = \frac{P(D|h)\,P(h)}{P(D)}$$

Keterangan :

h : Hipotesis data dari suatu kelas yang telah ditentukan

D : Data yang belum ada kelasnya

P(h) : Probabilitas dari suatu analisis

P(D) : Probabilitas dari data D

P(h|D) : Probabilitas data h berdasarkan kondisi D

P(D|h) : Probabilitas D berdasarkan hipotesis A

2.9.2. Prinsip Kerja Algoritma Naïve Bayes

Prinsip kerja algoritma ini melibatkan beberapa langkah utama:

- Menghitung Probabilitas Prior: Probabilitas awal dari setiap kelas dihitung berdasarkan distribusi data dalam dataset.
- Menghitung Likelihood: Probabilitas fitur-fitur dalam data x muncul pada setiap kelas dihitung. Dalam data teks, hal ini biasanya dilakukan dengan menghitung frekuensi kata-kata tertentu.

- Menghitung Probabilitas Posterior: Probabilitas posterior dihitung menggunakan Teorema Bayes dengan menggabungkan prior dan likelihood.
- Prediksi Kelas: Data x diklasifikasikan ke dalam kelas dengan probabilitas posterior tertinggi.

Sebagai contoh, dalam kasus deteksi spam, algoritma ini akan menghitung probabilitas sebuah email tergolong spam atau bukan spam berdasarkan kata-kata yang ada di dalamnya.

2.4.3. Keunggulan dan Kelemahan Algoritma Naïve Bayes

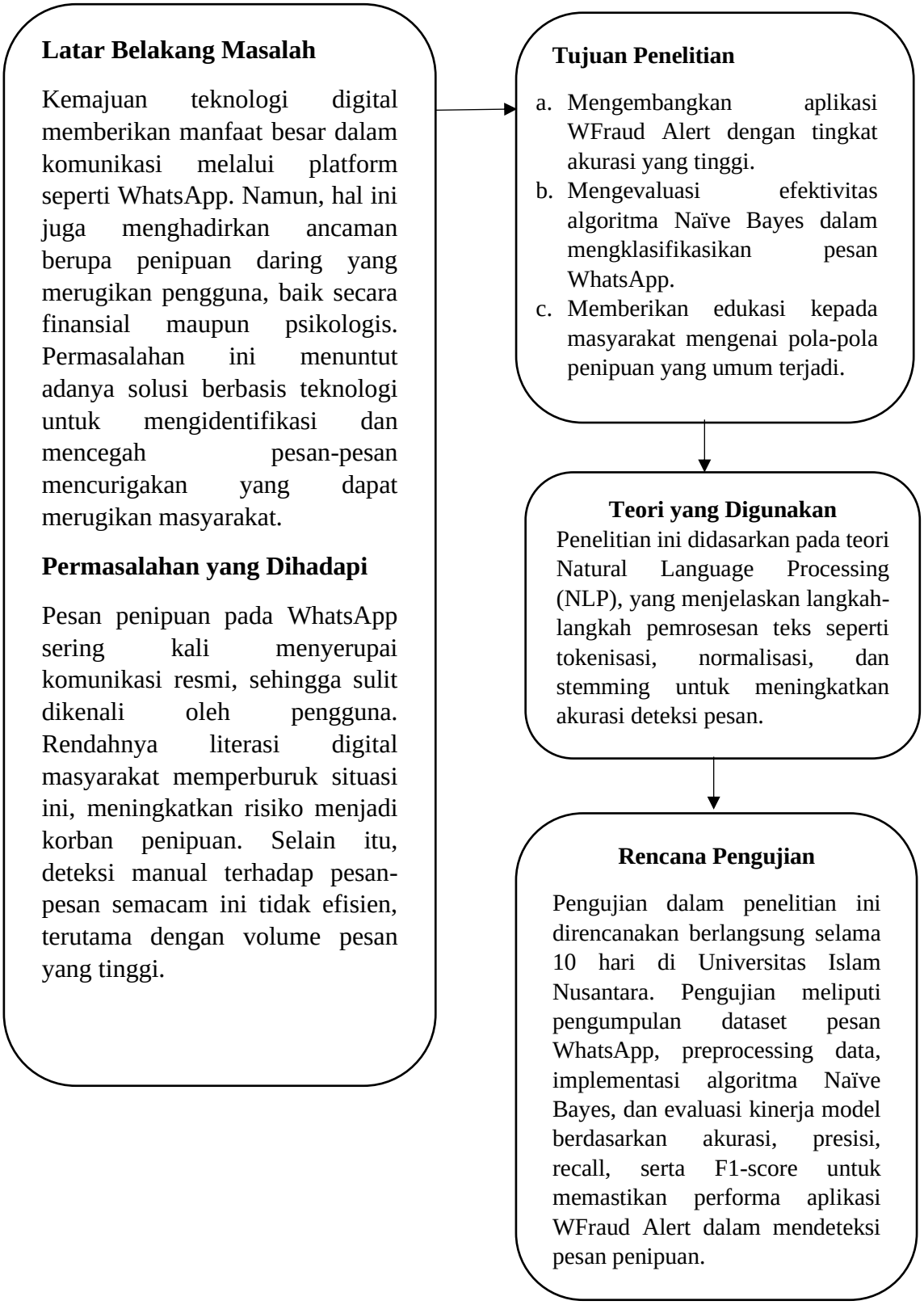
a. Keunggulan:

- Efisiensi: Algoritma ini memiliki waktu komputasi yang cepat, sehingga sangat cocok untuk dataset berukuran besar.
- Kesederhanaan: Implementasi algoritma ini relatif sederhana dibandingkan dengan metode lain.
- Kinerja yang Baik pada Data Teks: Algoritma ini banyak digunakan dalam klasifikasi teks, seperti analisis sentimen dan deteksi spam.

b. Kelemahan:

- Asumsi Independensi: Algoritma ini mengasumsikan bahwa fitur-fitur bersifat independen, yang seringkali tidak realistis dalam data dunia nyata.
- Sensitivitas terhadap Data Kosong: Jika suatu fitur tidak pernah muncul dalam data pelatihan untuk suatu kelas, probabilitasnya akan menjadi nol, yang dapat merusak prediksi. Masalah ini biasanya diatasi dengan smoothing, seperti Laplace smoothing.

2.10. Kerangka Berpikir



BAB 3 PERANCANGAN SISTEM DAN IMPLEMENTASI SISTEM

3.1. Metode Penelitian

3.1.1. Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif, yang bertujuan untuk mengembangkan dan menguji efektivitas algoritma Naïve Bayes dalam klasifikasi pesan WhatsApp. Pendekatan kuantitatif dipilih karena bersifat sistematis, terukur, dan berbasis pada data numerik, sehingga hasil yang diperoleh dapat diinterpretasikan dengan jelas. Pendekatan ini mendukung validasi hasil penelitian melalui analisis berbasis data, yang memungkinkan peneliti untuk mengevaluasi performa model yang dikembangkan dalam mendeteksi pola-pola tertentu dalam teks pesan WhatsApp.

Dalam penelitian ini, algoritma Naïve Bayes digunakan untuk mengklasifikasikan pesan WhatsApp ke dalam tiga kategori utama, yaitu pesan normal, pesan penipuan, dan pesan promosi judi online. Ketiga kategori ini dipilih berdasarkan relevansinya dengan kasus-kasus yang sering ditemukan dalam penggunaan WhatsApp sebagai platform komunikasi digital. Pendekatan ini memungkinkan penerapan algoritma pembelajaran mesin yang dapat memberikan hasil yang akurat meskipun data yang tersedia relatif terbatas.

Keunggulan pendekatan kuantitatif dalam penelitian ini terletak pada kemampuannya untuk mengolah data dalam jumlah besar secara efisien. Dengan demikian, penelitian ini tidak hanya berfokus pada pengembangan algoritma yang andal, tetapi juga pada bagaimana algoritma tersebut dapat diterapkan dalam skala yang lebih luas untuk memberikan solusi praktis bagi pengguna WhatsApp dalam mengenali pesan mencurigakan.

3.1.2. Jenis Penelitian

Jenis penelitian yang dilakukan adalah penelitian eksperimen terapan. Penelitian eksperimen terapan dipilih karena bertujuan untuk menghasilkan solusi praktis berupa aplikasi yang dapat membantu pengguna WhatsApp dalam mendeteksi pesan

mencurigakan. Dalam konteks ini, algoritma Naïve Bayes diterapkan untuk menyelesaikan masalah klasifikasi teks, yang merupakan inti dari penelitian ini.

Eksperimen dilakukan dengan menguji efektivitas algoritma Naïve Bayes dalam mengolah data teks yang telah dikategorikan sebelumnya. Pengujian dilakukan secara bertahap, mulai dari pengumpulan data, preprocessing, pelabelan, pelatihan model, hingga evaluasi hasil. Setiap tahap eksperimen dirancang untuk memastikan bahwa algoritma yang digunakan mampu mengklasifikasikan pesan WhatsApp dengan tingkat akurasi yang tinggi.

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi praktis dalam meningkatkan literasi digital masyarakat. Dengan adanya aplikasi WFraud Alert, pengguna WhatsApp diharapkan dapat lebih waspada terhadap pesan-pesan penipuan yang sering kali dirancang untuk menipu secara halus.

3.1.3 Populasi dan Sampel Penelitian

Populasi dalam penelitian ini mencakup seluruh pesan WhatsApp yang dapat dikelompokkan ke dalam tiga kategori utama, yaitu pesan normal, pesan penipuan, dan pesan promosi judi online. Pesan-pesan ini dipilih karena mencerminkan variasi konten yang sering muncul dalam komunikasi digital menggunakan WhatsApp.

Untuk memastikan hasil yang representatif, sampel penelitian dipilih menggunakan teknik purposive sampling. Teknik ini memungkinkan peneliti untuk memilih data yang sesuai dengan tujuan penelitian, yaitu melatih dan menguji model klasifikasi berbasis Naïve Bayes. Jumlah sampel yang digunakan dalam penelitian ini adalah 156 pesan, yang terdiri dari:

- 52 pesan normal, yang mewakili komunikasi biasa tanpa indikasi penipuan atau promosi.
- 52 pesan penipuan, yang berisi upaya manipulasi untuk mengambil keuntungan dari penerima pesan.
- 52 pesan promosi judi online, yang mencakup ajakan atau iklan terkait aktivitas perjudian.

Setiap kategori pesan dipilih untuk memastikan keberagaman data, sehingga model dapat mengenali pola-pola unik dari setiap jenis pesan. Proses pemilihan dan pelabelan data dilakukan secara manual untuk menjaga akurasi dan konsistensi.

3.1.4 Sumber Data

Data yang digunakan dalam penelitian ini adalah data primer, yang dikumpulkan langsung oleh peneliti dari berbagai sumber yang relevan. Data primer dipilih karena memberikan fleksibilitas dalam menentukan karakteristik pesan yang sesuai dengan tujuan penelitian.

Proses pengumpulan data dilakukan dengan hati-hati, mencakup pesan-pesan yang diterima secara pribadi maupun pesan yang diperoleh dari grup atau forum komunikasi digital. Data yang terkumpul kemudian diperiksa dan dikategorikan berdasarkan isi dan tujuan pesan. Kategori yang digunakan meliputi pesan normal, pesan penipuan, dan pesan promosi judi online.

Setelah dikategorikan, data disusun dalam format tabel menggunakan Microsoft Excel untuk mempermudah pengorganisasian. Data ini kemudian dikonversi ke dalam format CSV agar dapat diolah menggunakan perangkat lunak Python dalam tahap analisis lebih lanjut. Dengan cara ini, setiap pesan dapat diproses secara sistematis untuk menghasilkan model klasifikasi yang akurat.

Proses pelabelan data juga dilakukan secara manual oleh peneliti, yang memeriksa setiap pesan dengan cermat untuk memastikan bahwa label yang diberikan sesuai dengan karakteristik isi pesan. Hal ini dilakukan untuk menghindari kesalahan dalam pelatihan model, yang dapat memengaruhi performa algoritma Naïve Bayes.

3.1.5 Manfaat Pendekatan Penelitian

Pendekatan penelitian yang digunakan dalam penelitian ini memiliki beberapa manfaat utama. Pertama, pendekatan kuantitatif memungkinkan peneliti untuk menganalisis data secara objektif dan terukur. Kedua, pemilihan jenis penelitian eksperimen terapan memberikan kontribusi praktis berupa aplikasi yang dapat langsung digunakan oleh masyarakat. Ketiga, penggunaan data primer memastikan bahwa

penelitian ini relevan dengan kebutuhan nyata pengguna WhatsApp, yang sering kali menghadapi pesan-pesan mencurigakan dalam kehidupan sehari-hari.

Dengan mengikuti metode yang dirancang secara sistematis, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam bidang literasi digital dan keamanan siber. Hasil penelitian ini juga membuka peluang untuk pengembangan lebih lanjut, seperti penggunaan algoritma pembelajaran mesin lainnya atau penerapan pada dataset yang lebih besar.

3.1.6 Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data primer berupa pesan WhatsApp yang dikumpulkan langsung oleh peneliti. Proses pengumpulan data dilakukan menggunakan perangkat handphone milik peneliti, yang menjadi media utama untuk mengakses pesan-pesan yang relevan dengan kategori penelitian.

Proses pengumpulan data dimulai dengan mengidentifikasi pesan-pesan yang masuk di aplikasi WhatsApp. Pesan-pesan tersebut diperoleh dari berbagai sumber komunikasi, seperti pesan pribadi, grup, maupun broadcast yang diterima oleh peneliti. Pesan-pesan ini dipilih berdasarkan karakteristik tertentu yang mencerminkan salah satu dari tiga kategori utama, yaitu:

- Pesan Normal: Pesan yang tidak memiliki indikasi penipuan atau promosi, biasanya berisi komunikasi sehari-hari seperti salam, pertanyaan, atau informasi umum.
- Pesan Penipuan: Pesan yang didesain untuk menipu penerima dengan menggunakan taktik manipulatif, seperti permintaan uang mendesak, tautan berbahaya, atau informasi palsu yang mengatasnamakan lembaga tertentu.
- Pesan Promosi Judi Online: Pesan yang mengandung ajakan atau iklan terkait perjudian, biasanya dengan menawarkan keuntungan finansial yang tidak realistis.

Setelah pesan-pesan yang sesuai diidentifikasi, pesan-pesan tersebut direkapitulasi secara manual oleh peneliti. Setiap pesan dicatat dengan hati-hati, mencakup teks

lengkap, tanggal penerimaan, dan kategori pesan. Data yang terkumpul kemudian dipindahkan ke Microsoft Excel untuk proses pengorganisasian lebih lanjut.

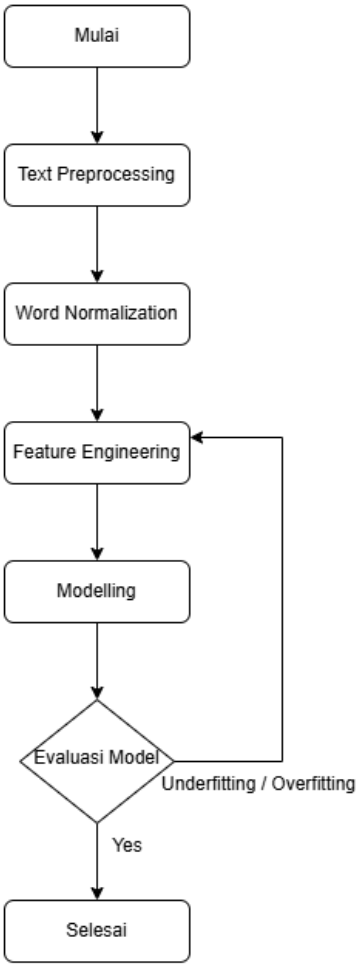
Pada tahap berikutnya, data yang telah direkap dikonversi ke dalam format CSV (Comma-Separated Values) untuk memudahkan proses pengolahan data menggunakan perangkat lunak Python. Format ini dipilih karena fleksibilitasnya dalam mendukung analisis data dengan berbagai library Python, seperti Pandas dan NumPy.

Selama proses pengumpulan data, peneliti juga memastikan bahwa pesan-pesan yang dikumpulkan adalah data nyata yang mencerminkan pola komunikasi yang sering ditemukan di WhatsApp. Hal ini penting untuk memastikan bahwa model yang dikembangkan memiliki relevansi praktis dan dapat diaplikasikan secara luas. Untuk menjaga keamanan dan privasi, semua informasi pribadi yang terdapat dalam pesan, seperti nama pengirim atau nomor telepon, dihapus atau disamarkan sebelum data digunakan dalam analisis.

Dengan metode pengumpulan data langsung dari handphone peneliti, penelitian ini dapat memastikan bahwa data yang digunakan benar-benar relevan dan mendukung tujuan penelitian, yaitu mengembangkan aplikasi yang mampu mendeteksi pesan penipuan di WhatsApp dengan akurasi tinggi.

3.2. Perencanaan Blok Diagram

Tahapan ini merupakan suatu pendekatan yang secara terperinci menjelaskan seluruh prosedur yang terlibat dalam pembuatan program yang bertujuan untuk melakukan prediksi pesan penipuan di platform Whatsapp. Proses ini melibatkan peneliti dalam mengumpulkan data yang relevan dan signifikan untuk kemudian diolah menjadi sebuah informasi yang dapat digunakan untuk pengklasifikasian pesan tersebut. Seluruh langkah-langkah dalam proses ini, sebagaimana diilustrasikan dalam Gambar 1 yang terdapat dalam flowchart penelitian.



Gambar 3.1. Flowchart Sistem

1. Mulai

Tahapan awal dari proses analisis dimulai dengan inisiasi proyek. Pada tahap ini, dilakukan penetapan tujuan analisis, termasuk menentukan apakah proses bertujuan untuk mendeteksi kecurangan dalam transaksi finansial, data teks, atau aktivitas lainnya. Tim juga mengidentifikasi sumber data yang diperlukan dan memastikan data yang dikumpulkan memenuhi kriteria kualitas, seperti akurasi, kelengkapan, dan relevansi.

Misalnya, jika data berasal dari transaksi, maka elemen-elemen seperti tanggal, nominal, jenis transaksi, dan identitas pengguna perlu diverifikasi terlebih dahulu.

2. Text Preprocessing

Text preprocessing merupakan langkah krusial untuk membersihkan data dari elemen-elemen yang tidak diperlukan. Data sering kali mengandung noise, seperti spasi berlebih, karakter asing, atau simbol yang tidak relevan untuk analisis. Dalam tahap ini, elemen-elemen tersebut dihapus atau diformat ulang agar data menjadi lebih terstruktur. Teknik-teknik seperti penghapusan stopwords, tokenisasi, dan stemming sering digunakan dalam tahap ini untuk mempersiapkan data teks yang bersih dan konsisten.

Contohnya, dalam data kecurangan berbasis ulasan pelanggan, preprocessing dapat membantu menghilangkan kata-kata umum seperti "dan," "atau," serta "yang" agar hanya kata-kata penting yang dianalisis lebih lanjut.

3. Word Normalization

Setelah teks diproses, dilakukan normalisasi kata. Proses ini bertujuan untuk menyamakan variasi ejaan atau bentuk kata yang berbeda tetapi memiliki arti yang sama. Sebagai contoh, dalam analisis data yang berbahasa Indonesia, kata-kata seperti “tdk” akan dinormalisasi menjadi “tidak.” Normalisasi ini bertujuan untuk mengurangi redundansi dalam data dan memastikan model dapat menangkap pola dengan lebih baik.

Normalisasi juga mencakup pengubahan kata-kata slang atau singkatan menjadi bentuk standar. Hal ini penting terutama dalam analisis data teks yang sering kali mengandung bahasa informal atau campuran.

4. Feature Engineering

Feature engineering merupakan inti dari proses analisis data. Dalam tahap ini, dilakukan transformasi data mentah menjadi fitur-fitur yang relevan dan bermakna bagi model. Proses ini bisa melibatkan pembuatan variabel baru berdasarkan data yang ada atau pemilihan atribut tertentu yang paling berpengaruh terhadap hasil analisis.

Sebagai contoh, dalam mendeteksi kecurangan kartu kredit, fitur seperti "jumlah transaksi dalam satu jam," "lokasi transaksi," atau "nominal transaksi di atas rata-rata" dapat menjadi indikator penting. Teknik seperti encoding untuk data kategorikal atau standardisasi untuk data numerik juga dilakukan untuk memastikan data siap digunakan oleh algoritma pemodelan.

5. Modeling

Pada tahap ini, data yang telah diproses dan direpresentasikan dalam bentuk fitur digunakan untuk membangun model. Pemilihan algoritma yang tepat menjadi sangat penting untuk menghasilkan model yang efektif. Algoritma yang sering digunakan dalam deteksi kecurangan mencakup metode supervised learning seperti Logistic Regression, Random Forest, dan Gradient Boosting, serta metode unsupervised seperti Clustering atau Autoencoder.

Proses ini juga mencakup pelatihan model menggunakan data historis yang mengandung pola kecurangan. Data ini digunakan untuk mengenali karakteristik anomali yang biasanya muncul dalam kasus kecurangan.

6. Evaluasi Model

Evaluasi model dilakukan untuk menilai seberapa baik model dalam mendeteksi kecurangan. Metode evaluasi ini mencakup pengukuran akurasi, precision, recall, dan F1-score untuk memahami keseimbangan antara deteksi kecurangan dan minimisasi kesalahan prediksi. Jika model menunjukkan indikasi underfitting atau overfitting, maka dilakukan revisi pada model atau pada data yang digunakan.

- Underfitting: Model tidak mampu menangkap pola dalam data, biasanya terjadi karena model terlalu sederhana atau data tidak cukup representatif.
- Overfitting: Model terlalu menyesuaikan dengan data pelatihan sehingga kurang efektif saat digunakan pada data baru.

Tahapan ini sangat penting untuk memastikan model tidak hanya optimal dalam pengujian tetapi juga dapat diandalkan dalam aplikasi nyata.

7. Iterasi

Jika hasil evaluasi menunjukkan bahwa model belum optimal, proses akan kembali ke tahap feature engineering atau bahkan preprocessing untuk memperbaiki kualitas data dan fitur. Iterasi ini dilakukan hingga model mampu mencapai performa yang diinginkan. Proses iterasi sering kali memakan waktu, tetapi merupakan langkah yang krusial untuk menghasilkan model yang benar-benar efektif.

8. Selesai

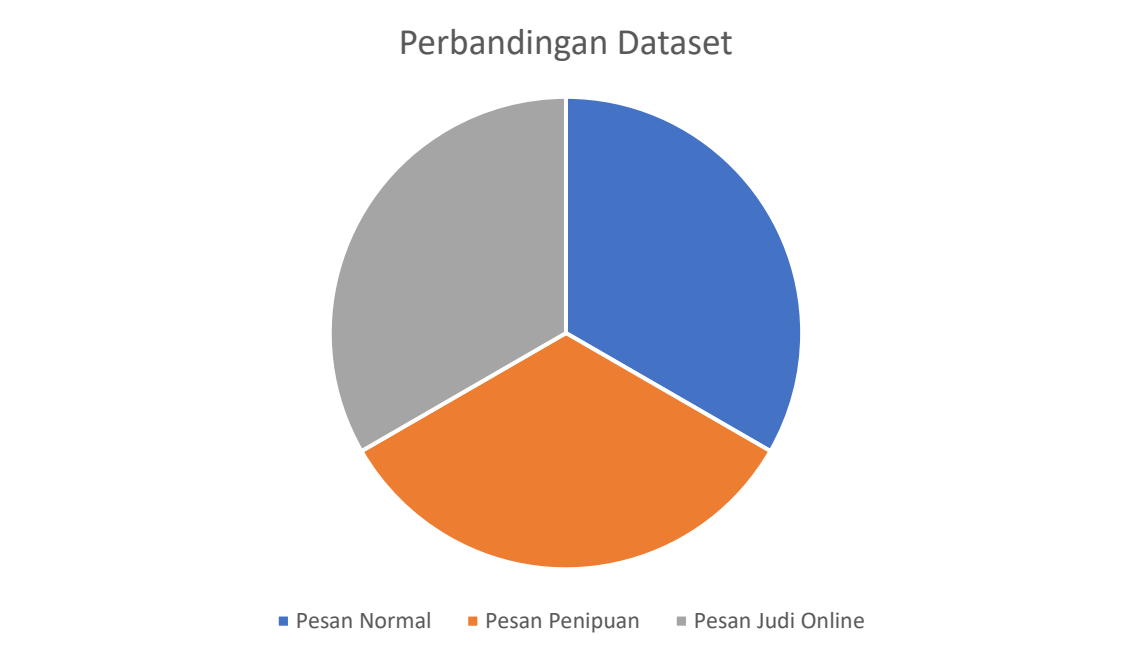
Setelah model dinyatakan memenuhi kriteria evaluasi, maka proses analisis dianggap selesai. Model ini kemudian dapat diimplementasikan dalam sistem untuk mendeteksi kecurangan secara otomatis pada data baru. Selain itu, hasil dari analisis ini dapat digunakan untuk membuat kebijakan yang lebih baik dalam mencegah kecurangan di masa mendatang.

Proses yang terstruktur seperti ini memastikan bahwa setiap langkah memiliki kontribusi terhadap keberhasilan proyek analisis. Dengan mengikuti tahapan ini, tim dapat menghasilkan model yang andal dan memberikan dampak nyata pada pengambilan keputusan.

BAB 4 PEMBAHASAN DAN ANALISIS

4.1. Dataset

Dataset yang digunakan dalam penelitian ini terdiri dari 156 pesan WhatsApp yang diklasifikasikan ke dalam tiga kategori utama: pesan normal, pesan promosi judi daring, dan pesan penipuan. Setiap kategori memiliki jumlah data yang seimbang, yaitu sebanyak 52 pesan per kategori. Data ini diperoleh langsung dari perangkat seluler peneliti melalui pengumpulan pesan yang secara aktif diterima selama kurun waktu tertentu. Penggunaan perangkat peneliti sebagai sumber data bertujuan untuk memastikan bahwa data yang digunakan relevan dengan konteks penelitian dan mencerminkan pola komunikasi yang nyata di platform WhatsApp.



Gambar 4.1. Perbandingan Dataset yang Digunakan

Kategori pesan normal mencakup berbagai jenis komunikasi sehari-hari yang tidak memiliki unsur kejahatan atau promosi ilegal. Contohnya termasuk pesan pribadi seperti ucapan selamat pagi, pesan dalam grup keluarga atau teman yang berisi diskusi ringan, hingga pemberitahuan rutin tanpa niat komersial atau manipulasi. Kategori promosi judi daring mencakup pesan-pesan yang dirancang untuk menarik perhatian

pengguna agar terlibat dalam aktivitas perjudian. Pesan-pesan dalam kategori ini sering kali berisi kata-kata promosi seperti “bonus besar” atau “peluang menang tinggi” disertai dengan tautan yang mengarahkan ke situs tertentu. Sementara itu, kategori pesan penipuan meliputi pesan-pesan dengan modus manipulasi psikologis, seperti permintaan transfer uang mendesak, pemberitahuan memenangkan hadiah palsu, atau tautan berbahaya yang dirancang untuk mencuri data pribadi pengguna.

Pesan Normal	Pesan Penipuan	Pesan Judi Online
Assalamualaikum kak kak bisa ngisi ulang galon sekarang?	Slamat"" Anda Terpilih Mendapatkan Cek Tunai Program 2021 PIN (479kl278) Unrtuk info Verifikasi Klik: bit.ly/eraswasembadapt	Situs betting On'line JOK3RB3T888 terbaik yang menyediakan jackpot terbesar hingga ratusan juta rupiah setiap harinya segera daftar di bit.ly/J0K3RBET888
Langsung bagi bagi aja ya sa pekerjaanya biar enak di zoom? Dari yg diatas itu?	INFORMASI TRAKHIR !!! N.O.Anda dpt Rp.175.000.000 dri program GIVE AWAY KED4S B34UTY Kode pin (CK805KU) Silahkan cocokkan pin anda di s.id/giveaway47	Bnyk promo main cardgames HKB SL0T-GAM3S terlengkap. Ada jg D1NG-D0NG T0-93L Sydney B0-L4 & C4-S1N0 SB0. Yuk d4ft4r P-K3RBOYA bit.ly/P-K3Rboya
Assalamualaikum Kak sorry ganggu malem2 gini. Ini penasaran aja sih . Gabut mau coba baca ulang buku di rak ini	Sya TKW Dri Hongkong Terlilit Htang Di Bantu Oleh Kyai H.Yusuf Alhamdulillah Htang Sya Terlunasi Smua Melalui	Bingung cari BO terpercaya? yuk gabung di jayapoker http://103.10.200.41 Nikmati Serunya permainan kartu yang

	Dana G4iB WA: +6285757647277	menarik yaitu : Ceme Texaspoker dan Domino QQ

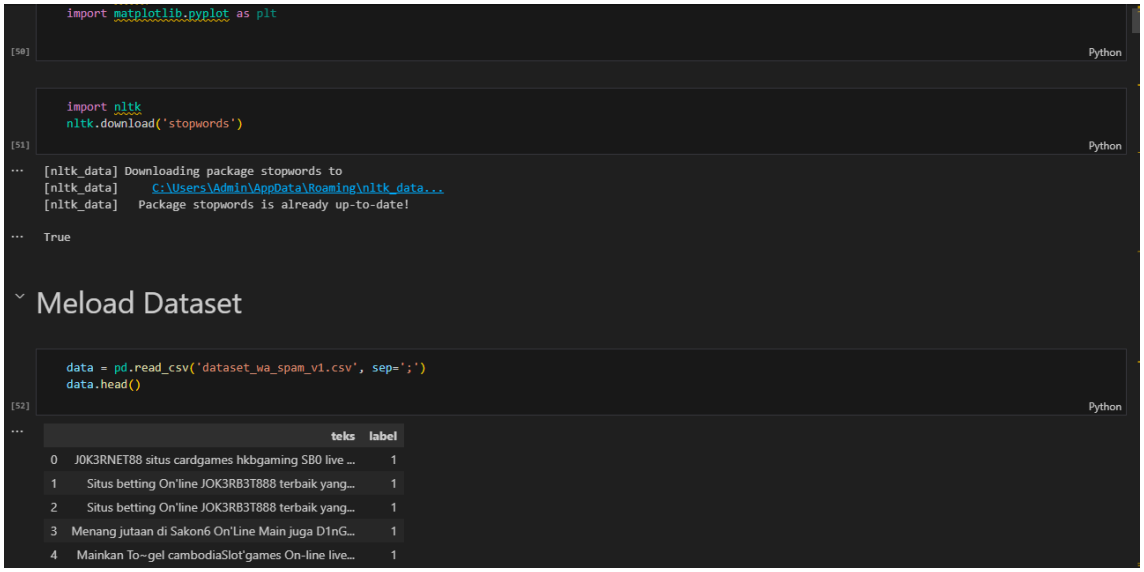
Tabel 4.1. Sampel Dataset yang digunakan

Data yang dikumpulkan melalui perangkat peneliti dipilih dengan hati-hati untuk memastikan keberagaman dalam setiap kategori. Seluruh pesan telah diverifikasi dan diseleksi secara manual untuk menghilangkan data yang ambigu atau tidak relevan. Selain itu, proses ini dilakukan dengan tetap mematuhi prinsip etika penelitian, di mana pesan yang digunakan tidak mengandung informasi sensitif atau pribadi dari individu lain.

Setelah data dikumpulkan, langkah awal yang dilakukan adalah memastikan format pesan seragam, sehingga memudahkan proses analisis. Dataset ini kemudian diproses melalui serangkaian tahap pra-pemrosesan untuk meningkatkan kualitas data, seperti case folding untuk menyamakan huruf, normalisasi kata untuk merapikan ejaan, penghapusan stopword untuk mengeliminasi kata-kata umum yang tidak relevan, dan stemming untuk mengembalikan kata-kata ke bentuk dasarnya.

Distribusi dataset yang seimbang di antara kategori pesan normal, promosi judi daring, dan penipuan merupakan langkah strategis untuk memastikan performa model yang optimal. Hal ini memungkinkan algoritma Naïve Bayes yang digunakan dalam penelitian ini dapat mengenali pola yang khas dari setiap kategori pesan dengan baik. Proses analisis ini diharapkan menghasilkan model klasifikasi yang mampu memberikan tingkat akurasi tinggi dalam mengidentifikasi jenis pesan WhatsApp.

Setelah data terkumpul, dataset ini diatur dalam format CSV untuk memudahkan proses analisis. Dataset kemudian di-load langsung ke dalam Visual Studio Code menggunakan pustaka Python, seperti pandas. Proses ini melibatkan pemanggilan file dataset melalui kode sederhana, seperti:



Gambar 4.2. Load Dataset yang digunakan

Penggunaan VS Code sebagai Integrated Development Environment (IDE) memudahkan peneliti untuk langsung memvisualisasikan data, mengidentifikasi potensi masalah seperti missing values atau duplikasi data, serta memastikan bahwa struktur dataset sesuai dengan kebutuhan analisis lebih lanjut. Langkah ini merupakan bagian awal dari tahapan pra-pemrosesan data, yang mencakup case folding, normalisasi kata, penghapusan stopword, dan stemming.

Selain itu, seluruh data telah diperiksa untuk memastikan tidak adanya informasi sensitif atau pribadi yang melanggar prinsip etika penelitian. Dataset juga dirancang untuk mendukung proses klasifikasi dengan distribusi kategori yang seimbang, sehingga model yang dibangun dapat mengenali pola karakteristik dari setiap jenis pesan secara optimal.

Proses pengolahan dataset yang terstruktur, mulai dari pengumpulan, validasi, hingga pengolahan awal di VS Code, memastikan bahwa data siap untuk digunakan dalam tahap pemodelan dengan algoritma Naïve Bayes. Dengan langkah-langkah ini, dataset yang dihasilkan menjadi salah satu elemen kunci dalam keberhasilan penelitian ini.

4.2. Case Folding

Case folding adalah salah satu tahap awal dalam proses pra-pemrosesan data teks yang sangat penting untuk menyederhanakan dan menyeragamkan format teks. Tujuan utama dari case folding adalah untuk menghilangkan perbedaan yang tidak relevan dalam teks, seperti penggunaan huruf kapital, simbol, angka, atau karakter non-alfabet lainnya, yang dapat mengganggu analisis dan menurunkan performa model klasifikasi.

Tahap ini sangat relevan dalam penelitian yang menggunakan teks, seperti pesan WhatsApp, karena pesan-pesan tersebut sering kali ditulis dengan format yang tidak standar. Dalam dunia nyata, pesan yang diterima cenderung bervariasi dalam penggunaan huruf besar, tanda baca, atau bahkan mengandung elemen-elemen tambahan seperti tautan dan angka yang tidak selalu memiliki nilai informasi signifikan dalam proses analisis teks.

Dataset penelitian ini berisi 156 pesan WhatsApp yang terdiri dari pesan normal, promosi judi daring, dan penipuan. Sebelum dilakukan case folding, teks pesan tersebut memiliki format yang tidak seragam, seperti contoh berikut:

Pesan Asli	*Case Folding*
MAM'PIR SK'RG JU'GA K MA'FIA CA'SH AD'A PR'OMO FR'EE CH'IP DE'PO 100 DPT 200 KA'KEK ZE'US L'G GA'COR B'GT DF'TR SK'RG bit.ly/3Y6MEV8	mampir skrg juga k mafia cash ada promo free chip depo dpt kakek zeus lg gacor bgt dftr skrg bitlyymev

Tabel 4.2. Proses case folding

Proses case folding dilakukan melalui beberapa langkah sistematis:

a. Mengubah huruf kapital menjadi huruf kecil:

Seluruh teks diubah menjadi huruf kecil untuk memastikan konsistensi dalam analisis. Contohnya, "MAM'PIR" diubah menjadi "mampir". Hal ini penting karena model analitik

memperlakukan "MAM'PIR" dan "mampir" sebagai entitas yang berbeda jika tidak dilakukan normalisasi.

b. Menghilangkan tanda baca dan simbol:

Tanda baca, seperti tanda petik ('), dihilangkan karena tidak memberikan nilai informasi yang signifikan untuk analisis. Sebagai contoh, "CA'SH" diubah menjadi "cash".

c. Menghapus angka:

Angka-angka, seperti "100" dan "200", dihapus dari teks. Angka-angka ini biasanya tidak relevan dalam proses analisis teks kecuali terdapat konteks khusus yang memerlukan perhatian terhadap numerik.

d. Membersihkan tautan dan elemen non-alfabet lainnya:

Tautan dalam teks, seperti "bit.ly/3Y6MEV8", disederhanakan menjadi "bitlyymev". Proses ini dilakukan untuk mengurangi kompleksitas data dan memastikan bahwa hanya informasi yang relevan yang dipertahankan.

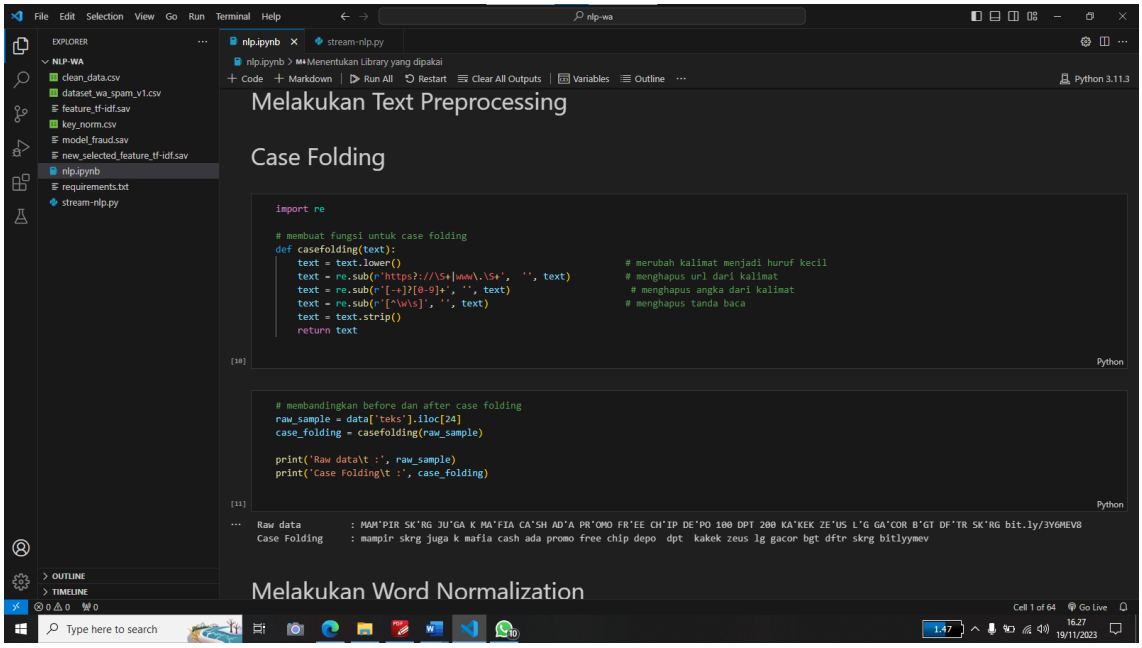
Proses case folding memberikan banyak manfaat dalam konteks penelitian ini:

- Penyederhanaan Data:

Dengan mengubah teks menjadi format yang seragam, case folding mengurangi kompleksitas data dan mempermudah algoritma untuk mengenali pola.

- Mengurangi Noise:

Tanda baca, simbol, dan angka yang tidak relevan dihapus, sehingga data lebih fokus pada informasi penting.



Gambar 4.3. Proses Case Folding

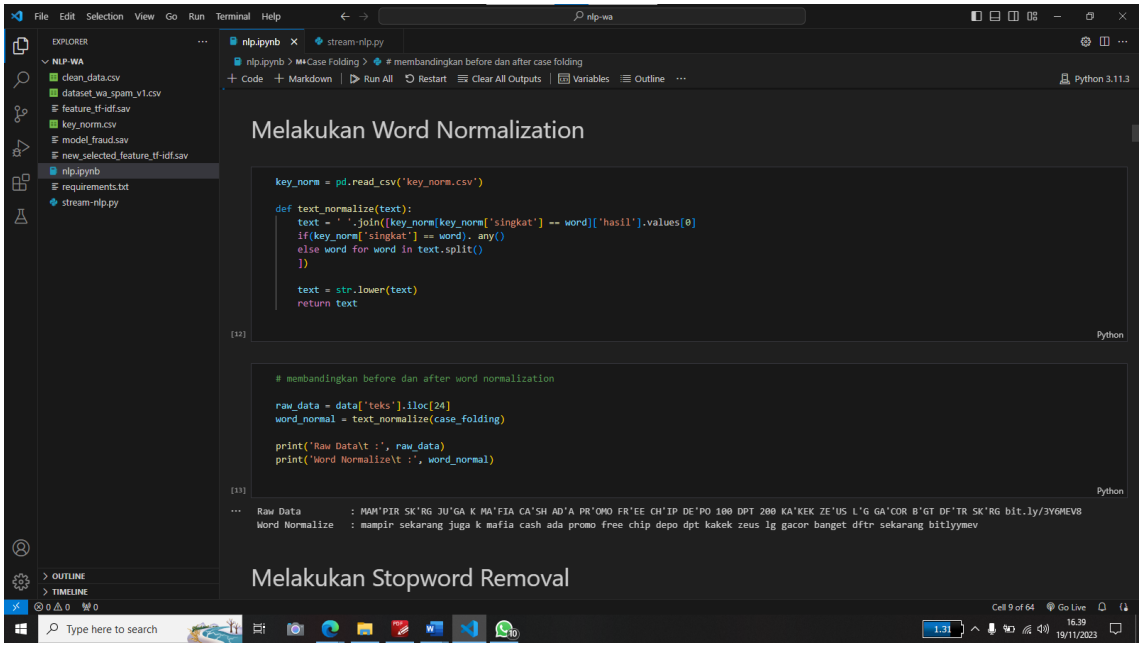
Tahapan ini merupakan langkah awal dari serangkaian proses pra-pemrosesan data. Setelah case folding, data yang telah seragam akan diproses lebih lanjut melalui tahapan seperti normalisasi kata, penghapusan stopword, dan stemming.

Dengan case folding, pesan seperti "MAM'PIR SK'RG JU'GA" dan "mampir skrg juga" akan diperlakukan sebagai teks yang sama oleh model klasifikasi, sehingga meningkatkan akurasi prediksi. Oleh karena itu, case folding tidak hanya menyederhanakan teks tetapi juga berperan penting dalam keberhasilan keseluruhan proses analisis data teks.

4.3. Word Normalization

Proses Word Normalization atau normalisasi kata merupakan salah satu tahapan penting dalam pengolahan data teks untuk memastikan bahwa teks yang digunakan dalam penelitian memiliki struktur yang konsisten dan dapat dipahami oleh mesin. Normalisasi kata bertujuan untuk mengubah kata-kata tidak baku, singkatan, dan istilah informal lainnya menjadi bentuk kata yang sesuai dengan standar bahasa baku. Langkah ini sangat relevan dalam penelitian berbasis pengolahan bahasa alami (Natural Language

Processing) karena data teks yang dihasilkan dari sumber seperti media sosial, pesan instan, atau percakapan sehari-hari sering kali mengandung variasi bahasa yang tidak terstandar.



Gambar 4.4. Proses Word Normalization

Pada penelitian ini, proses normalisasi kata dilakukan setelah tahap Case Folding, di mana seluruh teks telah diubah menjadi huruf kecil untuk memastikan keseragaman penulisan. Hal ini bertujuan untuk mengurangi keragaman yang disebabkan oleh penggunaan huruf besar atau kecil dalam teks, sehingga data menjadi lebih siap untuk diproses lebih lanjut. Setelah Case Folding, setiap kata pada data teks diperiksa untuk memastikan kesesuaiannya dengan aturan bahasa baku. Jika ditemukan kata-kata tidak baku, proses normalisasi dilakukan untuk menggantinya dengan kata baku yang sesuai.

4.3.1. Tujuan dan Manfaat Word Normalization

Tujuan utama dari proses Word Normalization adalah untuk meningkatkan kualitas data teks yang akan digunakan dalam analisis dan pemodelan. Data yang tidak dinormalisasi dapat menyebabkan model mengalami kesulitan dalam mengenali pola

yang ada, karena variasi dalam penulisan kata dapat memengaruhi hasil analisis. Oleh karena itu, normalisasi kata memberikan manfaat utama berikut:

a. Standarisasi Teks

Dengan mengubah kata-kata tidak baku menjadi kata baku, teks menjadi lebih seragam dan mudah dipahami, baik oleh manusia maupun mesin.

b. Pengurangan Variasi Data

Normalisasi mengurangi variasi yang tidak diperlukan dalam dataset, seperti penggunaan singkatan atau istilah slang.

c. Mempermudah Analisis Selanjutnya

Data yang sudah dinormalisasi lebih mudah digunakan untuk analisis statistik, visualisasi, atau pemodelan berbasis pembelajaran mesin.

Tahapan Word Normalization

Proses normalisasi kata pada penelitian ini dilakukan melalui beberapa tahapan berikut:

1. Tokenisasi

Data teks dipisahkan menjadi unit-unit kecil berupa kata-kata individu atau token untuk mempermudah proses identifikasi kata tidak baku.

2. Pencocokan dengan Kamus Normalisasi

Setiap token dicocokkan dengan daftar kata tidak baku yang telah disusun sebelumnya. Kamus normalisasi ini memuat pasangan kata tidak baku dan kata baku yang relevan.

3. Penggantian Kata Tidak Baku

Kata tidak baku yang ditemukan akan diganti dengan kata baku yang sesuai. Jika sebuah kata tidak ditemukan dalam kamus, kata tersebut akan dipertahankan dan dianalisis lebih lanjut.

4.3.2. Contoh Proses Word Normalization

Dataset yang digunakan dalam penelitian ini terdiri dari berbagai jenis teks yang diambil dari sumber pesan instan, seperti WhatsApp. Contoh pesan yang dianalisis adalah sebagai berikut:

Case Folding	*Word Normalization*
mampir skrg juga k mafia cash ada promo free chip depo dpt kakek zeus lg gacor bgt dftr skrg bitlyymev	mampir sekarang juga k mafia cash ada promo free chip depo dapat kakek zeus lagi gacor banget daftar sekarang bitlyymev

Tabel 4.3. Word Normalization

Pada tabel di atas, terlihat bahwa kata-kata tidak baku seperti "skrg," "k," "depo," "bgt," dan "dftr" diubah menjadi kata baku seperti "sekarang," "ke," "deposit," "banget," dan "daftar." Normalisasi ini dilakukan untuk memastikan bahwa data teks yang diolah **konsisten dengan aturan bahasa baku.**

4.3.3. Hasil dan Analisis Word Normalization

Proses normalisasi berhasil meningkatkan kualitas data teks dengan menghilangkan unsur-unsur tidak baku yang dapat memengaruhi hasil analisis. Dengan normalisasi yang baik, teks menjadi lebih bersih dan siap digunakan untuk tahapan analisis lebih lanjut, seperti pembuatan model atau evaluasi data. Beberapa manfaat yang diperoleh dari proses ini meliputi:

1. Meningkatkan Kesesuaian Data

Data teks yang dinormalisasi lebih sesuai dengan standar bahasa, sehingga mempermudah proses klasifikasi atau analisis berbasis pembelajaran mesin.

2. Mengurangi Noise pada Data

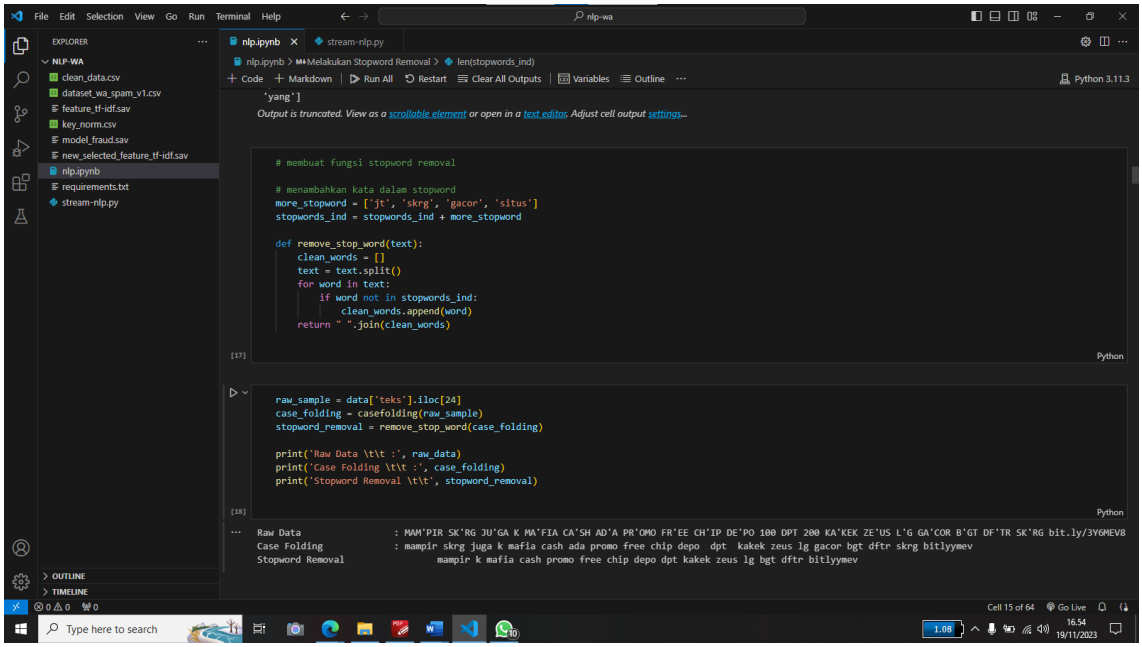
Kata-kata tidak baku yang dapat mengganggu proses analisis telah berhasil dihilangkan atau digantikan dengan kata baku yang relevan.

3. Tantangan pada Normalisasi

Salah satu tantangan utama adalah menangani kata-kata baru atau istilah slang yang tidak terdapat dalam kamus normalisasi. Oleh karena itu, diperlukan pembaruan kamus secara berkala untuk menjaga akurasi proses normalisasi.

4.4. Stopword Removal

Stopword Removal adalah salah satu tahap penting dalam pengolahan teks pada Natural Language Processing (NLP). Stopword merujuk pada kata-kata umum yang sering digunakan dalam bahasa sehari-hari namun memiliki makna yang minim atau bahkan tidak signifikan untuk tujuan analisis data teks. Contohnya termasuk kata-kata seperti "dan," "ke," "di," "adalah," dan lain-lain. Dalam konteks analisis data teks, stopword sering kali dianggap sebagai "noise" atau gangguan yang dapat mengurangi efisiensi pemrosesan data serta memengaruhi hasil analisis. Oleh karena itu, penghapusan stopwords adalah langkah penting untuk meningkatkan kualitas data yang akan dianalisis.



Gambar 4.4. Stopword Removal

Proses Stopword Removal dilakukan dengan menghapus kata-kata yang dianggap tidak relevan untuk menganalisis pola atau informasi dalam teks. Kata-kata yang dihapus ini biasanya tidak berkontribusi terhadap interpretasi semantik utama dari teks dan lebih sering digunakan untuk membentuk struktur gramatikal dalam kalimat. Dengan menghapus stopwords, teks menjadi lebih ringkas, fokus, dan siap untuk diproses lebih lanjut, misalnya dalam tahap pemodelan atau analisis statistik.

4.4.1. Manfaat Stopword Removal

Penghapusan stopwords memberikan banyak manfaat, antara lain:

1. Meningkatkan Efisiensi Proses: Dengan mengurangi jumlah kata yang perlu dianalisis, algoritma dapat bekerja lebih cepat dan efisien.
2. Menyederhanakan Teks: Teks yang telah dihapus stopwords menjadi lebih bersih dan terfokus pada informasi utama.
3. Mengurangi Dimensi Data: Stopwords sering kali mengisi sebagian besar dari data teks, sehingga penghapusan stopwords membantu mengurangi dimensi data secara signifikan.
4. Mempermudah Visualisasi dan Interpretasi: Teks yang lebih ringkas mempermudah proses interpretasi manusia maupun algoritma untuk memahami konteks utama dari pesan.

4.4.2. Langkah-Langkah Proses Stopword Removal

Tahapan dalam Stopword Removal mencakup beberapa proses yang sistematis:

a. Tokenisasi Teks: Teks pertama-tama dipecah menjadi unit-unit terkecil berupa kata-kata atau token. Proses tokenisasi memastikan bahwa setiap kata dapat dianalisis secara terpisah.
b. Identifikasi Stopwords: Setiap token dibandingkan dengan daftar kata yang termasuk dalam kategori stopwords. Daftar ini biasanya telah didefinisikan sebelumnya dan mencakup kata-kata seperti "dan," "yang," "ke," "dari," dan lain-lain.

c. Penghapusan Stopwords: Token yang cocok dengan daftar stopwords akan dihapus dari teks. Kata-kata lain yang tidak termasuk dalam daftar tetap dipertahankan.

Pada penelitian ini, daftar stopwords disesuaikan dengan konteks dataset yang digunakan. Penyesuaian ini dilakukan untuk memastikan bahwa hanya kata-kata yang benar-benar tidak relevan yang dihapus, sedangkan kata-kata penting tetap dipertahankan untuk keperluan analisis.

4.4.3. Contoh Proses Stopword Removal

Proses ini dapat dijelaskan lebih lanjut melalui sebuah contoh konkret. Berikut adalah pesan teks yang digunakan dalam penelitian ini:

Pesan Asli	*Filtering*
MAM'PIR SK'RG JU'GA K MA'FIA CA'SH AD'A PR'OMO FR'EE CH'IP DE'PO 100 DPT 200 KA'KEK ZE'US L'G GA'COR B'GT DF'TR SK'RG bit.ly/3Y6MEV8	mampir k mafia cash promo free chip depo dpt kakek zeus lg bgt dftr bitlyymev

Tabel 4.4. Stopword Removal

Pada tabel di atas, kata-kata seperti "k" dan "ada" dihapus karena termasuk dalam daftar stopwords yang telah ditentukan sebelumnya. Kata-kata ini tidak memiliki makna penting dalam konteks analisis pesan dan dianggap tidak relevan untuk proses pemodelan selanjutnya.

4.4.4. Analisis Hasil Stopword Removal

Setelah melalui proses Stopword Removal, pesan yang dihasilkan menjadi lebih ringkas dan relevan untuk analisis lebih lanjut. Kata-kata yang tetap dipertahankan, seperti "promo," "chip," "depo," dan "gacor," mencerminkan informasi penting yang dapat digunakan untuk mengidentifikasi pola dalam teks.

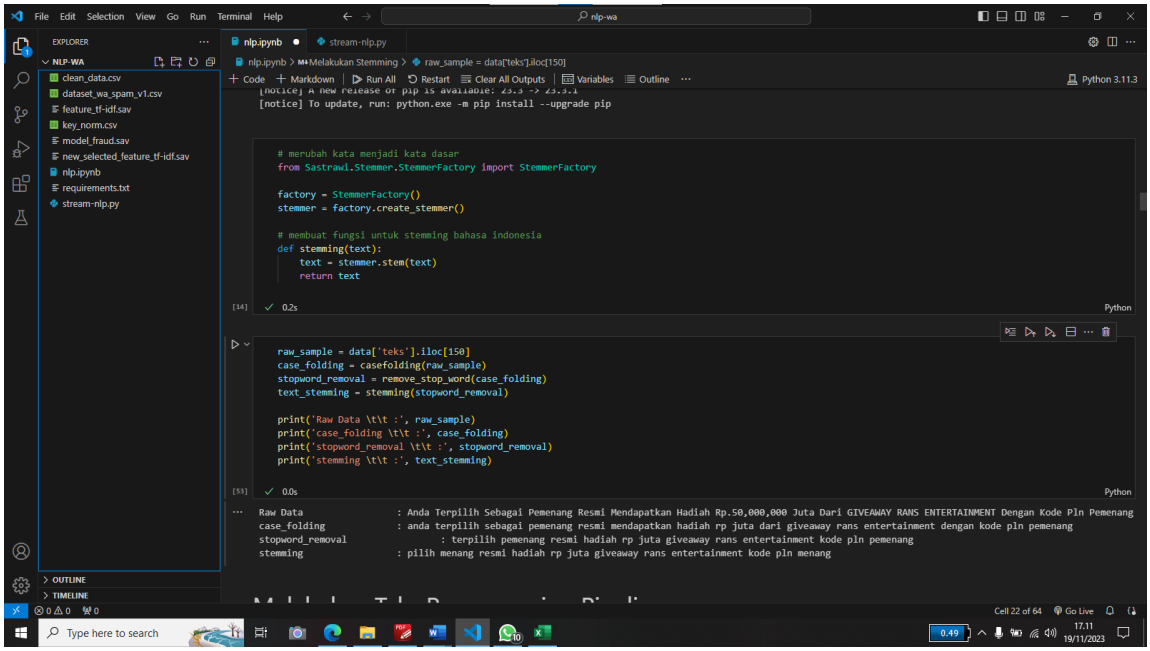
4.4.5. Tantangan Stopword Removal

Beberapa tantangan yang dihadapi dalam proses Stopword Removal adalah:

- Penyesuaian Daftar Stopwords: Tidak semua daftar stopwords dapat digunakan secara universal. Setiap dataset memiliki karakteristik unik yang membutuhkan penyesuaian daftar stopwords.
- Konteks Pesan: Kata-kata tertentu yang biasanya dianggap stopwords bisa menjadi penting dalam konteks tertentu, sehingga membutuhkan evaluasi lebih lanjut sebelum dihapus.
- Kata yang Mirip: Beberapa kata mungkin memiliki bentuk mirip dengan stopwords namun memiliki arti berbeda, sehingga diperlukan teknik tambahan untuk menghindari penghapusan yang tidak sengaja.

4.5. Stemming

Stemming adalah proses dalam Natural Language Processing (NLP) yang bertujuan mengubah kata turunan menjadi bentuk dasarnya atau akar kata (root word). Proses ini penting untuk mengurangi variasi kata yang muncul akibat penggunaan imbuhan dalam bahasa Indonesia, seperti awalan (prefix), sisipan (infix), akhiran (suffix), maupun gabungan dari ketiganya.



Gambar 4.5. Proses Stemming

Dalam analisis teks, bentuk turunan sering kali menyebabkan redundansi data dan dapat mengurangi efisiensi pemrosesan. Dengan menggunakan teknik stemming, kata-kata seperti "menang," "memilih," dan "pemenang" diubah menjadi akar katanya, yaitu "pilih" dan "menang." Hal ini memungkinkan analisis data teks yang lebih efisien dan konsisten.

4.5.1. Manfaat Stemming

1. Mengurangi Variasi Kata:

Mengelompokkan berbagai bentuk kata turunan ke dalam akar kata yang sama.

2. Efisiensi Proses Komputasi:

Dengan menyederhanakan data, waktu dan sumber daya yang diperlukan untuk analisis menjadi lebih efisien.

3. Mempermudah Interpretasi Data:

Teks yang telah distem menjadi lebih sederhana dan memungkinkan algoritma atau manusia memahami pola lebih mudah.

4.5.2. Proses Stemming

Untuk menggambarkan proses stemming, berikut adalah contoh yang diambil dari data pesan teks:

Pesan Asli	Teks Setelah Case Folding	Teks Setelah Stemming
Anda Terpilih Sebagai Pemenang Resmi Mendapatkan Hadiah Rp.50,000,000 Juta Dari GIVEAWAY RANS ENTERTAINMENT Dengan Kode Pln Pemenang	anda terpilih sebagai pemenang resmi mendapatkan hadiah rp juta dari giveaway rans entertainment dengan kode pln pemenang	pilih menang resmi hadiah rp juta giveaway rans entertainment kode pln menang

Tabel 4.5 Proses Stemming

Penjelasan proses:

a. Pesan Asli:

Pesan awal berisi teks dalam format aslinya, termasuk huruf kapital, tanda baca, dan struktur kalimat yang kompleks.

b. Case Folding:

Teks diubah menjadi huruf kecil seluruhnya untuk memastikan konsistensi.

c. Stemming:

- Kata "terpilih" diubah menjadi "pilih."
- Kata "pemenang" menjadi "menang."
- Kata "mendapatkan" menjadi "dapat."

4.5.3. Langkah-Langkah Penerapan Stemming

1. Identifikasi Pola Kata:

Sistem mengenali pola kata yang mengandung imbuhan.

2. Penghapusan Imbuhan:

Menghilangkan awalan, akhiran, atau sisipan untuk menemukan akar kata.

3. Validasi Kata Dasar:

Setelah penghapusan imbuhan, hasilnya divalidasi dengan kamus bahasa untuk memastikan bahwa kata yang dihasilkan benar-benar memiliki makna.

4.5.4. Studi Kasus Lainnya

Proses stemming juga diterapkan pada pesan-pesan berikut untuk menunjukkan fleksibilitas teknik ini:

Pesan Asli	Teks Setelah Stemming
Selamat! Anda Mendapatkan Kesempatan Eksklusif Untuk Menjadi Bagian Dari Program Kami Secara Gratis!	selamat dapat kesempatan eksklusif jadi bagian program gratis
Kami Mengundang Anda Untuk Bergabung Bersama Kami Dalam Acara Spesial Akhir Tahun Ini, Segera Daftar Sekarang!	undang bergabung acara spesial akhir tahun daftar sekarang
Promo Khusus Bulan Ini! Dapatkan Diskon Hingga 50% Untuk Semua Produk Kami!	promo khusus bulan dapat diskon hingga produk

Tabel 4.6. Contoh Lain Stemming

4.5.5. Pentingnya Stemming dalam Analisis Teks

Dengan mengurangi variasi bentuk kata, stemming mempermudah proses pemodelan data dan analisis semantik. Sebagai contoh:

- Dalam sistem pencarian (search engine), stemming memungkinkan hasil pencarian yang lebih relevan meskipun kata kunci yang digunakan dalam pencarian berbeda bentuk.

- Dalam klasifikasi teks, model dapat lebih fokus pada kata-kata penting dan mengabaikan variasi turunan yang tidak signifikan.

4.6. Modelling

Tahap modelling merupakan salah satu langkah paling krusial dalam pengembangan sistem WFraud Alert. Tahap ini tidak hanya berperan sebagai dasar teknis dalam menciptakan model klasifikasi, tetapi juga menjadi landasan utama untuk memastikan bahwa sistem dapat menjalankan fungsinya dengan efisien dan akurat. Tujuan utama dari tahap ini adalah membangun model prediktif yang mampu mengklasifikasikan pesan WhatsApp ke dalam tiga kategori utama, yaitu:

- Pesan Normal: Pesan yang tidak mengandung unsur mencurigakan, seperti percakapan pribadi atau pesan biasa dari pengguna.
- Pesan Penipuan: Pesan yang bertujuan untuk menipu penerima, misalnya dengan iming-iming hadiah palsu, meminta data pribadi, atau mengarahkan penerima untuk melakukan tindakan tertentu.
- Pesan Promosi Judi Online: Pesan yang mempromosikan aktivitas perjudian melalui media daring, sering kali menggunakan bahasa persuasif untuk menarik perhatian penerima.

Melalui pendekatan ini, sistem WFraud Alert diharapkan dapat memberikan manfaat praktis dalam mendeteksi dan mengantisipasi potensi ancaman yang mungkin tersembunyi dalam pesan-pesan tersebut. Sistem ini dirancang untuk berfungsi sebagai alat pencegah awal terhadap kejahatan siber yang semakin meningkat melalui media komunikasi digital.

4.6.1. Tahapan Modelling dan Dataset yang Digunakan

Dataset yang digunakan dalam penelitian ini terdiri dari 156 pesan WhatsApp yang telah dikategorikan secara manual ke dalam tiga kelas utama. Pesan normal mencakup pesan-pesan biasa yang tidak memiliki indikasi penipuan atau promosi judi daring. Contohnya adalah pesan pribadi antar pengguna, seperti “Halo, apa kabar?” atau “Jangan lupa besok ada rapat.” Pesan penipuan mencakup teks yang dirancang untuk

menipu penerima. Pesan jenis ini sering kali menggunakan pola bahasa yang manipulatif, seperti “Selamat Anda mendapatkan hadiah Rp50 juta, klik tautan berikut untuk klaim hadiah Anda.” Sementara itu, pesan promosi judi daring mengandung konten yang mempromosikan aktivitas perjudian melalui media digital, seperti “Daftar sekarang dan dapatkan bonus 100% di situs kami, kakek Zeus lagi gacor!”

Dataset ini telah dikategorikan secara manual oleh ahli untuk memastikan validitas data. Kategori ini penting untuk membantu model memahami pola dan karakteristik unik dari setiap jenis pesan.

Untuk memastikan bahwa model dapat belajar dengan optimal, dataset dibagi menjadi dua kelompok menggunakan metode stratified sampling. Teknik ini memastikan distribusi kategori dalam data latih dan data uji tetap seimbang. Sebanyak 80% dari total dataset dialokasikan untuk pelatihan (training data), sementara 20% sisanya digunakan untuk pengujian (testing data). Dengan distribusi ini, data latih terdiri dari 124 pesan (41 pesan per kategori), sementara data uji mencakup 32 pesan (11 pesan per kategori).

4.6.2. Pemilihan Algoritma dan Metode Transformasi Data

Algoritma yang digunakan dalam penelitian ini adalah Multinomial Naive Bayes. Algoritma ini dipilih karena kemampuannya yang andal dalam menangani tugas klasifikasi teks dengan efisiensi yang tinggi. Multinomial Naive Bayes bekerja berdasarkan prinsip probabilistik, di mana algoritma menghitung kemungkinan suatu pesan termasuk dalam kategori tertentu berdasarkan pola distribusi kata dalam data latih.

Keunggulan utama algoritma ini terletak pada:

- Kesederhanaan Implementasi: Algoritma ini mudah diterapkan, bahkan untuk dataset dengan ukuran yang relatif kecil.
- Efisiensi Komputasi: Multinomial Naive Bayes dikenal cepat dalam proses pelatihan dan prediksi, menjadikannya ideal untuk tugas klasifikasi teks.
- Kemampuan Generalisasi: Model ini mampu mengenali pola-pola umum dalam data, bahkan jika terdapat variasi penulisan dalam teks.

Agar data teks dapat diolah oleh algoritma, setiap pesan diubah menjadi representasi numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF merupakan teknik penghitungan bobot kata yang memberikan nilai lebih tinggi untuk kata-kata yang relevan dalam suatu kategori dan mengurangi bobot kata-kata yang terlalu sering muncul di seluruh kategori. Dengan demikian, metode ini membantu model untuk fokus pada kata-kata yang memiliki makna signifikan untuk klasifikasi.

Kata-kata yang sering muncul dalam kategori tertentu, seperti "promo" dalam pesan promosi judi daring atau "hadiah" dalam pesan penipuan, diberi bobot lebih tinggi. Sebaliknya, kata-kata umum seperti "dan" atau "ke" diberi bobot lebih rendah karena kemunculannya yang konsisten di seluruh kategori.

4.6.3. Proses Pelatihan Model

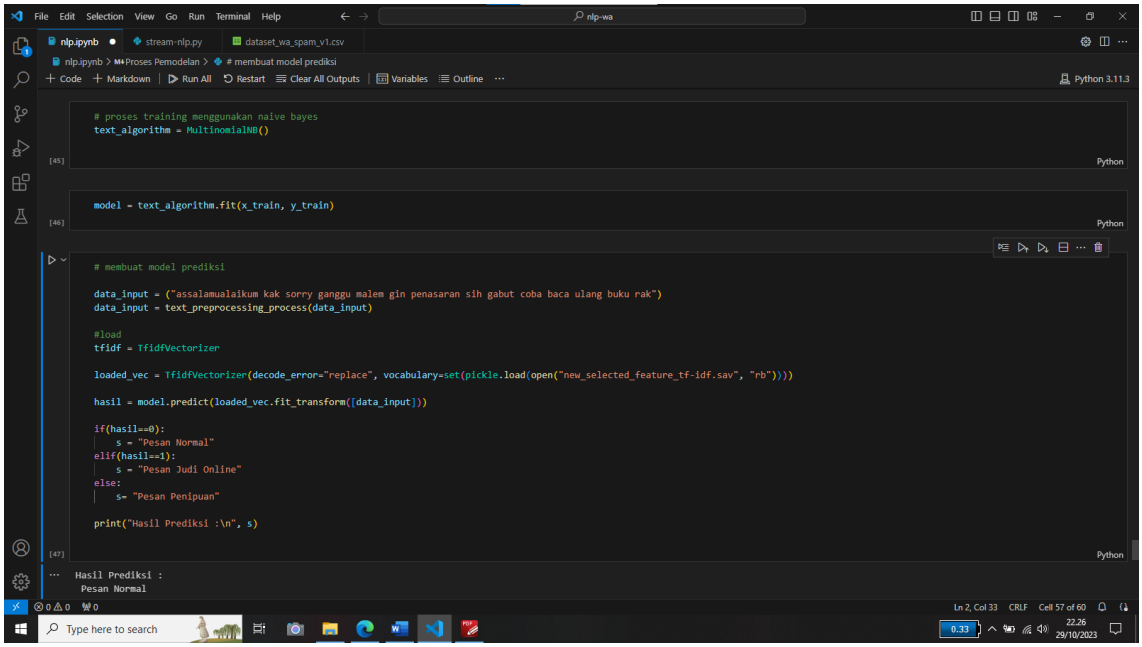
Tahap pelatihan model dimulai dengan memasukkan data latih ke dalam algoritma Multinomial Naive Bayes. Pada proses ini, model mempelajari pola distribusi kata-kata dalam setiap kategori pesan dan menghitung probabilitas kemunculan kata-kata tertentu berdasarkan frekuensi dalam data latih. Model kemudian menggunakan probabilitas ini untuk membangun aturan klasifikasi yang dapat memprediksi kategori pesan baru dengan akurat.

Sebagai ilustrasi, jika sebuah kata seperti “promo” muncul dengan frekuensi tinggi dalam pesan promosi judi daring tetapi jarang ditemukan dalam pesan normal atau pesan penipuan, maka model akan memberikan bobot probabilitas yang lebih tinggi untuk kategori "Pesan Promosi Judi Daring." Sebaliknya, kata-kata seperti "mendapatkan" mungkin memiliki probabilitas tinggi dalam pesan penipuan karena sering digunakan untuk menarik perhatian penerima.

Model yang dihasilkan dari tahap pelatihan ini dirancang untuk mengenali pola-pola unik dalam teks dan menggunakannya untuk membuat prediksi yang akurat tentang kategori pesan. Proses ini merupakan langkah awal yang penting sebelum model diuji dan diimplementasikan untuk prediksi pesan baru.

4.7. Evaluasi Model

Evaluasi model merupakan tahap penting dalam penelitian ini untuk menilai kinerja model yang diterapkan pada aplikasi WFraud Alert. Tujuan dari evaluasi ini adalah untuk memastikan bahwa model dapat mengklasifikasikan pesan WhatsApp secara akurat dan konsisten ke dalam tiga kategori utama, yaitu pesan normal, pesan penipuan, dan pesan promosi judi daring. Proses evaluasi dilakukan menggunakan data uji yang telah dipisahkan sebelumnya, sehingga memberikan gambaran yang objektif mengenai performa model pada data yang belum pernah dilihat.



Gambar 4.6. Evaluasi Model

4.7.1. Hasil Evaluasi Model

Hasil pengujian menunjukkan bahwa aplikasi WFraud Alert berhasil memberikan performa yang sangat baik dalam mengklasifikasikan pesan WhatsApp. Beberapa metrik evaluasi yang digunakan untuk menilai kinerja model meliputi:

a. Presisi (Precision)

Presisi model mencapai 88%, yang berarti sebagian besar pesan yang diklasifikasikan sebagai penipuan benar-benar merupakan pesan penipuan. Nilai ini mencerminkan

keandalan prediksi model dalam memberikan hasil yang tepat pada kategori yang ditentukan.

b. Recall

Recall model juga berada pada tingkat 90%, menunjukkan bahwa sebagian besar pesan penipuan yang ada dalam dataset berhasil terdeteksi oleh model. Nilai recall yang tinggi ini menunjukkan kemampuan model dalam menangkap semua pesan relevan dari kategori tertentu, sehingga meminimalkan kemungkinan pesan penipuan yang tidak terdeteksi.

c. F1-Score

F1-score model, yang merupakan rata-rata harmonis antara presisi dan recall, mencapai 87%. Nilai ini menunjukkan bahwa model memiliki keseimbangan yang baik antara akurasi dalam prediksi positif dan cakupan deteksi dari setiap kategori.

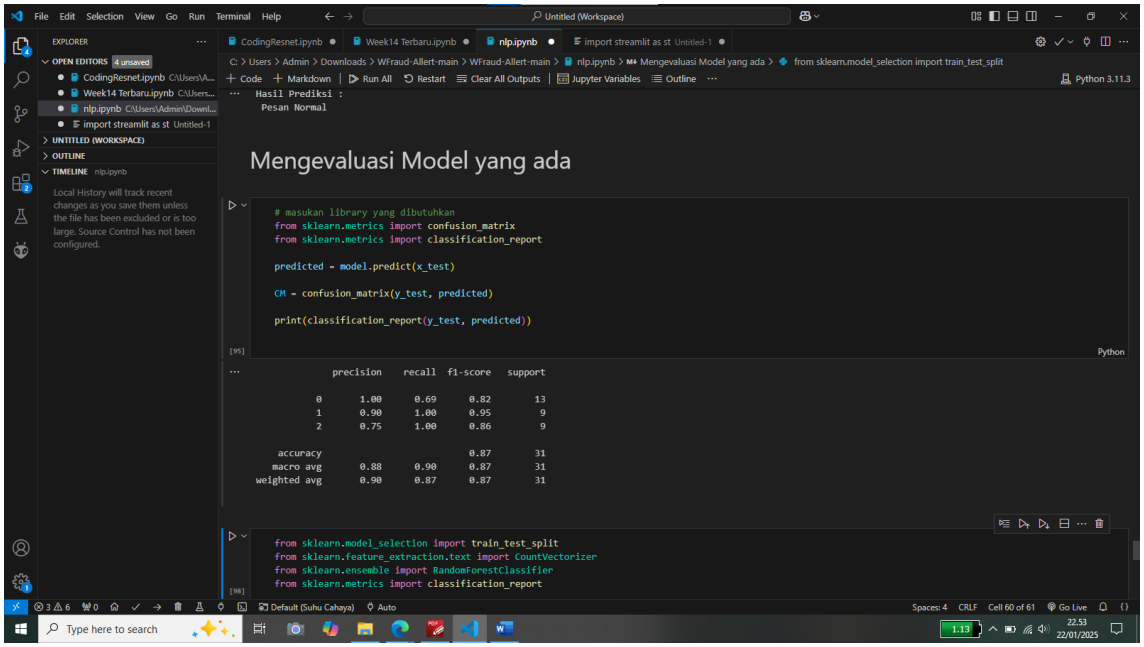
Hasil ini menunjukkan bahwa aplikasi WFraud Alert memiliki kemampuan yang kuat dalam mendeteksi pesan mencurigakan, baik yang mengandung unsur penipuan maupun promosi judi daring. Dengan tingkat presisi, recall, dan F1-score yang sama-sama tinggi, model dapat diandalkan untuk digunakan dalam mendeteksi ancaman penipuan pada platform WhatsApp.

4.7.2. Analisis Hasil Evaluasi

Model yang dikembangkan memiliki tingkat akurasi yang sangat baik, terutama dalam mendeteksi pesan-pesan penipuan dan promosi judi daring. Hal ini menunjukkan bahwa algoritma Multinomial Naive Bayes bekerja secara efektif dalam mengidentifikasi pola-pola kata yang relevan untuk kategori tersebut. Misalnya, kata-kata seperti “promo,” “hadiah,” dan “bonus” sering muncul dalam pesan promosi judi daring dan pesan penipuan, sehingga membantu model dalam membedakan kategori ini dari pesan normal.

Namun, hasil evaluasi juga mengungkapkan beberapa aspek yang dapat ditingkatkan, terutama dalam mengenali pesan normal yang memiliki variasi pola bahasa lebih luas dibandingkan dengan kategori lainnya. Kesalahan prediksi pada kategori pesan normal sebagian besar disebabkan oleh kemiripan struktur atau kata-kata tertentu dengan

pesan penipuan. Hal ini menunjukkan bahwa meskipun performa model secara keseluruhan sudah sangat baik, masih ada peluang untuk meningkatkan akurasi pada kategori tertentu melalui penambahan data atau penyempurnaan proses pra-pemrosesan.



Gambar 4.7. Hasil Evalusi Model

4.7.3. Peluang Pengembangan Lebih Lanjut

Untuk mengatasi beberapa keterbatasan yang ditemukan selama evaluasi, beberapa langkah berikut dapat diambil sebagai strategi pengembangan:

a. Penambahan Data Latih

Penambahan dataset, terutama untuk kategori pesan normal, dapat membantu model mempelajari lebih banyak variasi pola bahasa, sehingga mengurangi kemungkinan kesalahan klasifikasi.

b. Optimalisasi Pra-Pemrosesan

Meningkatkan proses pra-pemrosesan teks, seperti penghapusan kata-kata yang tidak relevan atau penyesuaian kamus normalisasi, dapat membantu model lebih fokus pada fitur-fitur yang signifikan.

c. Penerapan Algoritma Ensembel

Menggunakan kombinasi beberapa algoritma machine learning dapat meningkatkan performa model secara keseluruhan. Teknik seperti Random Forest atau Voting Classifier dapat digunakan untuk memperkuat kelemahan algoritma Naive Bayes.

d. Penggunaan Teknik Augmentasi Data

Teknik augmentasi data dapat diterapkan untuk memperluas variasi pola dalam dataset tanpa harus menambah jumlah data secara manual.

Hasil evaluasi menunjukkan bahwa aplikasi WFraud Alert memiliki performa yang sangat baik dalam mengklasifikasikan pesan WhatsApp. Dengan tingkat presisi 88%, recall 90%, dan F1-score 87%, aplikasi ini mampu memberikan prediksi yang akurat dan konsisten. Keandalan model ini menjadikannya solusi yang efektif untuk mendeteksi pesan penipuan dan promosi judi daring, sehingga dapat membantu pengguna WhatsApp melindungi diri dari ancaman kejahatan siber yang terus berkembang.

Meskipun hasil evaluasi sangat memuaskan, penelitian ini juga mengidentifikasi beberapa aspek yang dapat ditingkatkan, terutama dalam pengenalan kategori pesan normal. Dengan komitmen untuk terus mengembangkan aplikasi ini, WFraud Alert memiliki potensi besar untuk menjadi alat yang lebih canggih dalam melindungi pengguna dari ancaman penipuan melalui media komunikasi digital.

4.8. Confussion Matrix

Confusion Matrix adalah salah satu metode evaluasi performa model klasifikasi yang memetakan hasil prediksi model terhadap data aktual. Dalam penelitian yang menggunakan aplikasi WFraud Alert, Confusion Matrix digunakan untuk mengevaluasi efektivitas algoritma Naive Bayes dalam mengklasifikasikan pesan WhatsApp ke dalam tiga kategori utama: pesan normal, pesan penipuan, dan pesan judi daring.

4.8.1. Konsep Dasar Confusion Matrix

Confusion Matrix adalah tabel matriks yang menampilkan hasil prediksi model berdasarkan empat komponen utama:

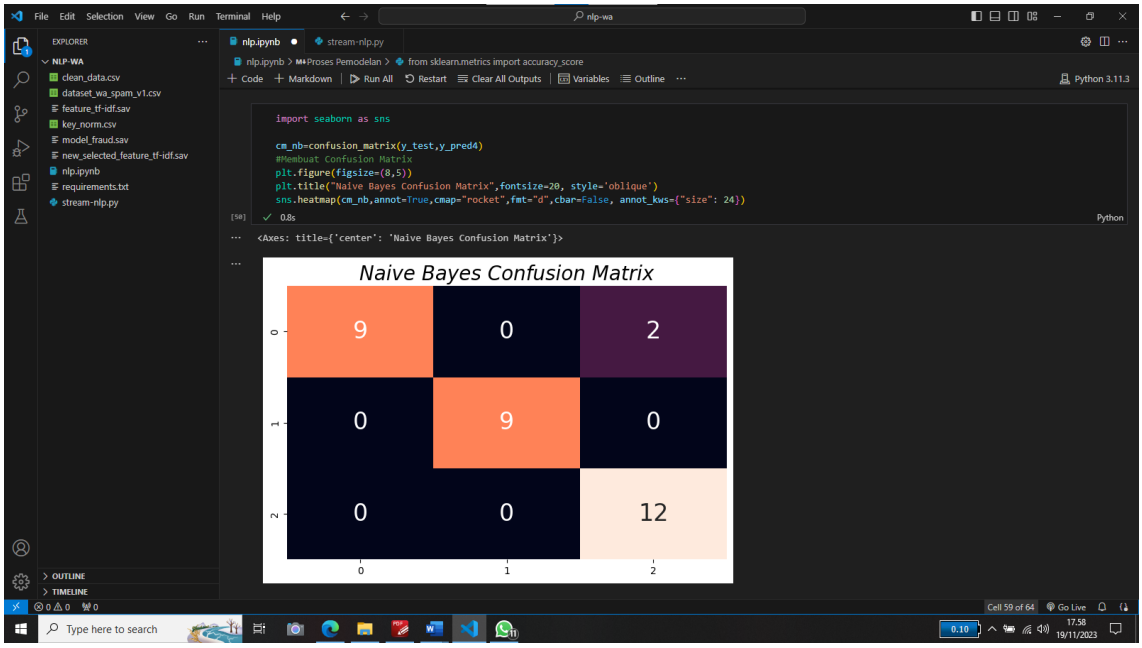
1. True Positive (TP): Jumlah data yang benar-benar positif dan berhasil diklasifikasikan dengan benar oleh model.
2. True Negative (TN): Jumlah data yang benar-benar negatif dan berhasil diklasifikasikan dengan benar oleh model.
3. False Positive (FP): Jumlah data negatif yang salah diklasifikasikan sebagai positif oleh model.
4. False Negative (FN): Jumlah data positif yang salah diklasifikasikan sebagai negatif oleh model.

Rumus confusion matrix untuk menghitung accuracy, precision, dan recall keadaan 3 kelas seperti berikut:

$$Accuracy = \frac{Rasio\ True\ Positive\ (TP)}{Total\ jumlah\ data}$$

$$Precision = \frac{Rasio\ True\ Positive\ (TP)}{TP + FP}$$

$$Recall = \frac{Rasio\ True\ Positive\ (TP)}{TP + FN}$$



Gambar 4.8. Confusion Matrix

Gambar di atas menggambarkan hasil evaluasi model Naive Bayes yang diterapkan pada aplikasi WFraud Alert melalui Confusion Matrix. Matriks ini menunjukkan hubungan antara prediksi model dengan data aktual pada tiga kategori utama: pesan normal (kategori 0), pesan penipuan (kategori 1), dan pesan judi daring (kategori 2). Dari hasil visualisasi, terlihat bahwa model menunjukkan performa yang cukup baik dengan sebagian besar data berhasil diklasifikasikan dengan benar.

Pada baris pertama, yang mewakili kategori pesan normal, terdapat 9 prediksi yang benar, menunjukkan bahwa pesan-pesan normal ini berhasil diidentifikasi dengan tepat. Namun, terdapat 2 kesalahan klasifikasi, di mana pesan normal salah diidentifikasi sebagai pesan judi daring. Hal ini menunjukkan adanya pola atau karakteristik tertentu dalam pesan normal yang dapat menyerupai pesan judi daring, sehingga menyebabkan model membuat kesalahan.

Pada baris kedua, yang merepresentasikan kategori pesan penipuan, seluruh pesan dalam kategori ini diklasifikasikan dengan benar oleh model, menghasilkan 9 prediksi yang akurat tanpa adanya kesalahan klasifikasi. Ini menunjukkan bahwa model memiliki kemampuan yang baik dalam mengenali pola teks yang terkait dengan pesan penipuan.

Pada baris ketiga, yang mewakili kategori pesan judi daring, semua pesan dalam kategori ini juga diklasifikasikan dengan benar. Sebanyak 12 pesan judi daring berhasil diidentifikasi dengan tepat sebagai pesan judi daring tanpa adanya kesalahan. Hal ini mencerminkan kemampuan model yang sangat baik dalam mendeteksi kategori ini.

Secara keseluruhan, Confusion Matrix ini menunjukkan bahwa model Naive Bayes mampu mengklasifikasikan data dengan tingkat akurasi yang tinggi. Namun, masih terdapat ruang untuk perbaikan, terutama pada kesalahan klasifikasi yang terjadi antara pesan normal dan pesan judi daring. Hal ini dapat ditingkatkan dengan memperluas dataset pelatihan atau menggunakan teknik pengolahan teks yang lebih canggih untuk menangkap pola yang lebih kompleks.

4.9. Deployment

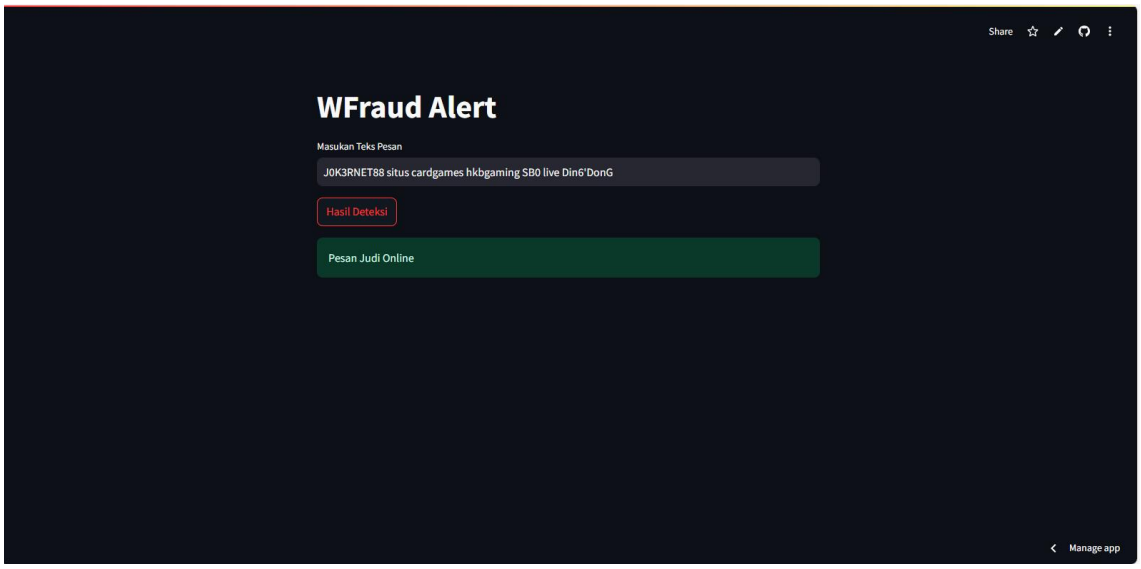
Aplikasi WFraud Alert telah berhasil diimplementasikan dan dideploy menggunakan Streamlit, sebuah framework Python yang mempermudah pembuatan antarmuka pengguna berbasis web untuk aplikasi machine learning. Aplikasi ini dirancang untuk mendeteksi dan mengklasifikasikan pesan WhatsApp ke dalam tiga kategori utama: pesan normal, pesan penipuan, dan pesan judi daring. Berikut adalah penjelasan masing-masing bagian aplikasi berdasarkan gambar yang diberikan:

1. Antarmuka Input dan Output

Pada setiap gambar, terlihat antarmuka aplikasi yang sederhana namun efektif. Pengguna hanya perlu memasukkan teks pesan yang ingin diperiksa pada kolom input yang disediakan. Setelah teks dimasukkan, pengguna dapat melihat hasil deteksi yang langsung ditampilkan di bawahnya.

1. Kolom Input: Pengguna dapat mengetik atau menempelkan teks pesan yang ingin dideteksi. Misalnya, pesan-pesan yang sering kali mencurigakan atau tidak jelas sumbernya.
2. Tombol Deteksi: Setelah memasukkan teks, aplikasi akan memproses pesan tersebut melalui model yang telah dilatih, lalu memberikan hasil klasifikasinya.

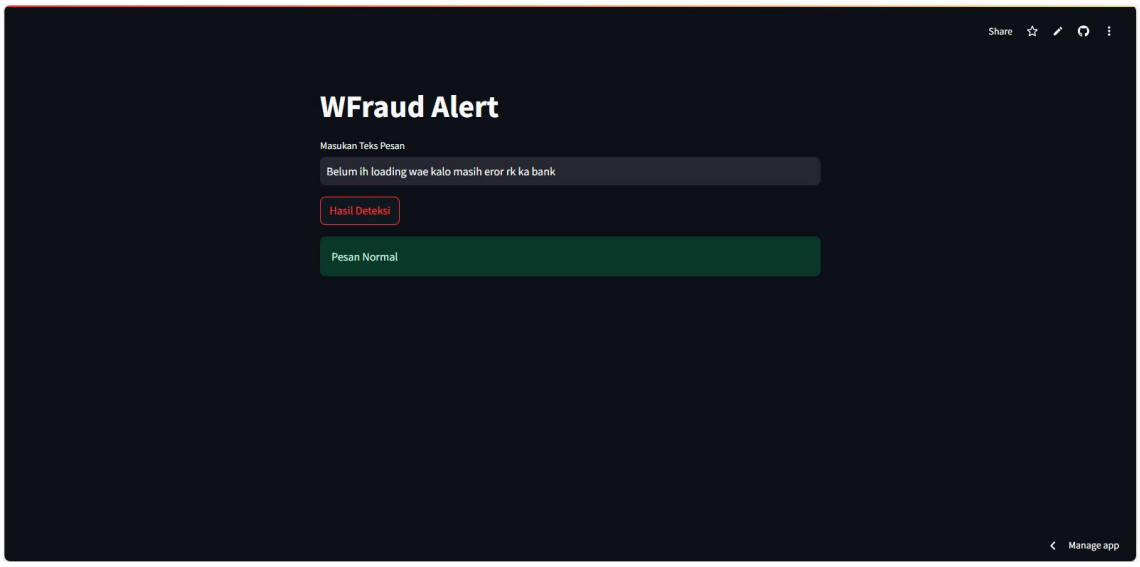
3. Hasil Deteksi: Hasil klasifikasi pesan ditampilkan dengan label yang jelas, seperti "Pesan Judi Online," "Pesan Normal," atau "Pesan Penipuan." Hasil ini didukung oleh warna latar belakang yang membantu membedakan jenis pesan.



Gambar 4.9. Deployment Pesan untuk Judi Online

Pada gambar pertama, aplikasi menerima masukan berupa teks "J0K3RNET88 situs cardgames hkbgaming SBO live Din6’DonG"

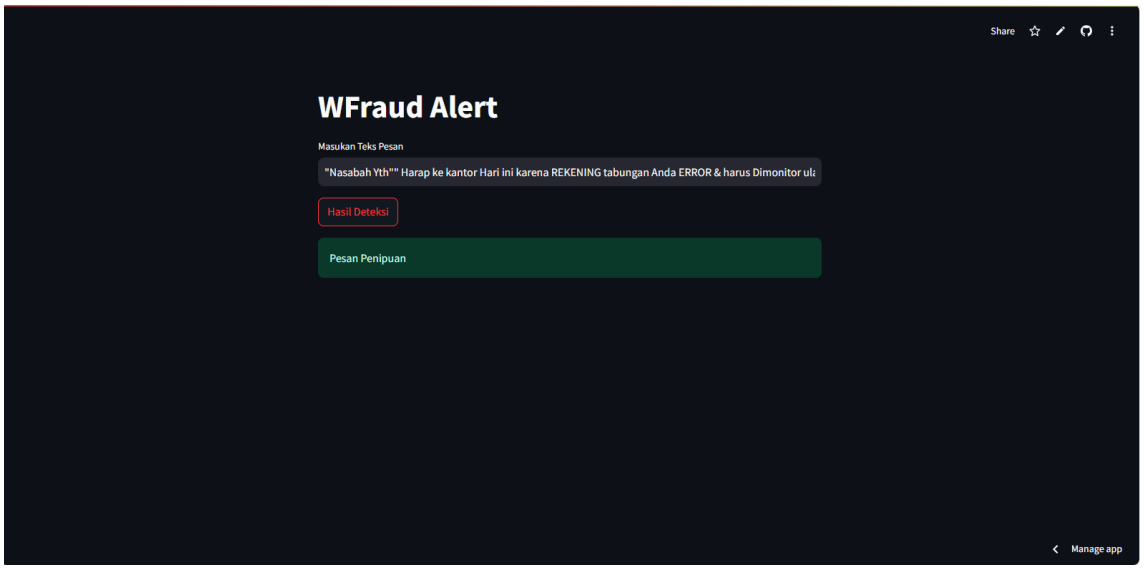
Aplikasi mendeteksi bahwa teks ini termasuk dalam kategori Pesan Judi Online. Hasil deteksi ditampilkan dalam kotak dengan latar belakang hijau bertuliskan "Pesan Judi Online." Kategori ini menunjukkan bahwa teks tersebut mengandung pola yang sering ditemukan pada pesan promosi judi daring.



Gambar 4.10. Deployment Pesan untuk Pesan Normal

"Belum ini loading wae kalo masih error rk ka bank"

Aplikasi mendeteksi pesan ini sebagai Pesan Normal. Hasil ini ditampilkan dalam kotak hijau bertuliskan "Pesan Normal." Deteksi ini menunjukkan bahwa pesan tersebut tidak mengandung elemen penipuan atau promosi yang mencurigakan, sehingga masuk dalam kategori pesan biasa.



Gambar 4.11. Deployment Pesan untuk Pesan Penipuan

Pada gambar ketiga, aplikasi menerima masukan berupa teks "Nasabah Yth** Harap ke kantor Hari ini karena REKENING tabungan Anda ERROR & harus Dimonitor ulang"

Aplikasi mendeteksi pesan ini sebagai Pesan Penipuan. Hasil deteksi ditampilkan dalam kotak hijau bertuliskan "Pesan Penipuan." Pesan ini mengandung pola yang sering digunakan oleh penipu, seperti klaim adanya masalah pada rekening bank untuk memancing respons korban.

4.9.1. Teknologi Streamlit untuk Deployment

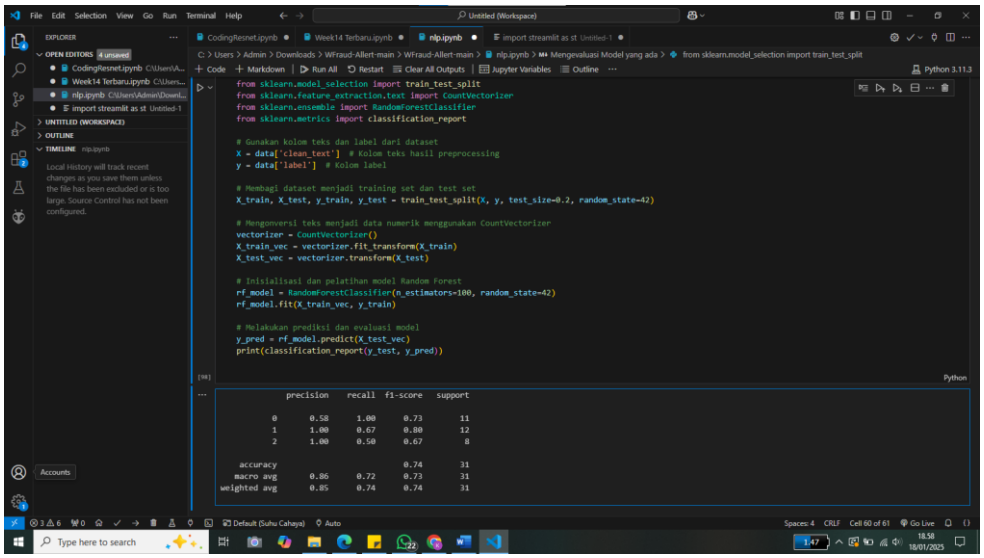
Framework: Aplikasi ini dibangun dengan Streamlit, yang memungkinkan pengembang membuat aplikasi berbasis web dengan cepat tanpa memerlukan banyak pengaturan frontend.

- Simplicity: Antarmuka aplikasi ini dirancang dengan fokus pada kesederhanaan dan kemudahan penggunaan, sehingga pengguna awam sekalipun dapat memanfaatkannya dengan mudah.
- Real-Time Processing: Aplikasi memproses pesan secara real-time menggunakan model Naive Bayes yang telah dilatih sebelumnya untuk mendeteksi dan mengklasifikasikan jenis pesan.
- Deployment: Aplikasi dapat diakses melalui browser setelah di-deploy, menjadikannya solusi yang praktis untuk digunakan oleh masyarakat luas.

Aplikasi WFraud Alert menunjukkan bagaimana teknologi machine learning dapat diintegrasikan dengan framework Streamlit untuk menciptakan solusi yang berguna dalam mendeteksi penipuan pada platform komunikasi digital. Dengan antarmuka yang ramah pengguna dan hasil klasifikasi yang jelas, aplikasi ini memberikan kontribusi nyata dalam meningkatkan literasi digital masyarakat dan melindungi mereka dari ancaman pesan penipuan serta promosi judi daring.

Penelitian tentang klasifikasi teks menggunakan algoritma Naïve Bayes telah banyak dilakukan, termasuk studi oleh Utami et al. (2021) yang membandingkan algoritma Naïve Bayes dan Support Vector Machine dalam menganalisis pesan SMS untuk mendeteksi

spam. Penelitian tersebut menunjukkan bahwa Naïve Bayes mampu mencapai akurasi hingga 90% dalam klasifikasi pesan, meskipun dataset yang digunakan adalah pesan SMS dan bukan pesan instan seperti WhatsApp. Selain itu, Wahyudin et al. (2024) menyoroti berbagai modus penipuan di platform WhatsApp, seperti phishing, sniffing, dan tautan palsu. Studi ini menekankan pentingnya deteksi dini untuk mengurangi risiko kejahatan siber, yang menjadi dasar bagi penelitian dalam deteksi pesan penipuan. Penelitian ini menempatkan dirinya sebagai pengembangan lebih lanjut dari studi-studi tersebut dengan fokus pada pesan WhatsApp, yang saat ini menjadi salah satu platform komunikasi paling populer di dunia.



Gambar 4.12. Evaluasi Model Random Forest

Metode Naïve Bayes dipilih dalam penelitian ini karena memiliki keunggulan dalam mengolah data teks, terutama untuk dataset yang relatif kecil seperti yang digunakan dalam penelitian ini (156 pesan WhatsApp). Naïve Bayes dikenal karena efisiensi komputasinya, kemampuannya menangani data teks secara sederhana, dan keakuratannya yang kompetitif dalam klasifikasi berbasis probabilistik, seperti yang ditunjukkan dalam penelitian Utami et al. (2021) yang mencapai akurasi 90% pada data SMS

BAB 5 KESIMPULAN DAN SARAN

5.1. Kesimpulan

Setelah melakukan serangkaian proses penelitian dan pengembangan aplikasi WFraud Alert, diperoleh hasil sebagai berikut:

a. Aplikasi Deteksi Penipuan

Aplikasi WFraud Alert berhasil dikembangkan dengan kemampuan mengklasifikasikan pesan WhatsApp ke dalam tiga kategori utama: pesan normal, pesan penipuan, dan pesan promosi judi daring. Tingkat akurasi model yang diperoleh berkisar antara 87% hingga 90% berdasarkan evaluasi menggunakan metrik presisi, recall, dan F1-score.

b. Efektivitas Preprocessing Data

Tahapan preprocessing data, seperti case folding, normalisasi kata, penghapusan stopwords, dan stemming, terbukti mampu meningkatkan performa model Naïve Bayes secara signifikan. Penghapusan noise pada data teks menghasilkan representasi data yang lebih akurat untuk analisis.

c. Penerapan Algoritma Naïve Bayes

Algoritma Naïve Bayes menunjukkan efektivitas yang baik dalam menangani tugas klasifikasi teks berbahasa Indonesia. Dengan pendekatan probabilistik, algoritma ini mampu mengenali pola-pola dalam dataset dengan ukuran terbatas.

d. Manfaat Aplikasi

Aplikasi ini dapat membantu pengguna WhatsApp untuk mendeteksi dan menghindari pesan penipuan serta memberikan edukasi tentang pola-pola umum dalam kejahatan siber.

Pada proses pengembangan dan implementasi aplikasi WFraud Alert, terdapat beberapa tantangan yang memengaruhi hasil penelitian, yaitu:

- **Keterbatasan Dataset**

Dataset yang digunakan relatif kecil sehingga kemampuan model dalam generalisasi pola penipuan dapat terbatas.

- **Keberagaman Format Pesan**

Pesan-pesan WhatsApp memiliki variasi bahasa dan struktur yang kompleks, sehingga dapat memengaruhi performa model dalam beberapa kasus.

- **Kemiripan Pola pada Pesan**

Pesan promosi judi daring dan pesan penipuan kadang memiliki pola serupa, yang dapat menyebabkan kesalahan klasifikasi.

Dengan hasil penelitian ini, aplikasi WFraud Alert diharapkan dapat memberikan kontribusi positif dalam melindungi masyarakat dari ancaman penipuan daring serta meningkatkan literasi digital di Indonesia.

5.2. Saran

Untuk meningkatkan performa dan manfaat aplikasi WFraud Alert, penulis memberikan beberapa saran yang berfokus pada pengembangan fitur, penyesuaian metode, serta pengoptimalan model. Adapun saran-saran tersebut adalah sebagai berikut:

1. Pengembangan Fitur Aplikasi

Beberapa fitur tambahan yang disarankan untuk meningkatkan kemudahan penggunaan dan fleksibilitas aplikasi adalah:

a. Notifikasi Real-Time: Memberikan peringatan langsung kepada pengguna jika terdeteksi pesan mencurigakan.
b. Dashboard Edukasi Pengguna: Menyediakan informasi mengenai pola-pola umum penipuan untuk meningkatkan literasi digital.
c. Fitur Pelaporan Pesan: Menambahkan fitur bagi pengguna untuk melaporkan pesan mencurigakan yang belum terdeteksi oleh sistem.
d. Integrasi Multiplatform: Memungkinkan penggunaan aplikasi di berbagai perangkat, termasuk desktop dan mobile.

2. Pengoptimalan Algoritma dan Dataset

a. Ekspansi Dataset: Mengumpulkan lebih banyak data pesan WhatsApp dari berbagai sumber untuk meningkatkan kemampuan generalisasi model.
b. Eksplorasi Algoritma Lanjutan: Menggunakan algoritma pembelajaran mendalam, seperti LSTM atau transformer, untuk menangani kompleksitas pola bahasa yang lebih tinggi.
c. Penyempurnaan Preprocessing Data: Mengembangkan metode preprocessing yang dapat menangani variasi format pesan dan bahasa secara lebih adaptif.

3. Penyesuaian pada Infrastruktur Aplikasi

a. Penambahan Sistem Cache: Mengurangi waktu pemrosesan dengan menyimpan data sementara untuk pesan yang sering muncul.
b. Pengoptimalan Sistem Penyimpanan: Menggunakan cloud storage untuk menyimpan log deteksi pesan penipuan agar dapat diakses dengan mudah.

4. Pengembangan dan Pengujian Lanjutan

a. Evaluasi Berbasis Kasus Nyata: Menggunakan aplikasi pada data pesan nyata secara langsung untuk memastikan akurasi dan efektivitas dalam kondisi operasional.
b. Penyesuaian untuk Multibahasa: Mengembangkan aplikasi agar mampu mendeteksi penipuan dalam berbagai bahasa yang sering digunakan di WhatsApp.

DAFTAR PUSTAKA

[1] Veronika, "Perkembangan Teknologi Informasi dan Komunikasi: Sejarah, Dampak, Tantangan, dan Peluang. ," Kompasiana.

[2] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *TEKNO KOMPAK*, vol. 15, no. 1, 2019.

[3] Lenny, "Kominfo Catatkan 1.730 Kasus Penipuan Online, Kerugian Ratusan Triliun," Kata Data.

[4] LERAVIO, "Naive Bayes: Pengertian, Kelebihan, dan Implementasinya," ADMIN LERAVIO.

[5] R. Rachman, R. N. Handayani, and I. Artikel, "Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM," *JURNAL INFORMATIKA*, vol. 8, no. 2, 2021, [Online]. Available: http://ejournal.bsi.ac.id/ejurnal/index.php/ji

[6] R. S. Stmik and N. Mandiri, "Komparasi Algoritma Support Vector Machine, Naïve Bayes Dan C4.5 Untuk Klasifikasi SMS," *IJCIT (Indonesian Journal on Computer and Information Technology*, vol. 2, no. 2, 2017.

[7] D. Delvia Arifin and Ma. Bijaksana, "SMS Filtering Menggunakan Naive Bayes Classifier dan FP-Growth Algorithm Frequent Itemset," *e-Proceeding of Engineering*, 2019, [Online]. Available: http://www.ranks.nl/stopwords.

[8] R. Dwiyansaputra, G. Satya Nugraha, F. Bimantoro, and A. Aranta, "DETEKSI SMS SPAM BERBAHASA INDONESIA MENGGUNAKAN TF-IDF DAN STOCHASTIC GRADIENT DESCENT CLASSIFIER (Indonesian SMS Spam Detection using TF-IDF and Stochastic Gradient Descent Classifier)," *JTIKA*, 2021, [Online]. Available: http://jtika.if.unram.ac.id/index.php/JTIKA/

[9] L. D. Utami, L. Yusuf, and D. Nurlaela, “Komparasi Algoritma Naïve Bayes dan Support Vectors Machine pada Analisis Sentimen SMS HAM dan SPAM,” *Jurnal Informatika dan Teknologi*, vol. 4, no. 2, 2021, doi: 10.29408/jit.v4i2.3665.

[10] U. Banten Jaya, J. Syeh Nawawi Albantani, and S. -Banten, “PERBANDINGAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE (SVM) DALAM KLASIFIKASI SMS SPAM BERBAHASA INDONESIA,” *SAINTEK | Jurnal Sains & Teknologi* , 2019.

[11] H. Herwanto, N. L. Chusna, and M. S. Arif, “Klasifikasi SMS Spam Berbahasa Indonesia Menggunakan Algoritma Multinomial Naïve Bayes,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 4, p. 1316, Oct. 2021, doi: 10.30865/mib.v5i4.3119.

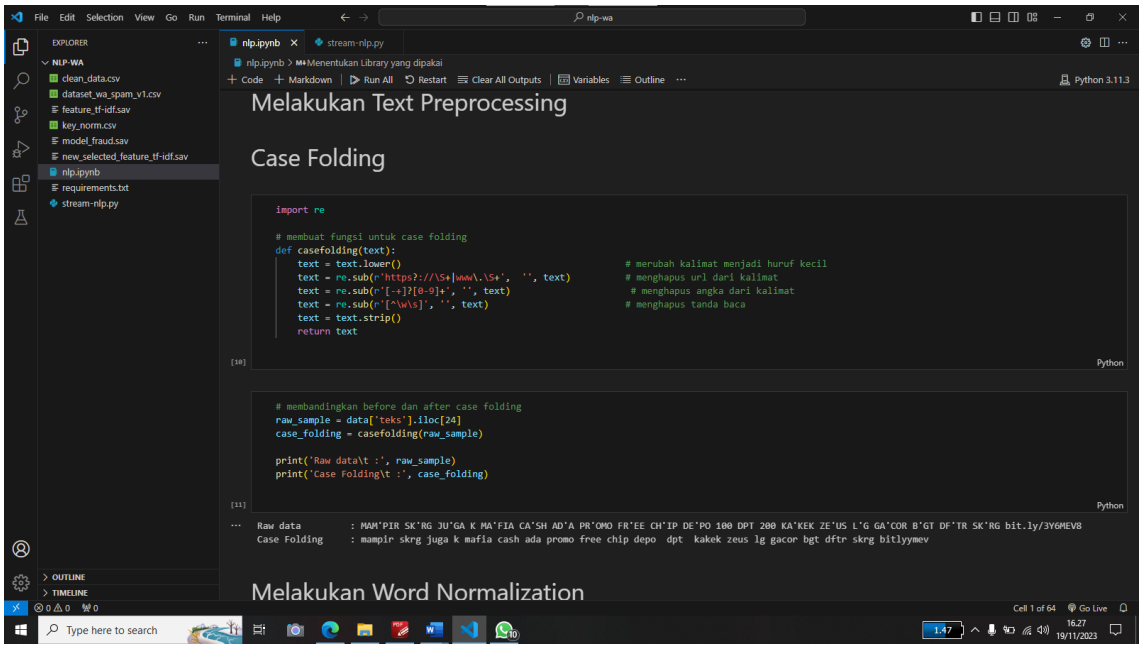
[12] Zia Ayu Nuansa Gumilang, “IMPLEMENTASI NAÏVE BAYES CLASSIFIER DAN ASOSIASI UNTUK ANALISIS SENTIMEN DATA ULASAN APLIKASI E-COMMERCE SHOPEE PADA SITUS GOOGLE PLAY,” YOGYAKARTA, 2018.

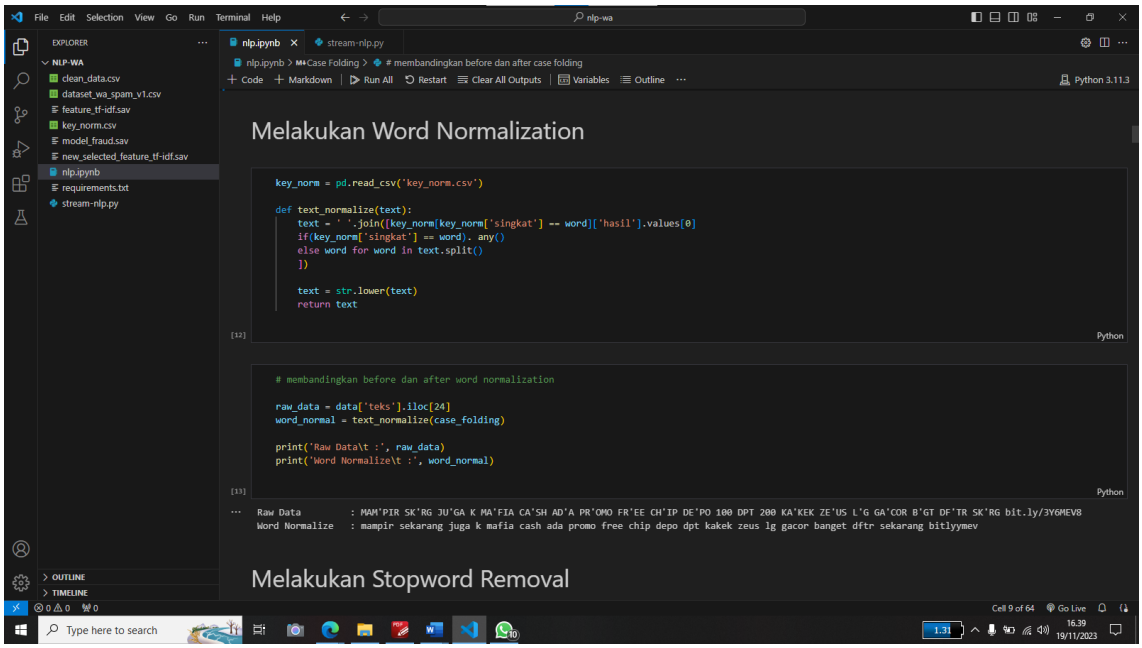
[13] M. Ibrahim, E. Bu, and I. Lubis, “RESOLUSI : Rekayasa Teknik Informatika dan Informasi Penerapan Algoritma Naive Bayes Classifier Untuk Mendeteksi Tingkat Krediblitas Hoax News/ Fake News Pada Sosial Media Di Indonesia Berbasis Android (Studi Kasus : Kantor Tribun Medan),” *Media Online)*, vol. 1, no. 1, 2020, [Online]. Available: https://djournals.com/resolusi

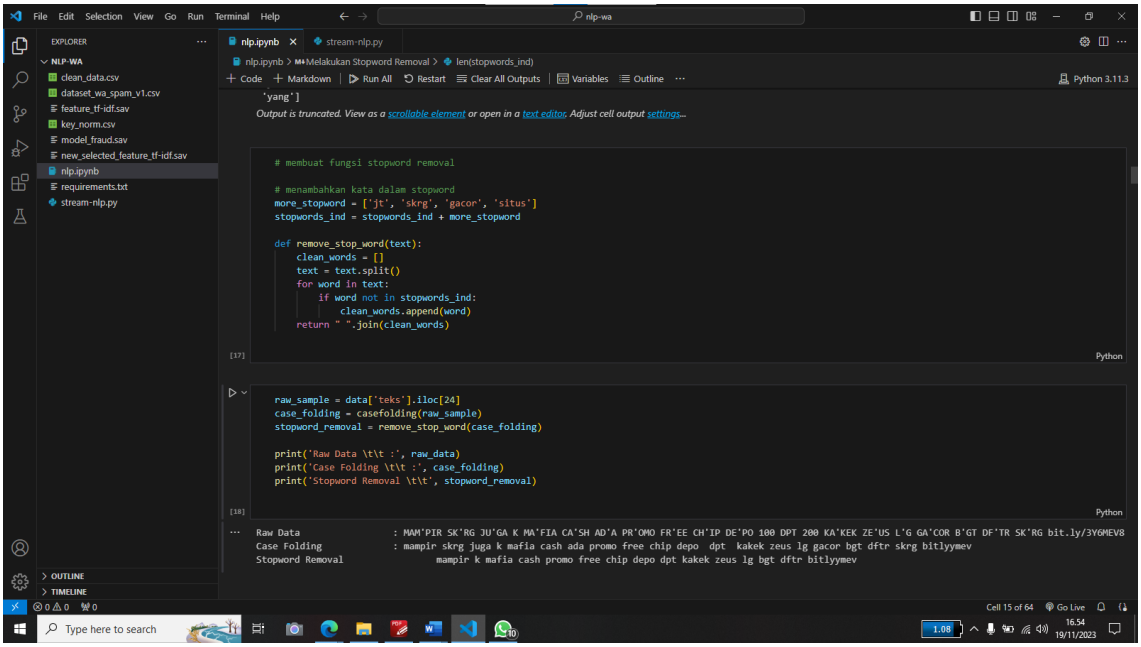
[14] D. Normawati and S. A. Prayogi, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” 2021.

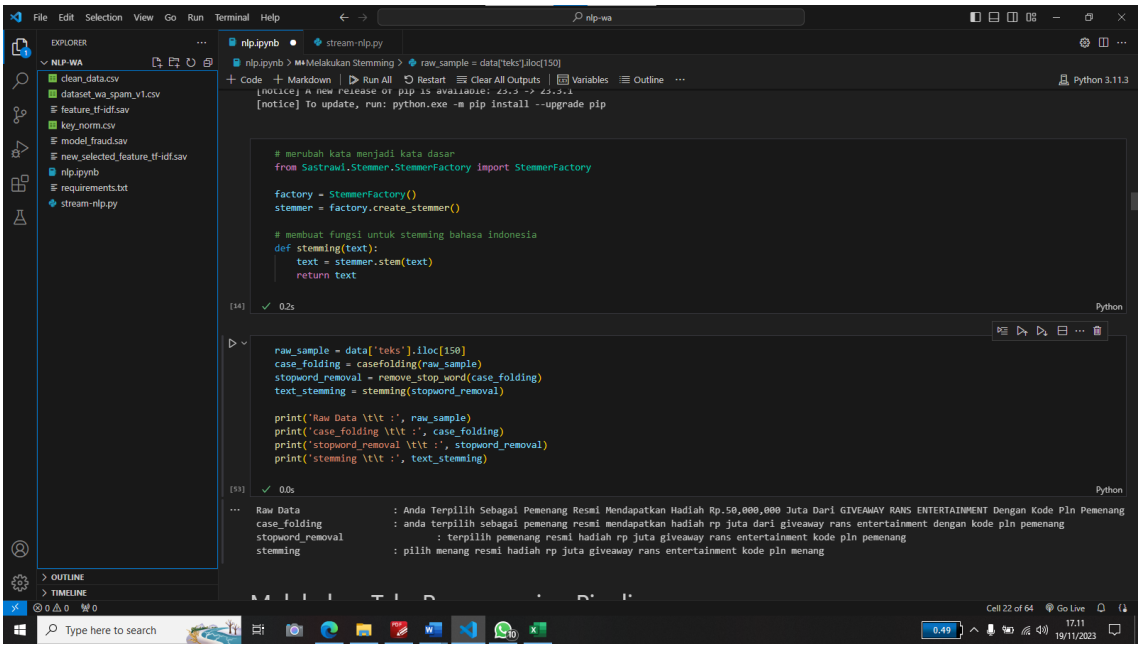
LAMPIRAN

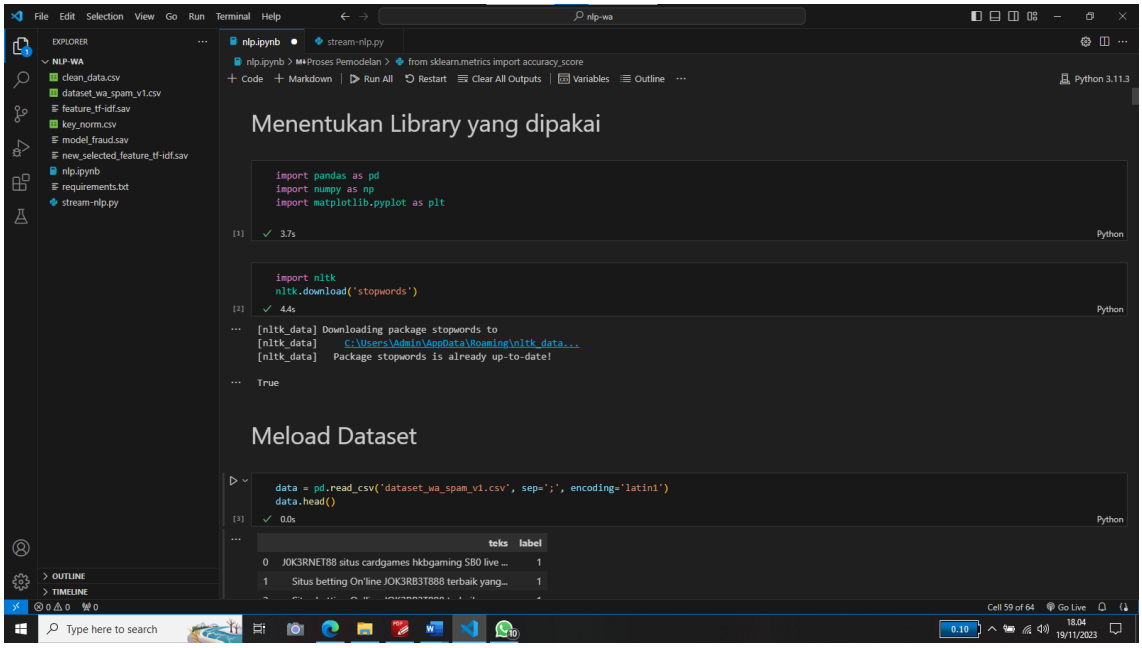
Foto Pembuatan Machine Learning

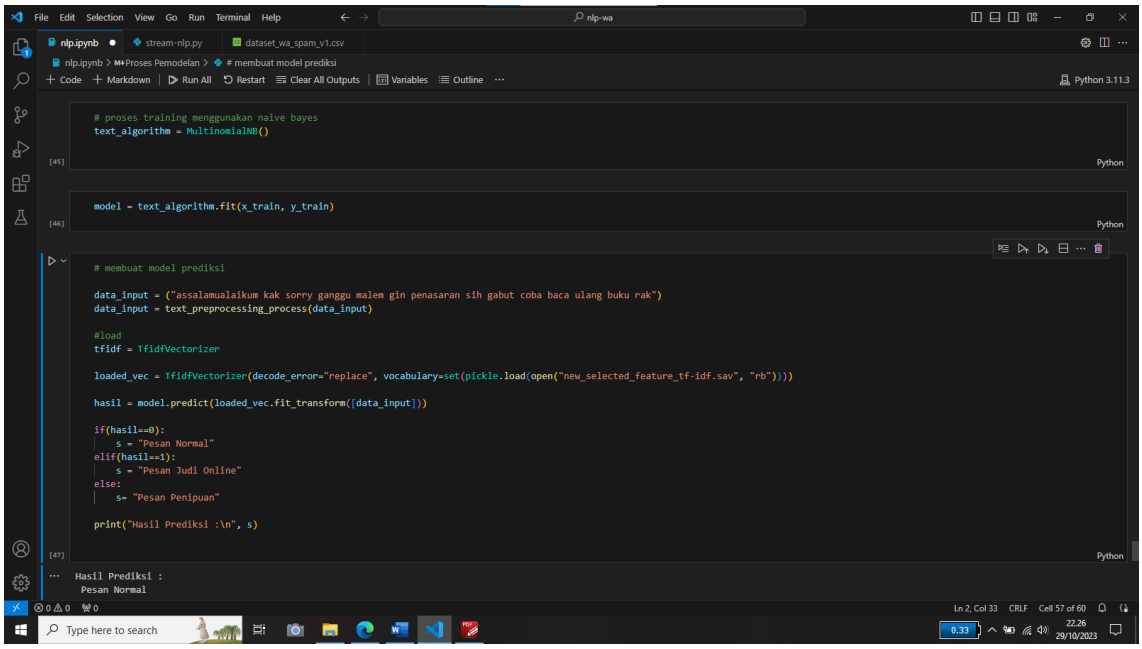


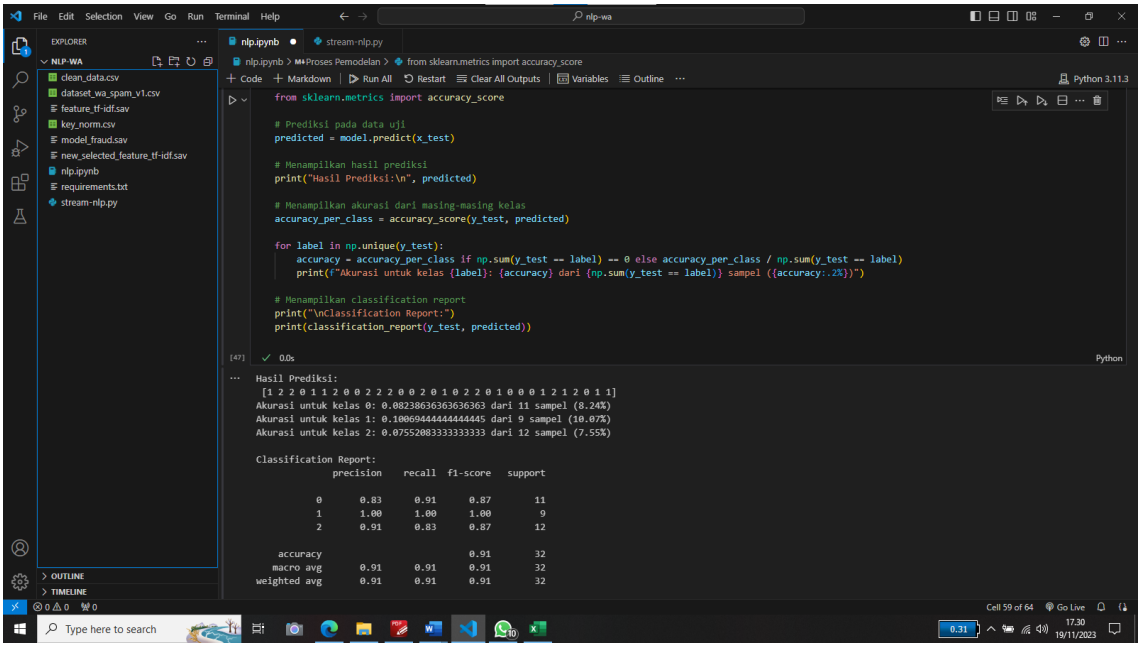


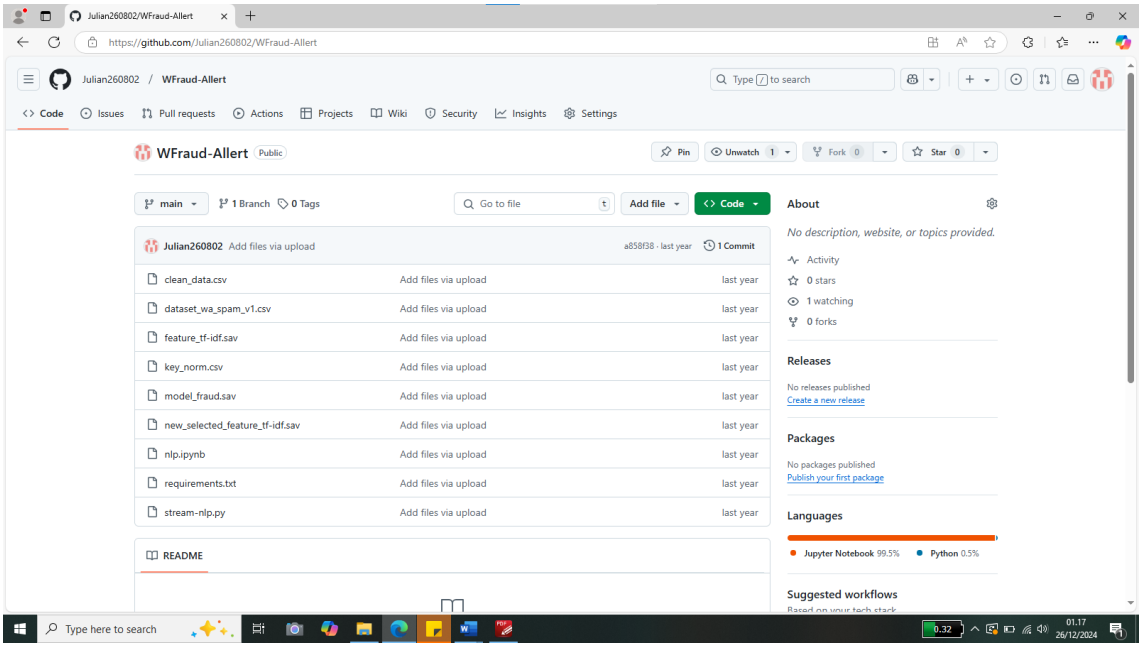

'yang']
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
# membuat fungsi stopword removal
# menambahkan kata dalam stopword
more_stopword = ['jt', 'skrg', 'gacor', 'situs']
stopwords_ind = stopwords_ind + more_stopword
def remove_stop_word(text):
clean_words = []
text = text.split()
for word in text:
if word not in stopwords_ind:
clean_words.append(word)
return " ".join(clean_words)
raw_sample = data['teks'].iloc[24]
case_folding = casefolding(raw_sample)
stopword_removal = remove_stop_word(case_folding)
print('Raw Data \t\t :', raw_data)
print('Case Folding \t\t :', case_folding)
print('Stopword Removal \t\t', stopword_removal)
Raw Data : MAM'PIR SK'RG JU'GA K MA'FIA CA'SH AD'A PR'OMO FR'EE CH'IP DE'PO 100 DPT 200 KA'KEK ZE'US L'G GA'COR B'GT DF'TR SK'RG bit.ly/3Y6MEV8
Case Folding : mampir skrg juga k mafia cash ada promo free chip depo dpt kakek zeus lg gacor bgt dftr skrg bitlyymev
Stopword Removal mampir k mafia cash promo free chip depo dpt kakek zeus lg bgt dftr bitlyymev


[notice] To update, run: python.exe -m pip install --upgrade pip
# merubah kata menjadi kata dasar
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
factory = StemmerFactory()
stemmer = factory.create_stemmer()
# membuat fungsi untuk stemming bahasa indonesia
def stemming(text):
text = stemmer.stem(text)
return text
raw_sample = data['teks'].iloc[150]
case_folding = casefolding(raw_sample)
stopword_removal = remove_stop_word(case_folding)
text_stemming = stemming(stopword_removal)
print('Raw Data \t\t :', raw_sample)
print('case_folding \t\t :', case_folding)
print('stopword_removal \t\t :', stopword_removal)
print('stemming \t\t :', text_stemming)
Raw Data : Anda Terpilih Sebagai Pemenang Resmi Mendapatkan Hadiah Rp.50,000,000 Juta Dari GIVEAWAY RANS ENTERTAINMENT Dengan Kode Pln Pemenang
case_folding : anda terpilih sebagai pemenang resmi mendapatkan hadiah rp juta dari giveaway rans entertainment dengan kode pln pemenang
stopword_removal : terpilih pemenang resmi hadiah rp juta giveaway rans entertainment kode pln pemenang
stemming : pilih menang resmi hadiah rp juta giveaway rans entertainment kode pln menang


Menentukan Library yang dipakai
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import nltk
nltk.download('stopwords')
[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\Admin\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
True
Meload Dataset
data = pd.read_csv('dataset_wa_spam_v1.csv', sep=';', encoding='latin1')
data.head()
teks label
0 J0K3RNET88 situs cardgames hkbgaming SB0 live ... 1
1 Situs betting On'line JOK3R83T888 terbaik yang... 1


# proses training menggunakan naive bayes
text_algorithm = MultinomialNB()
model = text_algorithm.fit(x_train, y_train)
# membuat model prediksi
data_input = ("assalamualaikum kak sorry ganggu malem gin penasaran sih gabut coba baca ulang buku rak")
data_input = text_preprocessing_process(data_input)
#load
tfidf = TfidfVectorizer
loaded_vec = TfidfVectorizer(decode_error="replace", vocabulary=set(pickle.load(open("new_selected_feature_tf-idf.sav", "rb"))))
hasil = model.predict(loaded_vec.fit_transform([data_input]))
if(hasil==0):
s = "Pesan Normal"
elif(hasil==1):
s = "Pesan Judi Online"
else:
s= "Pesan Penipuan"
print("Hasil Prediksi :\n", s)
Hasil Prediksi :
Pesan Normal


File Edit Selection View Go Run Terminal Help
nlp-wa
EXPLORER
NLP-WA
clean_data.csv
dataset_wa_spam_v1.csv
feature_tf-idf.sav
key_norm.csv
model_fraud.sav
new_selected_feature_tf-idf.sav
nlp.ipynb
requirements.txt
stream-nlp.py
nlp.ipynb
stream-nlp.py
nlp.ipynb > Proses Pemodelan > from sklearn.metrics import accuracy_score
Code Markdown Run All Restart Clear All Outputs Variables Outline
Python 3.11.3
from sklearn.metrics import accuracy_score
# Prediksi pada data uji
predicted = model.predict(x_test)
# Menampilkan hasil prediksi
print("Hasil Prediksi:\n", predicted)
# Menampilkan akurasi dari masing-masing kelas
accuracy_per_class = accuracy_score(y_test, predicted)
for label in np.unique(y_test):
accuracy = accuracy_per_class if np.sum(y_test == label) == 0 else accuracy_per_class / np.sum(y_test == label)
print(f"Akurasi untuk kelas {label}: {accuracy} dari {np.sum(y_test == label)} sampel ({accuracy:.2%})")
# Menampilkan classification report
print("\nClassification Report:")
print(classification_report(y_test, predicted))
[47] 0.0s
Python
Hasil Prediksi:
[1 2 2 0 1 1 2 0 0 2 2 2 0 0 2 0 1 0 2 2 0 1 0 0 0 1 2 1 2 0 1 1]
Akurasi untuk kelas 0: 0.08238636363636363 dari 11 sampel (8.24%)
Akurasi untuk kelas 1: 0.10069444444444445 dari 9 sampel (10.07%)
Akurasi untuk kelas 2: 0.07552083333333333 dari 12 sampel (7.55%)
Classification Report:
precision recall f1-score support
0 0.83 0.91 0.87 11
1 1.00 1.00 1.00 9
2 0.91 0.83 0.87 12
accuracy 0.91 32
macro avg 0.91 0.91 0.91 32
weighted avg 0.91 0.91 0.91 32
OUTLINE
TIMELINE
Cell 59 of 64
Go Live
Type here to search
17.30
19/11/2023


Julian260802/WFraud-Allert
https://github.com/Julian260802/WFraud-Allert
Julian260802 / WFraud-Allert
Type / to search
Code Issues Pull requests Actions Projects Wiki Security Insights Settings
WFraud-Allert Public
Pin Unwatch 1 Fork 0 Star 0
main 1 Branch 0 Tags
Go to file
Add file
Code
Julian260802 Add files via upload
a858f38 · last year
1 Commit
clean_data.csv Add files via upload last year
dataset_wa_spam_v1.csv Add files via upload last year
feature_tf-idf.sav Add files via upload last year
key_norm.csv Add files via upload last year
model_fraud.sav Add files via upload last year
new_selected_feature_tf-idf.sav Add files via upload last year
nlp.ipynb Add files via upload last year
requirements.txt Add files via upload last year
stream-nlp.py Add files via upload last year
README
About
No description, website, or topics provided.
Activity
0 stars
1 watching
0 forks
Releases
No releases published
Create a new release
Packages
No packages published
Publish your first package
Languages
Jupyter Notebook 99.5%
Python 0.5%
Suggested workflows
Type here to search
01.17
26/12/2024

Coding Machine Learning

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
data = pd.read_csv('dataset_wa_spam_v1.csv', sep=';')
data.head()
import re
# membuat fungsi untuk case folding
def casefolding(text):
    text = text.lower()                                 # merubah kalimat menjadi huruf kecil
    text = re.sub(r'https?://\S+|www\.\S+',  '', text)          # menghapus url dari kalimat
    text = re.sub(r'[-+]?[0-9]+', '', text)                    # menghapus angka dari kalimat
    text = re.sub(r'[^\w\s]', '', text)                     # menghapus tanda baca
    text = text.strip()
    return text
# membandingkan before dan after case folding
raw_sample = data['teks'].iloc[151]
case_folding = casefolding(raw_sample)
print('Raw data\t :', raw_sample)
print('Case Folding\t :', case_folding)
key_norm = pd.read_csv('key_norm.csv')
def text_normalize(text):
    text = ' '.join([key_norm[key_norm['singkat'] == word]['hasil'].values[0]
```

```
    if(key_norm['singkat'] == word). any()
    else word for word in text.split()
    ])
    text = str.lower(text)
    return text
# membandingkan before dan after word normalization
raw_data = data['teks'].iloc[32]
word_normal = text_normalize(case_folding)
print('Raw Data\t :', raw_data)
print('Word Normalize\t :', word_normal)
from nltk.tokenize import sent_tokenize, word_tokenize
from nltk.corpus import stopwords
stopwords_ind = stopwords.words('indonesian')
len(stopwords_ind)
# melihat daftar stopword dari nltk
stopwords_ind
# menambahkan kata dalam stopword
more_stopword = ['gacor', 'juta', 'rb', 'situs']
stopwords_ind = stopwords_ind + more_stopword
def remove_stop_word(text):
    clean_words = []
    text = text.split()
    for word in text:
```

```
        if word not in stopwords_ind:
            clean_words.append(word)
    return " ".join(clean_words)
raw_sample = data['teks'].iloc[151]
case_folding = casefolding(raw_sample)
stopword_removal = remove_stop_word(case_folding)
print('Raw Data \t\t :', raw_data)
print('Case Folding \t\t :', case_folding)
print('Stopword Removal \t\t', stopword_removal)
# merubah kata menjadi kata dasar
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
factory = StemmerFactory()
stemmer = factory.create_stemmer()
# membuat fungsi untuk stemming bahasa indonesia
def stemming(text):
    text = stemmer.stem(text)
    return text
raw_sample = data['teks'].iloc[151]
case_folding = casefolding(raw_sample)
stopword_removal = remove_stop_word(case_folding)
text_stemming = stemming(stopword_removal)
print('Raw Data \t\t :', raw_sample)
print('case_folding \t\t :', case_folding)
```

```
print('stopword_removal \t\t :', stopword_removal)
print('stemming \t\t :', text_stemming)
# membuat fungsi untuk menggabungkan seluruh langkah text preproceessing
def text_preprocessing_process(text):
    text = casefolding(text)
    text = text_normalize(text)
    text = remove_stop_word(text)
    text = stemming(text)
    return text
%%time
data['clean_text']= data['teks'].apply(text_preprocessing_process)
# menyimpan data yang sudah di preprocessing ke dalam file csv
data.to_csv('clean_data.csv')
# pisahkan kolom feature daan target
x = data['clean_text']
y = data['label']
# save model
import pickle
#TF-IDF
from sklearn.feature_extraction.text import TfidfVectorizer
#Unigram
vec_TF_IDF = TfidfVectorizer(ngram_range=(1,1))
vec_TF_IDF.fit(x)
```

```
x_tf_idf = vec_TF_IDF. transform(x)
pickle.dump(vec_TF_IDF.vocabulary_,open("feature_tf-idf.sav", "wb"))
# melihat jumlah feature
print(len(vec_TF_IDF.get_feature_names_out()))
x1 = vec_TF_IDF.transform(x).toarray()
data_tabular_tf_idf =
pd.DataFrame(x1,columns=vec_TF_IDF.get_feature_names_out())
data_tabular_tf_idf
x_train = np.array(data_tabular_tf_idf)
y_train = np.array(y)
from sklearn.feature_selection import SelectKBest
from sklearn.feature_selection import chi2
chi2_features = SelectKBest(chi2, k=800)
x_kbest_features = chi2_features.fit_transform(x_train, y_train)
# untuk reduced features
print('Original Feature Number', x_train.shape[1])
print('Reduced feature Number', x_kbest_features.shape[1])
Data = pd.DataFrame(chi2_features.scores_,columns=['Nilai'])
# menampilkan feature beserta nilainya
feature = vec_TF_IDF.get_feature_names_out()
feature
```

```
Data['Fitur'] = feature
# mengurutkan nilai feature terbaik
Data.sort_values(by='Nilai', ascending=False)
# menampilkan fitur yang terpilih berdasarkan nilai mask atau nilai tertinggi yang sudah ditetapkan pada chi square
new_feature=[]
for bool, f in zip(mask, feature):
    if bool :
        new_feature.append(f)
    selected_feature=new_feature
selected_feature
# membuat vocabulary baru berdasarkan fitur yang terseleksi
new_selected_feature = {}
for (k,v) in vec_TF_IDF.vocabulary_.items():
    if k in selected_feature:
        new_selected_feature[k]=v
new_selected_feature
pickle.dump(new_selected_feature,open("new_selected_feature_tf-idf.sav","wb"))
# menampilkan fitur-fitur yang sudah diseleksi
data_selected_feature = pd.DataFrame(x_kbest_features, columns=selected_feature)
data_selected_feature
selected_x = x_kbest_features
selected_x
```

```
# import library
import random
from sklearn.model_selection import train_test_split
# import algoritma naive bayes
from sklearn.naive_bayes import MultinomialNB
x = selected_x
y = data.label
x_train, x_test, y_train, y_test = train_test_split(x,y, test_size=0.2, random_state=0)
# menampilkan jumlah data training dan data testing
print('Banyaknya X_train :', len(x_train))
print('Banyaknya X_test :', len(x_test))
print('Banyaknya Y_train :', len(y_train))
print('Banyaknya Y_test :', len(y_test))
# proses training menggunakan naive bayes
text_algorithm = MultinomialNB()
model = text_algorithm.fit(x_train, y_train)
# membuat model prediksi
data_input = ("assalamualaikum kak sorry ganggu malem gin penasaran sih gabut coba baca ulang buku rak")
data_input = text_preprocessing_process(data_input)
#load
tfidf = TfidfVectorizer
```

```
loaded_vec = TfidfVectorizer(decode_error="replace",
vocabulary=set(pickle.load(open("new_selected_feature_tf-idf.sav", "rb"))))
hasil = model.predict(loaded_vec.fit_transform([data_input]))
if(hasil==0):
   s = "Pesan Normal"
elif(hasil==1):
   s = "Pesan Judi Online"
else:
   s= "Pesan Penipuan"
print("Hasil Prediksi :\n", s)
# masukan library yang dibutuhkan
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
predicted = model.predict(x_test)
CM = confusion_matrix(y_test, predicted)
print(classification_report(y_test, predicted))
# menyimpan model
pickle.dump(model,open("model_fraud.sav","wb"))
```