

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Penelitian Terdahulu

Berikut merupakan beberapa referensi jurnal penelitian yang digunakan pada penelitian ini.

Tabel 2. 1 Penelitian Terdahulu

No	Penulis	Judul	Hasil
1	Amru Setio Wibowo & Andri	<i>Dashboard Business Intelligence</i> Visualisasi Data Akreditasi Sekolah Pada SMP Negeri 1 Sembawa	Penelitian ini menghasilkan <i>dashboard monitoring</i> untuk menunjang akreditasi pada smp negeri 1 sembawa dengan menggunakan teknologi <i>businessintelligence</i> (Wibowo & Andri, 2021).
2	Evan Himawan Saragih, I Putu Agung Bayupati, Gusti Agung Ayu Putri	Pengembangan <i>Business Intelligence Dashboard</i> Untuk <i>Monitoring</i> Aktivitas Pariwisata (Studi Kasus: Dinas Pariwisata Provinsi Bali)	Perancangan <i>business intelligence dashboard</i> untuk <i>monitoring</i> aktivitas pariwisata studi kasus Dinas Pariwisata Provinsi Bali dapat

Tabel 2. 1 Penelitian Terdahulu (Lanjutan)

No	Penulis	Judul	Hasil
			<p>digunakan untuk memajemen data dengan menggantikan pemakaian kertas menjadi media komputer serta manajemen data agar data tidak hilang begitu saja, namun digunakan sebagai acuan dalam menentukan keputusan (Saragih et al., 2021).</p>
3	Tommy, Amir Mahmud Husein	<p>Model Prediksi Prestasi Mahasiswa</p> <p>Berdasarkan Evaluasi Pembelajaran Menggunakan Pendekatan <i>Data Science</i></p>	<p>berdasarkan hasil evaluasi dari kelima model yang diusulkan <i>DecisionTree</i> menghasilkan akurasi sebesar 68 % lebih baik dari model lain sehingga model ini telah di-deploy dan dapat beroperasi dengan baik sehingga dapat digunakan untuk</p>

Tabel 2. 1 Penelitian Terdahulu (Lanjutan)

No	Penulis	Judul	Hasil
			kebutuhan prediksi (Tommy & Husein, 2021).
4	Deo Haganta Depari, Yuni Widiastiwi, Mayanda Mega Santoni	Perbandingan Model <i>Decision Tree</i> , <i>Naïve Bayes</i> Dan <i>Random Forest</i> Untuk Prediksi Klasifikasi Penyakit Jantung.	metode <i>decision tree</i> pada kasus ini merupakan model terbaik dikarenakan memiliki akurasi sebesar 72% dan dengan waktu eksekusi sebesar 0.118 detik (Depari et al., 2022).
5	Aida Meimela	Prediksi Jumlah Kunjungan Wisatawan Mancanegara Ke Indonesia	Metode prediksi dengan tingkat kekuratan yang baik untuk meramalkan jumlah kunjungan wisatawan mancanegara ke Indonesia adalah <i>Seasonal ARIMA</i> dengan model $(0,1,1)(0,1,1)_{12}$ (Meimela, 2021).

Tabel 2.1 tersebut menggambarkan berbagai penelitian dalam domain teknologi informasi, khususnya terkait penerapan *business intelligence* (BI) dan prediksi data. Pada penelitian pertama yang terdahulu ini fokus pada pembuatan *dashboard* BI untuk memantau akreditasi di SMP Negeri 1 Sembawa, menggunakan BI sebagai alat visualisasi (Wibowo & Andri, 2021).

Selanjutnya, penelitian kedua terdahulu membahas pengembangan BI *dashboard* untuk memonitor aktivitas pariwisata di Dinas Pariwisata Provinsi Bali, bertujuan menggantikan penggunaan kertas dengan media komputer untuk memudahkan manajemen data (Saragih et al., 2021).

Penelitian terdahulu ketiga berfokus pada pengembangan model prediksi prestasi mahasiswa dengan *Data Science. Decision Tree*, dengan akurasi 68%, dipilih sebagai model yang baik untuk memprediksi prestasi mahasiswa (Tommy & Husein, 2021).

Penelitian terdahulu keempat, kunci penelitian tersebut adalah membandingkan model *Decision Tree, Naïve Bayes, dan Random Forest* untuk prediksi klasifikasi penyakit jantung. Hasilnya menunjukkan bahwa *Decision Tree* dengan akurasi 72% dan waktu eksekusi singkat adalah pilihan terbaik (Depari et al., 2022).

Dan untuk penelitian terdahulu yang terakhir membahas prediksi jumlah kunjungan wisatawan mancanegara ke Indonesia dengan menggunakan *Seasonal ARIMA (0,1,1)(0,1,1)12*, yang diklaim memiliki tingkat keakuratan yang baik. Secara keseluruhan, tabel mencerminkan beragam aplikasi teknologi BI dan analisis data dalam berbagai konteks (Meimela, 2021).

Dari hasil refrensi penelitian terdahulu yang ada, penelitian ini dilakukan dikarenakan tidak adanya/masih belum ada pengembangan *business intelligence* dengan penggunaan model *machine learning* untuk melakukan sebuah prediksi menjadi satu aplikasi atau tambahan dalam fitur aplikasi.

2.2 Kajian Teoritis

2.2.1 Business Intelligence

Business intelligence (BI) adalah suatu pendekatan yang menggabungkan teknologi, proses, dan metode analisis untuk mengubah data menjadi informasi yang bermakna dan *actionable* bagi perusahaan (Tumini & Subekti, 2023). *Business intelligence* (BI) merupakan sistem yang digunakan untuk menganalisis data dan mengolahnya menjadi informasi untuk membantu pengambilan keputusan proses bisnis (Prasena, 2020).

Pemanfaatan *business intelligence* dalam melakukan analisis data dalam jumlah yang besar, sangat efektif dipakai untuk mengetahui keadaan perusahaan sehingga menghasilkan informasi yang berguna bagi setiap penggunanya (*stakeholders*) yaitu manajemen, staf, konsumen, mitra bisnis, pemilik perusahaan, dan pihak lain yang bersangkutan (Saragih et al., 2021). BI dapat dianggap sebagai pendorong Sistem Manajemen Informasi, dan di sektor pendidikan tinggi, BI menyajikan peluang untuk memberikan informasi yang bermakna tentang kinerja (Trisnawarman & Imam, 2020).

2.2.2 Framework

Framework merupakan sebuah kerangka kerja yang dibuat untuk memudahkan pembuatan sebuah *website*. *Framework* memiliki komponen dan variabel yang dibutuhkan para perancang *website* untuk memudahkan pembacaan kode, perencanaan, pengujian dan pemeliharaan (Prasena, 2020).

2.2.3 Modelling

Modelling adalah proses pembuatan model atau representasi dari suatu sistem atau objek yang akan digunakan untuk memahami, memprediksi, atau mengoptimalkan kinerja sistem atau objek tersebut. Berikut beberapa contoh *modelling* yang sering dilakukan.

Modelling digunakan untuk menganalisis keputusan apa saja yang ada di dalam proses bisnis manajemen proyek (Permana et al., 2021). Untuk menganalisis titik-titik keputusan ini diperlukan BPMN (*Business Process*

Model Notation) dari proses-proses yang ada. Dan pada penelitian (Ismiati et al., 2020), modelling digunakan untuk membuat suatu sistem pendukung keputusan pemilihan restoran di kota Palembang.

2.2.4 Hyperparameter

Hyperparameter merujuk pada parameter-parameter yang tidak dapat diperbarui selama pelatihan dalam *machine learning*. Mereka dapat terlibat dalam pembentukan struktur model, seperti jumlah lapisan tersembunyi dan fungsi aktivasi, atau dalam menentukan efisiensi dan akurasi pelatihan model, seperti tingkat pembelajaran (*learning rate*) dari *stochastic gradient descent* (SGD), ukuran *batch*, dan *optimizer* (hyp) (Yu & Zhu, 2020).

2.2.5 Decision Tree Regression

Decision tree regressor adalah model yang digunakan dalam prediksi data dan pembelajaran mesin. Model ini menggambarkan struktur berbentuk pohon untuk memprediksi nilai variabel kontinu (seperti harga rumah) berdasarkan fitur-fitur atau atribut tertentu. Pohon keputusan ini dibangun dengan membagi *dataset* menjadi subset-subset yang lebih kecil secara bertahap. Setiap kali data dibagi, sebuah simpul keputusan baru ditambahkan ke dalam pohon. Hasil akhirnya adalah pohon dengan simpul daun yang digunakan untuk membuat prediksi nilai, dan simpul pemilihan yang mengarahkan bagaimana data dibagi menjadi subset-subset yang lebih kecil (Ravi Teja Reddy & Malathi, 2022).

Peran *decision tree regression* akan terus melakukan segmentasi data sampai dapat menyesuaikan model regresi linier. *Decision tree regression* menggunakan deviasi kuadrat rata-rata minimum untuk menentukan pembagian optimal dari pohon regresi, dan kriteria partisi adalah untuk meminimalkan varians kesalahan sub-pohon setelah partisi yang diharapkan (Yao, 2023).

2.2.6 Random Search

Random search sendiri merupakan perbaikan dasar dari *grid search*. Ini mengindikasikan pencarian acak atas *hyperparameter* dari distribusi tertentu dari nilai-nilai parameter yang mungkin. Proses pencarian berlanjut hingga anggaran yang telah ditentukan habis, atau hingga akurasi yang diinginkan tercapai. *Random Search* mungkin akan lebih baik terutama ketika beberapa *hyperparameter* tidak didistribusikan secara merata (Yu & Zhu, 2020).

2.2.7 Tableau

Tableau adalah *tools/platform* untuk membuat visualisasi data lebih interaktif, mudah dibaca, dan mudah dianalisa. Visualisasi yang dilakukan adalah mengubah data tabel yang kaku menjadi grafik, diagram, *geo mapping*, dan sebagainya yang mampu memperlihatkan perubahan dan perbedaan data menjadi lebih jelas. Alhasil, tampilan data hasil analisis statistik biasa menjadi lebih cantik dengan visual yang *eye catching* menggunakan pilihan dashboard interaktif yang disediakan *tableau* (Saepuloh, 2020).

Tableau juga dapat mempermudah pembuatan analisis visual interaktif dalam bentuk *dashboard*. Adapun pengertian *tableau* lain yaitu *tableau* adalah software yang mendukung visualisasi data secara kolaboratif bagi seseorang yang bekerja dalam menganalisis informasi (Marvaro & Sefina Samosir, 2021).

2.2.8 Streamlit

Streamlit adalah kerangka kerja (*framework*) *open source* gratis yang dapat diakses oleh semua orang. Masih banyak fitur-fitur *Streamlit* yang perlu dijelajahi dan diteliti lebih lanjut. Hal ini karena *Streamlit* dapat memberikan tampilan yang menarik dan interaktif. *Streamlit* memungkinkan pengembang untuk membangun *dashboard* interaktif dan aplikasi web berbasis pembelajaran mesin. *Streamlit* dapat dikategorikan sebagai *scripting* sisi server (Jaelani & Pangestu, 2023).

Keuntungannya adalah bahwa *Streamlit* memiliki bahasa yang mudah dipahami tanpa perlu tahu cara menampilkan pembuatan *web* menggunakan

HTML, *CSS*, atau *JavaScript*. Namun, *streamlit* juga memiliki kelemahan, yaitu dalam membuat aplikasi *streamlit*, sensitivitas dalam membuat skrip disebabkan karena *streamlit* menggunakan bahasa pemrograman *python* yang merupakan bahasa pemrograman yang peka terhadap huruf besar-kecil (*case-sensitive*) dan bahasa tingkat tinggi (Jaelani & Pangestu, 2023).

2.2.9 Agile

Metodologi *Agile* membedakan diri dari pendekatan pengembangan tradisional dengan menekankan pentingnya pengiriman nilai yang berkelanjutan dan responsif terhadap perubahan (Ahmad et al., 2022).

Metodologi *Agile* ini memiliki beberapa tahapan atau fase yang dilakukan untuk mengembangkan sebuah aplikasi atau *software*. tahap perencanaan (*planning*), tahap implementasi (*implementation*), tahap pengujian perangkat lunak (*testing*), tahap pembuatan dokumentasi (*documentation*), tahap implementasi sistem (*deployment*), dan tahap pemeliharaan (*maintenance*) (M. Asep Rizkiawan et al., 2023).

2.2.10 Mean Absolute Error

Mean Absolute Error (MAE) adalah rata-rata nilai aktual dan nilai prediksi yang bernilai mutlak positif (Ihzaniah et al., 2023). Persamaan 2.1 tersebut merupakan cara perhitungan untuk menghitung MAE

$$\frac{1}{\Pi} \sum_t^{\Pi} = 1 |At - Ft| \quad (2.1)$$

Berikut keterangan dari Persamaan 2. 1,

Π : Jumlah Data

Σ : Penjumlahan

At : Nilai Aktual

Ft : Nilai Prediksi

2.2.1 Mean Absolute Percentage Error

Mean Absolute Percent Error (MAPE) merupakan rata-rata absolut antara nilai peramalan dan nilai aktual yang dinyatakan sebagai persentase. MAPE adalah pengukuran statistik tentang akurasi perkiraan (prediksi) pada metode peramalan. MAPE digunakan untuk menghitung persentase kesalahan rata-rata secara mutlak (absolut) (Nurani et al., 2023). Yang dimana semakin kecil persentase yang dihasilkan pada rumus ini maka semakin baik model. Persamaan 2.2 tersebut merupakan cara perhitungan untuk menghitung MAPE

$$\frac{100}{\Pi} \sum_{t=1}^{\Pi} \left| \frac{At - Ft}{At} \right| \quad (2.2)$$

Berikut keterangan dari Persamaan 2. 2,

Π : Jumlah Data

Σ : Penjumlahan

At : Nilai Aktual

Ft : Nilai Prediksi

Tabel 2. 2 Range Nilai MAPE

Nilai Mape	Akurasi Prediksi
< 10 %	Sangat Baik
10 – 20 %	Baik
20 – 50 %	Cukup
> 50 %	Buruk

Sumber : Nurani et al., 2023

Pada Tabel 2.3, Jika MAPE kurang dari 10%, prediksinya sangat akurat. Antara 10% sampai 20%, prediksinya baik. Antara 20% sampai 50%, prediksinya masih cukup bagus. Di atas 50%, prediksinya rendah. Idealnya, usahakan MAPE di bawah 10%.(Nurfadilah et al., 2022).