

BAB I

PENDAHULUAN

1.1 Latar Belakang

Mahasiswa merupakan calon intelektual dalam suatu lapisan masyarakat yang berkaitan dengan berbagai predikat. Mahasiswa dapat diartikan sebagai seseorang yang sedang dalam proses menimba ilmu dan sedang menjalani pendidikan pada salah satu bentuk perguruan tinggi, yang terdiri dari akademik, politeknik, sekolah tinggi, institut dan universitas (Hartaji & Sedjo, 2012).

Sesuai dengan Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia (Permendikbud) No. 49 tahun 2014 tentang Standar Nasional Pendidikan Tinggi dalam pasal 17 ayat 2 berbunyi “144 sks untuk diploma empat dan program sarjana” dan ayat 3 berbunyi “4 sampai 5 tahun untuk program diploma empat dan program sarjana”. Hal ini sesuai dengan Akademik Universitas Muhammadiyah Jember “Program Sarjana (S1) merupakan jenjang pertama program akademik yang mempunyai beban studi 144 sampai 160 SKS, dijadwalkan sekurang-kurangnya 8 semester dan selama-lamanya 14 semester”.

Data yang digunakan yaitu data IPS (Indeks Prestasi Semester) dari semester 1 sampai semester 6, yang akan dijadikan sebagai dasar perhitungan prediksi yang dilakukan dengan harapan dapat menemukan informasi mahasiswa yang terindikasi lulus tepat waktu dan mahasiswa yang lulus tidak tepat waktu. Untuk memprediksi masa studi mahasiswa maka perlu dipersiapkan sebuah data training, dimana data tersebut diambil dari data sampel kelulusan mahasiswa Teknik Informatika sebelumnya dan mengetahui parameter apa saja yang mempengaruhi seorang mahasiswa dapat menyelesaikan studi tepat waktu. Sehingga perlu dibuatkan model algoritma yang digunakan untuk mendukung prediksi masa studi mahasiswa adalah algoritma *Modified K-Nearest Neighbor* merupakan pengembangan dari algoritma *K-Nearest Neighbor*.

Modified K-Nearest Neighbor merupakan sebuah algoritma yang dikembangkan dari algoritma KNN atau *K-Nearest Neighbor* yang dimodifikasi dengan adanya penambahan proses baru yakni, perhitungan validitas dan *weight*

voting. Jadi seluruh data latih harus divalidasi satu kali terlebih dahulu, yang dihitung sesuai dengan tetangga terdekatnya kemudian dilakukan perhitungan *weight voting* menggunakan nilai validitas tersebut (Parvin dkk., 2010).

Sebagai pendukung proses olah data, terdapat beberapa penelitian sebelumnya yang dilakukan oleh (Amelia dkk, 2017) Prediksi masa studi mahasiswa dengan menggunakan Algoritma *Naive Bayes*. Dalam penelitian ini diperoleh tingkat akurasi sebesar 85,17% sedangkan nilai akurasi tertinggi sebesar 88,96%. Penelitian selanjutnya (Orpa dkk, 2019) Model prediksi awal masa studi mahasiswa menggunakan Algoritma *Decision tree C4.5*. Dalam penelitian ini diperoleh rata-rata nilai *Precision*, *Recall* dan *Accuracy* diatas 50%. Penelitian selanjutnya (Hermawanti dkk, 2019) Implementasi Algoritma *C4.5* untuk prediksi kelulusan tepat waktu (Studi Kasus : Program Teknik Informatika). Dalam penelitian ini pola yang dibentuk mempunyai akurasi kecocokan sebesar 68,42% nilai presisi sebesar 42,86, dan penarikan sebesar 60%. Penelitian selanjutnya (Abidin dkk, 2017) Penerapan Metode *K-Nearest Neighbor* dalam Memprediksi Masa Studi Mahasiswa (Studi Kasus : Mahasiswa STIKOM Dinamika Bangsa) dalam penelitian ini untuk 6 semester memprediksi masa studi mahasiswa dengan tingkat keberhasilan 83%. Penelitian selanjutnya (Muhammad, 2015) dengan judul *Modified K-Nearest Neighbor* untuk Prediksi Curah Hujan. Dari hasil penelitian dapat disimpulkan bahwa akurasi terbaik dihasilkan oleh BPNN-*lm* sebesar 82,46%. Pada penelitian selanjutnya yang dilakukan oleh (Fitri et al., 2017) yang berjudul Deteksi Penyakit Kucing dengan menggunakan *Modified K-Nearest Neighbor Teroptimasi*. Pada penelitian ini di dapatkan hasil akurasi dari metode *Modified K-Nearest Neighbor* menggunakan algoritma genetika untuk *k* optimal 1 sebesar 100%. Dan pada penelitian yang dilakukan oleh (Imanda et al., 2018) dengan judul Klasifikasi Kelompok Varietas Unggul Padi Menggunakan *Modified K-Nearest Neighbor*. Dari hasil penelitian ini didapatkan nilai rata-rata akurasi tertinggi sebesar 79,96% dan nilai akurasi terendah sebesar 51,2%. Melihat dari penelitian – penelitian sebelumnya diharapkan pada penelitian ini mampu menghasilkan akurasi yang lebih baik.

Berdasarkan paparan di atas, penulis tertarik untuk mengklasifikasikan lama studi mahasiswa melalui penelitian dengan judul “Penerapan Algoritma *Modified K-Nearest Neighbour (MKNN)* Pada Klasifikasi Masa Studi Mahasiswa Teknik Informatika”. Penggunaan metode *Modified K-Nearest Neighbor* diharapkan dapat menghasilkan nilai akurasi, presisi dan recall yang tinggi.

1.2 Rumusan Masalah

Adapun rumusan masalah dari penelitian tugas akhir ini yakni Berapa tingkat akurasi, presisi dan *recall* pada Penerapan Algoritma *Modified K-Nearest Neighbour (MKNN)* pada klasifikasi masa studi mahasiswa ?

1.3 Batasan Masalah

1. Data yang digunakan merupakan data Mahasiswa Jurusan Teknik Informatika dari angkatan 2010 - 2016 sebanyak 260 data.
2. Dataset diambil dari Pusat Data dan Informasi Universitas Muhammadiyah Jember.
3. Dengan data yang dijadikan kriteria meliputi : NIM, Tahun Lulus, IPS dari Semester 1 – Semester 6.
4. Tools yang digunakan yaitu *Jupyter Notebook*.
5. K yang digunakan yaitu $K = 3$, $K = 5$ dan $K = 7$ dan $k = 9$, karena jika digunakan nilai genap maka kemungkinan yang dihasilkan nilainya sama.
6. K *Fold* yang digunakan k -fold = 2, k -fold = 4, k -fold = 5, k -fold = 10, digunakan nilai k -fold dengan membagi sampel data secara acak dengan dan mengelompokkan data tersebut sebanyak nilai K k -fold, salah satu kelompok dijadikan sebagai data uji sedangkan sisa kelompok lain akan dijadikan data latih (Jiang & Chen, 2016).

1.4 Tujuan Penelitian

Tujuan dari penelitian tugas akhir ini adalah untuk mengetahui berapa tingkat akurasi, presisi dan *recall* pada Penerapan Algoritma *Modified K-Nearest Neighbour* (MKNN) pada klasifikasi masa studi mahasiswa Teknik Informatika.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian tugas akhir ini, yaitu :

1. Untuk mengetahui nilai akurasi, presisi, dan recall pada Algoritma *Modified K-Nearest Neighbor*.
2. Dapat dijadikan bahan kajian mengenai Algoritma *Modified K-Nearest Neighbor* pada klasifikasi masa studi mahasiswa Teknik Informatika.

