

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Program Jalur Cepat (JC) atau dalam bahasa Inggris disebut *Fast Track* merupakan sebuah program yang di adakan oleh Perguruan Tinggi (PT). PT yang telah mengadakan program JC adalah PT Universitas Andalas (UNAND). Program JC ini diperuntukkan bagi mahasiswa yang sedang menempuh Program Sarjana untuk menuju Program Magister dengan menambah 1 (satu) tahun dari program studi Sarjana. Program Sarjana (S1) merupakan pendidikan akademik yang diperuntukan bagi lulusan pendidikan menengah sederajat sehingga mampu mengamalkan ilmu pengetahuan dan teknologi melalui penalaran ilmiah. Sedangkan, Program Magister (S2) merupakan pendidikan akademik yang diperuntukan bagi lulusan program sarjana atau sederajat sehingga mampu mengamalkan dan mengembangkan ilmu pengetahuan dan teknologi melalui penalaran dan penelitian ilmiah. Program JC bertujuan untuk menghasilkan mahasiswa lulusan dengan keterampilan dan pengetahuan yang sama dengan rekan-rekan mereka yang melakukan program studi 4 tahun tetapi dalam separuh waktu (Aldiabat dkk, 2021).

Program JC memberikan kesempatan kepada mahasiswa untuk dapat menyelesaikan S1 dan S2 selama 10 semester dan juga kesempatan kepada mahasiswa berkemampuan akademik yang sangat baik mengembangkan dirinya secara optimal dan dalam waktu yang relatif lebih cepat. Dalam mengembangkan kemampuan akademik, mahasiswa mendapatkan proses pembelajaran digital yang semakin cepat melalui dukungan teknologi internet sehingga dapat mengikuti perkuliahan dengan baik. Selain itu, mahasiswa dapat memantau dan mengevaluasi metode pembelajarannya sehingga tingkat penyerapan ilmunya lebih tinggi (Gunawan dkk, 2021).

Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA), salah satu fakultas di bawah naungan PT UNAND, merupakan fakultas yang telah melaksanakan program JC, untuk meningkatkan mutu input pendidikan S2 dan memberikan peluang bagi mahasiswa S1 berprestasi akademik tinggi menempuh dan memperoleh gelar Magister Sains (M.Si.) dalam waktu lebih pendek di FMIPA UNAND. Pada pelaksanaannya, FMIPA UNAND perkuliahan program JC S1-S2 dirancang dengan total waktu studi 10 (sepuluh) semester atau 5 (lima) tahun, terdiri dari maksimal 4 (empat) tahun program S1 dan minimal 1 (satu) tahun program S2. Semua mata kuliah pada kurikulum S1 sampai semester 8 wajib dipenuhi dengan beban minimal 144 sks. Beban sks per semester pada program JC S1-S2 selama berstatus mahasiswa S1 merupakan gabungan mata kuliah S1 dan S2 maksimal 18 sks dan perkuliahan dilaksanakan bersama dengan program S2 yang sudah berjalan di kelas reguler dan tidak diadakan kelas khusus untuk program JC S1-S2. Selanjutnya FMIPA UNAND terus melakukan pemantauan berkelanjutan dan pada akhirnya mahasiswa dengan kemampuan akademik dan kemampuan bekerja mandiri yang tinggi yang diharapkan akan terus melanjutkan studi hingga program magister. Mahasiswa S1 dinyatakan secara efektif menjalankan program JC S1 - S2 mulai semester 7 (tujuh) jika mahasiswa telah mengisi formulir minat mengikuti program JC dan pengajuannya disetujui oleh Dekan Fakultas. Namun, jika melewati waktu yang ditetapkan maka dinyatakan mengundurkan diri dan dengan sanksi yang diberikan yaitu semua mata kuliah S2 yang pernah diambil tidak dapat diakui dan yang bersangkutan tidak dapat melanjutkan ke S2.

Tahap proses pada penerimaan program JC yaitu dari Ketua Jurusan mengajukan nama mahasiswa yang akan mengikuti program JC ke Wakil Dekan I, kemudian data akan diproses lalu akan dibuatkan surat keterangan daftar mahasiswa yang akan mengikuti program JC, pada tahap penyeleksian yang dilakukan oleh Ketua Jurusan adalah dengan menetapkan beberapa kriteria yakni data indeks prestasi mahasiswa semester 1-6, nilai mata kuliah semester 1-6, jumlah sks sampai semester 6, dan nilai *TOEFL*. Bayangkan proses seleksi penyeleksian mahasiswa yang dapat menerima program JC oleh Ketua Jurusan dilakukan secara manual maka, akan memakan waktu yang sangat lama dan memungkinkan juga terjadi kesalahan manusia. Teknik klasifikasi diperlukan untuk menyeleksi mahasiswa yang dapat menerima program JC agar lebih optimal. Penyeleksian penerimaan program JC merupakan proses penting dalam menentukan penerima program JC yang sesuai pada perguruan

tinggi. Mengantisipasi terjadinya kegagalan oleh penerima program JC dalam pelaksanaannya maka perlu diperhatikan tingkat dan teknik klasifikasi penyeleksian penerimaan agar sesuai dengan kriteria (Karo dkk, 2021). Berdasarkan permasalahan tersebut dapat diperbaiki dengan dioptimalkan menggunakan Sistem Pendukung Keputusan (SPK) dengan menggunakan metode algoritma *K-Nearest Neighbor* (K-NN).

Pada penelitian ini, dilakukan optimalisasi dengan menerapkan algoritma K-NN untuk menyeleksi mahasiswa yang layak menerima program JC (Fitriyadi & Muqorobin, 2021). Algoritma K-NN adalah algoritma pembelajaran terawasi yang didasarkan pada pencarian kedekatan atribut penelitian dan mengklasifikasikannya berdasarkan suara terbanyak dari tetangga yang ditemukan (Pujiyanto dkk, 2021). K-NN mengasumsikan kedekatan antara data baru dan data yang tersedia dan menempatkan data baru ke dalam kategori yang paling dekat dengan kategori yang tersedia. Ini menyimpan semua data yang tersedia dan mengklasifikasikan titik data baru berdasarkan kedekatan (Geetha dkk, 2021). Kelebihan dari algoritma K-NN adalah pelatihan sangat cepat, sederhana, dan mudah dipelajari, tahan terhadap *data training* yang memiliki *noise*, dan efektif jika *data training*-nya besar (Roviani dkk, 2021). Algoritma K-NN dapat memberikan kinerja yang baik karena algoritma tersebut tahan terhadap *noise* data (Winarti dkk, 2021). K-NN termasuk kedalam metode *machine learning* karena melibatkan data masa lalu untuk memprediksi data masa depan. Dengan ini, garis keputusan kelas yang dihasilkan oleh K-NN bisa sangat fleksibel dan *nonlinier*. Dalam hal ini, untuk mengklasifikasikan data, metode K-NN dapat melakukannya dengan tepat (Rukmawan dkk, 2021). Proses K-NN tidak hanya untuk peramalan tetapi untuk memprediksi klasifikasi dan asosiasi dalam menemukan pola. K-NN merupakan hasil improvisasi lanjutan dari teknik klasifikasi Nearest Neighbor (NN). Ini menyimpulkan karena setiap data baru dapat dilatih oleh banyak  $k$  tetangga, dimana  $k$  adalah bilangan bulat positif, dengan sejumlah kecil data. Tahapnya, K-NN melakukan data latih dan data uji berdasarkan kategori pada sampel atau pada data masa lalu dan sesuai dengan sampel latih  $k$  yang merupakan tetangga terdekat dengan sampel uji, kemudian masuk ke dalam kategori yang memiliki kategori probabilitas terbesar (Khowarizmi dkk, 2021).

Algoritma K-NN dipilih sebagai algoritma dalam klasifikasi karena memiliki tingkat akurasi yang relatif tinggi (Effendy dkk, 2022). Algoritma K-NN merupakan algoritma *Nonparametric*. Analisis pembelajaran dan prediksi dilakukan berdasarkan

masalah atau kumpulan data yang diberikan (Theerthagiri dkk, 2021). Berbagai macam pengklasifikasi telah diselidiki sepanjang tahun namun, klasifikasi K-NN terbukti memiliki kinerja yang unggul (Hatem, 2022). Regresi dapat dilakukan dengan menggunakan K-NN tetapi sebagian besar cocok untuk klasifikasi (Geetha dkk, 2021). Hal ini dibuktikan dengan beberapa penelitian sebelumnya, di antaranya penelitian yang telah dilakukan Hanes, dkk (2021) yaitu perbandingan algoritma *C4.5*, *K-Nearest Neighbor*, *Naïve Bayes* dan *Random Forest* dalam memprediksi mahasiswa wisuda tepat waktu, dengan data yang diolah adalah data lulusan Program Studi S1 Teknik Informatika dan Sistem Informasi STMIK Mikroskil Medan periode 2011-2014 sejumlah 2,022 data, atribut yang digunakan meliputi jenis kelamin, daerah asal, waktu studi, nilai ujian masuk, dan Indeks Prestasi Kumulatif (IPK) dengan hasil akhir, algoritma *K-Nearest Neighbor* dan *Random Forest* memiliki akurasi tertinggi sebesar 72,651%, disusul oleh algoritma *C4.5* 72,453%, dan algoritma *Naïve Bayes* 71,860% (Gunawan dkk, 2021). Kelebihan lainnya, Algoritma K-NN memodelkan proses pelatihan yang sangat cepat, sederhana, memiliki kinerja yang baik terhadap data pelatihan yang bising, dan efektif dengan data pelatihan yang besar (Pujianto dkk, 2021). Pendekatan dalam algoritma K-NN adalah dengan cara menghitung kedekatan antara kasus baru dengan kasus yang lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah atribut yang ada. Algoritma K-NN dapat mengelompokkan banyak data secara efisien dan dalam waktu yang cepat, namun memiliki kelemahan yaitu dapat menjadi jauh lebih lambat dengan jumlah data yang semakin banyak (Priyasadie & Isa, 2021). Penelitian lainnya dilakukan Effendy, dkk (2021) pada perbandingan algoritma *Naïve Bayes Classifier* dan K-NN dalam klasifikasi Deposito Bank, data yang digunakan adalah data nasabah bank yang terdiri dari 4,521 *record* dengan 17 faktor variabel, dan hasilnya metode K-NN lebih baik daripada metode *Naïve Bayes Classifier*. Dimana klasifikasi terbaik dari metode K-NN adalah pada proses *training-testing* sebesar 70%:30% yang memiliki tingkat akurasi sebesar 89,23%. Keakuratan klasifikasi keseluruhan dari algoritma K-NN sangat tergantung pada pilihan jumlah nilai tetangga terdekat ( $k$ ).

*Data Mining* adalah proses mengekstrak atau menambang data dan informasi yang besar, yang sebelumnya tidak diketahui, tetapi dapat dipahami dan berguna dari *database* yang besar dan digunakan untuk membuat keputusan bisnis yang sangat penting. Konsep yang digunakan pada *Data Mining* yaitu untuk menemukan nilai tambah yang tersembunyi dalam *database* dengan tujuan mengekstraksi dan

mengidentifikasi informasi pengetahuan yang potensial dan berguna (Siahaan & Turnip, 2022). *Data Mining* digunakan untuk mengolah data besar yang tersimpan dalam *database* yang telah tertumpuk hingga menghasilkan data baru yang memberikan informasi dan pengetahuan baru. *Data Mining* menggambarkan sekumpulan teknik dengan tujuan untuk menemukan pola yang tidak diketahui pada data yang telah dikumpulkan. *Data Mining* memungkinkan pengguna untuk menemukan pengetahuan dalam *database* data yang tidak mungkin bagi pengguna untuk mengetahui keberadaannya (Fitriyadi & Muqorobin, 2021).

Misalkan akan diidentifikasi seleksi penerimaan seorang mahasiswa program JC yang masih semester 6 dengan menggunakan data dari mahasiswa yang telah diseleksi sebelumnya. Untuk mengidentifikasi seleksi penerimaan mahasiswa tersebut, dihitung kedekatan kasus mahasiswa yang akan diseleksi dengan semua kasus mahasiswa yang telah diseleksi sebelumnya dengan kedekatan terbesar lah yang akan diambil sebagai solusi pada kasus mahasiswa yang akan diseleksi tersebut.

Pada penelitian ini data yang digunakan adalah data mahasiswa program JC FMIPA UNAND. Dan beberapa atributnya seperti data indeks prestasi mahasiswa semester 1-6, jumlah sks sampai semester 6, dan keterangan status mahasiswa *fast-track* dan bukan *fast-track* dari mahasiswa angkatan 2014-2019. Kebanyakan *dataset* memiliki kepadatan dan distribusi kelas yang tidak merata. Distribusi densitas dapat dikatakan rapat jika jarak antar titik data sangat dekat, dan renggang jika titik data berjauhan. Dekat atau jauh jarak suatu titik dari tetangganya dapat dihitung dengan menggunakan jarak *Euclidean*. Fungsi jarak *Euclidean* populer digunakan oleh sebagian besar algoritma K-NN untuk mencari jarak antar titik data (Onyezewe dkk, 2021). Jarak *Euclidean* digunakan dalam pengklasifikasi ini dalam titik pengujian dan semua titik pelatihan untuk menyimpulkan K jarak *Euclidean* terkecil sebagai tetangga terdekat (Abuzaraida dkk, 2021). Dalam pencarian nilai K-NN, rumus yang digunakan berdasarkan dengan perhitungan jarak (*euclidean*) dan data lama (*data training*) yang digunakan adalah seluruh calon mahasiswa *Fast-Track* yang sudah mempunyai keterangan seleksi diterima atau ditolak *Fast-Track*, sedangkan data baru (*data testing*) menggunakan mahasiswa angkatan 2019 yang diambil secara acak.

## 1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, permasalahan penelitian ini dirumuskan sebagai berikut:

1. Bagaimana proses metode K-NN digunakan untuk mengidentifikasi seleksi mahasiswa program JC?
2. Bagaimana identifikasi penerimaan program JC menjadi lebih optimal?

## 1.3 Batasan Masalah

Untuk mencapai tujuan penelitian ini, peneliti memfokuskan batasan penelitian sebagai berikut:

1. Data yang digunakan merupakan data mahasiswa FMIPA UNAND.
2. Atribut yang digunakan yaitu NIM, Nama, Indeks Prestasi Semester 1-6, jumlah SKS sampai semester 6, dan keterangan status mahasiswa *fast-track* atau bukan *fast-track*.
3. Metode yang digunakan adalah K-NN.

## 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Memahami cara kerja metode K-NN dalam identifikasi penerimaan mahasiswa program JC.
2. Menganalisa keakuratan metode K-NN dalam identifikasi penerimaan mahasiswa program JC.
3. Merancang SPK yang dapat membantu penerimaan mahasiswa program JC.
4. Menerapkan metode K-NN dengan menggunakan rumus perhitungan jarak *Euclidean*.
5. Menguji hasil dari sistem yang telah dibangun.

## 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

1. Dapat memberikan informasi kelayakan mahasiswa untuk menerima program JC.
2. Dapat membantu mengevaluasi penerimaan mahasiswa program JC pada sistem akademik yang berjalan di perguruan tinggi.

## 1.6 Sistematika Penulisan

Sistematika penulisan merupakan sebuah kerangka atau urutan dalam sebuah penelitian agar lebih terstruktur. Untuk pembahasan lebih rinci, maka dalam penulisan tesis ini akan diurutkan menjadi beberapa bab, dimana satu sama lain saling berhubungan sesuai dengan ruang lingkup masalah. Secara umum gambaran urutan dan isi dari masing-masing bab 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 latar belakang, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian yang diakhiri dengan sistematika penulisan.

### **BAB II     Landasan Teori**

Pada bab ini dijelaskan teori-teori mengenai teknik untuk menunjang penulisan tesis. Adapun dibahas dalam bab ini adalah teori yang berkaitan tentang *Data Mining* dan metode K-NN.

### **BAB III    Metodologi Penelitian**

Bab ini membahas tentang kerangka penelitian dan penggunaan secara matematis metode K-NN.

### **BAB IV    Analisa dan Perancangan**

Bab ini membahas analisa dari sistem lama dengan membuktikan hasil scan bentuk laporan (*layout*) yang ada dan menyajikan data-data yang akan diolah.

**BAB V Implementasi dan Hasil**

Bab ini menguji hasil implementasi metode K-NN pada data penerimaan mahasiswa program JC FMIPA UNAND.

**BAB VI Kesimpulan dan Saran**

Bab ini membuat kesimpulan dan hasil penelitian keunggulan metode K-NN dalam identifikasi penerimaan mahasiswa program JC pada FMIPA UNAND.