

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Berkembangnya jaman sekarang ini peran teknologi informasi sangatlah penting, dimana kita sebagai pengguna dapat memanfaatkan teknologi informasi dengan hanya sekali klik. Informasi yang diolah tidak hanya text, gambar, video namun bisa berbentuk multimedia, kita dapat memanfaatkan komputer dengan berbagai cara, mulai menulis, menggambar, edit foto, memutar video/lagu sampai dengan menganalisis data penelitian maupun mengatasi masalah-masalah lainnya. Sejak awal 2018 hingga sekaranglah diperkirakanlah merupakan zaman revolusi industri 4.0. Revolusi Industri 4.0 adalah revolusi industri yang menggabungkan teknologi otomatisasi dengan teknologi cyber. Ini merupakan tren otomatisasi dan pertukaran data dalam teknologi manufaktur. Pada era ini, industri mulai menyentuh dunia virtual, membentuk konektivitas antar manusia, mesin dan data, yang dikenal dengan nama Internet of Things (IoT) (Taufik, Ahmad. Pengantar Teknologi Informasi. Jawa Tengah: Jakarta, 2022. Pena Persada.)

Politeknik LP3I Kampus Padang merupakan salah satu perguruan tinggi dengan jenjang pendidikan Diploma Tiga (D-III) yang bernaung di bawah bendera LP3I Group. Kehadirannya telah memberikan kontribusi Sumber Daya Manusia yang unggul bagi berbagai dunia industri, dalam meningkatkan kualitas pembelajaran Politeknik LP3I Kampus Padang terus melakukan upaya perkembangan seperti pembangunan infrastruktur sebagai penunjang bagi kegiatan perkuliahan dan perbaikan sarana prasarana. Sumber dana terbesar yang digunakan adalah keuangan mahasiswa yaitu melalui pembayaran uang kuliah. Keterlambatan pembayaran uang kuliah dapat mengakibatkan terhambatnya kegiatan operasional dan mempengaruhi progress

akademik karena tidak dapat mengurus Kartu Rencana Studi (KRS) pada awal semester.

Penelitian tentang prediksi keterlambatan pembayaran uang kuliah yang pernah dilakukan oleh beberapa peneliti diantaranya Muhammad Rhosyid Akhmad dan Taghfirul Azhima Yoga Siswa Memprediksi Keterlambatan Pembayaran Biaya Kuliah Di Perguruan Tinggi pada Universitas Muhammadiyah Kalimantan Timur tahun ajaran 2020/2021. Selanjutnya Muhammad Rhosyid dan Taghfirul melakukan prediksi keterlambatan uang kuliah menggunakan algoritma Naive Bayes. Adapun penelitian yang telah dilakukan oleh Muqorobin, dkk tentang Optimasi Metode Naive Bayes Dengan Feature Selection Information Gain Untuk Prediksi Keterlambatan Pembayaran Sumbangan Pembinaan Pendidikan Sekolah dimana sistem prediksi tersebut dibuat sebagai solusi dalam menangani permasalahan keterlambatan pembayaran SPP sekolah. Maka pihak sekolah dapat melihat informasi mengenai perkiraan siswa-siswi yang akan membayar tepat waktu atau terlambat. Sehingga dengan adanya hasil informasi prediksi tersebut maka pihak sekolah akan dapat melakukan tindakan berupa pembinaan terhadap siswa-siswi yang diperkirakan akan terlambat dalam melakukan pembayaran SPP Sekolah (Muqorobin et al., 2019).

Feature selection adalah salah satu fokus yang ada di data mining. Feature selection merupakan suatu proses pemilihan bagian dari variable dari semua variable yang ada di dataset. Salah satu metode feature selection yaitu forward selection (Theopilus Bayu Sasongko, Oki Arifin, 2019). Metode Forward Selection merupakan salah satu metode yang biasa digunakan untuk melakukan seleksi fitur dari dataset yang memiliki banyak atribut. Seleksi fitur sendiri merupakan proses yang biasa digunakan di machine learning yang mana sekumpulan dari fitur/atribut yang dimiliki oleh data digunakan untuk proses pembelajaran algoritma (Ika Nur Fajri, Femi Dwi Astuti, 2022).

Algoritma Naive Bayes menghitung probabilitas kejadian masa datang dari kejadian sebelumnya di mana masing- masing variable dianggap tidak saling tergantung. Algoritma Forward Selection merupakan salah satu seleksi fitur yang dilakukan sebelum pemrosesan dalam klasifikasi yang terbukti efektif digunakan untuk memecahkan masalah pemilihan fitur yang relevan dalam suatu data Forward Selection dijalankan menggunakan pemilihan fitur yang relevan untuk data yang berpengaruh terhadap hasil klasifikasi (M. Rudi Fanani, 2020).

Dunia saat ini tengah waspada dengan penyebaran sebuah virus yang dikenal dengan virus corona. Pada penelitian ini peneliti telah melakukan penerapan algoritma Naive Bayes untuk memprediksi tingkat penyebaran COVID-19 di Indonesia dengan mengimplementasikan Algoritma Naive Bayes, yang berguna sebagai langkah antisipasi pandemic COVID-19. Penelitian dilakukan secara bertahap dimulai dari perencanaan, menentukan fokus penelitian, waktu penelitian, pengumpulan data, analisis, dan penyajian hasil penelitian. Selanjutnya pengujian dilakukan untuk mengetahui apakah penelitian yang dilakukan telah sesuai dengan tujuan yang diharapkan yaitu untuk memprediksi tingkat penyebaran COVID-19 di Indonesia. Berdasarkan data Kasus COVID-19 Per Provinsi yang dijadikan data training, metode Naive Bayes berhasil mengklasifikasikan 16 data dari 33 data yang diuji. Sehingga metode Naive Bayes berhasil memprediksi besarnya Kasus COVID-19 Per Provinsi dengan persentase keakuratan sebesar 48,4848% (Alvina Felicia Watratan, 2020).

Kegiatan bimbingan dan konseling di sekolah adalah cara membantu siswa dalam mengembangka kepribadian, sikap sosial, pembelajaran, pengembangan karir dan perencanaan. Data mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data warehouse, atau penyimpanan informasi lainnya. Naïve Bayes adalah perhitungan probabilitas dengan metode pengklasifikasian. Forward selection adalah salah satu prosedur bertahap yang bertujuan untuk menambah variabel yang dikendalikan satu per satu ke dalam persamaan. Penelitian ini menggunakan data sekunder. Data sekunder berupa data bimbingan konseling yang dilakukan guru BK Bawon Sulastri yang telah berpengalaman dan berkompeten di bidangnya selama 8 tahun, tetapi tidak diterbitkan. Dari jumlah 329 record, 102 record menghasilkan label Diharuskan Konseling dan 227 record menghasilkan Tidak diharuskan konseling. Hasil eksperimen dan evaluasi penelitian dapat disimpulkan bahwa akurasi metode Naïve Bayes pada bimbingan konseling siswa mencapai 94.55%. setelah dilakukan penambahan fitur seleksi untuk memilih atribut yang relevan terhadap hasil klasifikasi didapatkan nilai akurasi yang lebih tinggi yaitu 94.84%. (M. Rudi Fanani, 2020).

Naïve Bayes Classification merupakan salah satu algoritma klasifikasi yang sangat efektif (mendapatkan hasil yang tepat) dan efisien (proses penalaran dilakukan

memanfaatkan input yang ada dengan cara yang relatif cepat). Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Keuntungan penggunaan Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naive Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan (Michal Dennis et al., 2022).

Adapun kekurangan pada algoritma Naïve Bayes ini pada saat melakukan proses training dan testing klasifikasi tidak optimal jika tidak disimpan pada dataset untuk itu perlu disimpan hasil klasifikasi pada dataset (Muslehatin, 2017). Langkah bagaimana untuk mengoptimalkan dari masing-masing algoritma maka ditambahkan teknik Optimasi Forward Selection pada algoritma klasifikasi data mining untuk mengurangi tingkat kompleksitas, meningkatkan akurasi pada algoritma klasifikasi, juga melihat atribut yang berpengaruh terhadap tingkat akurasi (Supriyanti, 2018).

Berdasarkan berlimpahnya data mahasiswa dan data jumlah kelulusan mahasiswa, informasi yang tersembunyi dapat diketahui dengan cara melakukan pengolahan terhadap data mahasiswa sehingga berguna bagi pihak universitas. Pengolahan data mahasiswa perlu dilakukan untuk mengetahui informasi penting berupa pengetahuan baru (knowledge Discovery). Penelitian ini akan melakukan pengklasifikasian berdasarkan dataset IAsol yang didapat dari Universitas Abadi Karya Indonesia khususnya di Fakultas Ilmu Komputer pada tahun ajaran 2008 sampai 2011. Penelitian kali ini selain mendapatkan nilai akurasi yang baik juga bertujuan mendapatkan model atribut yang berpengaruh dengan cara menerapkan Feature Selection. Sumber data pada penelitian ini adalah dataset yang didapat dari IAsol UNAKI khususnya di Fakultas Ilmu Komputer pada tahun ajaran 2008 sampai 2011. Tahap modeling untuk mengklasifikasikan status kelulusan dengan menggunakan dua metode yaitu algoritma Naïve Bayes dan Forward Selection-Naïve Bayes. Forward Selection berbasis Naive Bayes lebih akurat dan efektif dalam mengklasifikasikan status kelulusan mahasiswa dari dataset yang bersifat class imbalance dengan data yang besar dengan hasil akurasi 99.17% dan termasuk dalam kategori “Kappa excellent”. Hasil dari penelitian ini adalah memperoleh atribut yang berpengaruh yaitu: Kelompok, IP

Semester 1, Semester 3, IP Semester 9. Dibanding dengan menggunakan algoritma Naive Bayes saja dengan hasil akurasi 95.83% (Bondhan Arya Purnanditya et al., 2020).

Penelitian ini menguji algoritma Naive Bayes dalam melakukan prediksi keterlambatan pembayaran uang kuliah di perguruan tinggi, agar dapat dilakukan pencegahan dan penanganan sedini mungkin atas kasus yang sama di masa mendatang. Berdasarkan uraian di atas dapat dilakukan penelitian yang berupa tesis dengan judul **“PENERAPAN FITUR SELEKSI FORWARD SELECTION MENGGUNAKAN ALGORITMA NAÏVE BAYES UNTUK MENENTUKAN ATRIBUT PADA PREDIKSI KETERLAMBATAN PEMBAYARAN UANG KULIAH MAHASISWA”**.

1.2 Perumusan Masalah

Rumusan masalah adalah pengarah tujuan dari sebuah tulisan ilmiah agar fokus terhadap pembahasan hal tertentu. Berdasarkan latar belakang yang telah dipaparkan, maka didapatkan rumusan masalah pada penelitian ini yaitu :

1. Bagaimana memahami Algoritma Naive Bayes dapat digunakan dalam memprediksi keterlambatan pembayaran uang kuliah?
2. Bagaimana menerapkan algoritma Naive Bayes dalam memprediksi keterlambatan pembayaran uang kuliah berdasarkan data Bagian Administrasi Akademik dan data keterlambatan pembayaran uang kuliah?
3. Bagaimana menguji penerapan algoritma Naive Bayes dalam memprediksi keterlambatan pembayaran uang kuliah?

1.3 Batasan Masalah

Batasan masalah adalah ruang lingkup masalah yang ingin dibatasi oleh peneliti yang disebabkan masalah yang terlalu luas atau lebar yang bisa mengakibatkan penelitian itu tidak bisa fokus. Supaya lingkup permasalahan tidak melebar, maka penulis membatasi penyelesaian permasalahan sebagai berikut:

1. Data yang digunakan merupakan data mahasiswa yang membayar biaya

kuliah dari tahun 2020 – 2021.

2. Mengolah dan menganalisa data keterlambatan biaya kuliah mahasiswa Politeknik LP3I Kampus Padang.
3. Penelitian ini dirancang dengan menggunakan Bahasa Pemrograman PHP dan MySQL.

1.4 Tujuan Penelitian

Tujuan dari suatu penelitian dapat untuk mengidentifikasi atau menggambarkan suatu konsep atau untuk menjelaskan atau memprediksi suatu situasi atau solusi untuk suatu situasi yang mengindikasikan jenis studi yang akan dilakukan. Adapun tujuan yang akan dari penelitian ini adalah :

1. Menganalisa dan menentukan atribut yang mempengaruhi keterlambatan mahasiswa dalam membayar uang kuliah.
2. Merancang menggunakan Algoritma Naive Bayes dalam memprediksi mahasiswa yang terlambat dalam pembayaran uang kuliah.
3. Implementasi algoritma Naive Bayes untuk memprediksi keterlambatan mahasiswa dalam melakukan pembayaran uang kuliah di Politeknik LP3I Kampus Padang.

1.5 Manfaat Penelitian

Manfaat penelitian adalah penegasan dan harapan peneliti , tentang hasil yang didapatkan dalam penelitian memberi manfaat atau kegunaan secara akademik dan praktis. Adapun manfaat yang didapat dalam penelitian ini adalah :

1. **Bagi Peneliti** : pengimplementasian ilmu pengetahuan Programming (ilmu komputer) yang telah didapatkan dan dipelajari pada penggunaan algoritma Naive Bayes sehingga dapat berguna dan bermanfaat bagi masyarakat.
2. **Bagi Referensi** : dengan adanya penelitian diharapkan dapat menambah

referensi sebagai bahan penelitian dimasa mendatang, terutama penelitian-penelitian yang berkaitan dengan penerapan fitur seleksi forward selection dan algoritma Naive Bayes.

3. **Bagi Masyarakat** : dapat dijadikan sebagai bahan pertimbangan dalam memprediksi keterlambatan pembayaran uang kuliah pada masa yang akan datang.

1.6 Sistematika Penulisan

Sistematika penulisan merupakan sebuah metode atau urutan dalam menyelesaikan sebuah riset, penelitian, maupun karya tulis. Sistematika disesuaikan dengan template yang diatur dalam tata penulisan program studi masing-masing, seperti:

Bab I : Pendahuluan

Berisi Latar Belakang, Perumusan Masalah, Batasan Masalah, Tujuan Penelitian, Manfaat Penelitian, dan Sistematika Penelitian.

Bab II : Landasan Teori

Pada bab ini dijelaskan teori tentang *Fitur Seleksi Forward Selection* mengenai dan juga teori-teori yang berhubungan dengan algoritma *Naive Bayes*.

Bab III : Metodologi Penelitian

Bab ini membahas tentang analisa dan penggunaan secara matematis algoritma *Naive Bayes* dalam menentukan atribut pada prediksi keterlambatan pembayaran uang kuliah mahasiswa.

Bab IV : Analisa dan Perancangan

Bab ini membahas hasil implementasi dari algoritma *Naive Bayes* pada data uji dalam menentukan atribut pada prediksi keterlambatan pembayaran uang kuliah mahasiswa.

Bab V : Implementasi dan Hasil

Bab ini akan dilakukan implementasi dan pengujian Penerapan Fitur

Seleksi *Forward Selection* dalam menentukan atribut pada prediksi keterlambatan pembayaran uang kuliah mahasiswa menggunakan algoritma *Naive Bayes*.

Bab : VI Kesimpulan dan Saran

Bab ini membuat kesimpulan dan hasil penelitian dari penggunaan ataupun penerapan fitur seleksi *forward selection* yang diimplementasikan pada prediksi keterlambatan pembayaran uang kuliah mahasiswa dan memberikan saran pada peneliti lain dalam mengembangkan sistem kedepannya.