

BAB III

METODE PENELITIAN

3.1. Metode Penelitian

Penelitian ini menerapkan metodologi *CRISP-DM* karena pengolahan transkrip mengenai budidaya tanaman cabai memerlukan langkah-langkah yang teratur, sistematis, dan terstruktur dengan baik. *CRISP-DM* terdiri dari enam tahapan utama, yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, serta penerapan. Pendekatan ini sangat sesuai untuk menjamin bahwa setiap proses pengelolaan data berlangsung secara jelas, terorganisir, dan dapat dipertanggungjawabkan. Tahapan pemahaman data dan persiapan data akan menjadi fokus utama dalam mengumpulkan serta membersihkan transkrip video, sementara tahapan pemodelan dan evaluasi akan dimanfaatkan untuk membangun serta menguji kualitas soal *multiple choice questions* yang dihasilkan. Metode ini memberikan fleksibilitas yang mirip dengan prinsip *Agile*, tetapi dengan penekanan lebih kuat pada aspek data. Hal tersebut memungkinkan pengembangan sistem dilakukan secara bertahap, dengan kemampuan melakukan penyesuaian berdasarkan hasil evaluasi di setiap siklus. Dengan demikian, kualitas serta relevansi soal dapat terus dijaga agar tetap sesuai dengan kebutuhan pembelajaran di Desa Pete.

3.1.1 Perbandingan Metode

Tabel 3. 1 Tabel Perbandingan Metode

Aspek	Crisp DM [33]	SEMMA [51]	KDD [52]
Definisi	Pendekatan standar dalam proses penambangan data yang terdiri atas enam tahapan utama: pemahaman konteks bisnis, pemahaman data, persiapan data,	Metode yang mengintegrasikan struktur graf dengan elemen semantik teks untuk memprediksi hubungan antar entitas dalam Knowledge Graphs.	Proses yang meliputi eksplorasi data mendalam, seleksi fitur, pembuatan model pembelajaran, serta penilaian dan analisis

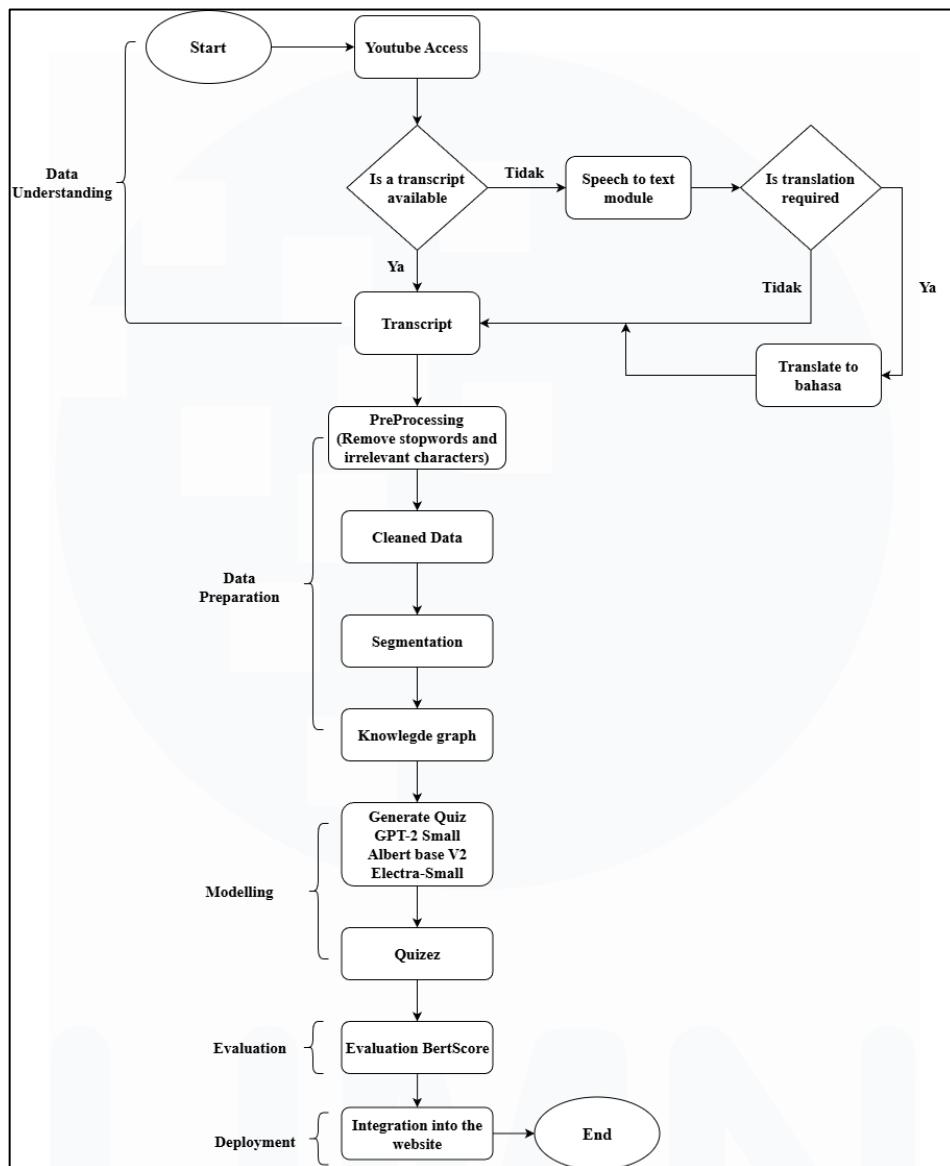
Aspek	Crisp DM [33]	SEMMA [51]	KDD [52]
	pemodelan, evaluasi, serta penerapan hasil.		terhadap hasil akhir.
Fokus Utama	Menyediakan kerangka kerja yang terstruktur untuk eksplorasi serta analisis data dalam lingkup yang lebih komprehensif.	Menekankan pada penggabungan semantik dalam struktur graf serta pembelajaran zero-shot untuk prediksi hubungan.	Berfokus pada penemuan pengetahuan dari dataset besar melalui penggunaan model pembelajaran mesin serta teknik optimasi.
Keterlibatan Pengguna	Pengguna aktif terlibat di setiap tahap dengan menyampaikan masukan serta keputusan untuk menjamin kecocokan data dengan tujuan.	Manfaatkan teknik berbasis teks untuk memperdalam pemahaman tanpa memerlukan interaksi pengguna secara intensif.	Pengguna berperan dalam pengumpulan data serta memberikan umpan balik selama tahap evaluasi dan implementasi.
Kecepatan Pengembangan	Bersifat iteratif dan bertahap, dengan kemungkinan penyesuaian berdasarkan masukan pengguna, sehingga memungkinkan proses pengembangan yang relatif cepat.	Lebih mengutamakan pengembangan mendalam pada aspek struktur dan teks untuk mencapai prediksi yang lebih presisi.	Kecepatan tergantung pada efisiensi metode pembelajaran dan evaluasi yang diterapkan.
Skala	Dapat diterapkan pada proyek dengan berbagai ukuran, mulai dari skala kecil hingga besar, dengan penekanan pada	Dirancang khusus untuk sistem berskala besar yang memerlukan pengolahan data dalam volume	Umumnya digunakan untuk analisis data berskala besar dengan penekanan pada pemodelan yang kompleks.

Aspek	Crisp DM [33]	SEMMA [51]	KDD [52]
	proses yang terorganisir.	yang sangat tinggi.	
Penerapan	Telah banyak digunakan di berbagai sektor industri untuk analisis serta pemahaman data secara umum.	Lebih sesuai untuk aplikasi yang melibatkan analisis Knowledge Graph serta sistem rekomendasi berbasis data.	Sering dipakai dalam analisis data yang bersifat teknis dan mendalam, seperti di bidang kesehatan atau ilmu pengetahuan.
Respon terhadap Perubahan	Mampu beradaptasi dengan perubahan kebutuhan atau data baru melalui pendekatan iteratif yang fleksibel.	Memungkinkan penyesuaian serta adaptasi terhadap data yang bersifat dinamis, namun tetap lebih terfokus pada struktur data.	Cenderung lebih kaku dalam menangani perubahan secara langsung, meskipun terdapat fleksibilitas pada tahap evaluasi.

Kesimpulan dari perbandingan antara *CRISP-DM*, *SEMMA*, dan *KDD* menunjukkan bahwa *CRISP-DM* memiliki keunggulan lebih besar dalam hal fleksibilitas serta kemampuan beradaptasi terhadap perubahan, sehingga menjadi pilihan yang paling sesuai untuk penelitian ini yang bertujuan menyusun soal pilihan ganda berdasarkan transkrip budidaya tanaman cabai. Pendekatan ini mendukung proses yang bersifat iteratif, terstruktur, dan melibatkan partisipasi aktif pengguna pada setiap tahapannya, yang sangat krusial untuk menjamin bahwa soal yang dihasilkan benar-benar sesuai dengan kebutuhan proses pembelajaran. Walaupun *SEMMA* dan *KDD* lebih unggul dalam penanganan dataset berskala besar serta analisis prediktif, *CRISP-DM* justru menyediakan kerangka kerja yang lebih relevan untuk penelitian yang mengutamakan evaluasi berkelanjutan serta penyesuaian secara terus-menerus terhadap data dan konteks lokal.

3.2 Tahapan Penelitian

Selanjutnya, berikut adalah tahapan *CRISP-DM* yang akan diterapkan dalam penelitian ini:



Gambar 3. 1 Alur Penelitian

Pada Gambar 3.1, diagram alur tersebut mengilustrasikan keseluruhan langkah pembuatan soal secara otomatis dengan menerapkan kerangka *CRISP-DM*. Proses diawali dengan mengambil materi dari video YouTube sebagai sumber utama. Kemudian, dilakukan pengecekan ketersediaan transkrip; apabila sudah ada, proses langsung dilanjutkan ke tahap pra-pemrosesan, yaitu pembersihan data melalui penghapusan *stopwords* serta elemen-elemen yang tidak diperlukan. Setelah data berhasil dibersihkan, tahap berikutnya adalah segmentasi teks, di mana teks dipisahkan menjadi paragraf-paragraf yang lebih

terstruktur. Selanjutnya, dibuat *knowledge graph* guna memvisualisasikan hubungan antar konsep penting dalam teks tersebut. Setelah data terorganisir dengan baik, model *Small Language Model* seperti *GPT-2* dimanfaatkan untuk menghasilkan soal MCQ yang didasarkan pada pemahaman mendalam terhadap teks yang telah diringkas. Soal-soal yang tercipta kemudian dinilai menggunakan metrik *BERTScore* untuk memverifikasi kualitas serta kesesuaianya. Seluruh rangkaian ini mencakup keseluruhan tahapan *CRISP-DM*, mulai dari pemahaman data, persiapan data, pembuatan model, hingga evaluasi, dengan mengikuti enam fase pengembangan yang dimiliki metodologi *CRISP-DM*.

3.2.1 Business Understanding

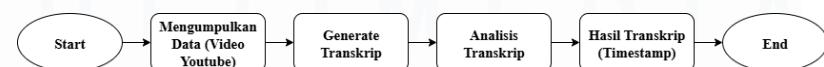
Pada tahap *Business Understanding* ini, penelitian difokuskan untuk memahami sasaran utama dari sistem yang akan dibangun, yakni menciptakan soal MCQ yang relevan dan selaras dengan kurikulum pendidikan di Desa Pete. Dari perspektif bisnis, tujuan utamanya adalah meningkatkan mutu pembelajaran di desa melalui alat evaluasi yang mampu mengukur sekaligus mengembangkan kemampuan berpikir kritis pada siswa sekolah dasar. Untuk mewujudkan hal tersebut, sistem harus mampu mengonversi transkrip materi pembelajaran, seperti konten budidaya tanaman cabai, menjadi pertanyaan yang menantang serta merangsang kemampuan analisis siswa. Dengan memahami secara mendalam kebutuhan pendidikan setempat beserta berbagai hambatan yang ada, tahap ini memastikan bahwa solusi yang dihasilkan tidak hanya memenuhi target pembelajaran, tetapi juga dapat diterima dan dimanfaatkan secara optimal oleh para guru serta siswa di lapangan.

3.2.2 Data Understanding

Tahap *Data Understanding* bertujuan untuk mendapatkan pemahaman yang komprehensif mengenai data yang menjadi bahan utama penelitian ini, yaitu tiga buah video YouTube yang membahas teknik budidaya tanaman cabai. Ketiga video tersebut masing-masing berdurasi 7

menit 8 detik, 9 menit 36 detik, dan 15 menit 22 detik. Durasi keseluruhan transkrip yang dihasilkan sudah memadai untuk membentuk dataset dasar. Proses pembuatan transkrip dilakukan dengan bantuan ekstensi *Sider AI* pada peramban *Google Chrome*. Peneliti membuka video, mengaktifkan ekstensi tersebut, lalu menyalin teks transkrip yang muncul secara langsung beserta *timestamp*-nya. Hasil transkrip dari ketiga video tersebut menghasilkan sekitar 2.000 baris teks dengan lebih dari 1.700 kata unik, yang mencakup berbagai topik seperti pemilihan media tanam, pemupukan, teknik penyiraman, pengendalian hama, serta perawatan tanaman cabai secara keseluruhan.

Setelah transkrip diperoleh, peneliti melakukan eksplorasi awal dan pembersihan data guna memastikan kualitas serta relevansi informasi. Dari keseluruhan teks tersebut, dipilih 10 kalimat yang mengandung informasi kunci sebagai dasar penyusunan soal. Kalimat-kalimat terpilih ini kemudian dipasangkan secara manual dengan pertanyaan untuk membentuk pasangan konteks-pertanyaan, sehingga dijamin bahwa soal yang dihasilkan benar-benar menguji kemampuan analisis dan pemahaman mendalam siswa. Dalam setiap penggunaan, *generator* dirancang untuk memproduksi maksimal lima soal MCQ sesuai dengan jenjang kelas dan topik yang dipilih. Evaluasi secara langsung oleh para guru terhadap soal-soal yang dihasilkan akan dilakukan pada saat presentasi sistem di Desa Pete. Semua soal akan dinilai berdasarkan aspek kualitas, kejelasan, serta kesesuaian dengan materi lokal. Dengan demikian, tahap *Data Understanding* menjamin bahwa data yang digunakan relevan, berkualitas tinggi, dan siap mendukung proses pemodelan untuk menghasilkan soal MCQ yang kontekstual dan bermakna.

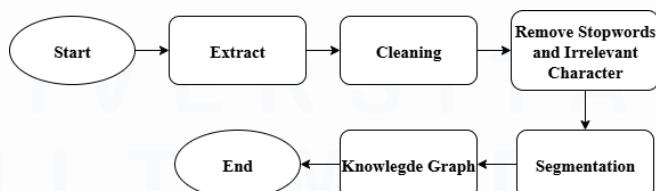


Gambar 3. 2 Alur Data Understanding

3.2.3 Data Preparation

Pada tahap *Data Preparation*, data yang telah terkumpul dibersihkan dan disiapkan secara optimal untuk tahap pemodelan selanjutnya. Pada fase ini, transkrip video tentang budidaya tanaman cabai akan dibersihkan dengan menghilangkan *stopwords*, karakter tidak relevan, serta memperbaiki ketidakkonsistenan teks seperti kesalahan penulisan atau format kalimat yang tidak seragam. Tujuan utama proses ini adalah memastikan hanya informasi bernilai tinggi dan relevan yang akan diproses lebih lanjut, seperti penjelasan mengenai media tanam berupa tanah hitam, sekam, pupuk kandang, jenis pupuk NPK Mutiara, teknik penyiraman, serta pengendalian hama. Setelah teks berhasil dibersihkan, dilakukan normalisasi data, misalnya dengan menyeragamkan istilah-istilah tertentu seperti “cabai rawit” atau “media tanam”. Selanjutnya, teks dipisah menjadi segmen-segmen yang lebih terfokus, berupa paragraf-paragraf yang memisahkan topik seperti pemilihan bibit, media tanam, atau teknik perawatan, sehingga mempermudah pemrosesan oleh model.

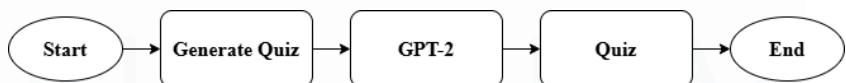
Di samping itu, dibuat pula *Knowledge Graph* yang menggambarkan keterkaitan antar konsep utama yang terdapat dalam transkrip, contohnya hubungan antara “media tanam”, “pemupukan”, “penyiraman”, serta “pengendalian hama”. *Knowledge Graph* ini berfungsi sebagai representasi visual dari keterkaitan antar topik penting dalam budidaya cabai, yang membantu model memahami konteks secara lebih baik sehingga dapat menghasilkan soal MCQ yang berkualitas tinggi dan sesuai dengan kebutuhan pembelajaran di Desa Pete.



Gambar 3. 3 Alur Data Preparation

3.2.4 Modelling

Pada tahap *Modelling*, proses dimulai dengan pemilihan *Small Language Model* yang paling tepat untuk menghasilkan soal secara otomatis berdasarkan transkrip video budidaya tanaman cabai yang telah melalui tahap pemrosesan sebelumnya. Beberapa model seperti *GPT-2 Small*, *Electra-Small* dan *Albert-Base-V2* dapat dimanfaatkan untuk memahami isi teks serta menciptakan soal MCQ yang didasarkan pada paragraf-paragraf terstruktur dan *Knowledge Graph* yang telah disusun pada tahapan terdahulu. Di fase ini, dilakukan *fine-tuning* atau pelatihan ulang terhadap model agar dapat menghasilkan pertanyaan yang selaras dengan sasaran serta konteks pembelajaran di Desa Pete. Model SLM seperti *GPT-2 Small* akan dilatih untuk menangkap konteks dari transkrip, mengenali konsep-konsep inti seperti media tanam berupa tanah hitam, sekam, pupuk kandang, jenis pemupukan seperti NPK Mutiara, teknik penyiraman, serta pengendalian hama, kemudian menghasilkan soal MCQ yang relevan dan efektif untuk mengukur tingkat pemahaman siswa SD. Soal-soal yang dihasilkan dirancang agar mencakup beragam tingkat kognitif sesuai *taksonomi Bloom*, mulai dari level mengingat (contoh: dosis pemupukan yang tepat) hingga level analisis (contoh: pengaruh media tanam terhadap pertumbuhan tanaman cabai), sehingga mampu merangsang kemampuan berpikir kritis pada siswa. Model ini juga dioptimalkan untuk menghasilkan soal pilihan ganda sederhana yang sesuai dengan tingkat pemahaman anak SD, memastikan bahwa soal tetap menarik, mudah dipahami, serta relevan dengan pengalaman belajar langsung di kebun edukasi.

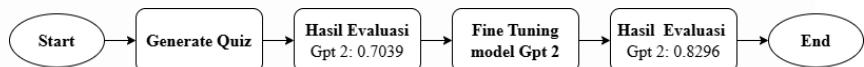


Gambar 3. 4 Alur Modelling

3.2.5 Evaluation

Setelah model berhasil menghasilkan soal-soal, tahap evaluasi segera dilaksanakan. Pada fase ini, dilakukan penilaian terhadap kualitas soal

MCQ yang dihasilkan dari transkrip video budidaya tanaman cabai guna memastikan kesesuaian dengan tujuan pembelajaran serta standar soal yang diharapkan di Desa Pete. Evaluasi dilakukan dengan menggunakan metrik seperti *BERTScore* untuk mengukur tingkat relevansi, kesulitan, serta kecocokan soal dengan materi, termasuk konsep media tanam (tanah hitam, sekam, pupuk kandang), pemupukan, penyiraman, dan pengendalian hama. Proses penilaian dilakukan dengan membandingkan soal hasil model terhadap ringkasan referensi dari transkrip, sehingga dapat dipastikan kejelasan, ketepatan, serta keselarasan dengan kurikulum anak SD. Apabila soal yang dihasilkan belum memenuhi standar kualitas, seperti tingkat relevansi rendah atau pilihan penganggu yang kurang logis, maka dilakukan penyesuaian pada model atau proses fine-tuning tambahan untuk meningkatkan hasil. Evaluasi ini memastikan bahwa soal yang dihasilkan benar-benar efektif dalam mengukur pemahaman mendalam siswa terhadap materi budidaya cabai dalam kerangka pembelajaran berbasis eduwisata di Desa Pete.



Gambar 3. 5 Alur Evaluation

3.2.6 Deployment

Pada tahap *Deployment*, sistem *generator* soal otomatis yang telah dibangun akan diintegrasikan langsung ke dalam website resmi Desa Pete sebagai bagian dari kerja sama dengan tim pengembang. Sistem ini akan diterapkan pada berbagai halaman pendidikan di situs desa, seperti halaman pembuatan soal, halaman ujian, serta halaman analisis hasil ujian, guna mendukung proses pembelajaran siswa SD tentang budidaya tanaman cabai. Soal-soal MCQ yang dihasilkan secara otomatis dirancang untuk mengukur pemahaman siswa terhadap materi, termasuk pemilihan media tanam, pemupukan, penyiraman, serta pengendalian hama. Soal-soal tersebut disusun agar mencakup berbagai tingkat kognitif sesuai

taksonomi Bloom, mulai dari level mengingat hingga memahami dan menganalisis, yang disesuaikan dengan tingkat kemampuan kognitif anak SD. Halaman-halaman tersebut akan saling terhubung untuk mempermudah interaksi antara guru, siswa, dan orang tua, sehingga mereka dapat mengakses soal-soal yang relevan dengan kurikulum serta pengalaman belajar langsung di kebun edukasi Desa Pete. Sistem ini juga akan terus diperbaiki dan diperbarui agar tetap efektif dalam mendukung pembelajaran berbasis teknologi di desa, dengan selalu memperhatikan masukan dari para pengguna guna meningkatkan relevansi serta mutu soal yang dihasilkan.

3.3 Teknik Pengumpulan Data

Penelitian ini menerapkan beragam teknik pengumpulan data guna memperoleh informasi yang mendalam tentang kebutuhan serta potensi Desa Pete sebagai kawasan wisata edukasi yang berbasis teknologi. Metode-metode yang dipakai meliputi wawancara, observasi langsung, dan studi pustaka. Ketiga teknik tersebut secara bersama-sama memberikan gambaran yang lebih lengkap mengenai kondisi riil di lapangan serta kebutuhan sistem yang akan dibuat untuk mendukung proses pembelajaran bagi siswa SD di desa tersebut.

3.3.1 Wawancara

Wawancara telah dilaksanakan bersama pengelola Badan Usaha Milik Desa (BUMDes) di Desa Pete dengan tujuan menggali informasi seputar kebutuhan dan harapan mereka dalam mengembangkan wisata edukasi. Dalam sesi wawancara tersebut, kami menanyakan berbagai aktivitas yang biasanya dilakukan oleh pengunjung, termasuk urutan kunjungan yang melewati pemandian Syekh Mubarok sebelum menunaikan shalat di Masjid Syekh Mubarok. Dari hasil wawancara ini, kami mendapatkan data penting mengenai pola kunjungan wisatawan serta cara mengubah pengalaman tersebut menjadi materi pembelajaran yang menarik. Kami juga menanyakan kebutuhan terkait kebun edukasi yang difokuskan untuk siswa SD, termasuk jenis kegiatan apa saja yang cocok dilakukan anak-

anak di kebun tersebut dan unsur-unsur edukasi apa yang harus disertakan agar pengunjung, terutama anak-anak, memperoleh manfaat belajar yang optimal. Hasil wawancara ini menjadi masukan yang sangat berharga dalam merancang sistem pembelajaran yang benar-benar sesuai dengan kebutuhan desa wisata dan para pengunjungnya.

3.3.2 Observasi

Selain wawancara, telah dilakukan observasi langsung di lokasi desa wisata untuk menilai kondisi yang ada secara nyata. Tujuan utama observasi ini adalah mendapatkan pemahaman yang lebih mendalam tentang bagaimana objek wisata dan kebun edukasi tersebut dapat diintegrasikan dengan teknologi pembelajaran berbasis kecerdasan buatan. Selama kegiatan observasi, kami mengunjungi beberapa titik penting seperti pemandian Syekh Mubarok, Masjid Syekh Mubarok, serta area kebun yang akan dimanfaatkan untuk kegiatan edukasi. Mengevaluasi kondisi infrastruktur yang tersedia serta mengidentifikasi potensi dan berbagai kendala yang mungkin timbul dalam pengembangan sistem berbasis *Small Language Model*. Observasi ini memberikan gambaran yang jelas mengenai penerapan teknologi di lingkungan wisata yang ada serta bagaimana pengalaman wisata dapat ditingkatkan melalui bantuan teknologi.

3.3.3 Studi Pustaka

Untuk memperkaya pemahaman tentang solusi yang telah ada, sudah melakukan studi pustaka yang mencakup berbagai penelitian terdahulu mengenai penerapan teknologi AI dalam bidang pariwisata dan pembelajaran interaktif berbasis teknologi. Kajian pustaka ini membantu kami mengenali tren terkini dalam pengembangan sistem chatbot interaktif, analisis sentimen, serta pembuatan konten secara otomatis yang dapat diterapkan pada konteks desa wisata. Di samping itu, penulis juga mempelajari sejumlah penelitian terkait pendekatan pembelajaran berbasis wisata, khususnya untuk anak SD, serta bagaimana teknologi dapat

memperkaya pengalaman belajar mereka di luar ruang kelas. Temuan dari studi pustaka ini memberikan landasan teori yang kuat untuk merancang sistem berbasis *Small Language Model* yang benar-benar sesuai dengan kebutuhan dan karakteristik desa wisata.

3.4 Teknik Analisis Data

Dalam penelitian ini, pengolahan dan analisis data dilakukan dengan memanfaatkan bahasa pemrograman *Python*, yang dipilih karena tingkat fleksibilitasnya yang tinggi serta ketersediaan berbagai pustaka pendukung seperti *Pandas*, *NLTK*, dan *PyTorch*. Pustaka-pustaka tersebut sangat memudahkan dalam memproses data berbentuk teks serta menerapkan berbagai model pembelajaran mesin. *Python* dikenal luas karena kemampuannya menangani beragam jenis data dan model secara efisien, terutama dalam pengolahan teks, bila dibandingkan dengan bahasa pemrograman lain seperti R. Berikut adalah tabel 3.2 perbandingan antara R dan Python yang menyoroti kelebihan serta kekurangan masing-masing dalam konteks analisis data.

Tabel 3. 2 Tools

Kriteria	R [53]	Python [54]
Kelebihan	<ul style="list-style-type: none"> - Memiliki pustaka statistik serta analisis data yang sangat kuat - Sangat mudah untuk membuat visualisasi statistik - Banyak digunakan dalam analisis data ekonomi dan sosial - Dikenal karena kemudahan penggunaannya untuk analisis statistik 	<ul style="list-style-type: none"> - Sangat fleksibel dan dapat diterapkan pada berbagai jenis aplikasi - Didukung pustaka machine learning serta AI yang lengkap (contoh: Pandas, TensorFlow) - Lebih efisien dalam mengelola data besar serta data teks - Memiliki komunitas pengguna yang besar beserta dokumentasi yang lengkap dan terperinci
Kekurangan	<ul style="list-style-type: none"> - Cakupannya lebih terbatas pada analisis statistik serta aplikasi ekonomi - Dukungan untuk machine learning tidak sekuat Python 	<ul style="list-style-type: none"> - Kurang optimal untuk analisis ekonomi atau statistik murni - Memerlukan konfigurasi tambahan untuk analisis data lanjutan tertentu

Kriteria	R [53]	Python [54]
	<ul style="list-style-type: none"> - Kurang cocok untuk pengembangan aplikasi atau integrasi sistem yang kompleks - Sedikit lebih sulit untuk analisis statistik tingkat lanjut 	

R memang unggul dalam bidang statistik serta ekonometrika, terutama untuk analisis data statistik yang lebih spesifik dan terfokus. Namun, *Python* memberikan fleksibilitas yang jauh lebih besar karena didukung oleh beragam pustaka untuk *machine learning* dan *deep learning*, seperti *Pandas*, *Matplotlib*, serta *PyTorch*, yang sangat dibutuhkan dalam penelitian ini guna melakukan analisis data dan menerapkan model pembelajaran mesin. *Python* juga mempermudah integrasi dengan berbagai sistem eksternal serta *API*. Oleh karena itu, *Python* menjadi pilihan utama dalam penelitian ini berkat kemampuannya yang lebih luas dan fleksibel dalam mengelola *model machine learning* untuk penyusunan soal berbasis teknologi. Dengan mempertimbangkan keunggulan Python yang lebih serbaguna dalam pengembangan sistem berbasis kecerdasan buatan serta dukungan pustaka yang sangat lengkap, pemilihan Python untuk menganalisis data transkrip, membuat ringkasan, serta mengimplementasikan model pembelajaran mesin merupakan keputusan yang sangat tepat.

3.5 Keterbatasan Model SLM

Model SLM yang diterapkan dalam penelitian ini memiliki sejumlah keterbatasan yang cukup signifikan, terutama dalam konteks penerapan pada materi budidaya tanaman cabai serta pengembangan wisata edukasi di Desa Pete, Kecamatan Tigaraksa, Kabupaten Tangerang. Keterbatasan-keterbatasan tersebut adalah sebagai berikut:

1. Kemampuan memahami konteks lokal yang masih terbatas dapat memengaruhi tingkat ketepatan soal MCQ yang dihasilkan, khususnya karena model SLM belum cukup terlatih dengan data khusus mengenai budidaya cabai yang bersifat unik di wilayah setempat, seperti variasi jenis media tanam atau teknik pemupukan yang khas digunakan masyarakat desa.

2. Kapasitas pembuatan soal yang terbatas pada tingkat kerumitan teks yang berasal dari transkrip video YouTube dapat menghambat penyusunan soal MCQ yang lebih mendalam terkait aspek teknis budidaya cabai, misalnya analisis pengaruh penyiraman terhadap pertumbuhan tanaman, strategi pengendalian hama yang optimal, serta kaitannya dengan potensi wisata edukasi yang dimiliki Desa Pete.
3. Ketergantungan terhadap kualitas hasil transkrip yang dihasilkan dari teknologi *Speech-to-Text* berpotensi menimbulkan kesalahan dalam penangkapan konsep utama, seperti teknik pemeliharaan tanaman cabai atau informasi sejarah terkait makam Syekh Mubarok, sehingga dapat menurunkan tingkat relevansi dan ketepatan soal yang dihasilkan oleh sistem.