

BAB II

TINJAUAN PUSTAKA

2.1 Teori Tentang Topik

2.1.1 Pendidikan

Pendidikan adalah proses pembelajaran yang membantu seseorang mengembangkan potensi diri, baik dalam aspek pengetahuan, keterampilan, maupun sikap [12]. Melalui pendidikan, individu dapat belajar memahami dirinya sendiri dan lingkungannya, belajar berpikir logis, serta membentuk kepribadian yang positif [12]. Pendidikan dimulai dari lingkungan keluarga, kemudian berlanjut ke sekolah, dan berlangsung sepanjang hidup. Di sekolah, siswa tidak hanya belajar mata pelajaran, tetapi juga belajar berinteraksi, bekerja sama, dan menyelesaikan masalah [13]. Oleh karena itu, pendidikan memiliki peran penting dalam membentuk karakter dan kesiapan seseorang menghadapi masa depan.

Lebih dari itu, pendidikan juga menjadi salah satu cara paling efektif untuk meningkatkan kualitas hidup seseorang. Pendidikan tidak hanya memberikan pengetahuan dan keterampilan, tetapi juga membuka akses terhadap berbagai peluang sosial dan ekonomi. Seseorang yang memiliki latar belakang pendidikan yang baik umumnya memiliki peluang lebih besar untuk memperoleh pekerjaan yang lebih baik, penghasilan yang lebih layak, serta kehidupan yang lebih stabil dan sejahtera [14]. Selain itu, pendidikan yang berkualitas berperan penting dalam mengurangi kesenjangan sosial, mendorong mobilitas sosial, dan menciptakan masyarakat yang lebih adil dan inklusif. Oleh karena itu, perhatian terhadap sistem pendidikan, baik di sekolah negeri maupun swasta, menjadi hal yang sangat penting. Perbedaan dalam hal fasilitas, kualitas tenaga pengajar, metode pengajaran, serta lingkungan belajar yang tersedia di masing-masing jenis sekolah dapat berdampak signifikan terhadap hasil belajar siswa. Dalam konteks ini, perlu dilakukan kajian secara mendalam untuk mengevaluasi dan membandingkan kualitas pendidikan di sekolah negeri dan

swasta guna memastikan bahwa setiap siswa, tanpa memandang latar belakang sosial atau ekonomi, memiliki kesempatan yang sama untuk berkembang secara optimal [12].

2.1.2 Sosial Ekonomi

Sosial ekonomi merupakan kedudukan individu dalam struktur masyarakat yang dipengaruhi oleh kombinasi faktor sosial dan ekonomi. Faktor sosial mencakup aspek budaya, usia, tingkat pendidikan, dan jenis kelamin, sementara faktor ekonomi melibatkan pendapatan, jenis pekerjaan, tingkat pendidikan, dan kepemilikan aset atau investasi [15]. Kedudukan sosial ekonomi ini tidak hanya menentukan peran seseorang dalam masyarakat, tetapi juga memengaruhi akses terhadap sumber daya, peluang, serta hak dan kewajiban yang harus dijalankan dalam kehidupan sehari-hari [16].

Kondisi sosial ekonomi seseorang sangat memengaruhi kualitas hidup dan kemampuannya memenuhi kebutuhan dasar, baik fisik maupun psikologis. Akses terhadap layanan kesehatan, pendidikan, perumahan, dan air bersih menjadi indikator utama kesejahteraan [15]. Pekerjaan yang layak juga mendukung partisipasi aktif dalam masyarakat. Selain itu, tingkat sosial ekonomi turut membentuk aspirasi, gaya hidup, dan pandangan seseorang terhadap masa depan [15].

2.1.3 Korelasi

Korelasi adalah suatu konsep statistik yang menggambarkan tingkat hubungan atau keterkaitan antara dua variabel [17]. Jika dua variabel memiliki korelasi, maka perubahan pada salah satu variabel cenderung diikuti oleh perubahan pada variabel lainnya, baik secara searah (positif) maupun berlawanan arah (negatif). Korelasi tidak mengindikasikan sebab-akibat, melainkan hanya menunjukkan bahwa terdapat pola keterkaitan antara dua variabel yang diamati [17]. Misalnya, ketika nilai salah satu variabel meningkat dan variabel lainnya juga ikut meningkat, maka dikatakan terdapat korelasi

positif. Sebaliknya, jika satu variabel naik sementara yang lain turun, maka disebut korelasi negatif.

Dalam analisis data kuantitatif, korelasi diukur menggunakan koefisien korelasi, seperti Pearson Product Moment (untuk data interval/rasio yang berdistribusi normal) atau Spearman Rank (untuk data ordinal atau yang tidak memenuhi asumsi normalitas) [17]. Koefisien korelasi memiliki rentang nilai dari -1 hingga +1. Nilai yang mendekati +1 mengindikasikan hubungan yang sangat kuat dan searah antar variabel, sedangkan nilai yang mendekati -1 menunjukkan hubungan yang sangat kuat namun berlawanan arah. Sementara itu, nilai yang mendekati 0 menunjukkan bahwa tidak terdapat hubungan yang signifikan antara variabel-variabel tersebut [18].

2.1.4 Sumber Data

Sumber data dalam penelitian ini terdiri dari data primer dan data sekunder. Data primer diperoleh melalui penyebaran kuesioner yang disusun oleh peneliti bersama tim, dan dibagikan kepada mahasiswa Universitas Multimedia Nusantara secara daring melalui Google Form selama periode 2 Mei 2025 hingga 30 Mei 2025. Kuesioner ini mencakup informasi sosial ekonomi mahasiswa, seperti pendidikan dan pekerjaan orang tua, pendapatan keluarga, serta penanggung biaya pendidikan. Dari proses ini, berhasil dikumpulkan sebanyak 65 responden mahasiswa aktif Universitas Multimedia Nusantara dengan data yang valid.

Selain data primer, data sekunder juga digunakan sebagai bahan pendukung dalam penelitian ini. Data sekunder diperoleh dari data historis sistem internal Universitas Multimedia Nusantara tepatnya dari Lab Fakultas Teknik Informatika Universitas Multimedia Nusantara, seperti laporan akademik dan data internal yang relevan dengan topik penelitian. Periode data tersebut mencakup tahun 2019 hingga 2024 dan berjumlah 961 baris dan 426 kolom. Kombinasi antara data primer dan sekunder memberikan gambaran yang lebih utuh dalam menganalisis hubungan antara kondisi sosial ekonomi

mahasiswa dan prestasi akademik, serta meningkatkan akurasi dan validitas hasil penelitian.

2.2 Teori Tentang Framework dan Metode

2.2.1 CRISP-DM

CRISP-DM (Cross-Industry Standard Process for Data Mining) adalah metode yang banyak digunakan dalam analisis data secara sistematis. Metodologi ini terdiri dari enam tahap utama, yaitu Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment. Keenam tahap tersebut saling berkaitan dan sering dilakukan secara berulang untuk memastikan hasil yang optimal [19]. Pendekatan ini sangat cocok diterapkan dalam penelitian yang bertujuan untuk menemukan pola atau hubungan antar variabel, seperti hubungan antara kondisi sosial ekonomi dengan prestasi akademik [19].



Gambar 2. 1 Alur CRISP-DM

Sumber : [20]

Berikut penjelasan dari tiap tahap dalam CRISP-DM dan kaitannya dengan penelitian ini:

1. Business Understanding

Tahap pertama ini berfokus pada pemahaman terhadap tujuan dan kebutuhan dari sisi pengguna atau stakeholder. Di sini, peneliti atau analis berusaha merumuskan permasalahan bisnis ke dalam bentuk tujuan analisis data yang jelas [21]. Pemahaman konteks yang tepat sangat penting agar proses selanjutnya relevan dengan kebutuhan yang ingin dicapai [19].

2. Data Understanding

Setelah permasalahan terdefinisi dengan baik, langkah berikutnya adalah mengumpulkan dan mengeksplorasi data awal. Tahap ini mencakup kegiatan seperti identifikasi tipe data, pemeriksaan kelengkapan data, dan pemahaman awal terhadap pola-pola dasar dalam data [22]. Tujuannya adalah untuk mengenali potensi masalah dan menemukan wawasan awal sebelum proses persiapan data dilakukan.

3. Data Preparation

Proses ini difokuskan pada penyesuaian dan pengolahan data agar dapat diimplementasikan dalam tahap pemodelan selanjutnya. Prosesnya meliputi pembersihan data (cleaning), transformasi, penggabungan, pemilihan atribut penting, mengatasi data yang terdistribusi tidak normal, serta penanganan data yang tidak konsisten atau hilang. Tahap ini sering kali menjadi salah satu yang paling memakan waktu karena kualitas data akan sangat memengaruhi hasil model nantinya [22].

4. Modeling

Di tahap ini, teknik analisis yang digunakan difokuskan pada metode yang dapat mengukur hubungan antar variabel, seperti analisis korelasi. Proses modeling melibatkan pemilihan metode yang tepat untuk menentukan sejauh mana variabel sosial ekonomi berkaitan dengan prestasi akademik mahasiswa. Selain itu, penyesuaian parameter dilakukan untuk memastikan hasil korelasi yang akurat dan signifikan secara statistik. Model

yang dibangun kemudian dievaluasi untuk melihat kekuatan dan arah hubungan antar variabel, sehingga dapat memberikan pemahaman yang lebih jelas tentang pola dan hubungan yang ada dalam data penelitian [22].

5. Evaluation

Setelah model dibentuk, tahap selanjutnya adalah mengevaluasi hasil model tersebut. Tujuannya adalah memastikan bahwa model yang dibangun mampu menjawab tujuan awal dari analisis data. Evaluasi tidak hanya mencakup performa teknis model, tetapi juga sejauh mana model tersebut relevan dan bermanfaat dalam konteks masalah yang telah didefinisikan [21].

6. Deployment

Tahap terakhir dalam CRISP-DM adalah penyampaian hasil analisis. Bentuk implementasi dapat beragam, mulai dari pembuatan laporan, presentasi, hingga integrasi model ke dalam sistem yang lebih besar. Dalam konteks akademik atau penelitian, tahap ini biasanya direpresentasikan melalui penarikan kesimpulan, penyusunan rekomendasi, dan dokumentasi hasil penelitian [21].

2.2.2 Spearman Corellation

Korelasi Spearman adalah metode statistik yang digunakan untuk mengetahui seberapa kuat hubungan antara dua variabel, terutama jika data yang digunakan berbentuk peringkat (ranking) atau tidak mengikuti distribusi normal [23]. Metode ini pertama kali dikenalkan oleh Charles Spearman pada tahun 1904, sehingga dinamakan Korelasi Spearman [24].

Kelebihan dari korelasi ini adalah bisa digunakan meskipun data berasal dari skala pengukuran yang berbeda, misalnya satu data berskala ordinal dan satu lagi berskala numerik [23]. Selain itu, metode ini cocok untuk data yang tidak memenuhi syarat-syarat statistik parametrik seperti normalitas. Dalam korelasi Spearman, setiap nilai dalam variabel akan diberi peringkat [11]. Setelah itu, dihitung selisih peringkat antara pasangan data dari kedua variabel. Selisih tersebut kemudian dikuadratkan dan dijumlahkan untuk

mendapatkan nilai korelasi [23]. Rumus yang digunakan adalah sebagai berikut:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2-1)} [25]$$

Keterangan:

ρ = Nilai koefisien korelasi Spearman [25]

d_1 = Selisih peringkat antara dua variabel pada setiap observasi [25]

n = Jumlah pasangan data [25]

2.3 Teori Tentang Tools dan Teknik

2.3.1 Apps Script

Google Apps Script adalah platform berbasis JavaScript dari Google yang digunakan untuk mengotomatisasi tugas dan memperluas fungsi layanan Google Workspace seperti Google Sheets, Docs, Forms, dan Gmail [26]. Dengan Apps Script, pengguna bisa membuat skrip untuk mengelola data secara otomatis, misalnya mengambil data dari Google Form ke Google Sheets, membersihkan data, atau menghitung statistik secara langsung [26]. Selain itu, Apps Script juga bisa digunakan untuk menjadwalkan tugas seperti mengirim email pengingat atau membuat laporan rutin dari data yang ada.

Keunggulan lain dari Apps Script adalah kemudahannya untuk diakses langsung dari browser tanpa perlu instalasi perangkat lunak tambahan, serta integrasinya yang mendalam dengan ekosistem Google [27]. Hal ini menjadikannya alat yang ideal untuk mendukung efisiensi kerja dalam proyek penelitian, terutama dalam tahap pengumpulan data dan pelaporan awal secara otomatis dan terstruktur [26].

2.3.2 Python

Python merupakan bahasa pemrograman yang populer karena kesederhanaannya dan kemampuannya yang serbaguna. Bahasa ini kerap digunakan dalam berbagai bidang, termasuk pengembangan web, otomatisasi,

analisis data, serta pembuatan model machine learning [28]. Salah satu keunggulan utama Python adalah sifatnya yang open-source, sehingga dapat digunakan dan dimodifikasi secara bebas tanpa perlu membayar lisensi [28].

Dalam konteks penelitian kuantitatif seperti analisis hubungan antara faktor sosial ekonomi dan prestasi akademik, Python menyediakan beragam pustaka yang sangat bermanfaat. Misalnya, Pandas dan NumPy digunakan untuk pengolahan dan manipulasi data, sementara Scikit-learn dan SciPy dapat dimanfaatkan untuk menghitung korelasi Spearman. Untuk visualisasi hubungan antarvariabel, Matplotlib dan Seaborn memberikan dukungan grafis yang kuat [28].

2.3.3 Jupyter

Jupyter merupakan perangkat lunak open-source yang dirancang untuk mendukung komputasi interaktif dengan berbagai bahasa pemrograman seperti Python, R, dan Julia [29]. Salah satu komponen utamanya adalah Jupyter Notebook, sebuah platform yang memungkinkan pengguna menyusun dokumen yang memadukan kode program, penjelasan naratif, visualisasi data, dan notasi matematis dalam satu file [29]. Dalam konteks analisis korelasi antara faktor sosial ekonomi dan prestasi akademik, Jupyter Notebook sangat berguna karena memungkinkan peneliti menuliskan kode Python untuk mengolah data, menghitung korelasi Spearman, serta menampilkan hasil visualisasi secara terpadu dan berurutan [30].

Lebih jauh, ekosistem Jupyter juga meliputi alat-alat tambahan seperti JupyterLab yang menyediakan antarmuka kerja lebih fleksibel, nbconvert untuk mengubah dokumen menjadi format lain, misalnya PDF atau HTML. Serta Jupyter Widgets yang memungkinkan interaktivitas dalam penyajian data [30]. Dengan fitur-fitur ini, Jupyter menjadi sarana yang sangat mendukung transparansi, dokumentasi, dan kolaborasi dalam proses analisis data sosial-educational secara menyeluruh.

2.3.4 Github

GitHub adalah platform hosting berbasis web yang memudahkan pengelolaan dan kolaborasi kode sumber menggunakan sistem kontrol versi Git. Dengan GitHub, pengguna dapat menyimpan, melacak perubahan, serta mengelola berbagai versi kode program secara terstruktur dan terorganisir [31]. Fitur seperti branching, pull request, dan issue tracking mendukung kerja sama tim dalam pengembangan perangkat lunak maupun penelitian yang melibatkan kode pemrograman [32]. Dalam konteks penelitian ini, GitHub digunakan untuk menyimpan dan mengelola kode program yang berkaitan dengan pengolahan data dan analisis, sehingga memudahkan dokumentasi serta kolaborasi pada saat proses deployment ke dalam streamlit.

Selain itu, penggunaan GitHub meningkatkan transparansi dan reproducibility penelitian karena semua versi kode dapat diakses dan diperiksa oleh peneliti lain atau pihak yang berkepentingan. GitHub juga memungkinkan integrasi dengan berbagai layanan lain, seperti platform continuous integration dan deployment, yang membantu mempercepat proses pengembangan dan validasi kode dalam proyek penelitian [31]. Hal ini membuat GitHub menjadi alat penting dalam manajemen proyek berbasis data dan pemrograman.

2.3.4 Streamlit

Streamlit adalah framework open-source yang dirancang untuk mempermudah pembuatan aplikasi web interaktif khususnya untuk visualisasi data dan pengembangan prototipe model korelasi menggunakan bahasa Python [33]. Dengan Streamlit, peneliti dapat dengan cepat mengubah skrip analisis data menjadi aplikasi web yang mudah digunakan, tanpa perlu menguasai bahasa pemrograman front-end seperti HTML, CSS, atau JavaScript [34]. Fitur interaktif seperti slider, tombol, dan grafik dinamis membuat aplikasi yang dibuat menggunakan Streamlit sangat cocok untuk menyajikan hasil penelitian secara langsung dan interaktif kepada pengguna atau pemangku kepentingan [34].

Dalam penelitian ini, Streamlit dapat digunakan untuk membuat dashboard yang menampilkan hasil analisis korelasi antara kondisi sosial ekonomi dan prestasi akademik mahasiswa. Dengan demikian, hasil penelitian tidak hanya disajikan dalam bentuk laporan statis, tetapi juga dalam format yang mudah diakses dan dipahami secara visual oleh berbagai pihak. Penggunaan Streamlit membantu memperluas dampak penelitian dengan menghadirkan data yang dapat dieksplorasi secara real-time dan interaktif.

2.4 Penelitian Terdahulu

Penelitian sebelumnya memberikan gambaran mengenai berbagai pendekatan yang digunakan dalam menganalisis hubungan antara faktor sosial ekonomi dan pencapaian akademik mahasiswa. Setiap penelitian menawarkan perspektif unik berdasarkan variabel sosial ekonomi yang digunakan, metode analisis yang diterapkan, serta hasil yang diperoleh terkait kontribusi faktor-faktor tersebut terhadap nilai akademik. Berikut adalah rangkuman dari beberapa penelitian terdahulu yang relevan dengan topik ini:

Tabel 2. 1 Penelitian Terdahulu

No	Penulis	Sumber	Algoritma	Kontribusi	Hasil
1	Depi Pramika et al. [10]	Hubungan antara Status Sosial Ekonomi dengan Prestasi Belajar Mahasiswa, 2018	Korelasi Spearman dan Kendall's Tau	Memberi wawasan empiris mengenai hubungan status ekonomi dengan prestasi akademik mahasiswa	Tidak ditemukan hubungan signifikan antara status sosial ekonomi dan prestasi belajar mahasiswa.
2	Diesebeil L. Obina et al. [8]	Potential Relationship of Socio-Economic	Korelasi Spearman Rho	Menjelaskan pengaruh multiple intelligences	Status sosial ekonomi tidak signifikan terhadap prestasi,

		Status and Multiple Intelligences to Academic Performance, 2024		terhadap prestasi akademik serta kaitannya dengan latar belakang sosial ekonomi	tetapi multiple intelligences sangat berpengaruh ($p = 0.70428$).
3	Emmanuel Byiringiro [7]	Influence of Family Socio-Economic Status on Mathematics Achievement in Secondary School, 2024	Korelasi Pearson, Spearman Rho, dan Regresi Linier	Menunjukkan peran penting tipe keluarga dan pendidikan orang tua dalam pencapaian prestasi matematika	Tipe keluarga berpengaruh signifikan ($r = 0.610$) terhadap prestasi matematika; dua orang tua menunjukkan hasil lebih baik.
4	Ronald Floriano Rodríguez et al. [35]	Influence of Digital Skills on the Academic Performance of University Students: A Socioeconomic Approach, 2024	Korelasi Spearman Rho	Menganalisis pengaruh kompetensi digital dan faktor sosial ekonomi terhadap prestasi akademik mahasiswa..	Kompetensi digital berpengaruh sebesar 30,5% terhadap prestasi akademik ($r_s = 0,546$; $p = 0,000$), sisanya 69,5% dipengaruhi faktor sosial ekonomi seperti latar belakang keluarga, pekerjaan, dan waktu belajar.
5	Niko Putra Indika [36]	Dampak Sosial dan Ekonomi Keberadaan Perguruan	Korelasi Spearman Rho, TCR	Memberikan gambaran empiris tentang tingkat	Meskipun responden memiliki capaian sangat tinggi, tidak ada

		Tinggi di Kelurahan Sumber Sari, Kota Malang, 2024		pengaruh universitas terhadap kondisi sosial ekonomi lokal di Sumber Sari	hubungan signifikan antara keberadaan kampus dan perubahan sosial ekonomi Masyarakat.
--	--	--	--	---	---

Tabel 2.1 merangkum berbagai studi terdahulu yang membahas hubungan antara faktor sosial ekonomi dan prestasi akademik. Beragam pendekatan analisis digunakan, seperti korelasi Spearman, Pearson, dan regresi linier untuk mengidentifikasi pola hubungan antara latar belakang keluarga dan capaian akademik mahasiswa atau siswa [7], [10], [35]. Beberapa penelitian menemukan bahwa status sosial ekonomi tidak memiliki pengaruh signifikan terhadap prestasi belajar menunjukkan bahwa faktor lain seperti kemampuan individu atau lingkungan belajar mungkin lebih dominan. Di sisi lain, multiple intelligences justru menunjukkan kontribusi besar terhadap prestasi siswa, meskipun latar belakang ekonomi tidak signifikan [8]. Selain itu, struktur keluarga dengan kehadiran dua orang tua juga berdampak positif terhadap prestasi akademik, khususnya dalam mata pelajaran seperti matematika [7].

Faktor kompetensi digital turut menjadi perhatian, di mana keterampilan mahasiswa dalam menggunakan teknologi secara efektif memberikan pengaruh signifikan terhadap kinerja akademik, meskipun faktor sosial ekonomi turut dipertimbangkan [35]. Sementara itu, ada juga penelitian yang menyoro dampak keberadaan perguruan tinggi terhadap perubahan sosial ekonomi masyarakat sekitar, namun hasilnya tidak menunjukkan hubungan yang signifikan [36]. Secara keseluruhan, studi-studi tersebut memberikan landasan penting bahwa hubungan antara kondisi sosial ekonomi dan prestasi akademik bersifat kompleks dan multidimensional, sehingga memerlukan pendekatan yang menyeluruh.