

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Indeks prestasi merupakan nilai rata-rata dari seluruh matakuliah yang sudah diambil oleh para mahasiswa. Indeks prestasi sendiri dibedakan menjadi Indeks Prestasi Semester (IPS) dan Indeks Prestasi Kumulatif (IPK). Indeks Prestasi Semester (IPS) merupakan hasil belajar mahasiswa tiap semester yang dihitung berdasarkan jumlah SKS yang diambil. Indeks Prestasi Semester (IPS) ini sangat berpengaruh terhadap nilai Indeks Prestasi Kumulatif (IPK) mahasiswa. Tingkat keberhasilan mahasiswa sendiri dapat dilihat dari prestasi yang mereka capai, melalui Indeks Prestasi Semester (IPS) dapat dilihat apakah mahasiswa mengalami perubahan tiap semesternya perubahan dapat menjadi lebih baik, kurang baik dan tetap. Prediksi Indeks Prestasi mahasiswa juga dapat digunakan untuk mengevaluasi pembelajaran di program studi serta dapat memonitoring proses pembelajaran dan melakukan kebijakan selanjutnya yang berguna untuk meningkatkan prestasi mahasiswa. Namun, memprediksi indeks prestasi mahasiswa merupakan masalah yang tidak mudah, disebabkan banyaknya atribut – atribut yang mempengaruhi Indeks Prestasi mahasiswa baik secara internal maupun eksternal (Efendi et al., 2018).

Hasudungan dan Pranoto (Hasudungan & Pranoto, 2021) memprediksi Prestasi Mahasiswa menggunakan Teorema Naive Bayes dengan total 40 data mahasiswa yang menghasilkan tingkat akurasi sebesar 77,5%. (Efendi et al., 2018) menggunakan *Rough Set* dan Jaccard untuk melakukan seleksi atribut Indeks Prestasi Mahasiswa yang dapat disimpulkan banyak bacaan Al-Qur'an, jumlah unduhan ilmiah dan durasi memainkan gadget sedikit banyak memang berpengaruh pada Indeks Prestasi Mahasiswa.

Naive Bayes merupakan metode klasifikasi probabilitas sederhana yang dapat menghitung satu set probabilitas dengan menghitung kombinasi nilai dalam kumpulan data yang sudah diberikan. Klasifikasi Naive Bayes diasumsikan ada atau

tidak berdasarkan ciri tertentu dari kelas yang tidak ada hubungannya dengan kelas lainnya, jumlah data yang dibutuhkan oleh metode Naïve Bayes sedikit dan ini merupakan keunggulan dari metode ini (Hasudungan & Pranoto, 2021). Naive Bayes sendiri memiliki kekurangan karena naive bayes tidak dapat memilih atribut terbaik dikarenakan naive bayes menganggap semua atribut yang ada memiliki tugas yang sama. Sedangkan pada data yang ada terdapat banyak atribut yang berlebihan dan kadang berpengaruh terhadap hasil klasifikasi, sehingga diperlukan metode untuk melakukan seleksi atribut yang terbaik dan meningkatkan hasil klasifikasi.

Rough Set merupakan salah satu metode matematika untuk menangani data yang tidak konsisten dan samar serta tidak lengkap. *Rough Set* digunakan dalam seleksi atribut seperti (Hasudungan & Pranoto, 2020) menggunakan *Rough Set* maximum dependency of attributes untuk memilih atribut terbaik pada hasil Classification for Students Performance Prediction. Dari pembahasan, *Rough Set* memiliki kemampuan dalam melakukan seleksi atribut terbaik dan mengeliminasi atribut yang berlebihan, sehingga *Rough Set* dapat meningkatkan akurasi naive bayes dalam melakukan prediksi indeks prestasi mahasiswa.

1.2 Rumusan Masalah

Berdasarkan latar belakang Prediksi Indeks Prestasi mahasiswa dapat digunakan untuk melihat proses pembelajaran di program studi serta melihat keberhasilan mahasiswa dalam melakukan studi. Sehingga rumusan masalah pada penelitian ini ialah Prediksi Indeks Prestasi Mahasiswa dengan Metode *Maximum Dependency Of Attributes* dan *Naïve Bayes*.

1.3 Tujuan Penelitian

Untuk menyelesaikan rumusan masalah, terdapat lima tujuan dari penelitian ini yakni :

1. Menentukan parameter/atribut yang berpengaruh terhadap indeks prestasi mahasiswa.
2. Mengkoleksi data dari mahasiswa Universitas Muhammadiyah Kalimantan Timur.

3. Mengimplementasikan *Rough Set MDA* untuk memilih atribut/parameter terbaik.
4. Mengimplementasikan *Naïve Bayes* dalam klasifikasi dan prediksi Indeks Prestasi Mahasiswa.
5. Mengevaluasi dan mengkomparasi hasil dari *Naïve Bayes* dan *MDA* melalui pengukuran akurasi.

1.4 Skop Penelitian / Batasan Penelitian

1. Data yang digunakan merupakan data mahasiswa fakultas Sains dan Teknologi angkatan 2017 - 2018 di Universitas Muhammadiyah Kalimantan Timur.
2. IP Semester yang diambil hanya sampai Semester 6 dengan jumlah IPK Keseluruhan dari IP Semester 1 sampai 6.
3. Pemilihan atribut hanya menggunakan *Rough Set Maximum Dependency of Attribute*.
4. Data diambil melalui SIKAD dan proses pemograman python di jalankan menggunakan *Google Colab*