

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Tinjauan Teori Umum**

Di dalam sub bab berikut, penulis akan membahas mengenai teori-teori umum yang berhubungan dengan pelaksanaan penelitian ini yang berupa pengertian-pengertian secara umum.

##### **2.1.1 Berita**

Berita adalah informasi tentang peristiwa, fakta, atau kejadian terbaru yang relevan dan menarik bagi masyarakat. Berita disajikan dalam format teks, gambar, audio, atau video melalui berbagai media, seperti surat kabar, televisi, radio, dan platform berita daring. Tujuannya adalah untuk memberikan informasi aktual

kepada publik tentang hal-hal yang terjadi di sekitar mereka, termasuk isu politik, sosial, ekonomi, budaya, dan lainnya. Berita bertujuan untuk memberikan pemahaman dan pengetahuan kepada pembaca atau penonton agar mereka dapat terinformasi secara akurat tentang dunia di sekitar mereka (Syaifullah and Anggraini 2022).

##### **2.1.2 Berita Hoax**

Berita *hoax* adalah informasi palsu atau tidak benar yang disajikan dalam bentuk berita dengan tujuan untuk menyesatkan atau memanipulasi pembaca atau

penonton. Berita *hoax* sering kali dirancang sedemikian rupa sehingga terlihat seperti berita yang sah dan dapat dengan mudah menipu orang yang membacanya. Tujuan dari berita *hoax* bisa bermacam-macam, termasuk menyebarkan propaganda, menciptakan kebingungan, mempengaruhi opini publik, atau bahkan mencari keuntungan finansial.

Berita *hoax* dapat disebarluaskan melalui berbagai media, termasuk surat kabar palsu, situs web berita palsu, media sosial, pesan berantai, dan lainnya. Konten berita *hoax* seringkali tidak didukung oleh fakta atau bukti yang valid, dan sumbernya dapat tidak dapat dipercaya atau tidak jelas.

Dalam era digital dan media sosial, penyebaran berita *hoax* dapat menjadi masalah serius karena dapat dengan cepat menyebar luas dan memengaruhi persepsi dan tindakan masyarakat. Oleh karena itu, penting untuk dapat mengenali dan memverifikasi kebenaran informasi sebelum menyebarkannya atau mengambil tindakan berdasarkan berita yang diterima (Winarta, Aqsha, and Anoraga 2022).

### **2.1.3 Identifikasi**

Identifikasi adalah proses mengenali, memahami, atau membedakan sesuatu berdasarkan ciri-ciri, karakteristik, atau tanda-tanda khusus yang dimilikinya. Tujuan dari identifikasi adalah untuk mengenali suatu objek, entitas, atau fenomena dengan cara membandingkan ciri-ciri yang ada dengan informasi yang telah diketahui sebelumnya. Identifikasi dapat dilakukan dalam berbagai konteks, termasuk dalam ilmu pengetahuan, teknologi, keamanan, dan lainnya.

Dalam konteks penelitian atau analisis data, identifikasi mengacu pada proses mengenali atau mengklasifikasikan suatu objek atau entitas berdasarkan ciri-ciri atau fitur-fitur tertentu yang dimilikinya. Misalnya, dalam penelitian mengenai deteksi berita *hoax*, identifikasi dapat merujuk pada proses mengenali apakah suatu berita merupakan berita *hoax* atau bukan berdasarkan analisis fitur-fitur teks atau pola tertentu.

Identifikasi juga dapat melibatkan penggunaan teknologi seperti pengenalan pola, algoritma klasifikasi, atau metode analisis lainnya untuk memproses data dan mengambil keputusan tentang apa yang sedang diidentifikasi. Tujuan akhir dari identifikasi adalah untuk memberikan pemahaman atau informasi yang lebih baik tentang objek atau fenomena yang diamati.

#### **2.1.4 *Web Scraping***

*Web scraping* adalah proses ekstraksi informasi atau data dari halaman web secara otomatis. Teknik ini melibatkan penggunaan perangkat lunak atau skrip khusus untuk mengakses dan mengambil data dari situs web, kemudian mengubahnya ke dalam format yang dapat diproses atau dianalisis lebih lanjut, seperti dalam bentuk tabel atau *file teks* (Priadana and Murdiyanto 2020).

Proses *web scraping* dapat melibatkan berbagai langkah, seperti:

1. *HTTP Request*: Perangkat lunak atau skrip akan mengirim permintaan *HTTP* ke server situs web untuk mengakses halaman web yang diinginkan.

2. Pengambilan Konten: Setelah halaman web diakses, perangkat lunak akan mengambil konten *HTML* atau data lain yang ada di dalamnya. Konten ini dapat berupa teks, gambar, tabel, atau elemen-elemen lain.
3. Ekstraksi Data: Data yang diperoleh dari halaman web kemudian diekstraksi menggunakan teknik pemrosesan teks atau analisis struktur *HTML*. Data yang diambil bisa berupa teks, *url*, gambar, tanggal, harga, dan lain-lain.
4. Pembersihan Data: Data yang diambil mungkin perlu diolah lebih lanjut untuk membersihkan atau merapihkannya agar sesuai dengan format yang diinginkan.
5. Penyimpanan Data: Hasil ekstraksi dapat disimpan dalam berbagai format, seperti *file csv*, *excel*, atau *database*.

*Web scraping* memiliki berbagai aplikasi, termasuk:

1. Pengumpulan Data: Mendapatkan informasi seperti harga produk, ulasan pengguna, atau berita dari berbagai situs web.
2. Penelitian: Menarik data untuk tujuan penelitian, analisis pasar, atau riset kompetitor.
3. Pemantauan: Melacak perubahan dalam konten situs web, seperti pemantauan harga, perubahan stok, atau berita terbaru.
4. Integrasi Data: Mengintegrasikan data dari situs web ke dalam aplikasi atau layanan lain.

### 2.1.5 *Dataset Berita*

*Dataset* berita adalah kumpulan data atau informasi yang terkumpul dari berbagai sumber berita, seperti surat kabar, situs web berita, jurnal *online*, atau media lainnya. *Dataset* ini biasanya berisi *teks* berita, termasuk judul, konten, tanggal publikasi, penulis, dan mungkin informasi tambahan seperti kategori berita, *tag*, atau sumber asal (Tambunan et al. 2021) .

*Dataset* berita memiliki tujuan untuk mendukung analisis, riset, pengembangan model, atau aplikasi lain yang memerlukan data berita dalam skala besar. Data ini dapat digunakan untuk berbagai keperluan, seperti:

1. Analisis Sentimen: Menganalisis pandangan publik atau opini dalam berita terkait suatu topik atau peristiwa.
2. Deteksi Berita *Hoax*: Mengembangkan model *machine learning* untuk mengenali atau mengklasifikasikan berita *hoax* dari berita yang sah.
3. Pemantauan Berita: Melakukan pemantauan tren berita, perubahan dalam pemberitaan, atau fokus berita dalam periode waktu tertentu.
4. Klasifikasi Berita: Mengelompokkan berita ke dalam kategori atau topik tertentu, seperti politik, teknologi, olahraga, dll.
5. Analisis Konten: Menganalisis pola, struktur, atau gaya penulisan berita.

### 2.1.6 *Python*

*Python* merupakan bahasa pemrograman tingkat tinggi yang serba guna dan mendukung pemrograman berorientasi objek. Dikenal dengan sintaksisnya yang

mudah dibaca dan dipahami, *Python* dirancang untuk menekankan penulisan kode yang lebih ringkas dibandingkan bahasa pemrograman lainnya. Salah satu ciri utama *python* adalah penggunaan indentasi sebagai bagian integral dari *sintaksisnya*, memberikan kejelasan dan kemudahan dalam membaca kode.

Keunggulan lainnya adalah interpretasinya *python* dapat dieksekusi tanpa proses kompilasi terlebih dahulu. Selain itu, Python mendukung pemrograman berorientasi objek, memungkinkan pembuatan dan penggunaan objek untuk mengelompokkan data dan fungsionalitas terkait. Keberagaman pustaka dan modul Python menjadikannya pilihan yang kuat untuk berbagai keperluan, termasuk pemrosesan *teks*, analisis data, pengembangan web, kecerdasan buatan, dan lainnya. Dengan portabilitasnya yang luas, *Python* dapat dijalankan di berbagai platform seperti *Windows*, *macOS*, dan distribusi *Linux*. Terakhir, dukungan yang besar dan aktif dari komunitas pengembang *python* menyediakan sumber daya, tutorial, dan bantuan yang melimpah (K. R. 2017).

## **2.2 Tinjauan Teori Khusus**

Di dalam sub bab berikut, penulis akan membahas mengenai teori-teori khusus yang berhubungan dengan pelaksanaan penelitian ini yang berupa pengertian-pengertian secara umum.

### **2.2.1 *Machine learning***

*Machine learning* adalah cabang dari kecerdasan buatan (*artificial intelligence*) yang berfokus pada pengembangan algoritma dan model komputer

yang dapat belajar dari data serta melakukan tugas-tugas tertentu tanpa perlu secara eksplisit diprogram. Dalam konsep tradisional pemrograman, akan memberikan aturan dan instruksi kepada komputer untuk melakukan suatu tugas tertentu. Namun, dalam *machine learning*, komputer belajar dari data yang diberikan dan dapat meningkatkan performanya seiring bertambahnya data (Diana, Warni, and Sutabri 2017).

Proses belajar dalam *machine learning* melibatkan identifikasi pola-pola, struktur, dan hubungan dalam data yang dapat digunakan untuk membuat prediksi, pengambilan keputusan, atau tindakan lainnya. Ada beberapa jenis pendekatan dalam *machine learning*, termasuk:

1. *Supervised Learning*: Model *machine learning* dibimbing oleh contoh data yang berisi pasangan input-output. Tujuannya adalah untuk mempelajari hubungan antara input dan output sehingga model dapat melakukan prediksi pada data baru yang belum pernah dilihat.
2. *Unsupervised Learning*: Dalam metode ini, model belajar dari data input tanpa adanya keluaran yang diawasi. Tujuan utamanya adalah untuk mengidentifikasi pola-pola tersembunyi dalam data, seperti pengelompokan (*clustering*) atau reduksi dimensi (*dimensionality reduction*).
3. *Reinforcement Learning*: Model dalam reinforcement learning belajar melalui interaksi dengan lingkungan dan menerima umpan balik dalam bentuk reward atau punishment. Tujuannya adalah untuk memaksimalkan reward dalam pengambilan keputusan berurutan.

4. *Semi-Supervised Learning*: Pendekatan ini menggabungkan unsur-unsur supervised dan unsupervised learning, di mana model memanfaatkan data terlabel dan tidak terlabel untuk meningkatkan performa.
5. *Deep Learning*: Ini adalah subdomain *machine learning* yang menggunakan *neural networks* dengan banyak lapisan (*layers*) untuk mengidentifikasi pola yang kompleks dalam data. *Deep learning* telah menghasilkan kemajuan besar dalam bidang pengenalan gambar, pemrosesan bahasa alami, dan sejumlah aplikasi lainnya.

### **2.2.2 Pustaka Visualisasi Data**

Pustaka visualisasi data (atau library visualisasi data) merujuk pada kumpulan alat, fungsi, dan komponen yang digunakan untuk membuat representasi grafis dari data. Tujuan utama dari pustaka visualisasi data adalah untuk membantu mengubah data yang kompleks menjadi gambar-gambar yang lebih mudah dimengerti dan dapat memberikan wawasan baru. Visualisasi data memungkinkan untuk mengidentifikasi pola, tren, hubungan, dan anomali dalam data dengan lebih jelas (Santosa 2022).

Dalam dunia pemrograman dan analisis data, pustaka visualisasi data memainkan peran penting dalam menerjemahkan informasi yang tersembunyi dalam dataset menjadi bentuk yang dapat diinterpretasikan oleh manusia. Ini membantu dalam eksplorasi data, presentasi hasil, dan pengambilan keputusan.

Pustaka visualisasi data biasanya menyediakan berbagai jenis plot dan grafik, seperti *scatter plots*, *bar plots*, *line plots*, *pie charts*, *histogram*, *heatmap*, dan banyak lagi. Selain itu, mereka sering menyediakan opsi untuk menyesuaikan tampilan visual, seperti pemilihan warna, label, anotasi, dan skala.

Beberapa contoh pustaka visualisasi data populer adalah *Matplotlib*, *Seaborn*, *Plotly*. Penggunaan pustaka-pustaka ini memungkinkan para analis data dan pengembang untuk membuat visualisasi data yang informatif dan mengesankan tanpa harus membangun semuanya dari awal.

### **2.2.3 Paket *Machine learning***

Paket *machine learning* adalah kumpulan perangkat lunak atau modul yang menyediakan alat, fungsi, dan algoritma untuk memfasilitasi pengembangan dan penerapan model pembelajaran mesin. Paket-paket ini dirancang untuk membantu pengembang atau ilmuwan data dalam merancang, melatih, dan menguji model pembelajaran mesin untuk berbagai tugas seperti klasifikasi, *regresi*, pengelompokan, dan lain-lain.

Paket *machine learning* biasanya menyediakan:

1. Algoritma Pembelajaran Mesin: Paket ini menyediakan berbagai algoritma pembelajaran mesin yang telah diimplementasikan, termasuk di dalamnya algoritma klasifikasi, regresi, pengelompokan, deteksi anomali, dan lain-lain.
2. Pustaka Matematika dan Statistik: Banyak paket *machine learning* juga

menyediakan fungsi matematika dan statistik yang diperlukan dalam pemrosesan data dan pembuatan model.

3. Fungsi Prapemrosesan Data: Ini meliputi fitur-fitur untuk normalisasi data, pengisian nilai yang hilang, ekstraksi fitur, dan transformasi data lainnya.
4. Peralatan Validasi Model: Paket ini sering menyertakan alat untuk membagi data menjadi set pelatihan, validasi, dan pengujian, serta metrik evaluasi model seperti akurasi, presisi, *recall*, dan lainnya.
5. Kemampuan Visualisasi: Beberapa paket memiliki kemampuan untuk membuat visualisasi yang membantu dalam pemahaman dan interpretasi model.
6. Kemampuan Optimasi: Beberapa paket juga menyertakan alat untuk melakukan optimasi parameter model guna meningkatkan performa.

#### **2.2.4 Kaggle**

*Kaggle* menjadi platform yang sangat diminati di kalangan data *scientist*, peneliti, dan pengembang untuk berkolaborasi, berbagi, dan menjelajahi proyek data. Salah satu fitur utamanya adalah menyelenggarakan berbagai kompetisi data yang menantang para peserta untuk menyelesaikan analisis data atau membangun model prediksi dengan imbalan hadiah menarik. Platform ini juga menawarkan lingkungan kerja berbasis web untuk membuat dan berbagi *Jupyter Notebooks*, memfasilitasi kolaborasi dan eksplorasi data interaktif.

*Kaggle* tidak hanya menawarkan kompetisi, tetapi juga menyediakan akses ke ribuan dataset publik yang dapat digunakan untuk proyek-proyek individu. Ini mengatasi tantangan pengumpulan data, memungkinkan pengguna untuk fokus pada analisis atau eksplorasi. *Kaggle* juga menyediakan lingkungan eksekusi, disebut *Kernels*, di mana pengguna dapat menjalankan dan berbagi kode secara gratis. Forum diskusi di *Kaggle* menjadi tempat bagi pengguna untuk berinteraksi, berbagi pengetahuan, dan berpartisipasi dalam diskusi seputar data science dan *machine learning*. Terdapat pula *Dataset Playground* untuk eksplorasi pada *dataset* kecil sebelum terlibat dalam proyek lebih besar, serta *Kaggle Courses* yang memberikan kursus online gratis dalam berbagai topik data science dan *machine learning*, membantu pengguna meningkatkan keterampilan mereka. Dengan *dataset*, *kernels*, dan beragam fitur lainnya, *Kaggle* menjadi pusat sumber daya dan kolaborasi yang berharga bagi para profesional data.

### **2.2.5 Algoritma Klasifikasi**

Algoritma klasifikasi dalam *machine learning* merupakan teknik yang digunakan untuk memprediksi kelas atau label suatu data berdasarkan fitur-fitur yang dimilikinya. Tujuan utamanya adalah mengelompokkan data ke dalam kelas atau kategori tertentu berdasarkan pola-pola yang terdapat dalam data pelatihan. Proses klasifikasi melibatkan pembelajaran dari data pelatihan yang sudah dikategorikan sebelumnya, dan kemudian pengetahuan yang diperoleh digunakan untuk mengklasifikasikan data baru yang belum pernah dilihat sebelumnya.

Algoritma ini mencari pola atau relasi dalam data untuk membuat keputusan tentang kategori mana yang paling cocok untuk data baru.

Aplikasi dari algoritma klasifikasi mencakup pengenalan tulisan tangan, analisis sentimen, deteksi penyakit melalui gambar medis, pemisahan spam pada email, dan prediksi harga saham. Beberapa algoritma klasifikasi populer meliputi *Logistic Regression*, *Decision Trees*, *Gradient Boosting* dan *Random Forest* (Sarker 2021).

### 2.2.6 Platform Cloud

Platform cloud untuk *machine learning* adalah infrastruktur dan lingkungan yang disediakan oleh penyedia layanan cloud untuk mendukung pengembangan, pelatihan, dan penerapan model pembelajaran mesin secara efisien dan skalabel. Platform ini memungkinkan pengembang atau ilmuwan data untuk mengakses sumber daya komputasi, alat-alat analisis data, dan lingkungan pengembangan tanpa harus mengelola fisik atau *virtual server* secara langsung (Nafis Alfarizi and Heidiani Ikasari 2023).

Ciri-ciri utama dari platform cloud untuk *machine learning* meliputi:

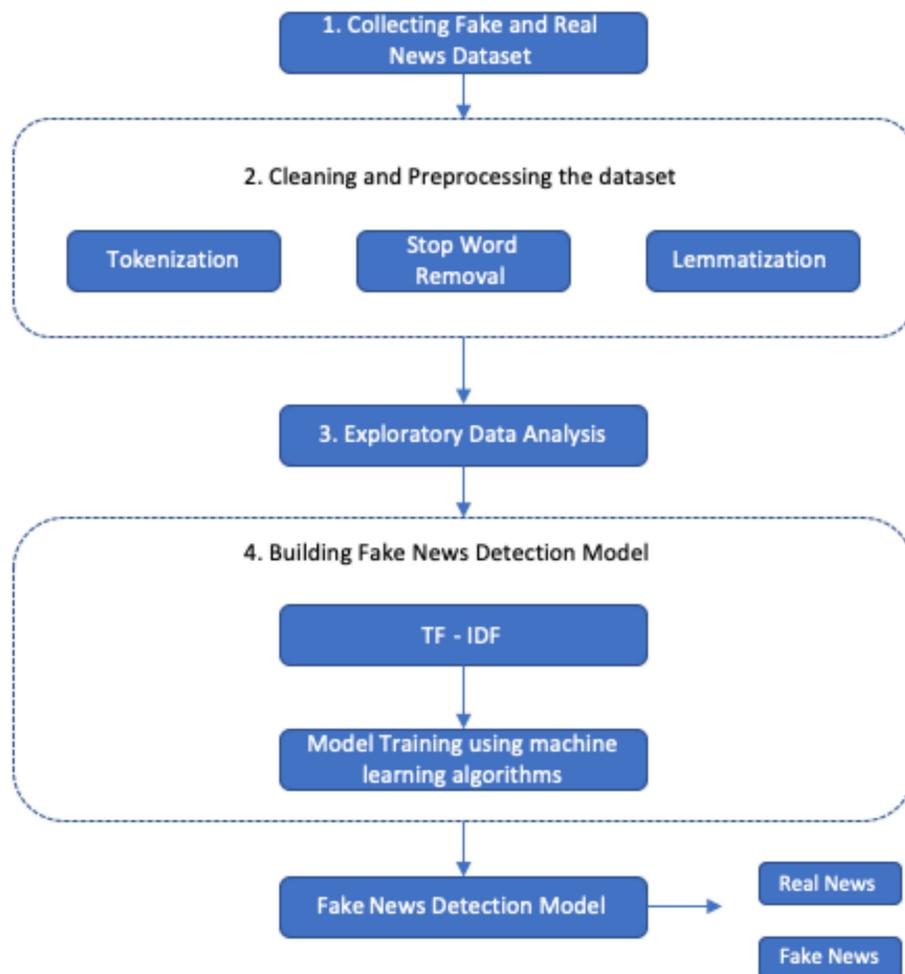
1. Sumber Daya Komputasi: Platform ini menyediakan akses ke sumber daya komputasi yang kuat, seperti *CPU*, *GPU*, atau *TPU*, yang diperlukan untuk melatih model *machine learning* dengan cepat dan efisien.
2. Lingkungan Pengembangan: Platform cloud biasanya menyediakan

lingkungan pengembangan berbasis web yang memungkinkan pengembang untuk menulis, menguji, dan menjalankan kode *machine learning* tanpa perlu mengatur lingkungan lokal.

3. Pustaka dan Framework: Banyak platform cloud memiliki dukungan untuk pustaka dan framework *machine learning* populer seperti *TensorFlow*, *PyTorch*, *Scikit-Learn*, dan lain-lain.
4. Alat Pemodelan: Platform ini seringkali memiliki alat bawaan untuk merancang, melatih, dan mengevaluasi model *machine learning*. Ini bisa termasuk antarmuka *drag-and-drop* untuk merancang alur kerja, alat validasi model, dan alat untuk mengukur kinerja model.
5. Pengelolaan Data: Platform *cloud* menyediakan alat untuk mengelola data pelatihan dan pengujian, serta untuk mengintegrasikan data dari berbagai sumber.

### **2.3 Kerangka Berfikir**

Proses yang umumnya diikuti dalam pengembangan model deteksi berita palsu menggunakan metode *Machine learning*, Berikut adalah penjelasan singkat untuk setiap langkah:



**Gambar 2 1** Kerangka Berfikir

1. Pengumpulan *Dataset* Berita Palsu dan Nyata
  - Ini adalah langkah awal dalam proyek deteksi berita palsu. Mengumpulkan *dataset* yang mencakup berita palsu dan nyata dengan label yang jelas. Pastikan dataset tersebut seimbang dan representatif agar model dapat belajar dengan baik.
2. Pembersihan dan *Pra-pemrosesan Dataset*

- *Tokenization*: Memecah teks menjadi kata-kata atau token.
  - *Stop Word Removal*: Menghapus kata-kata umum yang tidak memberikan banyak informasi seperti "dan", "atau".
  - *Lemmatization*: Mengubah kata-kata ke bentuk dasar untuk mengurangi dimensi dan mempertahankan makna.
3. Analisis Data Eksploratif (*Exploratory Data Analysis - EDA*)
- Melibatkan pemahaman mendalam tentang dataset, distribusi kelas (berita palsu dan nyata), analisis statistik sederhana, dan visualisasi untuk mendapatkan wawasan awal.
4. Pembangunan Model Deteksi Berita Palsu
- *TF-IDF (Term Frequency-Inverse Document Frequency)*: Teknik ekstraksi fitur yang memberikan bobot pada kata-kata berdasarkan seberapa sering mereka muncul dalam dokumen dan seberapa uniknya di seluruh *dataset*.
  - *Model Training Using Machine learning Algorithms*: Memilih dan melatih model *machine learning* seperti *Logistic Regression*, *Decision Tree*, *Gradient Boosting*, dan *Random Forest* menggunakan *dataset* yang telah diproses.
5. Deteksi Model: *Real News* atau *Fake News*
- Setelah model dilatih, mengujinya pada data uji untuk memprediksi apakah suatu berita termasuk dalam kategori berita nyata atau palsu.
  - Model memberikan output berupa probabilitas atau kelas prediksi yang

dapat diinterpretasikan (misalnya, kelas "*Real News*" atau "*Fake News*").

## 2.4 Penelitian Terdahulu

**Tabel 2 1** Penelitian Terdahulu

No	Judul Penelitian	Nama Peneliti	Hasil dari Penelitian
1.	Implementasi Algoritma <i>Naïve Bayes Classifier</i> untuk Mendeteksi Berita Palsu pada Sosial Media	Nova Agustina, Adrian, Mercy Hermawati	Penyebaran berita palsu di dunia maya mengancam tatanan demokratis dan stabilitas sosial. Menurut penelitian Masyarakat Telematika Indonesia, 44,3% responden mengaku sering menerima berita palsu. Kementerian Komunikasi dan Informatika mencatat hingga 11 Agustus 2021, terdapat 1848 laporan berita palsu terkait pandemi Covid-19 dan 290 laporan terkait Vaksin Covid-19. Untuk identifikasi berita palsu di media sosial, digunakan Metode Klasifikasi <i>Naïve Bayes</i> dengan data latihan 700 dari situs Kumparan. Uji coba model

			<p>menunjukkan akurasi 81%.</p> <p>Model diimplementasikan di platform <i>Heroku</i> untuk prediksi berita langsung melalui antarmuka pengguna.</p>
2.	<p>Klasifikasi Berita <i>Hoax</i> Dengan Menggunakan Metode <i>Naïve Bayes</i></p>	<p>Hery Mustofa, Adzhal Arwani Mahfudh</p>	<p><i>Hoax</i> adalah informasi yang tidak akurat atau tidak berdasar. Penyebarannya melalui internet sangat luas, didorong oleh perkembangan teknologi informasi. Untuk mengatasi hal ini, dibutuhkan sistem cerdas yang dapat mengklasifikasikan konten berita <i>hoax</i>. Proses klasifikasi <i>hoax</i> melibatkan preprocessing, pembobotan kata, dan klasifikasi menggunakan metode <i>naive bayes</i>. Pengukuran dilakukan dengan <i>10-fold cross-validation</i>, dan hasilnya menunjukkan tingkat keakuratan</p>

			<p>tertinggi pada <i>fold</i> ke-6 sebesar 85,28%. Dalam uji tersebut, 307 dokumen terklasifikasi sebagai relevan dan 53 dokumen tidak relevan, dengan tingkat kesalahan sebesar 14,72%. Nilai rata-rata precision untuk berita hoax dan benar masing-masing sebesar 0,896 dan <i>recall</i> sebesar 0,853.</p>
3.	<p>Sistem Identifikasi “Fake News” menggunakan Metode <i>Multinomial Naïve Bayes</i></p>	<p>Alifia Zahra, M. Nurkamal Fauzan</p>	<p>Informasi melalui media sosial berkembang pesat, termasuk di platform global seperti <i>Twitter</i>. Meski penyebarannya cepat, berita palsu atau hoax sering muncul. Melalui uji coba menggunakan metode <i>Multinomial Naïve Bayes</i> dengan data dari <i>Twitter</i> (500 tweet), model ini mencapai akurasi 83%, dengan akurasi pelatihan 94%.</p>

4.	Perbandingan Klasifikasi dengan Pendekatan Pembelajaran Mesin untuk Mengidentifikasi <i>Tweet Hoaks</i> di Media Sosial Twitter	Shanto Moyrano Tambunan, Yessica Nataliani, Elizabeth Sri Lestari	<i>Twitter</i> aktif sebagai platform pertukaran informasi. Penelitian ini menggunakan metode pembelajaran mesin ( <i>Stochastic Gradient Descent, Naïve Bayes, Random Forest, dan Rocchio</i> ) untuk mengidentifikasi <i>tweet</i> hoax atau valid. Data terdiri dari 898 <i>tweet</i> dengan kata kunci terkait <i>Corona</i> . Hasil uji menunjukkan bahwa <i>Stochastic Gradient Descent</i> memiliki akurasi terbaik 84.92%. <i>Naïve Bayes</i> memberikan presisi tertinggi 82.95%, sementara <i>recall</i> dan <i>f1-score</i> terbaik berasal dari <i>Stochastic Gradient Descent</i> (85.05% dan 82.42%).
5	Deteksi Berita Palsu Menggunakan	Nur Ghaniaviyanto Ramadhan, faisal	Berita palsu adalah informasi yang disajikan dengan cara yang tidak benar atau menyesatkan.

	Metode <i>Random Forest</i> dan <i>Logistic Regression</i>	Dharma Adhinata, Alon Jaia Tirta Segara, Diovianto Putra Rakhmadani	Penyebaran terus dapat menyebabkan pemahaman yang keliru terhadap informasi. Penelitian ini menggunakan model <i>supervised learning random forest</i> pada dataset berita 6256 judul. Setelah tahap pembersihan, <i>tokenisasi</i> , dan <i>stemming</i> , model <i>random forest</i> mencapai tingkat akurasi 84%, lebih tinggi dari model <i>logistic regression</i> yang mencapai 77%.
6	Komunikasi Perubahan Sosial Masyarakat Dalam Menerima Berita <i>Hoax</i> Di Era Digital	Shofwan Dana Winarta, Gilang Jibriel Aqsha, dan Ficko Aryaduta Prasetya Anoraga	Penelitian ini bertujuan mengevaluasi kemampuan mahasiswa dalam menyaring berita valid di era digital. Dengan mudahnya penyebaran berita, masyarakat perlu mampu memilah informasi yang benar atau hoax. Metode penelitian kualitatif digunakan melalui

			<p>wawancara dengan mahasiswa, fokus pada respons terhadap berita <i>hoax</i>. Hasil wawancara menunjukkan variasi respons, dengan mayoritas mahasiswa cenderung mencari validitas berita sebelum menyebarkannya. Sebanyak 82,4% audiens mengakui pernah menjadi korban berita <i>hoax</i>. Pendapat mereka tentang mudahnya diterimanya berita <i>hoax</i> melibatkan faktor seperti pembacaan singkat, kurangnya daya pikir kritis, dan minimnya pengawasan pemerintah. Kesimpulan menunjukkan bahwa sebagian besar masyarakat masih perlu meningkatkan keterampilan memilah berita.</p>
--	--	--	--

7.	<i>International Journal On Informatics Visualization</i>	Liliek Triyono, Rahmat Gernowo, Prayitno, Mosiur Rahaman, Tri Raharjo Yudiantoro	<i>People need to communicate to meet their needs, and online news plays a crucial role in sharing information. However, filtering out inappropriate content can be challenging. Popular news websites require significant effort to deliver accurate news. To address this, data mining is explored as a solution to extract specific information based on defined attributes. This study aims to identify false news content in Indonesian popular news using various classification methods. The experiment compares five machine learning classifiers independently, with Support Vector Machine (SVM) having the highest accuracy at 83.55% and Decision Tree</i>
----	---	---	--

			<i>Classifier (DTC) the lowest at 75.33%.</i>
--	--	--	---