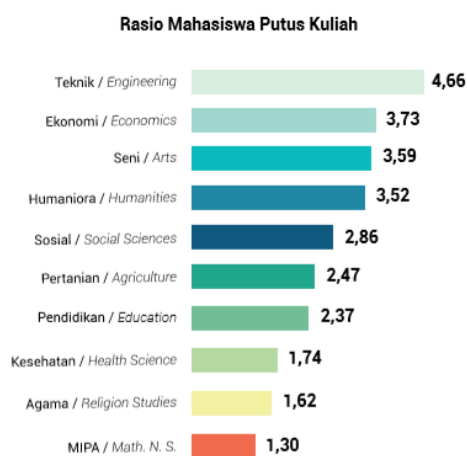


BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Menurut UU No. 20 Tahun 2003 Pasal 19 Ayat 1, Perguruan Tinggi (PT) merupakan jenjang pendidikan setelah pendidikan menengah mencakup program pendidikan diploma, sarjana, magister, spesialis, dan doktor yang diselenggarakan oleh PT. Berdasarkan bentuk penyelenggaranya PT yang ada di Indonesia terdiri dari Perguruan Tinggi Negeri (PTN) dan Perguruan Tinggi Swasta (PTS). Pada Tahun 2018, jumlah PT yang ada di Indonesia baik swasta maupun negeri sekitar 4.670 dan jumlah program studi (prodi) sekitar 27.779 (Ristekdikti, 2018). Bertambah maraknya jumlah PT disertai berbagai macam disiplin ilmu mengakibatkan tingkat persaingan sesama PT meningkat. Maka kualitas perguruan tinggi harus selalu ditingkatkan untuk menjadi yang terbaik dan tetap eksis di kalangan masyarakat. Menurut Purba et al., (2018) kualitas suatu perguruan tinggi dapat dilihat dari tingkat keberhasilan mahasiswa yang tinggi dan tingkat kegagalan mahasiswa yang rendah. Keberhasilan mahasiswa dapat dilihat dari prestasi keberhasilan mahasiswa di suatu perguruan tinggi yang dinaungi sedangkan kegagalan mahasiswa biasanya ditandai dari ketidakmampuan perguruan tinggi dalam menyelenggarakan pendidikan. Berdasarkan data statistik pendidikan tinggi tahun 2018 (Gambar 1.1), mahasiswa putus kuliah berasal dari kelompok bidang kesehatan dengan rasio 1,74 dari total mahasiswa bidang (Ristekdikti, 2018).



Gambar 1. 1 Rasio mahasiswa putus kuliah

STIKES Guna Bangsa Yogyakarta merupakan suatu institusi pendidikan swasta yang menyelenggarakan pendidikan tinggi di bidang kesehatan, salah satu program studinya yaitu D3 Teknologi Laboratorium Medik. Berdasarkan hasil wawancara dengan ketua program studi D3 Teknologi Laboratorium Medik, permasalahan yang sering terjadi di program studi Teknologi Laboratorium Medik STIKES Guna Bangsa Yogyakarta diantaranya: 1) mahasiswa tidak lulus tepat waktu, dan 2) mahasiswa lulus namun dengan nilai Indeks Prestasi Kumulatif (IPK) yang minimal (IPK=2.50). Dengan demikian, lulusan mahasiswa yang memiliki IPK minim tidak mampu bersaing di dalam dunia kerja khususnya menjadi Pegawai Negeri Sipil (PNS). Oleh karena itu untuk meningkatkan kualitas dan keberhasilan mahasiswa, STIKES Guna Bangsa Yogyakarta membutuhkan ramalan atau prediksi sedini mungkin untuk meminimalisir terjadi kegagalan pada mahasiswa yaitu dengan memprediksi kelulusan mahasiswa yaitu tepat waktu atau tidak tepat waktu (terlambat) di masa datang.

Seiring dengan cita-cita kampus STIKES Guna Bangsa Yogyakarta menjadi *Smart Campus*, *tools* untuk melakukan prediksi kelulusan mahasiswa sangat dibutuhkan. *Tools* tersebut dibutuhkan sebagai salah satu bahan pengingat (*alarm*) bagi mahasiswa untuk meningkatkan kinerja dan prestasinya serta sebagai bahan *monitoring* sedini mungkin bagi pengelola program studi D3 Teknologi Laboratorium Medik untuk membimbing mahasiswa yang berpotensi lulus tidak

tepat waktu (terlambat). Maka dari itu, peneliti mengusulkan untuk membuat pemodelan prediksi ketepatan waktu lulus mahasiswa yang nantinya dapat bermanfaat untuk pengelola program studi D3 Teknologi Laboratorium Medik. Diharapkan dengan adanya sistem yang dapat memprediksi ketepatan waktu lulus mahasiswa tersebut dapat meningkatkan kualitas program studi dan Perguruan Tinggi.

Teknologi Data mining (DM) telah banyak diadopsi untuk meningkatkan cakupan, kualitas, efisiensi dan pencapaian sistem pendidikan, sehingga disebut *education data mining* (EDM). Menurut F. Ahmad, Ismail, & Aziz (2015) Data Mining merupakan teknik yang kuat untuk berbagai bidang termasuk pendidikan. Secara khusus, pengguna semakin mendapatkan keuntungan dari teknologi yang menyediakan interpretasi data khusus dan teknik pemrosesan. Dengan kemampuan tersebut, DM kini dianggap sebagai salah satu teknologi pendidikan yang menjanjikan sehingga disebut sebagai *education data mining* (EDM) (Wook, Yusof, Zakree, & Nazri, 2015). Data Mining adalah suatu teknologi canggih yang dapat diartikan sebagai proses untuk mengekstraksi pengetahuan dan informasi yang berguna (El-Halees, 2015). Dengan demikian dari data historis tersebut dapat diambil informasi-informasi penting yang tersembunyi (Setiawan, 2016). Informasi penting tersebut nantinya dapat digunakan sebagai bahan evaluasi dan analisa bagi Perguruan Tinggi untuk meningkatkan kualitas pendidikan.

Dalam kasus prediksi, dibutuhkan sebuah metode atau model yang mampu menghasilkan pola klasifikasi data dengan tujuan akhir yaitu *forecasting*. Teknik atau metode yang akan digunakan untuk mengetahui informasi itu adalah dengan cara menambang data akademik mahasiswa. Algoritma klasifikasi untuk prediksi yang digunakan oleh peneliti sebelumnya diantaranya: C.45 (H. N. Ahmad et al., 2017) (Miguéis et al., 2018), *Random Forest*, *Bayesian Network* (Hussain et al., 2018), *Decision Tree*, *Naïve Bayes* (F. Ahmad et al., 2015); (Miguéis et al., 2018) dan SVM (Miguéis et al., 2018). Guna pengembangan penelitian terbaru, peneliti akan melakukan perbandingan (komparasi) dua metode untuk mencari model yang terbaik pada kasus prediksi ketepatan waktu lulus mahasiswa di program studi D3

Teknologi Laboratorium Medik STIKES Guna Bangsa Yogyakarta. Algoritma yang akan diterapkan yaitu *Naïve Bayes Classification* dan *K-Nearest Neighbors*.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian diatas maka diperoleh rumusan masalah sebagai berikut:

1. Apa saja atribut-atribut yang dapat mempengaruhi tingkat kelulusan mahasiswa untuk dapat digunakan sebagai variabel independen pada penelitian?
2. Bagaimana evaluasi implementasi algoritma *Naïve Bayes Classification* dan *K-Nearest Neighbors* dalam pemodelan prediksi ketepatan waktu lulus mahasiswa dengan mengukur tingkat akurasi, presisi, dan *recall* menggunakan teknik skenario uji *n-Folds Cross Validation*.
3. Algoritma mana yang memiliki tingkat akurasi, presisi, dan *recall* terbaik untuk kasus prediksi ketepatan waktu lulus mahasiswa tepat dan terlambat khususnya program studi D3 Teknologi Laboratorium Medik STIKES Guna Bangsa Yogyakarta?

1.3 Batasan Penelitian

Batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian adalah data mahasiswa program studi Teknologi Laboratorium Medik STIKES Guna Bangsa untuk Tahun Akademik (TA) 2015/2016 sampai dengan 2018/2019. Pertimbangan penggunaan keseluruhan data mahasiswa, baik yang sudah memiliki status akhir lulus maupun belum, dikarenakan agar sistem dapat menampilkan model prediksi yang akurat dengan tidak menghapus data sebenarnya.
2. Penentuan atribut-atribut yang mempengaruhi kelulusan mahasiswa didapatkan berdasarkan hasil studi literature dan *focus group discussion* yang selanjutnya divalidasi oleh ahli/*decision maker* dalam hal ini adalah pengelola Program Studi dan bagian admisi atau penerimaan mahasiswa

baru. Atribut-atribut tersebut kemudian dilakukan pembobotan dengan *weight chi square* untuk melihat tingkat variabel yang memiliki nilai pembobotan paling tinggi dan yang terendah sehingga dapat dijadikan variabel independen untuk prediksi ketepatan waktu lulus mahasiswa.

3. Penelitian ini tidak membahas masalah mengenai analisa pada kompleksitas algoritma.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengetahui atribut-atribut yang dapat mempengaruhi kelulusan mahasiswa yang nantinya akan digunakan sebagai variabel independen pada penelitian.
2. Mengevaluasi implementasi algoritma *Naive Bayes Classification* dan *K-Nearest Neighbors* dalam pemodelan prediksi ketepatan waktu lulus mahasiswa dengan mengukur tingkat akurasi, presisi, dan *recall* menggunakan teknik skenario uji *n-Folds Cross Validation*.
3. Mengetahui algoritma mana yang memiliki tingkat akurasi, presisi, dan *recall* yang lebih tinggi untuk kasus prediksi ketepatan waktu lulus mahasiswa tepat dan terlambat khususnya program studi D3 Teknologi Laboratorium Medik STIKES Guna Bangsa Yogyakarta.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberi manfaat berupa:

Manfaat yang diharapkan dapat diperoleh dari penelitian ini meliputi manfaat praktis dan manfaat teoritis. Manfaat praktis adalah manfaat langsung dirasakan oleh praktisi pengguna hasil penelitian. Sedangkan manfaat teoritis berarti manfaat yang didapatkan oleh pembaca penelitian. Adapun manfaat tersebut dijabarkan sebagai berikut:

1. Manfaat Praktis

Penelitian ini diharapkan dapat memberikan manfaat sebagai salah satu pertimbangan pengambilan keputusan untuk menindaklanjuti kinerja

akademik mahasiswa (berdasarkan prediksi kelulusan mahasiswa) D3
Teknologi Laboratorium Medik STIKES Guna Bangsa.

2. Manfaat Teoritis

Penelitian ini diharapkan dapat menjadi rujukan bagi penelitian-penelitian sejenis dan menambah keilmuan bagi penelitian-penelitian di bidang klasifikasi data mining.