

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Penggunaan teknologi informasi di sektor pariwisata dapat mendorong ekonomi daerah dan membuka peluang usaha bagi lingkungan masyarakat. Sebagai masyarakat dengan mengikuti perkembangan teknologi dapat mempermudah memperoleh informasi mengenai tempat wisata dan memahami nilai-nilai budaya disuatu daerah. Integrasi teknologi informasi sebagai strategi dalam pemasaran pariwisata akan menguntungkan baik penyedia layanan dan pengelola wisata [1].

Menurut Undang-undang No. 10 Tahun 2009 tentang Kepariwisata, pengertian wisata adalah kegiatan perjalanan yang dilakukan oleh seseorang atau sekelompok orang dengan mengunjungi tempat tertentu untuk tujuan rekreasi, pengembangan pribadi, atau mempelajari keunikan daya tarik wisata yang dikunjungi dalam jangka waktu sementara [2].

Berdasarkan hasil wawancara dengan pihak terkait pemasaran Wisata Kabupaten Rokan Hulu bahawasanya masing-masing tempat pariwisata mempunyai karakteristik yang berbeda-beda, sehingga hal ini akan menimbulkan hal tersendiri dalam berwisata. Untuk memahami karakteristik tempat wisata, pengelompokan tempat pariwisata sangat penting dalam industri pariwisata. Dengan memahami karakteristik dan referensinya para penyedia layanan pariwisata dapat mengembangkan destinasi wisata-wisata.

Hasil kuesioner penyebaran angket memberikan gambaran hasil dari responden yang diterima, terdapat 57.1% masyarakat tidak mengetahui apa itu pengelompokan wisata, terdapat 66.7% masyarakat tidak memahami pengelompokan wisata di Kabupaten Rokan Hulu. warga yang belum memahami secara jelas tentang konsep pengelompokan wisata menjadi isu yang patut diperhatikan dan terdapat 95.2% masyarakat yang percaya dengan adanya pengelompokan wisata akan membantu. Meskipun tempat wisata semakin bertambah dan variasinya pun semakin beragam, pengetahuan masyarakat umum tentang bagaimana mengelompokkan dan memahami jenis-jenis wisata masih minim. Ini akan berdampak terhambatnya pemahaman karakteristik tempat wisata. Mengingat semakin bertambah nya tempat wisata.

Agar pengelompokan dapat diketahui dan dipahami oleh masyarakat perlu adanya metode guna untuk membantu pemasaran wisata berdasarkan pengelompokan destinasi. Salah satu metode yang dapat digunakan untuk melakukan pengelompokan tempat wisata adalah *k-means Clustering*. Diantara sekian metode *Clustering* yang ada, *metode k-means Clustering* merupakan yang paling sederhana dan umum digunakan. Konsep dari algoritma ini yaitu obyek yang merupakan anggota dari suatu kumpulan Data, dikelompokkan menjadi sejumlah *k cluster* atau kelompok. Sehingga dalam aplikasinya, nilai *k* atau jumlah *clusternya* harus ditentukan terlebih dahulu. Secara detail, dalam penerapannya, ukuran ketidakmiripan biasa digunakan dalam rangka mengelompokkan objek-objek ini ke dalam suatu *cluster*.

Berdasarkan latar belakang yang telah dijabarkan diatas, maka akan dilakukan penelitian proposal ini berjudul **“Implementasi Data Mining**

Menggunakan *K-Means Clustering* Dalam Mendukung Pengelompokan Tempat Wisata Kabupaten Rokan Hulu ”.

1.2 Rumusan Masalah

Bersesuaian dengan masalah yang telah dijabarkan, maka rumusan masalah dalam skripsi ini adalah bagaimana menerapkan *k-means Clustering* untuk mengelompokkan tempat wisata pada Dinas Kebudayaan dan Pariwisata Kabupaten Rokan Hulu?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah merancang dan membangun untuk mendukung pengelompokan tempat wisata menggunakan metode *k-means Clustering* berdasarkan karakteristik masing-masing tempat wisata.

1.4 Batasan Masalah

Batasan masalah dilakukan agar pembahasan masalah tidak terlalu melebar. Berdasarkan pada rumusan masalah yang ada di atas maka batasan masalah dalam menerapkan *Data Mining* dengan memakai *k-means Clustering* terkait pengelompokan tempat wisata yang ada di Kabupaten Rokan Hulu yaitu:

1. Penelitian berfokus pada pengelompokan tempat wisata di Kabupaten Rokan Hulu guna mengetahui destinasi wisata yang berpotensi untuk dipahami sehingga dapat meningkatkan kontribusi sektor wisata. Data yang digunakan dalam proses ini adalah Data tahun 2024.
2. Perhitungan *K-Means Clustering* menggunakan *tools Python* dengan *Library Sk-Learn*.

3. pengelompokan tempat wisata berdasarkan Destinasi. Religi, Alam Dan Sejarah

1.5 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah untuk menerapkan metode *K-Means Clustering* untuk pengelompokan wisata Kabupaten Rokan Hulu.

1.6 Manfaat Penelitian

Penyusunan skripsi ini diharapkan memberikan manfaat, diantaranya yaitu: Guna mendukung terealisasinya hal tersebut perlu dibuat sebuah pengelompokan tempat wisata berdasarkan pengelompokan. Pengelompokan terbagi dari beberapa destinasi. yaitu Religi, Alam Dan Sejarah Menjadi media informasi olah Data untuk mengetahui kondisi tempat wisata di kabupaten Rokan Hulu pada tahun 2024

1.7 Sistematika Penulisan

Sistematika penulisan dari skripsi ini terdiri dari lima bagian utama sebagai berikut :

BAB 1 PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini berisi teori-teori yang digunakan pada penelitian ini. Teori-teori yang berhubungan dengan Data *Mining* metode *K-Means Clustering*.

BAB 3 METODOLOGI PENELITIAN

Bab ini berisi tahapan–tahapan dalam pengumpulan Data, perancangan sistem perumusan masalah dan analisa.

BAB 4 ANALISA DAN PERANCANGAN

Bab ini berisi analisa dan perancangan aplikasi Data *Mining* pengelompokan tempat wisata Pada Kabupaten Rokan Hulu Menggunakan Metode *K-means Clustering*.

BAB 5 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi implementasi dari analisa dan perancangan dan pengujian pada aplikasi yang berhasil dibangun.

BAB 6 PENUTUP

Bab ini berisi rangkuman dari hasil penelitian yang telah dilakukan dan saran – saran untuk pengembangan aplikasi atau penelitian selanjutnya.

BAB 2

LANDASAN TEORI

2.1. *Data Mining*

Data *mining* adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan dan pembelajaran mesin (*machine learning*) mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *Database* yang terkait [3].

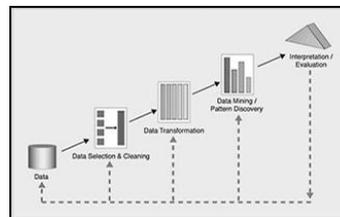
Data *mining* adalah pembelajaran mesin, pengenalan pola, *Database*, statistik, dan teknik visualisasi yang digunakan untuk memecahkan masalah penggalian informasi dari repository *Database* besar. Data mining sangat penting dalam proses penggalian Data secara manual dari kumpulan Data berupa pengetahuan yang tidak diketahui [4].

Data *Mining* merupakan suatu proses pencarian teknik analisis Data yang besar untuk proses yang diperoleh dari berbagai macam basis Data seperti Data relasional Data berorientasi objek dan Data transaksi untuk mengetahui informasi yang baru didalam *Database*. Data mining termasuk dalam proses *KDD* (*Knowledge Discovery Database*) yang dibantu oleh komputer untuk menggali dan menganalisis sejumlah besar himpunan Data dan mengekstrak informasi dan pengetahuan. Data *mining* terdiri dari beberapa fase yaitu Estimasi, *Clustering*, Prediksi, Klasifikasi dan Asosiasi [5].

2.2. *Knowledge Discovery in Database (KDD)*

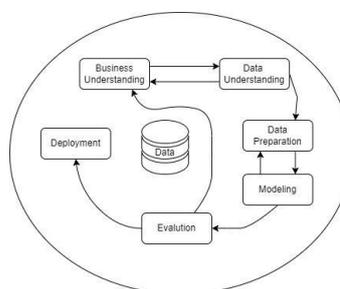
KDD adalah metode yang digunakan dalam penelitian ini. *KDD* adalah teknik untuk mendapatkan informasi berupa Data dari basis Data yang sudah ada.

Ada tabel yang terhubung satu sama lain dalam basis Data. Informasi dari Data yang dikumpulkan melalui proses *KDD* dapat berfungsi sebagai *Database* untuk pengambilan keputusan [6].



Gambar 2. 1 Knowledge Discovery in Database

Pada penerapannya, *Clustering* salah satu teknik *Data Mining*. *Clustering* merupakan teknik yang melakukan peng-*clusteran* Data sesuai tingkat kemiripan karakteristik antar Data. Pada penerapan *Data Mining*, perlu menjalankan suatu metode agar dapat diterapkan secara sistematis, salah satunya adalah metode *Cross-Standard Industry for Data Mining (CRISP-DM)*. metode *Cross-Standard Industry for Data Mining (CRISP-DM)* memiliki beberapa tahapan yaitu sebagai berikut:



Gambar 2. 2 Metode CRISP-DM

1. *Business Understanding*

Hal-hal yang dilaksanakan pada bagian ini meliputi pemahaman mengenai kebutuhan dan juga tujuan berdasarkan perspektif bisnis yang nantinya

akan dapat dimaknai pendefinisian masalah yang ada sebagai pengetahuan dalam bentuk di *Data Mining* yang nantinya dapat menetapkan rencana dan juga strategi dalam mencapai tujuan *Data Mining*.

2. *Data Understanding*

Langkah awal yang dilakukan yaitu mengumpulkan Data, diekspresikan Data kemudian dievaluasi kualitas Data.

3. *Data Preparation*

Pada bagian ini yaitu proses membangun Dataset akhir dari beberapa Data mentah. Hal-hal yang dilakukan yaitu mengenai proses pembersihan Data (*Data Cleaning*), berbagai atribut, *record*, memilih-milih Data (*Data Selection*), serta melakukan transformasi terhadap Data tersebut (*Data Transformation*) sehingga dapat diberikan masukan dalam proses pemodelan.

4. *Modeling*

Proses ini akan melibatkan machine learning secara langsung yang digunakan untuk menentukan cara *Data Mining*, dan juga membantu pembuatan algoritma *Data Mining*.

5. *Evaluation*

Tahap ini dilakukan dengan membuat interpretasi dari proses *Modelling* yang telah dilakukan.

6. *Deployment*

Pada proses ini perencanaan untuk *deployment* dimulai selama business

understanding dan harus menggabungkan tidak hanya bagaimana untuk menghasilkan nilai model, tetapi juga bagaimana mengkonversi skor keputusan/pengujian, dan bagaimana untuk menggabungkan Keputusan dalam sistem operasional. Namun pada penelitian ini untuk tahapan proses dari *CRISP-DM* hanya dilakukan sampai tahapan *Evaluation* saja karena untuk melanjutkan ke tahapan *Deployment* membutuhkan waktu lebih panjang yang akan dilakukan [5].

2.3. Implementasi

Secara etimologis pengertian implementasi menurut Kamus *Webster* berasal dari Bahasa Inggris yaitu *to implement*. Dalam kamus tersebut, *to implement* (mengimplementasikan) berarti *to provide the means for carrying out* (menyediakan sarana untuk melaksanakan sesuatu) dan *to give practical effect to* (untuk menimbulkan dampak/akibat terhadap sesuatu). Sementara itu dalam Kamus Besar Bahasa Indonesia, implementasi adalah pelaksanaan atau penerapan. Pengertian lain dari implementasi yaitu penyediaan sarana untuk melakukan sesuatu yang menimbulkan dampak atau akibat terhadap suatu hal. Pengertian dari implementasi juga bisa berbeda tergantung dari disiplin ilmunya [7].

Implementasi adalah bermuara pada aktivitas, aksi, tindakan atau adanya mekanisme suatu sistem. Implementasi bukan sekedar aktivitas tetapi suatu kegiatan yang terencana dan untuk mencapai tujuan kegiatan [8].

Pengertian implementasi dalam kamus besar Bahasa Indonesia(KBBI) secara umum adalah suatu tindakan atau pelaksanaan rencana yang telah disusun secara cermat dan rinci. Sedangkan menurut Totnanatzky dan Jhonshon

Implementasi sebagai *the translation of any tool technique process or method of doing from knowledge to practice*. Dari dua pendapat diatas implementasi dapat diartikan sebagai penerapan metode dari sebuah rencana menjadi tindakan [9].

2.4. Wisata

Di dalam kamus besar Bahasa Indonesia disebutkan bahwa destinasi merupakan kata benda yang berarti lokasi / tempat tujuan pengiriman. Istilah ini digunakan untuk menyebutkan suatu tempat signifikan yang akan dituju ketika seseorang dalam perjalanan. Pengertian wisata adalah suatu perjalanan yang dilakukan manusia baik perorangan maupun kelompok untuk mengunjungi destinasi tertentu dengan tujuan rekreasi, mempelajari keunikan daerah wisata, pengembangan diri dan sebagainya dalam waktu yang singkat atau sementara waktu [10].

Wisata adalah suatu proses bepergian yang bersifat sementara yang dilakukan seseorang untuk menuju tempat lain di luar tempat tinggalnya. Motif kepergiannya tersebut bisa karena kepentingan ekonomi, kesehatan, agama, budaya, social, politik, dan kepentingan lainnya. Menurut UU RI No 10 Tahun 2009 wisata merupakan kegiatan perjalanan yang dilakukan manusia baik perorangan maupun kelompok untuk mengunjungi destinasi tertentu dengan tujuan rekreasi, mempelajari keunikan daerah wisata, pengembangan diri dsb dalam kurun waktu yang singkat atau sementara waktu. Wisata merupakan suatu kegiatan perjalanan baik individu maupun grup dari tempat tinggal menuju suatu tempat tertentu untuk mendapatkan pengalaman di luar aktivitas kesehariannya

(seperti: bekerja, sekolah, mengurus rumah tangga) dalam waktu yang sementara [11].

Wisata adalah fenomena pergerakan manusia, barang, dan jasa, yang sangat kompleks. Ia terkait erat dengan organisasi, hubungan-hubungan kelembagaan dan individu, kebutuhan layanan, penyediaan kebutuhan layanan, dan sebagainya [12].

2.5. *K-Means*

Algoritma *K-means* merupakan algoritma yang relatif sederhana untuk mengklasifikasikan atau mengelompokkan sejumlah besar objek dengan atribut tertentu kedalam kelompok-kelompok (*cluster*) sebanyak K , pada algoritma *K-means* jumlah *cluster* K sudah ditentukan lebih dahulu [3].

Langkah 1: Menentukan jumlah *K-cluster* yang ingin dibentuk.

Langkah 2: Membangkitkan nilai acak untuk pusat *cluster* awal (*centroid*) sebanyak K .

Langkah 3: Menghitung jarak setiap Data input terhadap masing-masing centroid menggunakan rumus jarak Euclidean (*Euclidean Distance*) hingga ditemukan jarak yang paling dekat dari setiap Data dengan *centroid*.

Rumus *Euclidean Distance*:

$$d(x, \mu) = \sqrt{\sum_{i=1}^n (x_i - \mu_i)^2}$$

Langkah 4: Mengklasifikasikan setiap Data berdasarkan kedekatannya dengan *centroid* (jarak terkecil).

Langkah 5: Mengupdate nilai centroid. Nilai *centroid* baru diperoleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus:

$$\mu_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_i$$

Di mana μ_j adalah *centroid* baru untuk *cluster j*, N_j adalah jumlah Data dalam *cluster j*, dan x_i adalah Data dalam *cluster j*

Setelah langkah-langkah ini diulang sampai tidak ada perubahan yang signifikan dalam posisi *centroid* atau konvergensi tercapai, proses *Clustering* dianggap selesai.

2.6. MySQL

MySQL adalah salah satu jenis *Database* server yang sangat populer, hal ini disebabkan karena *MySQL* menggunakan *SQL* sebagai bahasa dasar untuk mengakses *Databasenya*. *MySQL* bersifat Open Source, Software ini dilengkapi dengan Source code (kode yang dipakai untuk membuat *MySQL* [13].

MySQL merupakan suatu jenis *Database* server yang sangat terkenal. *MySQL* termasuk jenis *RDBMS* (*Relational Database Management System*). *MySQL* mendukung bahasa pemrograman *PH*, bahasa permintaan yang terstruktur, karena pada penggunaannya *SQL* memiliki beberapa aturan yang telah distandarkan oleh asosiasi yang bernama ANSI. *MySQL* merupakan *RDBMS* (*Relational Database Management System*) server. *RDBMS* adalah program yang memungkinkan pengguna *Database* untuk membuat, mengelola, dan menggunakan Data pada suatu model *relational* [14].

MySQL merupakan *software* yang tergolong sebagai *DBMS* (*Database Management System*) yang bersifat *Open Source* [10]. *Open source* menyatakan

bahwa *software* ini dilengkapi dengan *source code* (kode yang dipakai untuk membuat *MySQL*), selain tentu saja bentuk *executable*-nya atau kode yang dapat dijalankan secara langsung dalam sistem operasi [15].

2.7. Hypertext Preprocessor (PHP)

PHP merupakan singkatan dari *PHP Hypertext Preprocessor*. *PHP* merupakan bahasa pemrograman script yang diletakkan dalam server yang biasa digunakan untuk membuat aplikasi web yang bersifat dinamis [13].

PHP adalah bahasa pelengkap *HTML* yang memungkinkan dibuatnya aplikasi dinamis yang memungkinkan adanya pengolahan Data dan pemrosesan Data. Semua syntax yang diberikan akan sepenuhnya dijalankan pada server sedangkan yang dikirimkan ke browser hanya hasilnya saja. Kemudian merupakan bahasa berbentuk script yang ditempatkan dalam server dan diproses di server. Hasilnya akan dikirimkan ke client, tempat pemakai menggunakan browser. *PHP* dikenal sebagai sebuah bahasa scripting, yang menyatu dengan tag-tag *HTML*, dieksekusi di server, dan digunakan untuk membuat halaman web yang dinamis seperti halnya *Active Server Pages (ASP)* atau *Java Server Pages (JSP)*. *PHP* merupakan sebuah *software Open Source* [14].

PHP (PHP: Hypertext Preprocessor) adalah suatu bahasa pemrograman yang digunakan untuk menterjemahkan basis kode program menjadi kode mesin yang dapat dimengerti oleh komputer yang bersifat *server side* yang ditambahkan ke *HTML*. *Hypertext Preprocessor (PHP)* merupakan bahasa pemrograman untuk pembuatan website dinamis, yang mampu berinteraksi dengan pengunjung atau penggunanya [15].

2.8. *CodeIgniter*

CodeIgniter adalah sebuah *Framework PHP* yang dapat membantu mempercepat *developer* dalam pengembangan aplikasi *web* berbasis *PHP* dibanding jika menulis semua kode program dari awal. *CodeIgniter* pertama kali dibuat oleh Rick Ellis, *CEO Ellislab, Inc.* (<http://ellislab.com>), sebuah perusahaan yang memproduksi *CMS (Content Management System)* yang cukup handal, yaitu *Expression Engine* (<http://www.expressionengine.com>). Saat ini, *CodeIgniter* dikembangkan dan dimainkan oleh *Expression Engine Development Team* [16].

CodeIgniter adalah Sebuah *Framework PHP* yang bersifat *open source* dan menggunakan metode *MVC (Model, View, Controller)* untuk memudahkan *developer* atau *programmer* dalam membangun sebuah aplikasi berbasis *web* tanpa harus membuatnya dari awal [17].

Tujuan utama dari *CodeIgniter* adalah untuk memudahkan programmer dalam mengembangkan aplikasi secara cepat tanpa harus melakukan pemrograman dari nol. Dengan demikian, programmer *PHP* yang telah memiliki dasar pemrograman, dapat bekerja lebih cepat menghasilkan aplikasi yang dibutuhkan. Berikut ini struktur file dalam *CodeIgniter Framework* [18].

2.9. *Unified Modelling Language (UML)*

Website merupakan sebuah kumpulan *Unified Modeling Language* atau *UML* merupakan suatu metode yang digunakan untuk pemodelan sistem yang mendukung proses pembuatan model abstrak dari suatu sistem berupa diagram, dimana setiap model memberikan pandangan yang berbeda dari sistem tersebut [19].

Diagram *UML* terdiri dari banyak tipe dan jenis, berikut adalah beberapa diagram *UML* :

1. Use Case Diagram

Use Case diagram menggambarkan interaksi antara sistem dengan sistem eksternal dan pengguna, siapa yang akan menggunakan sistem dan bagaimana cara pengguna berinteraksi dengan sistem.

2. Class Diagram

Class diagram digunakan untuk menggambarkan struktur objek-objek bernama kelas yang menyusun suatu sistem dan juga hubungan antara kelas-kelas tersebut.

3. Sequence Diagram

Sequence diagram menggambarkan kelakuan objek pada use case dengan mendeskripsikan waktu hidup objek dan pesan yang dikirimkan dan diterima antar objek. Gambaran *Sequence* diagram dibuat minimal sebanyak pendefinisian use case yang memiliki proses sendiri atau yang penting semua use case yang telah didefinisikan interaksi jalannya pesan sudah dicakup pada *Sequence* diagram sehingga semakin banyak use case yang didefinisikan, maka *Sequence* diagram yang harus dibuat juga semakin banyak.

4. Activity Diagram

Activity diagram digunakan untuk memberikan gambaran mengenai alur dari sebuah proses bisnis, langkah dalam sebuah use cases, dan metode dari sebuah objek [19].

Unified Modeling Language (UML) merupakan kesatuan struktur dan cara bagi pemodelan desain program berorientasi objek (*OOP*) serta aplikasinya. *UML* adalah metodologi untuk mengembangkan sistem *OOP* dan sekelompok perangkat *tool* untuk mendukung pengembangan sistem tersebut. *UML* mulai diperkenalkan oleh *Object Management Group*, sebuah organisasi yang telah mengembangkan model, teknologi, dan standar *OOP* sejak tahun 1980-an. Sekarang *UML* sudah mulai banyak digunakan oleh para praktisi *OOP*. *UML* merupakan dasar bagi perangkat (*tool*) desain berorientasi objek dari IBM. *UML* adalah suatu bahasa yang digunakan untuk menentukan, memvisualisasikan, membangun, dan mendokumentasikan suatu sistem informasi. *UML* dikembangkan sebagai suatu alat untuk analisis dan desain berorientasi objek oleh Grady Booch, Jim Rumbaugh, dan Ivar Jacobson. Namun demikian *UML* dapat digunakan untuk memahami dan mendokumentasikan setiap sistem informasi. Penggunaan *UML* dalam industri terus meningkat. Ini merupakan standar terbuka yang menjadikannya sebagai bahasa pemodelan yang umum dalam industri peranti lunak dan pengembangan sistem [20].

Unified Modelling Language (UML) adalah keluarga notasi grafis yang didukung oleh meta-model tunggal, yang membantu pendeskripsian dan desain

sistem perangkat lunak, khususnya sistem yang dibangun menggunakan pemrograman berorientasi objek [21].

2.10. BlackBox Testing

BlackBox Testing merupakan Teknik pengujian perangkat lunak yang berfokus pada spesifikasi fungsional dari perangkat lunak. *BlackBox Testing* bekerja dengan mengabaikan struktur kontrol sehingga perhatiannya difokuskan pada informasi domain. *BlackBox Testing* memungkinkan pengembang software untuk membuat himpunan kondisi input yang akan melatih seluruh syarat-syarat fungsional suatu program[25]. *BlackBox* merupakan salah satu pengujian yang lebih difokuskan pada spesifikasi fungsi-fungsi (modul) yang dikembangkan, baik dari sisi struktur Data, pengaksesan Data dalam *Database*, serta beberapa kesalahan performance yang diberikan. Beberapa keuntungan saat penggunaan *Black Box Testing* adalah pengguna tidak perlu memahami bahasa pemrograman yang digunakan dalam sistem [22].

2.11. Penelitian Terkait

Berikut adalah tabel penelitian terkait, penelitian mengenai perancangan sistem menggunakan metode *K-Means Clustering*

Tabel 2. 1 Penelitian Terkait

No	Nama	Tahun	Judul	Hasil

1.	L Fimawahib, E Rouza	2021	Penerapan <i>K-Means Clustering</i> pada Penentuan Jenis Pembelajaran di Universitas Pasir Pengaraian	Penelitian ini adalah menentukan strategi perkuliahan yang ada di Universitas Pasir Pengaraian dengan menerapkan algoritma <i>K-Means Clustering</i> . Jumlah k yang digunakan sebanyak 3 (tiga).
2.	H Kurniawan, S Defit	2020	Data Mining Menggunakan Metode <i>K-Means Clustering</i> Untuk Menentukan Besaran Uang Kuliah Tunggal	Sistem penentuan besaran UKT calon mahasiswa baru menggunakan algoritma <i>K-Means Clustering</i> dilakukan dengan mengikuti tahap-tahap Data Mining. Berikut adalah proses Data Mining dalam menentukan besaran UKT calon mahasiswa baru menggunakan metode <i>K Means Clustering</i> .
3.	D Ariyanto	2022	Data Mining Menggunakan Algoritma <i>K-Means</i> Untuk Klasifikasi Penyakit Infeksi Saluran Pernafasan Akut	Tahap awal dimulai dari pengumpulan Data yang diperoleh dari Klinik Rahmatan Lil Alamin, kemudian dilakukan tahap seleksi Data berdasarkan kebutuhan Data yang akan diteliti, Data yang telah lulus hasil seleksi selanjutnya dilakukan proses transformasi Data agar Data dapat diproses dengan menggunakan metode <i>K-Means</i> . Setelah Data ditransformasi ke dalam bentuk yang bisa di proses dengan metode

				<i>K-Means</i> selanjutnya Data siap diproses dimulai dengan menentukan berapa jumlah <i>cluster</i>
4.	WS Utami N Pratiwi F Muhamm ad	2023	Penerapan Data <i>Mining</i> Menggunakan Algoritma <i>K-Means</i> Untuk <i>Clustering</i> Perokok Usia Lebih dari 15 Tahun	Hasil dari penelitian ini diperoleh tiga <i>cluster</i> , C1 dengan kategori tinggi sebanyak 10 <i>cluster</i> , C2 dengan kategori rendah sebanyak 15 <i>cluster</i> , dan C3 dengan kategori rendah sebanyak 9 <i>cluster</i> .
5.	I Nozomi	2023	Penerapan Data <i>Mining</i> Untuk Peringatan Dini Banjir Menggunakan Metode <i>Clustering K-Means</i>	Berdasarkan hasil perhitungan baik secara manual ataupun dengan <i>software RapidMiner v.5.3</i> dengan menggunakan 18 sampel Data didapatkan hasil yang sama. Untuk <i>cluster</i> 0 terdiri dari 10 <i>record</i> Data yang masuk kategori tingkatan rendah dengan peringatan dini waspada, <i>cluster</i> 1 terdiri dari 1 <i>record</i> Data yang masuk kategori tingkatan sedang dengan peringatan dini siaga, <i>cluster</i> 2 terdiri dari 7 <i>record</i> Data yang masuk kategori tingkatan tinggi dengan peringatan dini awas. Sedangkan pengujian dengan 547 <i>record</i> , dengan <i>cluster</i> 0 terdiri dari 396 <i>record</i> , <i>cluster</i> 1 terdiri dari 145 <i>record</i> , <i>cluster</i> 2 terdiri dari 6 <i>record</i>

6.	RR Putra C Wadisman	2018	Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma <i>K-Means</i>	Data pelanggan yang potensial didapatkan setelah perhitungan algoritma <i>K-Means</i> selesai, Data dengan pusat <i>centroid</i> terbesar yang termasuk ke dalam pelanggan yang paling potensial.
7.	MR Alhapizi M Nasir I Effendy	2020	Penerapan Data Mining Menggunakan Algoritma <i>K-Means Clustering</i> Untuk Menentukan Strategi Promosi Mahasiswa Baru Universitas Bina Darma Palembang	Setelah dilakukan pengelompokan Data mahasiswa berdasarkan persebaran kabupaten, asal sekolah, penghasilan ayah dan informasi maka diperoleh suatu gagasan bahwa <i>cluster</i> yang jumlahnya terbanyak adalah strategi promosinya dengan menggunakan brosur dan promosi dari teman atau kerabat dekat. Fungsi media sosial belum begitu efektif dilihat dari frekuensi statistik calon mahasiswa baru dari instansi atribut ini.
8.	H Syahputra L Mayola D Guswandi	2022	<i>Clustering</i> Tingkat Penjualan Menu (<i>Food and Beverage</i>) Menggunakan Algoritma <i>K-Means</i>	Proses <i>clusterisasi</i> mampu menyajikan pengetahuan baru dari Data yang ada. Kasus tingkat penjualan menu pada restoran mama. Hasil tersebut mampu memberikan hasil yang cukup baik dalam meng <i>cluster</i> menu makanan yang terbaik.

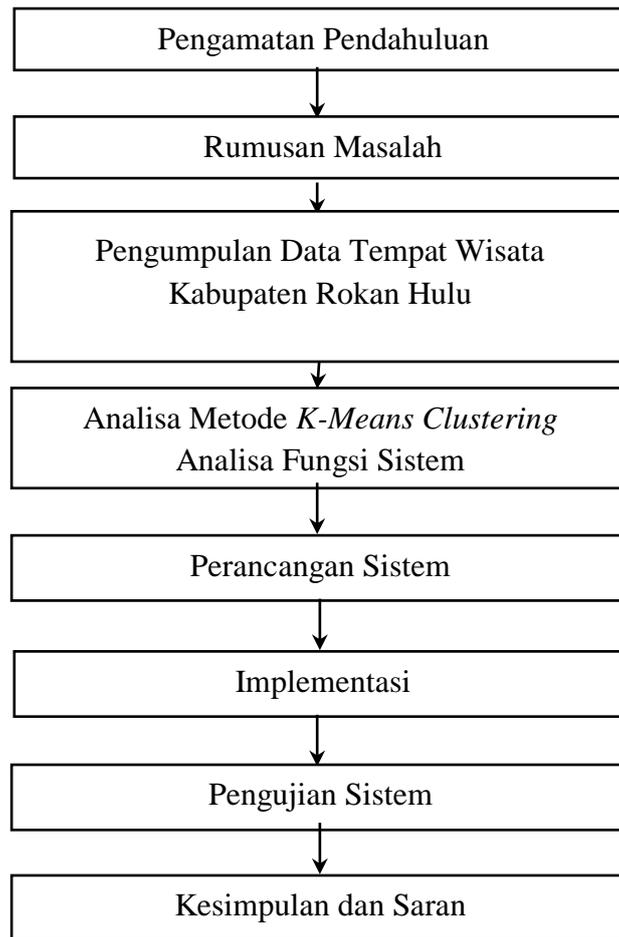
				Hasil penelitian ini juga menyajikan knowledge base untuk dijadikan dasar dalam manajemen penjualan yang ada di restoran mama.
9.	T Hidayat	2022	Klasifikasi Data Jamaah Umroh Menggunakan Metode K-Means <i>Clustering</i>	Setelah dilakukan perhitungan Data yang telah dilakukan melalui software RapidMiner diperoleh anggota kelompok sangat disukai dari rentang usia mulai 41 sampai 70 tahun, kelompok disukai dengan rentang usia mulai 21 sampai 40 tahun dan kelompok kurang disukai mulai usia 1 sampai 20 tahun. Dengan didapatkan hasil pengelompokan ini maka didapatkan pengetahuan baru yang dapat memudahkan manajer dalam strategi pemasaran pada AET Travel Indonesia
10	MA Sembiring RTA Agus	2021	Penerapan Metode Algoritma <i>K-Means Clustering</i> Untuk Pemetaan Penyebaran Penyakit Demam Berdarah Dengue (Dbd)	Dengan menggunakan Metode Algoritma <i>K-Means Clustering</i> , dapat ditemukan desa di Kecamatan Setia janji yang paling sering dan paling banyak masyarakatnya terjangkit penyakit DBD yaitu desa 2 dan 4. Berdasarkan devies bouldin index diketahui performance adalah sebesar 1.044 yang menunjukkan 3 <i>cluster</i> (K=3) adalah jumlah yang optimum dibandingkan dengan

				jumlah K lainnya.
--	--	--	--	-------------------

BAB 3

METODOLOGI PENELITIAN

Penerapan Data *Mining* dalam kajian ini terbagi ke dalam tiga proses yaitu proses awal, proses pengembangan dan proses akhir. Gambaran tahapan penelitian dapat pada gambar 3.1.



Gambar 3. 1 Tahapan Metodologi Penelitian

Pembuatan skripsi ini terbagi menjadi beberapa tahap pengerjaan yang tertera sebagai berikut:

3.1. Pengamatan Pendahuluan

Pengamatan pendahuluan merupakan tahapan awal yang dilakukan dalam penelitian ini, yang menggunakan metode *K-Means Clustering* yang dijadikan sebagai penelitian studi pustaka dalam penelitian skripsi ini.

3.2. Perumusan Masalah

Berdasarkan hasil dari tahapan pengamatan pendahuluan sebelumnya, maka tahapan selanjutnya adalah tahapan perumusan masalah. Pada tahapan perumusan masalah akan dirumuskan masalah yang dianggap sebagai penelitian dalam Skripsi ini. Permasalahan-permasalahan yang menjadi rumusan masalah dalam penelitian ini didapatkan dari penelitian dari penelitian terkait Data pengamatan pendahuluan sebelumnya. Solusi yang didapatkan pada tahapan perumusan masalah ini yang akan menjadi judul penelitian Skripsi ini “Implementasi Data Mining Menggunakan *K-Means Clustering* Dalam Mendukung Pengelompokan Tempat Wisata Kab Rokan Hulu (Studi Kasus : Dinas Pariwisata, Kecamatan Hutaraja Tinggi, Kabupaten Rokan Hulu, Provinsi Riau)”.

3.3. Pengumpulan Data

Pengumpulan Data adalah tahapan-tahapan yang bertujuan dalam memperoleh Data-Data informasi yang berhubungan dengan penelitian skripsi ini. Pada tahapan pengumpulan Data ini juga berguna untuk mengumpulkan semua kebutuhan Data yang akan diproses nantinya menggunakan metode “*The*

Technique for Order of Preference by Similarity to Ideal Solution (K-Means Clustering) ”

Dalam pengumpulan Data ini Data yang dikutip adalah sebagai berikut:

1. Wawancara

Wawancara dilakukan untuk mendapatkan Data penelitian. Dalam hal ini wawancara dilakukan kepada kepala Dinas Pariwisata Kab Rokan Hulu.

2. Analisis Kebutuhan Sistem

Dalam analisa sistem bertujuan mengidentifikasi sistem yang akan dirancang, yang meliputi perangkat lunak serta perangkat keras. Tahapan-tahapan yang menyusun analisa sistem ini adalah analisis Data yang dipakai, spesifikasi kebutuhan sistem, spesifikasi pengguna, perancangan basis Data, dan perancangan antarmuka.

3.4. Analisa Sistem

Tahapan selanjutnya adalah melakukan analisis metode sistem penelitian skripsi ini. Adapun tahapan analisa dalam penelitian ini sebagai berikut :

3.4.1 Analisa Metode K-Means Clustering

Tahapan ini adalah proses dimana langkah-langkah pengolahan Data menggunakan metode *K-Means Clustering* dijalankan.

3.4.2 Analisa Fungsi Sistem

Setelah melakukan tahapan analisis terhadap metode *K-Means Clustering*, maka selanjutnya adalah analisis fungsional yang akan dibangun. Adapun tahapan-tahapan analisis fungsional yaitu dalam pembuatan *Flowchart*, *Unified Modeling Language (UML)*, dan perancangan *user interface*.

3.5. Perancangan Sistem

Setelah tahapan analisis dilakukan, maka tahapan selanjutnya adalah perancangan sistem. Tahapan perancangan sistem terdiri dari :

- 1) Perancangan struktur menu yang akan digunakan pada sistem yang akan dibangun
- 2) Tahapan rancangan *Database* beserta atribut yang dibutuhkan.
- 3) Tahapan perancangan *user interface* atau antar muka pengguna terhadap sistem yang akan digunakan.

3.6. Pengujian

Pengujian merupakan sebuah tahapan yang memperlihatkan apakah prediksi tingkat akurasi dari penelitian sesuai dengan yang diinginkan atau tidak. Pengujian *black box*, digunakan untuk menguji tingkat kemampuan *user interface* terhadap sistem yang dibangun.

Metode *Black Box Testing* yaitu metode untuk menguji sebuah perangkat lunak tanpa harus memperhatikan hal detail perangkat lunak. Pengujian ini hanya memeriksa nilai keluaran berdasarkan nilai masukan masing- masing. Proses *Black Box Testing* dengan cara mencoba memasukan Data pada setiap formnya. Pengujian ini sangat diperlukan apakah program yang dibuat sesuai atau tidak sesuai[27].

Pengujian *Black Box* disebut juga sebagai *Behavioral Testing* yaitu metode pengujian perangkat lunak yang pengujinya tidak perlu mengetahui struktur

internal, desain aplikasi, dan implementasinya (pengkodean) dari perangkat lunak yang diuji. Pengujian *Black Box* ditujukan untuk menguji fungsionalitas dari perangkat lunak maupun non fungsionalitasnya, walaupun kadang hanya menguji fungsionalitasnya saja[28].

3.7. Kesimpulan dan Saran

Tahapan terakhir adalah menarik kesimpulan dari hasil penelitian yang didapatkan klasifikasi keluhan nasabah. Pada tahapan ini juga berisikan saran peneliti bagi pembaca untuk melakukan pengembangan terhadap penelitian ini kedepannya.