

BAB I PENDAHULUAN

1.1. Latar Belakang

Sebagai sumber daya yang penting, data merupakan aset yang sangat berharga bagi berbagai kepentingan terlebih di era transformasi dan digitalisasi informasi yang masif sekarang ini. Pembahasan dan penelitian tentang data, *big data*, dan *data mining* semakin gencar dilaksanakan oleh berbagai pihak, dari skala kecil individu untuk penelitian, bisnis, industri, sampai sektor pemerintahan, salah satunya sektor pariwisata di Kabupaten Pati yang menjadi objek dan studi kasus penelitian penulis. Berdasarkan pengamatan penulis, pariwisata kabupaten Pati semakin bertumbuh dan banyak jumlahnya dari tahun ke tahun. Dengan banyaknya pariwisata yang berkembang dan pulihnya operasional pasca pandemi covid-19 banyak data yang diperlukan untuk pengembangan wisata yang terstruktur dengan baik oleh pemerintah setempat.

Kabupaten Pati merupakan salah satu kabupaten yang berada di provinsi Jawa Tengah. Dengan letak strategis di pesisir utara pulau Jawa dan luas wilayah yang tercatat 150.368 Ha, menjadikan kabupaten Pati wilayah yang memiliki beberapa jenis pariwisata. Kabupaten Pati memiliki beberapa sistem informasi untuk setiap sektor pemerintahannya, termasuk sektor pariwisata. Akan tetapi pada sektor pariwisata pemanfaatan sistem informasi dan data pariwisata yang ada belum optimal. Tidak optimalnya pemanfaatan data tersebut diperburuk dengan kondisi pandemi COVID-19 di Indonesia yang berlangsung selama 2 tahun, sehingga data kunjungan wisata mengalami penurunan yang cukup drastis dibandingkan dengan data saat kondisi pra pandemi, hal ini menyebabkan pola data kunjungan wisatawan pada kondisi pra pandemi dan pasca pandemi cukup berbeda, sehingga perlakuan dalam memanfaatkan data-data tersebut juga menjadi berbeda.

Pandemi COVID-19 yang terjadi pada tahun 2020 telah berdampak besar pada sektor pariwisata di Indonesia, termasuk di Kabupaten Pati. Sektor pariwisata merupakan salah satu sumber penghasilan utama bagi masyarakat Kabupaten Pati dan juga salah satu sumber pendapatan daerah, sehingga dampak yang ditimbulkan oleh pandemi ini sangat merugikan bagi keberlangsungan hidup

masyarakat dan pendapatan wisata tersebut. Selain menyebabkan penurunan jumlah wisatawan yang datang ke Kabupaten Pati, pandemi ini juga menyebabkan terhambatnya kegiatan-kegiatan pariwisata yang sudah direncanakan sebelumnya. Oleh karena itu, penelitian ini bermaksud untuk mengkaji pemanfaatan data wisata yang ada di kondisi pasca pandemi ini dengan harapan dapat mengidentifikasi strategi yang dapat dilakukan untuk meningkatkan kembali daya tarik pariwisata di kabupaten Pati.

Dalam kondisi pasca pandemi, pola kunjungan wisatawan di Pati juga dipengaruhi kecenderungan sebagian besar wisatawan dalam mencari informasi pariwisata sejak operasional wisata kembali normal. Selama ini wisatawan atau masyarakat umum mengakses informasi wisata hanya melakukan pencarian tempat lokasi wisata yang sedang ramai dibicarakan atau tempat wisata yang mengadakan *event* tertentu yang dipromosikan, hal tersebut menjadikan jumlah kunjungan wisata di tempat tersebut akan lebih sering meningkat sedangkan jumlah kunjungan di tempat wisata yang tidak populer, tidak mengadakan event atau promosi akan semakin sepi. Dari faktor-faktor tersebut dapat dijadikan sebagai salah satu variabel pemanfaatan data pada penelitian ini, dimana dari data kunjungan yang ada dapat dianalisis untuk prediksi potensi kunjungan wisata pada periode waktu selanjutnya. Metode *data mining* diharapkan dapat membantu pemangku kepentingan pengelola pariwisata dalam mengambil keputusan dengan menganalisis potensi kunjungan setiap tempat wisata. Dataset kunjungan wisatawan kabupaten Pati yang terbentuk berdasarkan hasil pra-pemrosesan data oleh penulis menghasilkan data dengan bentuk kategorik dengan beberapa atribut parameter/prediktor dan label/target. Sedangkan metode yang digunakan adalah *Decision tree C4.5*

Metode *Decision tree C4.5* dipilih karena metode ini dikenal cukup baik menangani dataset awal yang digunakan penulis pada penelitian ini dengan melakukan transformasi data-data kunjungan wisata selama 2022 kedalam dataset yang memiliki bentuk parameter kategorik. *Decision tree C4.5* dikenal memiliki kemampuan dan akurasi yang baik dalam memodelkan data kategorik, baik data biner maupun non-biner. Pada proses transformasi penulis menambahkan parameter atau atribut berupa promosi, event, dan popularitas berdasarkan

observasi dan dokumentasi dalam proses pengumpulan dan transformasi data di objek penelitian. Berdasarkan hasil *preprocessing* dataset awal oleh penulis menggunakan *software* RapidMiner dengan metode *decision tree* C4.5 dan pustaka Scikit-Learn (sklearn-tree) dari bahasa pemrograman Python dengan *DecisionTreeClassifier* dan kriteria entropi, diperoleh hasil perhitungan skor akurasi prediksi di angka **96,42%**. Hasil perhitungan tersebut menunjukkan skor prediksi memiliki akurasi cukup baik dengan parameter yang diberikan.

Evaluasi model pada data penelitian penulis menggunakan metrik evaluasi yang lebih baik dibandingkan dengan beberapa penelitian prediksi dengan C4.5 sebelumnya mulai dari penelitian oleh (Purwanto, et al., 2020) yang menggunakan metrik *precision*, *recall*, *f-measure*, dan ROC tanpa metrik evaluasi nilai eror untuk 50 baris data pada penelitiannya yang memprediksi tingkat kasus pneumonia di Kabupaten Karawang. Kemudian penelitian oleh (Putri, et al., 2021) untuk prediksi penyakit diabetes yang hanya menggunakan metrik *accuracy* pada 49 baris data. Penelitian oleh (Suwandi, et al., 2022) pada prediksi stok barang PT Aerofood Indonesia menggunakan AUC, *recall*, dan *precision* tanpa menggunakan metrik nilai eror. Selanjutnya penelitian oleh (Damayanti, 2022), (Pa'o & Hendry, 2021), dan (Chan, 2018) ketiganya tidak mencantumkan hasil metrik evaluasi model atau nilai eror pada jurnalnya namun hanya menampilkan hasil perhitungan *gain* dan entropy pada atribut serta pembentukan rule pohon keputusan. Beberapa penelitian terkait tersebut dengan berbagai objek prediksi yang berbeda tidak menggunakan metrik yang sama dengan jumlah data bervariasi dan cenderung terlalu sedikit sehingga data awal penulis pada penelitian memiliki originalitasnya sendiri dan dapat diharapkan memberikan hasil prediksi yang lebih baik dari penelitian-penelitian sebelumnya.

Selain akurasi, penulis juga menghitung metrik evaluasi model yang lain pada dataset potensi kunjungan wisatawan yang terbentuk dengan metrik *Precision*, dan *Recall*. Hasilnya, metode *Decision tree* C4.5 memiliki evaluasi skor performa cukup baik untuk melakukan prediksi. Hasil dari perhitungan metrik evaluasi model dapat dilihat pada tabel berikut:

Tabel 1.1 Metriks evaluasi akurasi c4.5

Metrik	Skor Evaluasi
Akurasi	96,42%
<i>Precision</i>	96,42%
<i>Recall</i>	96,66%

Permasalahan pemanfaatan data pariwisata Kabupaten Pati pasca pandemi yang telah diuraikan mendorong penulis untuk berinisiatif melakukan penelitian dan selanjutnya pengembangan sistem yang bermanfaat untuk kebutuhan analisis data kunjungan pariwisata Pati. Berdasarkan akar permasalahan yang ada, maka penulis ingin menyelesaikan permasalahan ini dengan melakukan analisis potensi kunjungan wisatawan di kabupaten Pati dan membuat model untuk melakukan prediksi potensi kunjungan wisatawan pasca pandemi berdasarkan metode pemanfaatan data yang digunakan, yaitu *decision tree* atau pohon keputusan. Algoritma *decision tree C4.5* dalam penelitian dipilih karena data yang ada termasuk data kategorik dan dapat dilakukan *preprocessing* untuk menemukan pola klasifikasi 2 atau lebih variabel prediktor/independen dengan satu variabel label target/dependen untuk kemudian dilakukan perhitungan prediksi potensi kunjungan wisatawan objek wisata Pati di periode waktu kedepannya berdasarkan data historis yang sudah ada. Penelitian ini dirancang penulis guna bahan penelitian skripsi dengan mengajukan judul yaitu **“Prediksi Potensi Kunjungan Wisatawan Pasca Pandemi di Kabupaten Pati Menggunakan Metode *Decision tree C4.5*”**.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang terurai diatas, rumusan masalah pada penelitian ini yakni bagaimana membangun sebuah sistem berbasis website yang berfungsi untuk melakukan prediksi potensi tren kunjungan wisatawan pasca pandemi pada setiap objek wisata di Kabupaten Pati.

1.3. Batasan Masalah

Penulis membuat batasan masalah dalam penelitian ini agar dapat memberikan konteks yang jelas bagaimana arah pembahasan masalah yang telah dirumuskan. Batasan masalah dibuat agar rumusan masalah yang tercakup didalamnya tidak menyimpang dari tujuan awal penelitian ini sehingga beresiko mengurangi efektifitas pemecahannya. Maka dari itu, penulis menentukan batasan masalah sebagai berikut:

1. Sistem prediksi potensi kunjungan wisata dalam penelitian ini dirancang berbasis website.
2. Sistem ini mengelola data kunjungan wisatawan di setiap objek wisata yang dilaporkan oleh operator pariwisata pada sistem.
3. Sistem ini berfungsi mengekstrak informasi berupa hasil potensi prediksi kunjungan wisatawan dari dataset yang sudah diproses dalam sistem.
4. Fitur *data mining* dalam sistem yang dibangun pada penelitian ini dirancang hanya dapat digunakan pengguna dengan peran Kepala Bidang (Kabid) Pemasaran Pariwisata untuk melakukan pemrosesan dataset dan prediksi dengan metode tersebut.
5. Aktivitas pada sistem mulai dari proses CRUD (*create, read, update, delete*), validasi data pengguna, dan pemrosesan data wisata menggunakan penerapan metode klasifikasi pohon keputusan.
6. Fitur *data mining* dibuat menggunakan metode *decision tree* dengan algoritma C4.5.
7. Fitur *data mining* dalam aplikasi ini menghasilkan keluaran atau *output* berupa hasil prediksi potensi kunjungan objek wisata dengan label potensi: **Meningkat dan Menurun.**

1.4. Tujuan

Tujuan dari penelitian yang dilakukan penulis adalah sebagai berikut:

1. Membangun sistem yang dapat memprediksi potensi kunjungan wisatawan pada objek wisata Kabupaten Pati.
2. Menerapkan metode *data mining* dengan tepat untuk mendapatkan informasi berguna dari data wisata yang ada.

1.5. Sistematika Penulisan

BAB I: PENDAHULUAN

Bab ini memuat hal-hal meliputi latar belakang, perumusan masalah, batasan masalah, tujuan, manfaat penelitian, dan sistematika penulisan.

BAB II: TINJAUAN PUSTAKA

Bab ini memuat ringkasan penelitian-penelitian terdahulu mengenai topik yang sama dengan topik yang diangkat oleh penulis. Pada Bab II juga memuat landasan teori-teori yang terkait dengan topik bahasan penulis.

BAB III: METODOLOGI PENELITIAN

Bab ini memaparkan metode penelitian yang digunakan selama penelitian mulai dari tahap awal pengumpulan data hingga pembangunan sistem. Bab ini juga memaparkan perancangan sistem mulai dari analisa sistem hingga desain input output.

BAB IV: HASIL DAN PEMBAHASAN

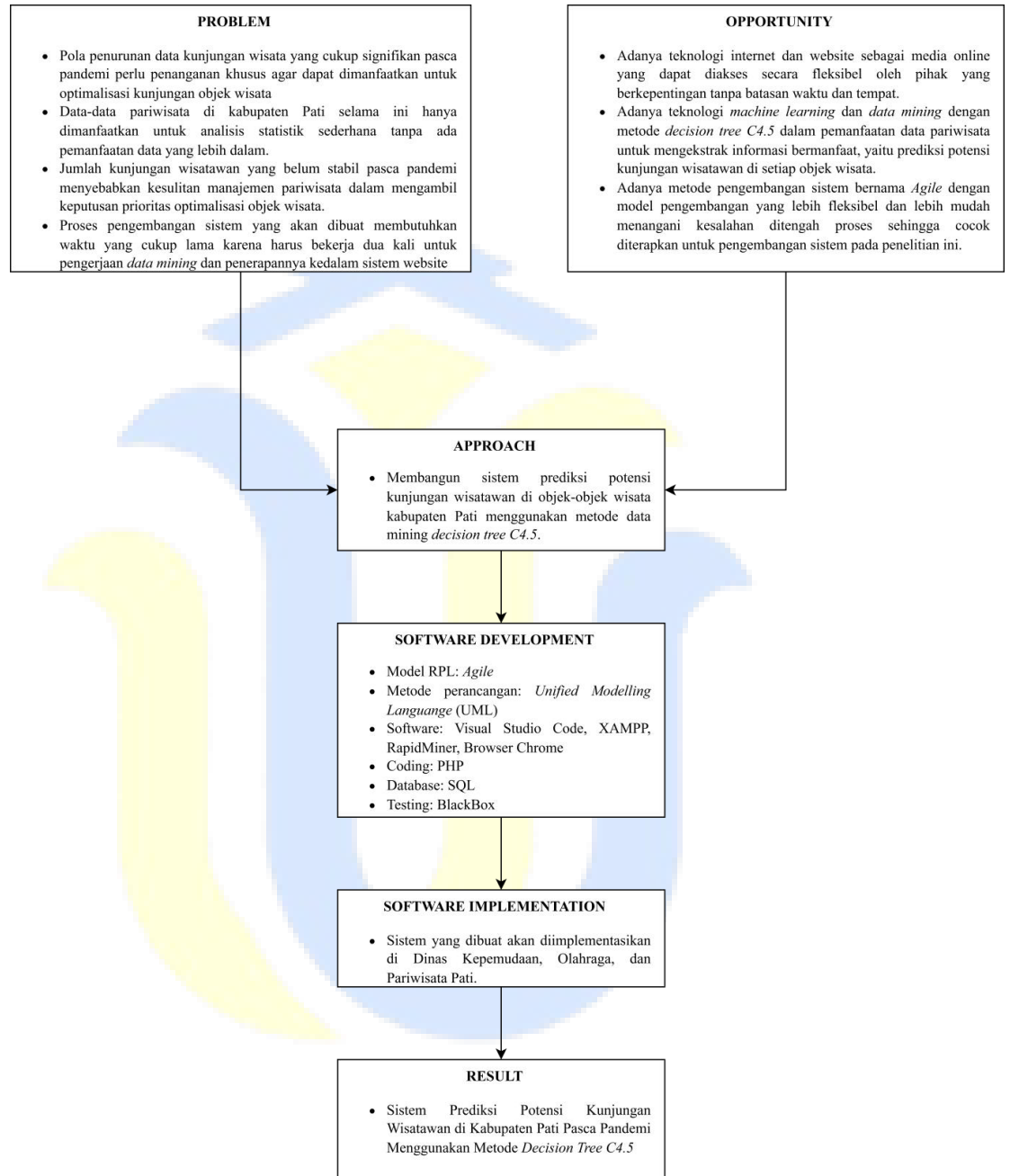
Pada bab ini berisi hasil implementasi dari sistem yang telah selesai, mulai dari tampilan interface hingga pengujian sistem.

BAB V: PENUTUP

Bab ini berisi kesimpulan kesimpulan dari penelitian yang telah dilaksanakan serta saran untuk penelitian selanjutnya.

1.6. Kerangka Pemikiran

Berikut adalah kerangka pemikiran yang digunakan dalam menyelesaikan penelitian ini, dapat dilihat pada gambar berikut.



Gambar 1.1 Kerangka pemikiran