

## **BAB 1**

### **PENDAHULUAN**

#### **1.1. Latar Belakang Masalah**

Ujian Nasional atau di singkat UN diatur dalam Undang – Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional / UU Sisdiknas yang mengatur evaluasi standar pendidikan secara nasional oleh Depdiknas (Undang Undang Republik Indonesia nomor 20 Tahun 2003)

Evaluasi standar nasional pendidikan bertujuan untuk meningkatkan kualitas pendidikan nasional dalam mencerdaskan dan membentuk kepribadian anak bangsa , serta Fungsi dari standar nasional pendidikan adalah sebagai dasar dalam perencanaan, pelaksanaan,dan pengawasan pendidikan nasional yang bermutu . (Murărescu, E. D., Ivan, L. U. M. I. N. I. Ţ. A., & Mihailovici, 2005).

Ujian Nasional merupakan salah satu tahap evaluasi standar pendidikan yang dapat meningkatkan mutu dan kualitas pendidikan . Mewujudkan mutu dan kualitas pendidikan di Sekolah Menengah Pertama se Jawa Tengah perlu adanya peningkatan Nilai Ujian Nasional . apabila di sekolah Menengah Pertama se Jawa Tengah memiliki Nilai Ujian Nasional yang tinggi dapat digunakan sebagai tolak ukur bahwa sekolah tersebut memiliki kualitas dan mutu pendidikan yang baik.

Dengan dilakukan pengolahan data Nilai Ujian Nasional Sekolah Menengah Pertama se Jawa Tengah dapat menghasilkan klaster dari Nilai Ujian Nasional Sekolah Menengah Pertama se Jawa Tengah , untuk mendapatkan tiap klaster Nilai Ujian Nasional Sekolah Menengah Pertama se Jawa Tengah maka perlu adanya data mining *Clustering* . Data mining *Clustering* dilakukan untuk mengelompokkan sekolah mana saja di Provinsi Jawa Tengah yang memiliki Nilai Ujian Nasional Tingkat Sekolah Menengah Pertama yang rendah , sedang , dan tinggi. Sehingga nantinya dapat dijadikan bahan evaluasi terhadap hasil ujian di tiap – tiap sekolah.

*Clustering* adalah pengelompokkan data yang menggunakan statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terikat dari data yang besar (Aditya et al., 2020) . salah satu algoritma yang digunakan adalah Algoritma *K – Means* dan Algoritma *K - Medoids* .

Penelitian ini menggunakan dua Algoritma *K – Means* dan Algoritma *K – Medoids* untuk membandingkan hasil klaster dari kedua algoritma tersebut sehingga mendapatkan hasil klaster dari Algoritma *K- Means* dan *K - Medoids*. data yang digunakan adalah data capaian Ujian Nasional Sekolah Menengah Pertama Tahun 2018/2019 se Jawa Tengah yang di peroleh dari website resmi Pusat Penilaian Pendidikan dan Kebudayaan Kementrian Pendidikan dan Kebudayaan Republik Indonesia.

## 1.2. Perumusan Masalah

Berdasarkan uraian dari Latar belakang membentuk perumusan masalah , bagaimana perbandingan hasil kluster nilai ujian nasional Sekolah Menengah Pertama se Jawa Tengah dengan menerapkan metode data mining *Clustering* menggunakan dua Algoritma *K – Means* dan *K – Medoids* pada data pencapaian Ujian Nasional Sekolah Menengah Pertama Tahun 2018 / 2019 se Jawa Tengah untuk mendapatkan tiap kluster Nilai Ujian Nasional Sekolah Menengah Pertama yang tinggi , sedang ,dan rendah .

## 1.3. Batasan Masalah

Agar penelitian ini tidak keluar dari permasalahan maka diperlukan batasan masalah . Batasan masalah penelitian ini adalah sebagai berikut :

1. Menggunakan data pencapaian Ujian Nasional Menengah Pertama Tahun 2018/2019 se Jawa Tengah yang di peroleh dari website resmi Pusat Penilaian Pendidikan dan Kebudayaan Kementrian Pendidikan dan Kebudayaan Republik Indonesia.
2. Penelitian menggunakan metode data mining *Clustering* menggunakan Algoritma *K – Means* dan Algoritma *K – Medoids*
3. Alat bantu yang digunakan adalah Microsoft Excel 2016 untuk perhitungan manual dan bahasa pemrograman R dengan software Rstudio untuk perhitungan komputasi.

#### **1.4. Tujuan dan Manfaat Penelitian**

Penelitian ini bertujuan untuk membandingkan hasil kluster Nilai Ujian Nasional Tingkat Sekolah Menengah Pertama menggunakan Algoritma K – *Means* dan Algoritma K – *Medoids* .

Manfaat Penelitian ini antara lain :

1. Mendapatkan hasil kluster tinggi , sedang rendah Nilai Ujian Nasional Sekolah Menengah Pertama se Jawa Tengah
2. Memperoleh algoritma yang cocok untuk digunakan dalam melakukan *clustering* data Nilai Ujian Nasional Sekolah Menengah Pertama se Jawa Tengah
3. Sebagai bahan evaluasi terhadap hasil ujian di tiap – tiap sekolah.

#### **1.5. Metodologi Penelitian**

##### **1.5.1. Objek Penelitian**

Penelitian ini menggunakan data Nilai Ujian Nasional Sekolah Menengah Pertama Tahun 2018 / 2019 se Jawa Tengah yang di peroleh dari website resmi Pusat Penilaian Pendidikan dan Kebudayaan Kementrian Pendidikan dan Kebudayaan Republik Indonesia.

##### **1.5.2. Metode Pengumpulan Data**

###### **1. Data Primer**

Data Primer berasal dari data Nilai Ujian Nasional Sekolah Menengah Pertama Tahun 2018 / 2019 se Jawa Tengah yang di peroleh dari

website resmi Pusat Penilaian Pendidikan dan Kebudayaan Kementerian Pendidikan dan Kebudayaan Republik Indonesia.

## **2. Data Sekunder**

Pengumpulan data diambil dari buku , jurnal , *study literature* dan referensi dari internet yang dapat membantu peneliti dalam menyelesaikan penelitian .

## **3. Studi Pustaka**

Pengumpulan informasi untuk membantu peneliti menyelesaikan masalah yang sedang dikerjakan dan mengumpulkan *Literature* yang relevan dengan masalah penelitian.

### **1.5.3. Metode Analisis Data**

Metode analisis data yang digunakan adalah CRISP-DM . CRISP-DM adalah keseluruhan tahapan penelitian ,yang digunakan untuk melakukan proses analisis dan prediksi dalam memecahkan suatu masalah penelitian atau bisnis . CRISP – DM (Indra Purnama,Ragil Saputra, Adi Wibowo 2015).

Tahapan – Tahapan CRISP-DM terdiri dari 6 fase :

- 1) *Business Understanding* atau pemahaman bisnis.

Pada fase ini menentukan tujuan peneliti. Tujuan dari penelitian ini adalah untuk menerapkan metode data mining menggunakan Algoritma *K- Means* dan Algoritma *K – Medoids* menggunakan data pencapaian Ujian Nasional

Sekolah Menengah Pertama Tahun 2018 / 2019 se Jawa Tengah selanjutnya menjelaskan tujuan yang diharapkan sehingga mendapatkan strategi awal untuk mencapai tujuan.

2) *Data Understanding* atau pemahaman data.

Pada fase ini dilakukan pengumpulan data awal , selanjutnya menganalisis data dan memahami isi seluruh data jika data ada yang kosong maka harus dilakukan pembersihan data . setelah data sudah di pahami secara menyeluruh lalu dilakukan proses pengolahan data .

3) *Data preparation* atau persiapan data.

Fase ini adalah tahapan untuk menyiapkan data yang sudah dipahami Dan melakukan proses pembersihan data, pemilihan atribut, tranformasi data dan standarisasi data. Semua data yang sudah siap akan siap untuk diproses.

4) *Modeling*.

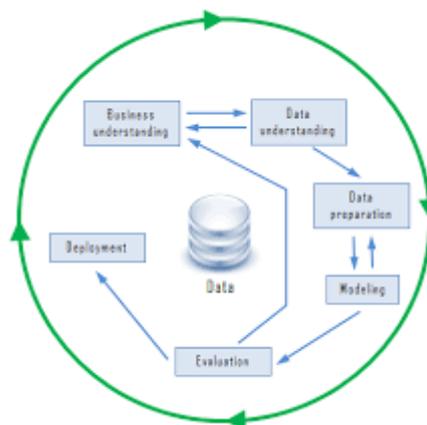
Fase ini adalah fase menentukan teknik permodelan yang sesuai dengan penelitian dan dapat menggunakan beberapa teknik pemodelan untuk menyelesaikan permasalahan. Pada penelitian ini menggunakan *Clustering* dan algoritma *K – Means* dan algoritma *K – Medoids* .

5) *Evaluation.*

Fase evaluasi adalah fase dimana dilakukan evaluasi terhadap teknik pemodelan yang digunakan , menentukan apakah Teknik pemodelan yang dipilih sudah sesuai dengan tujuan awal dan melihat apakah ada permasalahan penting dalam penelitian yang tidak terselesaikan dengan baik dan mengambil kesimpulan dari hasil yang sudah terselesaikan . Dalam penelitian ini tidak dilakukan tahap evaluasi.

6) *Deployment* atau penyebaran.

Pada fase ini adalah penyusunan laporan dan mempresentasikan hasil penelitian menggunakan teknik pemodelan yang sudah digunakan selanjutnya menjelaskan kepada pihak yang bersangkutan . Pada penelitian ini tidak dilakukan tahap penyebaran.



Gambar 1.1 Model Referensi Fase – Fase dari CRISP – DM

## **1.6. Sistematika Penulisan**

Sistematika penulisan disusun sebagai gambaran keseluruhan isi dari bab yang dibuat oleh peneliti. Sistematika penulisan penelitian ini adalah sebagai berikut :

### **BAB I PENDAHULUAN**

Membahas mengenai latar belakang permasalahan , rumusan masalah , batasan masalah , tujuan dan manfaat penelitian , metodologi penelitian dan sistematika penulisan .

### **BAB II TINJAUAN PUSTAKA**

Berisi tentang review dari hasil penelitian yang telah dilakukan sebelumnya yang berhubungan dengan masalah penelitian yang sedang dilakukan dan berisi materi atau metode yang membantu peneliti menyelesaikan permasalahan .

### **BAB III METODOLOGI PENELITIAN**

Membahas tentang metode penelitian yang digunakan dan proses penelitian yang dilakukan.

### **BAB IV HASIL PENELITIAN DAN PEMBAHASAN**

Berisi penjelasan hasil penelitian , analisa dan penjabaran Mengenai penelitian yang dilakukan .

## BAB V KESIMPULAN DAN SARAN

Tentang kesimpulan penelitian dan terdapat saran dari peneliti mengenai penelitian yang dibuat dan berisi solusi dari peneliti yang dapat dilanjutkan oleh penelitian selanjutnya

