

BAB III

METODOLOGI PENELITIAN

3.1 Gambaran Umum Objek Penelitian

Objek penelitian ini adalah komentar publik dari 20 kanal YouTube berbahasa Indonesia yang aktif, dipilih secara purposif berdasarkan popularitas dan keterkaitan dengan topik hiburan dan game, yang diketahui rentan menjadi target komentar spam promosi judi online. Dari masing-masing kanal, diambil 50 video terakhir hingga cutoff tanggal 9 Maret 2025, dengan total sebanyak 1.000 video. Dataset ini dipilih untuk merepresentasikan keberagaman konten dan gaya komunikasi pengguna dalam konteks lokal Indonesia. Menurut laporan We Are Social (2024), pengguna YouTube di Indonesia didominasi oleh kelompok usia 18–34 tahun, dengan tingkat keterlibatan komentar yang tinggi pada konten hiburan, musik, dan gaming. Target usia tersebut juga merupakan kelompok yang paling rentan terhadap paparan iklan atau promosi aktivitas ilegal seperti perjudian daring. Oleh karena itu, analisis komentar pada kanal-kanal dalam domain ini menjadi relevan baik dari sisi dampak sosial maupun urgensi moderasi. Dari sisi etika dan legalitas, YouTube secara eksplisit melarang konten dan komentar yang mempromosikan layanan perjudian melalui YouTube Community Guidelines. Komentar yang melanggar termasuk promosi situs taruhan, link referral, maupun komentar terselubung seperti penggunaan simbol (e.g., “s/l/o/t”). Penelitian ini tidak bertujuan untuk menyebarluaskan atau mereplikasi komentar tersebut, melainkan untuk mengembangkan sistem deteksi otomatis yang dapat membantu platform atau pengelola kanal dalam memoderasi komentar berisiko.

Dengan demikian, objek penelitian ini tidak hanya relevan secara teknis, tetapi juga penting dari sisi etika dan sosial dalam mendukung ekosistem digital yang aman dan sehat di Indonesia.

Komentar yang dimaksud dapat berupa ajakan bergabung ke situs perjudian, penyebaran tautan yang mengarah ke platform perjudian, maupun promosi hadiah dan bonus yang terkait dengan aktivitas tersebut. Jenis komentar ini secara langsung melanggar kebijakan komunitas YouTube dan berpotensi membahayakan pengguna lain, terutama kelompok yang rentan.

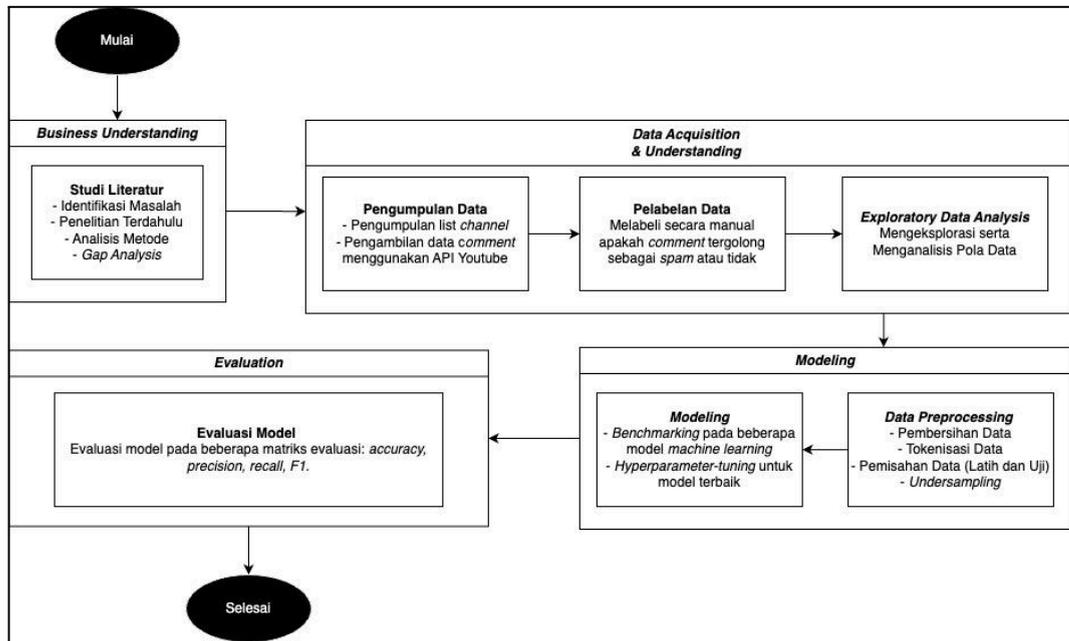
Moderasi komentar secara manual dalam skala besar menjadi tidak efektif, mengingat tingginya jumlah video serta volume komentar yang terus bertambah setiap hari. Oleh karena itu, diperlukan suatu sistem berbasis *machine learning* yang mampu melakukan moderasi komentar secara otomatis dan efisien. Penelitian ini difokuskan pada pengembangan model klasifikasi berbasis *machine learning* yang bertujuan untuk membedakan antara komentar yang mengandung unsur perjudian dan komentar normal. Model yang dikembangkan diharapkan dapat meningkatkan efektivitas moderasi komentar pada platform YouTube, sehingga menciptakan lingkungan digital yang lebih aman dan bersih dari aktivitas ilegal.

Dataset yang digunakan diperoleh melalui proses pengumpulan data (*crawling*) dari berbagai video YouTube yang relevan, dengan menggunakan kata kunci yang sering diasosiasikan dengan aktivitas perjudian seperti "slot", "jackpot", "casino", "bonus besar", dan istilah sejenis. Seluruh komentar yang dikumpulkan kemudian diberi label (*annotasi*) secara manual untuk membedakan antara komentar perjudian dan non-perjudian.

3.2 Metode Penelitian

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

3.2.1 Alur Penelitian



Gambar 3.1 Alur penelitian

Alur penelitian ini diawali dengan tahapan *business understanding* yang bertujuan untuk memahami permasalahan inti yang mendasari studi. Pada tahap ini, dilakukan studi literatur untuk mengidentifikasi isu utama terkait komentar spam pada platform YouTube. Proses ini mencakup peninjauan penelitian terdahulu, analisis terhadap pendekatan yang telah digunakan, serta penyusunan *gap analysis* untuk menemukan celah penelitian yang belum banyak dieksplorasi, sehingga dapat dijadikan landasan pengembangan metode yang lebih tepat sasaran.

Tahap berikutnya adalah *data acquisition & understanding*, yang mencakup proses pengumpulan dan pemahaman data. Data dikumpulkan dengan memanfaatkan API YouTube untuk mengakses komentar dari berbagai channel yang telah dipilih. Setelah data berhasil diperoleh, dilakukan pelabelan secara manual untuk menentukan apakah setiap komentar termasuk ke dalam kategori spam atau tidak. Selain itu, dilakukan analisis eksploratif terhadap data guna memahami pola distribusi, kecenderungan, serta karakteristik umum dari komentar yang terkumpul.

Setelah data dipahami, tahapan *modeling* dimulai dengan proses *data preprocessing*. Langkah ini meliputi pembersihan data, tokenisasi, serta pemisahan data menjadi subset pelatihan dan pengujian. Untuk mengatasi ketidakseimbangan data, diterapkan teknik *undersampling* pada kelas mayoritas. Selanjutnya, dilakukan *benchmarking* terhadap beberapa algoritma *machine learning* untuk menentukan pendekatan yang paling efektif. Proses *hyperparameter tuning* juga diterapkan guna mengoptimalkan performa dari masing-masing model yang diuji.

Tahap akhir adalah *evaluation*, di mana model yang telah dilatih dievaluasi menggunakan berbagai metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score* guna menilai sejauh mana model mampu mengklasifikasikan komentar secara akurat. Hasil evaluasi ini menjadi dasar dalam menentukan model terbaik yang dapat digunakan untuk mendeteksi komentar spam secara otomatis. Dengan mengikuti alur penelitian ini, diharapkan sistem yang dikembangkan mampu memberikan kontribusi nyata dalam penyaringan komentar spam di platform digital secara efektif dan efisien.

3.2.2 Metode Data Mining

TDSP (*Team Data Science Process*) dipilih sebagai kerangka kerja dalam penelitian ini karena kemampuannya dalam menyediakan pendekatan terstruktur dan kolaboratif untuk pengembangan solusi berbasis *machine learning*. TDSP dikembangkan oleh Microsoft sebagai metodologi yang mencakup seluruh siklus hidup proyek data science, mulai dari pemahaman masalah hingga implementasi model. Dibandingkan dengan metode lain seperti SEMMA atau KDD, TDSP tidak hanya menekankan aspek teknis, tetapi juga memfasilitasi integrasi antara kebutuhan bisnis, pengolahan data, dan deployment model secara praktis dalam lingkungan produksi. Dalam konteks penelitian ini, di mana tujuan utamanya adalah mengembangkan model klasifikasi untuk memoderasi komentar terkait judi online di platform YouTube, TDSP menjadi kerangka yang tepat karena mampu mengakomodasi kebutuhan eksplorasi data teks,

perancangan model klasifikasi, serta evaluasi dan penerapannya secara sistematis. Dengan tahapan-tahapan yang jelas dan terintegrasi, TDSP memastikan bahwa proses data mining dalam penelitian ini berjalan konsisten dan terarah. Berikut adalah penjabaran masing-masing tahapan TDSP sebagaimana diimplementasikan dalam penelitian ini.

3.2.2.1 Business Understanding

Tahap *Business Understanding* merupakan fondasi awal dalam proses pengembangan model klasifikasi berbasis *machine learning* pada kerangka TDSP untuk moderasi komentar promosi judi online di YouTube. Permasalahan utama yang diangkat dalam penelitian ini berkaitan dengan maraknya komentar yang mengandung unsur promosi judi online di kolom komentar video YouTube, yang tidak hanya meresahkan masyarakat tetapi juga berpotensi melanggar regulasi konten digital. Meskipun YouTube memiliki sistem moderasi otomatis, keberadaan komentar dengan penyamaran kata-kata dan pola bahasa yang tidak eksplisit sering kali lolos dari deteksi sistem bawaan platform tersebut. Hal ini menunjukkan perlunya pendekatan yang lebih adaptif dan spesifik dalam mengidentifikasi konten komentar semacam ini.

Untuk memahami permasalahan secara komprehensif, dilakukan studi literatur terhadap berbagai penelitian terdahulu yang membahas deteksi komentar spam dan klasifikasi teks berbasis machine learning. Studi ini mencakup pembahasan mengenai algoritma yang umum digunakan, teknik preprocessing data teks, serta strategi penanganan ketidakseimbangan kelas. Dari hasil kajian tersebut, ditemukan bahwa sebagian besar penelitian masih berfokus pada klasifikasi spam secara umum dan belum banyak yang menyoroti kasus khusus seperti promosi judi online yang memiliki karakteristik unik dalam penyamaran bahasa.

Sejalan dengan itu, dilakukan analisis metode terhadap beragam algoritma klasifikasi yang akan digunakan dalam penelitian, mencakup pendekatan linear seperti Logistic Regression dan Support Vector Machine, metode probabilistik seperti Naïve Bayes, serta pendekatan berbasis pohon keputusan dan ensemble seperti Decision Tree, Random Forest, XGBoost, CatBoost, dan AdaBoost. Selain itu, model berbasis jaringan saraf yaitu Multi-Layer Perceptron (MLP) juga dipertimbangkan untuk memberikan variasi pendekatan dari sisi deep learning. Pemilihan metode ini mempertimbangkan karakteristik data teks komentar yang bersifat tidak terstruktur serta potensi adanya distribusi kelas yang tidak seimbang.

Berdasarkan hasil kajian tersebut, dilakukan *gap analysis* untuk mengidentifikasi celah penelitian yang belum banyak dijelajahi. Ditemukan bahwa masih terbatasnya penelitian yang secara khusus menangani komentar promosi judi online di platform digital membuka peluang untuk mengembangkan solusi klasifikasi yang lebih kontekstual. Oleh karena itu, penelitian ini difokuskan pada pembangunan model klasifikasi yang tidak hanya mampu mengenali pola umum komentar spam, tetapi juga dapat mengidentifikasi komentar yang secara eksplisit maupun implisit memuat promosi judi online, dengan pendekatan berbasis data dan pengujian berbagai algoritma machine learning secara sistematis.

3.2.2.2 Data Acquisition & Understanding

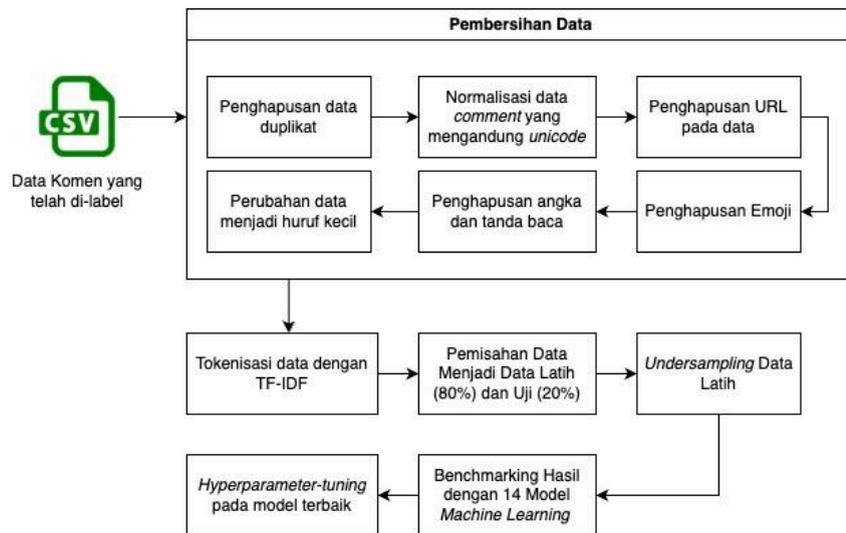
Tahapan *Data Acquisition & Understanding* dalam TDSP merupakan proses krusial dalam menyiapkan data yang akan digunakan sebagai fondasi bagi pengembangan model klasifikasi. Proses ini diawali dengan pengumpulan data, yang dilakukan dengan menyusun daftar channel YouTube sebagai objek observasi. Populasi dari penelitian ini mencakup seluruh channel YouTube yang terbuka untuk umum, namun untuk kepentingan efisiensi dan keterwakilan, dipilih

sebanyak 20 channel sebagai sampel. Dari masing-masing channel, diambil 50 video terbaru yang telah diunggah hingga tanggal 9 Maret 2025. Seluruh data komentar dikumpulkan secara otomatis menggunakan YouTube Data API, yang memungkinkan ekstraksi komentar dari video secara terstruktur dan efisien.

Setelah data berhasil dikumpulkan, dilakukan tahap pelabelan data secara manual. Setiap komentar yang diperoleh dianalisis dan diklasifikasikan ke dalam dua kategori, yaitu komentar spam (yang mengandung promosi judi online) dan non-spam. Proses pelabelan ini dilakukan secara cermat guna memastikan keakuratan klasifikasi dan validitas data latih yang akan digunakan pada tahap pemodelan. Kriteria pelabelan ditentukan berdasarkan keberadaan kata kunci tertentu, konteks komentar, serta pola bahasa yang biasa digunakan dalam promosi aktivitas perjudian secara daring.

Langkah selanjutnya adalah melakukan *exploratory data analysis* (EDA) terhadap komentar yang telah dilabeli. Tahapan ini bertujuan untuk mengeksplorasi dan memahami karakteristik dasar dari data yang diperoleh, termasuk distribusi kelas, frekuensi kata, panjang komentar, dan keberagaman struktur bahasa. Melalui proses ini, diperoleh wawasan awal mengenai pola-pola yang muncul pada komentar promosi judi online, serta potensi tantangan seperti ketidakseimbangan kelas (*class imbalance*) yang perlu ditangani secara khusus pada tahapan preprocessing dan modeling. Dengan demikian, tahapan ini memberikan dasar yang kuat bagi proses pengembangan model klasifikasi yang efektif dan relevan dengan konteks permasalahan.

3.2.2.3 Modeling



Gambar 3.2. Diagram Modelling

Tahapan ketiga dalam kerangka TDSP yaitu *Data Preprocessing & Modeling* yang merupakan inti dari proses pengembangan model klasifikasi berbasis machine learning dalam penelitian ini. Proses dimulai dengan pembersihan data teks komentar yang telah dilabeli. Tujuan utama dari tahap ini adalah untuk menghilangkan elemen-elemen yang tidak relevan dan berpotensi mengganggu performa model. Langkah-langkah pembersihan data meliputi penghapusan data duplikat, normalisasi karakter unicode, penghapusan URL, emoji, angka, dan tanda baca, serta konversi semua huruf menjadi huruf kecil agar tidak terjadi redundansi fitur akibat perbedaan kapitalisasi.

Setelah data dibersihkan, dilakukan proses tokenisasi menggunakan metode TF-IDF (Term Frequency–Inverse Document Frequency), yang berfungsi untuk mengubah data teks menjadi representasi numerik berbasis bobot kata. Representasi ini memungkinkan model machine learning untuk mengenali pentingnya sebuah kata dalam konteks lokal (komentar) dan global (seluruh dataset). Selanjutnya, data dibagi menjadi data latih (80%) dan data uji (20%) guna menjaga validitas pengujian model.

Menghadapi potensi ketidakseimbangan kelas dalam data latih, diterapkan teknik undersampling, yaitu pengurangan sampel dari kelas mayoritas agar distribusi antar kelas menjadi lebih seimbang. Langkah ini penting untuk memastikan bahwa model tidak bias terhadap kelas dominan dan mampu mengenali pola pada kelas minoritas secara efektif. Proses *modeling* dilakukan dengan pendekatan *benchmarking* terhadap 11 algoritma machine learning yang berbeda, termasuk metode linear, probabilistik, pohon keputusan, ensemble learning, dan neural network. Tujuannya adalah untuk mengevaluasi dan membandingkan performa masing-masing algoritma dalam menyelesaikan tugas klasifikasi komentar. Seluruh proses pelatihan dan evaluasi model dilakukan menggunakan platform Google Colab, yang memungkinkan pemrosesan berbasis cloud dengan dukungan GPU untuk efisiensi komputasi. Untuk meningkatkan kinerja model, dilakukan pula hyperparameter tuning pada algoritma terpilih, sehingga model akhir yang digunakan merupakan hasil optimasi dari serangkaian eksperimen yang sistematis.

3.2.2.4 Evaluation

Tahap *Evaluation* dalam kerangka TDSP bertujuan untuk mengukur kinerja model klasifikasi yang telah dibangun, guna memastikan bahwa model mampu membedakan komentar yang mengandung promosi judi online dengan akurasi dan konsistensi yang tinggi. Evaluasi dilakukan terhadap hasil prediksi model pada data uji yang telah dipisahkan sebelumnya, dengan menggunakan sejumlah metrik evaluasi standar dalam klasifikasi teks.

Metrik yang digunakan meliputi accuracy, yang menunjukkan proporsi prediksi yang benar terhadap keseluruhan data uji; precision, yang mengukur seberapa tepat model dalam mengidentifikasi komentar spam (berapa banyak dari komentar yang diprediksi spam benar-benar spam); serta recall, yang mencerminkan kemampuan model dalam

menangkap seluruh komentar spam yang sebenarnya. Selain itu, digunakan pula F1-score, yaitu rata-rata harmonis dari precision dan recall, untuk memberikan gambaran yang lebih seimbang terhadap performa model, khususnya dalam kondisi distribusi kelas yang tidak seimbang.

Pemilihan metrik-metrik ini didasarkan pada karakteristik data dan kebutuhan klasifikasi yang sensitif terhadap kesalahan klasifikasi komentar spam, sehingga evaluasi tidak hanya mempertimbangkan ketepatan umum (*accuracy*), tetapi juga memperhatikan dampak dari kesalahan dalam identifikasi kelas minoritas. Hasil evaluasi ini digunakan sebagai dasar dalam membandingkan performa antar model dan menentukan algoritma terbaik yang akan digunakan sebagai solusi akhir dalam sistem moderasi komentar YouTube.

3.3 Teknik Pengumpulan Data

Data komentar YouTube dikumpulkan menggunakan YouTube Data API v3 dengan bantuan skrip Python yang dibuat secara khusus. Komentar dikumpulkan dari 20 kanal YouTube Indonesia dengan masing-masing mengambil 50 video terbaru berdasarkan urutan waktu publikasi, hingga cutoff tanggal 9 Maret 2025. Untuk menjaga efisiensi dan mematuhi batas rate-limit API, digunakan token autentikasi OAuth 2.0 resmi dengan batas maksimum 10.000 unit kuota per hari. Setiap request terhadap endpoint komentar mengonsumsi sekitar 1 unit per video. Seluruh proses pengumpulan dilakukan melalui API resmi, bukan melalui scraping langsung ke halaman web, sehingga tetap mematuhi *Terms of Service* (ToS) yang ditetapkan oleh Google Developer Console dan YouTube.

Populasi dalam penelitian ini mencakup seluruh channel YouTube yang secara aktif mempublikasikan konten dan memungkinkan interaksi pengguna melalui kolom komentar. Namun, untuk menjaga fokus penelitian dan keterkelolaan data, dilakukan pemilihan sampel sebanyak 20 channel yang dipilih secara purposive berdasarkan tingkat interaksi, jumlah penonton, dan potensi

keberadaan komentar promosi judi online. Dari setiap channel yang telah di sampel, diambil 50 video terbaru yang tersedia per tanggal 9 Maret 2025 sebanyak 1.071.028 komentar sebagai objek observasi.

Seluruh komentar dari video-video tersebut dikumpulkan menggunakan skrip Python yang terhubung dengan API, sehingga proses pengambilan data dapat dilakukan secara efisien dan terstruktur. Komentar yang dikumpulkan mencakup komentar utama dan komentar balasan, untuk memastikan representasi bahasa dan konteks yang lebih lengkap. Dengan demikian, seluruh data yang digunakan dalam penelitian ini merupakan komentar yang tersedia hingga 9 Maret 2025, dan tidak mencakup komentar yang dipublikasikan setelah tanggal tersebut.

3.4 Teknik Analisis Data

Berdasarkan metodologi yang telah ditentukan dalam penelitian ini, diperlukan tools atau platform yang mampu mendukung proses implementasi model klasifikasi secara efisien dan terintegrasi. Dalam hal ini, teknik analisis data dilakukan dengan menggunakan Google Colab sebagai platform utama. Google Colab dipilih karena merupakan lingkungan pemrograman berbasis cloud yang mendukung eksekusi kode Python secara interaktif, serta menyediakan akses ke GPU tanpa perlu instalasi lokal, sehingga sangat sesuai untuk kebutuhan

pemrosesan data skala besar dan Analisis data dalam penelitian ini mencakup beberapa tahap, dimulai dari preprocessing, representasi teks, teknik balancing kelas, hingga evaluasi model klasifikasi. Langkah preprocessing meliputi normalisasi, penghapusan simbol, penghilangan stopword, dan tokenisasi. Data kemudian diubah ke dalam bentuk numerik menggunakan teknik TF-IDF (Term Frequency–Inverse Document Frequency), yang terbukti efektif dalam representasi fitur teks untuk klasifikasi spam [6].

Untuk mengatasi ketidakseimbangan kelas antara komentar spam judi dan non-spam, digunakan pendekatan undersampling terhadap kelas mayoritas. Teknik SMOTE (Synthetic Minority Oversampling Technique) dipertimbangkan, namun tidak digunakan karena rawan menghasilkan data sintesis yang tidak realistis dalam konteks bahasa alami dan dapat menimbulkan overfitting pada dataset

kecil. Pemilihan metrik evaluasi menggunakan F1-score makro dilakukan karena dataset memiliki distribusi kelas yang tidak seimbang, dan F1 makro memberi bobot setara ke setiap kelas, sehingga lebih adil dalam mengevaluasi performa deteksi terhadap kelas minoritas (spam judi). Selain F1, metrik akurasi, precision, recall, dan confusion matrix juga dianalisis untuk mendukung interpretasi performa model secara menyeluruh. Sebagai bagian dari validasi hasil, dilakukan uji signifikansi statistik menggunakan McNemar Test, untuk menguji apakah perbedaan performa antar model klasifikasi signifikan secara statistik atau hanya terjadi karena variasi data. Uji ini diterapkan pada hasil prediksi dua model teratas terhadap dataset uji yang sama.pelatihan model *machine learning*.

Penggunaan Google Colab memungkinkan peneliti untuk mengelola seluruh alur penelitian secara terpadu, mulai dari proses preprocessing data, pemodelan, evaluasi, hingga visualisasi hasil. Selain itu, Google Colab juga kompatibel dengan berbagai pustaka Python populer seperti Scikit-learn, Pandas, NumPy, dan Matplotlib, yang semuanya diperlukan dalam pelaksanaan analisis dan pengujian model. Keunggulan lain dari Google Colab adalah kemampuannya untuk terhubung langsung dengan Google Drive, yang sangat membantu dalam menyimpan dan mengelola dataset maupun model hasil pelatihan secara real time.

Dengan demikian, pemanfaatan Google Colab dalam penelitian ini tidak hanya memberikan efisiensi komputasi, tetapi juga mendukung keterulangan eksperimen dan dokumentasi proses secara sistematis. Hal ini menjadikan Google Colab sebagai pilihan yang tepat untuk melakukan analisis data dalam penelitian ini, khususnya dalam konteks implementasi individu yang memerlukan fleksibilitas tinggi dan aksesibilitas lintas perangkat.

Tabel 3.1 Perbandingan Alat Penelitian

Aspek Pembeding	Google Colab	Jupyter Notebook	Apache Zeppelin
Kemampuan untuk Diakses	Dapat diakses dari browser secara gratis tanpa instalasi	Perlu instalasi lokal atau konfigurasi server JupyterHub	Diperlukan instalasi dan setup server (lebih kompleks)
Kemampuan untuk Menjalankan Kode	Mendukung eksekusi kode Python secara interaktif + GPU/TPU	Mendukung eksekusi kode Python secara interaktif	Mendukung multi-bahasa (Python, Scala, SQL), interaktif
Sumber Daya Komputasi	Gratis GPU/TPU dari cloud Google (terbatas tapi memadai)	Bergantung pada spesifikasi perangkat lokal	Bergantung pada server/cluster yang dikonfigurasi sendiri
Kemampuan Eksekusi	Eksekusi cepat dan stabil untuk data skala menengah	Cepat jika spesifikasi lokal mendukung	Efisien untuk analisis terdistribusi, namun kompleks untuk setup awal

