

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data mining disebut juga dengan *knowledge discovery in database (KDD)* ataupun *pattern recognition*. Istilah KDD ini disebut penemuan pengetahuan data karena tujuan utama data mining adalah untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna oleh Fauziah Nur, et al. (2017). Sedangkan menurut Cipta Riang Sari (2016), KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar.

Data Mining merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data dengan ukuran yang cukup besar melalui serangkaian proses untuk mendapatkan informasi yang berharga dari data tersebut oleh Sulastrri dan Gufroni (2017). Aktivitas *data mining* dapat dipisahkan menjadi 6, berdasarkan tujuan dari analisis yaitu, Estimasi, Deskripsi, Prediksi, Klasifikasi, *Clustering*, Asosiasi oleh Waworuntu dan Amin (2018).

Data *Clustering* merupakan salah satu metode *Data Mining* yang bersifat tanpa arahan (*unsupervised*). Ada dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu *hierarchical* (hirarki) data clustering dan *non-hierarchical* (non hirarki) data *clustering* oleh Rosmini, et al. (2018). Menurut Syahdan dan Sindar (2018), *Clustering* lebih kearah pengelompokkan *record*, pengamatan, atau kasus dalam kelas yang memiliki kemiripan satu dengan yang lain dan memiliki ketidakmiripan dengan *record-record* dalam *cluster* yang lain. Algoritma *K-Means* merupakan salah satu algoritma dalam fungsi *clustering* atau pengelompokan oleh Asrul Sani (2018). Sedangkan menurut Irfiani dan Rani (2018), *K-Means* merupakan metode penganalisaan data pada Data Mining dimana proses

pemodelan tanpa supervisi dan merupakan salah satu metode yang mengelompokkan data secara partisi.

Salah satu metode pengelompokan ini untuk melihat hubungan kegiatan tahfiz dengan hasil belajar adalah pencapaian bentuk perubahan perilaku yang cenderung menetap dari ranah kognitif, afektif, dan psikomotorik dari proses belajar yang dilakukan dalam waktu tertentu oleh Hutauruk dan Simbolon (2018). Jadi penulis disini melihat hubungan antara kegiatan tahfiz dengan hasil belajar apakah kegiatan tahfiz tersebut mempengaruhi hasil belajar siswa tersebut semakin baik atau semakin menurun atau mengganggu hasil belajar siswa, Jadi untuk melihat hubungan kegiatan tahfiz dengan hasil belajar yang ada di sekolah Madrasah Aliyah Negeri (MAN) 1 Bukittinggi dengan menggunakan algoritma *k-means clustering* dengan cara mengelompokkan data kegiatan hafiz dan mengelompokkan juga data tentang hasil belajar siswa, setelah selesai di kelompokkan masing data maka dilihat data kegiatan tahfiz yang banyak hafalannya dan data yang nilai siswa dari yang tertinggi ke yang terendah, maka setelah selesai dikelompokkan masing data maka akan diambil kesimpulan apakah hubungan kegiatan tahfiz tersebut mempengaruhi hasil belajar siswa tersebut.

Salah satu solusi yang akan penulis lakukan untuk mengatasi permasalahan di atas adalah dengan menerapkan teknik data mining yaitu *clustering* menggunakan algoritma *k-means clustering*. Hasil penerapan algoritma *k-means clustering* adalah berupa cara menentukan kelompok atau *cluster* yang dapat membantu pihak sekolah MAN 1 Bukittinggi untuk melihat hubungan kegiatan tahfiz dengan hasil belajar. Berdasarkan latar belakang di atas, maka penulis mengajukan suatu penelitian dengan judul “Penerapan Algoritma *K-Means Clustering* Untuk Melihat Hubungan Kegiatan Tahfiz Dengan Hasil Belajar Siswa.

1.2 Perumusan Masalah

Berdasarkan latar belakang di atas maka dapat dirumuskan permasalahan dalam penelitian ini yaitu:

1. Bagaimana menerapkan algoritma *k-means clustering* untuk mengelompokkan siswa berdasarkan kegiatan tahfiz dengan hasil belajar?

2. Bagaimana merancang dan menerapkan aplikasi untuk analisa dan pertimbangan terhadap hasil *cluster* melihat hubungan kegiatan tahfiz dengan hasil belajar menggunakan algoritma *k-menas clustering*?

1.3 Batasan Masalah

Berdasarkan perumusan masalah di atas, penulis membatasi permasalahan sebagai berikut:

1. Teknik data mining yang digunakan adalah clustering menggunakan algoritma *k-means clustering*.
2. hubungan kegiatan tahfiz dengan hasil belajar yang akan dijadikan subjek pada penelitian ini.
3. Data yang diambil adalah data siswa yang hanya mengikuti kegiatan tahfiz dan data hasil belajar siswa kelas XI (sebelas) tahun ajaran 2018-2019 di MAN 1 Bukittinggi.

1.4 Tujuan Penelitian

Tujuan dari penelitian adalah untuk :

1. untuk melihat hubungan kegiatan tahfiz dengan hasil belajar siswa menggunakan algoritma *k-means clustering*.
2. Menerapkan aplikasi Rapid Miner untuk menghasilkan model dan pertimbangan kegiatan tahfiz dengan hasil belajar menggunakan teknik data mining dengan algoritma *k-means clustering*.
3. Menguji serta menganalisa kelayakan dalam menggunakan penerapan algoritma *k-means clustering* dalam melihat hubungan kegiatan tahfiz dengan hasil belajar siswa.

1.5 Manfaat Penelitian

Penelitian ini di harapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat bagi instansi yaitu untuk memberikan analisa dan pertimbangan terhadap hubungan kegiatan tahfiz dengan hasil belajar apakah berpengaruh atau tidak.
2. Manfaat bagi peneliti yaitu dapat mengimplementasikan ilmu yang telah di dapat dalam bangku perkuliahan.
3. Manfaat bagi pembaca pada umumnya yaitu sebagai referensi khususnya dalam memahami data mining dengan algoritma *k-means clustering*

1.6 Sistematika Penulisan

Sistematika penulisan yang terdapat dalam tesis ini adalah sebagai berikut:

Bab I : Pendahuluan

Bab ini berisikan pendahuluan dan latar belakang masalah dalam melakukan penelitian, gambaran umum tentang permasalahan yang akan dibahas pada bab-bab berikutnya, yaitu perumusan masalah, batasan masalah, tujuan penilitian, manfaat penilitian yang diakhiri dengan sistematika penulisan.

Bab II : Landasan Teori

Pada bab ini dijelaskan tentang beberapa hal yang mendasar tentang masalah yang akan di bahas, yang mana nantinya akan digunakan sebagai landasan dasar dalam melakukan pemecahan masalah pada penelitian yang dilakukan.

Bab III : Metode Penelitian

Pada bab ini berisikan tentang langkah kerja dan prosedur serta metode yang akan dilakukan dalam penelitian.

Bab IV : Analisa dan Hasil

Pada bab ini memuat tentang analisa data, calon penerima beasiswa, sesuai dengan langkah-langkah pada algoritma *k-means clustering*. Selain itu bab ini juga memuat hasil *clustering* kegiatan tahfiz dan hasil belajar dan analisa hasil *clustering* yang dapat dijadikan sebagai pertimbangan oleh pihak sekolah MAN 1 Bukittinggi

Bab V : Implementasi dan Pengujian

Pada bab ini berisi penerapan teknik *clustering* dengan algoritma *k-means clustering* menggunakan aplikasi Rapid Miner dan pengujian terhadap data sampel dan data siswa yang mengajukan beasiswa.

Bab VI : Kesimpulan dan Saran

Pada bab ini dijelaskan tentang kesimpulan dan saran-saran dalam melakukan penelitian dan pengembangan untuk selanjutnya.