

BAB I

PENDAHULUAN

1.1 Latar Belakang

Peningkatan data yang pesat menimbulkan penambahan data akademik yang sangat cepat dari masa ke masa dari kumpulan data telah menciptakan kondisi kaya akan data tetapi minim Informasi (Nurjoko, 2016), dimana membentuk sebuah tumpukan data yang sangat banyak. Namun pada kenyataannya, tumpukan data tersebut tidak dimanfaatkan secara maksimal bahkan tidak terpakai. Padahal dari kumpulan data yang ada bisa ditarik sebuah informasi yang bermanfaat didalamnya menggunakan teknik *Data Mining* (Kamagi D.H. dan Hansun S., 2018). *Data mining* adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu *database* atau gudang data (Vulandari R.T., 2017).

Penggunaan teknik *data mining* pada perguruan tinggi dapat berguna untuk mengolah informasi yang masih semu menjadi informasi yang berguna untuk menunjang kegiatan operasional akademik maupun untuk menentukan pengambilan keputusan strategis untuk masa yang akan datang menggunakan Algoritma KKN salah satu algoritma yang digunakan untuk melakukan klasifikasi dengan menggunakan data training dan data testing (Prasetya C.S.D., 2017).

Tumpukan Data akademik mahasiswa merupakan data yang dihimpun dari hasil kegiatan proses belajar mengajar selama mengikuti studi di suatu perguruan tinggi. Data tersebut antara lain: data pribadi mahasiswa, data rencana studi, dan data hasil studi (nilai dan indeks prestasi) . Jumlah data yang terakumulasi dari tahun ke tahun perlu dilakukan analisis untuk dapat membuka peluang dihasilkannya informasi yang berguna dalam pembuatan alternatif keputusan bagi manajemen perguruan tinggi (Totok., 2016). data yang digunakan merupakan data akademik di Jurusan Fisika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Andalas (UNAND) Padang Indonesia, yaitu data IP (Indeks Prestasi)

mahasiswa dari semester satu sampai semester enam khususnya Program Sarjana (S1) dengan menggunakan salah satu teknik *data mining* yaitu metode klasifikasi menggunakan algoritma *k-Nearest Neighbor* dengan harapan dapat menemukan informasi tingkat kelulusan dan persentase kelulusan mahasiswa sehingga dapat digunakan oleh pihak Jurusan untuk mencari solusi atau kebijakan dalam proses evaluasi pembelajaran di Jurusan Fisika Fakultas Mipa Universitas Andalas Padang Indonesia.

Universitas Andalas merupakan salah satu perguruan tinggi Negeri yang memiliki kemajuan sangat pesat. Jumlah mahasiswa masuk setiap tahunnya tergolong tinggi dibandingkan dengan perguruan-perguruan tinggi swasta lainnya. Hal ini tentunya menjadikan perlunya memperhatikan keseimbangan jumlah mahasiswa yang masuk dengan jumlah mahasiswa yang keluar. Selain memperhatikan kuantitas dari lulusan universitas juga perlu memperhatikan kualitas dari lulusan. Salah satu cara untuk melakukan evaluasi terhadap hasil kinerja sebuah universitas adalah dengan pendataan data lulusan atau tracer study, yang merupakan mekanisme yang diadakan oleh berbagai institusi pendidikan untuk mengeksplorasi alumni untuk mengukur relevansi tujuan pendidikan dengan kondisi saat ini alumni (Zahrotun L., 2017)

Teknik data mining ini telah banyak digunakan dalam memecahkan berbagai permasalahan di kehidupan nyata, Salah satunya Teknik data mining ini digunakan mengukur Kualitas air di daerah perkotaan di Surakarta telah menurun saat ini. Meningkatnya perkembangan industri, pengolahan limbah yang buruk, dan beberapa faktor lain menyebabkan perkotaan ini masalah. Hasil dari sistem pemantauan kualitas air dengan algoritma K-Nearest Neighbor pada penelitian ini tentu akan membantu tugas pekerja PDAM (PDAM) dalam menganalisis kualitas air. Untuk pertimbangan keluaran mayoritas dalam metode ini, sistem bekerja dengan mengambil jarak terdekat ke nomor K. data pelatihan yang ditugaskan untuk ini Penelitian diambil pada bulan Maret 2016 dari laporan hasil pemantauan air di laboratorium PDAM Surakarta. Hasil identifikasi dibagi menjadi memenuhi syarat (MS) dan tidak memenuhi syarat (TMS). Hasil data pengujian diterapkan dalam pengujian kinerja algoritma dengan confusion matrix yang memiliki tingkat akurasi 82,5%. Sebagai bahan pertimbangan dalam mengukur kualitas Air (Arnomo R. A., Saptomo W.L.Y. dan Harsadi P., 2018).

Demikian pula penerapan algoritma k-nearest neighbor untuk penentuan resiko kredit kepemilikan kendaraan bermotor. Sejalan dengan pertumbuhan bisnis, kredit merupakan masalah yang menarik untuk diteliti. Beberapa riset bidang komputer untuk mengurangi resiko kredit telah banyak dilakukan dalam rangka menghindari kehancuran suatu perusahaan pembiayaan. Paper ini membahas algoritma k-Nearest Neighbor (kNN) yang diterapkan pada data konsumen yang menggunakan jasa keuangan kredit kendaraan bermotor. Hasil testing untuk mengukur performa algoritma ini menggunakan metode Cross Validation, Confusion Matrix dan kurva ROC dan menghasilkan akurasi dan nilai AUC berturut-turut 81,46 % dan 0,984. Karena nilai AUC berada dalam rentang 0,9 sampai 1,0 maka metode tersebut masuk dalam kategori sangat baik (Leidiyana H., 2016).

Begitu juga Algoritma *K-Nearest Neighbor Classification* Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa suatu bentuk dari pencapaian hasil selama mengikuti kegiatan Akademik pada sebuah Perguruan Tinggi. Predikat prestasi mahasiswa diperoleh dari hasil sebuah prediksi. Proses prediksi dilakukan dengan menggunakan metode *K-Nearest Neighbor* (KNN). Atribut yang digunakan dalam proses prediksi adalah Jenis Kelamin, Jenis Tinggal, Umur, Jumlah Satuan Kredit Semester (SKS), dan Jumlah Nilai Mutu (NM), sehingga dengan menerapkan algoritma KNN dapat dilakukan sebuah prediksi berdasarkan kedekatan dari histori data lama (*training*) dengan data baru (*testing*). Penentuan atribut ini berdasarkan hasil penelitian terdahulu yang memiliki kesamaan dalam kasus prediksi mahasiswa yang selanjutnya divalidasi oleh bagian Akademik Fakultas Sains dan Teknologi. Proses prediksi dilakukan terhadap Mahasiswa Program Studi Sistem Informasi angkatan 2014/2015 sebagai data *testing* dengan jumlah 50 data, serta berdasarkan dari data angkatan 2012/2013 sebagai data *training* dengan jumlah 165 data yang menghasilkan pengujian akurasi sebesar 82%. Hasil dari perhitungan algoritma KNN diimplementasikan terhadap sebuah *Early Warning System* (EWS). *Output* dari sistem yang dibangun dapat dijadikan sebagai acuan bagi Mahasiswa untuk meningkatkan prestasi dan predikat perkuliahan dimasa yang akan datang (Mustakim. dan Oktaviani G., 2016).

Berdasarkan penelitian diatas,metode algoritma k-nearest neighborhood akan digunakan dalam penetian ini untuk memprediksi masa studi mahasiswa dengan memamfaatkan data IPK mahasiswa.

1.2 Perumusan Masalah

Berdasarkan latar belakang di atas, maka dapat dirumuskan permasalahan yang akan dibahas dalam penelitian ini adalah :

1. Bagaimana pihak akademik dapat melihat berapa banyak mahasiswa yang akan tamat tepat pada waktunya walaupun mahasiswa tersebut masih dalam masa kuliah dengan menggunakan metode algoritma KNN ?
2. Bagaimana memanfaatkan data akademik berupa kartu hasil studi menjadi sebuah informasi yang berguna dengan menggunakan metode algoritma KNN ?
3. Bagaimana pihak akademik mengevaluasi proses pembelajaran pada mahasiswa dalam perkuliahan apakah sudah baik atau belum?

1.3 Batasan Masalah

Pada tesis ini terdapat beberapa batasan masalah agar hasil dari tesis ini sesuai dengan tujuan. Adapaun beberapa batasan masalah pada tugas akhir ini, yaitu:

1. Studi Kasus tugas akhir ini yaitu data akademik Jurusan Fisika Fakultas Mipa Universitas Andalas Padang Indonesia .
2. Data yang digunakan merupakan data akademik Jurusan Fisika Fakultas Mipa Universitas Andalas Padang Indonesia tahun 2013-2016
3. Data yang digunakan hanya data Indeks Prestasi (IP) mahasiswa Jurusan Fisika Fakultas Mipa Universitas Andalas Padang Indonesia Strata 1 (S1) dari semester satu sampai semester enam.
4. Data yang digunakan telah dibersihkan secara manual sebelum diproses dalam data mining. Data yang tidak lengkap nilai Indeks Prestasinya tidak akan digunakan atau dihapus data record-nya

1.4 Tujuan Penelitian

1. Diharapkan teknik *data mining* ini dapat memprediksi masa studi mahasiswa sehingga dapat mengetahui tingkat kelulusan dan persentase kelulusan mahasiswa di Jurusan Fisika.
2. Diharapkan teknik ini dapat digunakan oleh pihak akademik untuk mengevaluasi proses pembelajaran apakah sudah efektif.

1.5 Manfaat Penelitian

1. Bagi Penulis
Dapat menambah pengetahuan dan wawasan serta dapat mensosialisasikan teori yang telah diperoleh selama perkuliahan.
2. Bagi Universitas Andalas
 - a. Dapat membantu pihak akademik dalam mengambil keputusan yang didapat dari proses *Data Mining*.
 - b. Sebagai langkah awal untuk mengevaluasi kinerja sistem perkuliahan yang ada apakah sudah berjalan secara maksimal atau belum.
 - c. Mendapatkan informasi yang berguna dan bermanfaat untuk mengambil keputusan terhadap hasil yang diperoleh.

1.6 Sistematika Penulisan

Sistematika disesuaikan dengan template yang diatur dalam tata penulisan program studi masing-masing. Seperti:

Bab I: Pendahuluan

Pada bab ini memuat tentang Latar Belakang dari penelitian, Perumusan Masalah, Batasan Masalah, Tujuan Penelitian, Manfaat Penelitian, dan Sistematika Penelitian.

Bab II : Landasan Teori

Pada bab ini dibahas tentang penelitian yang terdahulu sebagai tinjauan pustaka serta landasan teori yang digunakan dalam penyusunan penelitian ini.

Bab III : Metode Penelitian

Bab ini menyajikan metode penelitian yang dipakai dan juga menjelaskan kerangka kerja yang terdiri dari tahap-tahap penelitian, tujuan penelitian, software-software, alat untuk menganalisa data dan tempat penelitian.

Bab IV : Analisa dan Implementasi Sistem.

Pada bab ini berisi paparan hasil implementasi analisis terhadap permasalahan dan penyelesaian permasalahan dengan menggunakan teknik data mining dengan metode Algoritma k-nearest neighbor serta hasil perhitungan dari metode Algoritma k-nearest neighbor .

Bab V : Hasil dan Pengujian Metode

Bab ini akan membahas implementasi dari analisa yang didapatkan dari bab sebelumnya, kemudian seperti apa hasil terhadap pengujian dengan java netbean.

Bab VI : Kesimpulan dan Saran

Membahas kesimpulan dari seluruh hasil analisa dari penelitian ini, serta saran bagi para pembaca dan bagi para pengembang dimasa yang akan datang.