BAB I
PENDAHULUAN

1.1 Latar Belakang

Pertama kali pembelajaran dalam jaringan atau daring hadir akibat pengaruh perkembangan pembelajaran berbasis elektronik (*e-learning*) yang dikenalkan Universitas Illinois lewat sistem pembelajaran berbasis komputer [1]. Pembelajaran secara daring adalah sistem belajar yang menyediakan fasilitas tanpa batasan jarak, ruang, dan waktu untuk belajar dengan bantuan teknologi elektronika berupa komputer, laptop, tablet, atau *smartphone* [1][2]. Materi yang diajarkan disampaikan secara lebih bervariasi, bukan hanya verbal namun juga audio, visual, dan gerak [1].

Semenjak Pandemi *Corona Virus Disease* (COVID-19), pembelajaran daring mulai gencar diterapkan di Indonesia secara resmi pada bulan Maret tahun 2020 menurut surat edaran dari Menteri Pendidikan dan Kebudayaan Nomor 36962/MPK.A/HK/2020 dan Nomor 4 Tahun 2020 [3][4]. Pada jenjang pendidikan tinggi tercatat ada 98% mahasiswa sudah melakukan pembelajaran daring berdasarkan hasil survei Kemendikbud terhadap 237.193 responden dari banyaknya mahasiswa di Indonesia pada Pangkalan Data Pendidikan Tinggi (PPDikti) yaitu 8.325.013 orang [5]. Hasil survei tersebut menyatakan ada mahasiswa yang dapat mengikuti pembelajaran daring dan juga tidak. Oleh karena itu, pada masing-masing perguruan tinggi perlu mengadakan pendataan yang lebih mendalam terkait sistem pembelajaran ini.

Salah satu perguruan tinggi yang menerapkan sistem pembelajaran daring yaitu Institut Teknologi Telkom Purwokerto (ITTP). Berdasarkan survei kebutuhan dosen dalam pelaksanaan *E-Learning* yang dikumpulkan bagian Pengembangan Pembelajaran ITTP kendala yang dialami dosen yaitu sulitnya berinteraksi dengan mahasiswa. Akibatnya pengajar mengalami kesulitan dalam

menilai mana mahasiswanya yang sudah paham atau belum. Di samping itu, perolehan nilai tugas dan ujian juga belum tentu dapat menggambarkan tingkat pemahaman mahasiswa karena ada kemungkinan terjadi kecurangan dalam pengerjaannya sehingga tidak bisa dijamin keasliannya. Oleh karena itu, dibutuhkan alternatif lain untuk menentukan tingkat pemahaman mahasiswa yaitu melalui faktor kondisi kesehatan, motivasi, dan cara pengajaran menurut jurnal [6] dan pakar pembelajaran daring.

Alternatif penentu tingkat pemahaman mahasiswa yang telah disebutkan dapat membantu pengajar dalam menentukan mana mahasiswa yang sudah paham atau belum. Pengajar juga dapat mengukur keefektifan dari proses pembelajaran daring yang ada. Maka, dibutuhkan strategi yang tepat untuk mengoptimalkan pemahaman belajar mahasiswa saat sistem pembelajarannya daring. Salah satunya dengan mengetahui terlebih dahulu tingkat pemahaman mahasiswa. Berdasarkan pernyataan yang sudah disebutkan, proses klasifikasi dapat digunakan untuk menyelesaikan permasalahan terkait tingkat pemahaman mahasiswa [7].

Saat ini, dalam dunia pendidikan sudah banyak teknologi untuk menunjang proses pembelajaran, salah satunya *data mining* [8]. *Data mining* melakukan proses identifikasi pola dari kumpulan data yang bertujuan untuk memperoleh informasi tertentu yang belum diketahui sebelumnya [8][9]. Proses tersebut dapat dilakukan dengan teknologi komputer. Pada *data mining* ada beberapa teknik yaitu regresi, asosiasi, klaster, dan klasifikasi [8]. Klasifikasi adalah teknik pengolahan data berdasarkan kesamaan karakteristik sehingga objek yang belum memiliki label dapat diperkirakan termasuk kelas yang mana [8][9]. Beberapa algoritme klasifikasi diantaranya Naïve Bayes [10][11][12][13], *K-Nearest Neighbor* (KNN) [14][15], ID3 [16][17][18], C4.5 [19][9][17][20], dan *Support Vector Machine* (SVM) [12][13][11][21].

Menurut beberapa penelitian sebelumnya Naïve Bayes adalah algoritme sederhana dan lugas yang berdasar pada teorema Bayes untuk probabilitas bersyarat dengan pengklasifikasian data bergantung pada frekuensi pendeskripsian data dari set pelatihan [21][12]. KNN merupakan salah satu algoritme yang

memproses datanya dengan menentukan jarak data paling dekat [22]. ID3 adalah algoritme yang melakukan pencarian menyeluruh (*greedy*) pada semua kemungkinan pohon keputusan [23]. C4.5 yaitu algoritme yang membentuk pohon keputusan dengan setiap cabang terhubung ke *node* keputusan lain atau *node* daun yang mengakhiri [9]. SVM adalah algoritme yang sistem pembelajarannya menggunakan parameter kernel [24].

Berdasarkan penelitian sebelumnya, Naïve Bayes memiliki kelebihan yaitu dapat bekerja dengan jumlah data besar namun bekerja lebih baik dengan jumlah data yang sedikit untuk klasifikasi, sederhana, mudah untuk dibuat, dan kokoh pada atribut yang tidak relevan [25][15]. Algoritme Naïve Bayes selain dapat melakukan proses pengklasifikasian, algoritme tersebut juga dapat menghitung besar peluang pada suatu data untuk masuk ke kelas tertentu [26]. Lalu, C4.5 memiliki kelebihan yaitu dapat diinterpretasikan dalam bentuk pohon keputusan yang mudah dipahami, dapat mengatasi *noise* dan *missing value*, dapat melakukan pemangkasan cabang, dan dapat menghasilkan tingkat akurasi yang baik pada jumlah data besar namun tidak menutup kemungkinan juga dapat bekerja pada jumlah data kecil[19]. Menurut kelebihan dari kedua algoritme tersebut maka Naïve Bayes dan C4.5 dapat dijadikan bahan pada penelitian ini. Naive Bayes dan C4.5 akan digunakan untuk proses pengklasifikasian.

Penelitian ini akan menganalisis penggunaan Naïve Bayes dan C4.5 untuk klasifikasi pada studi kasus penentuan tingkat pemahaman mahasiswa. Kedua algoritme tersebut akan dibandingkan dari segi akurasi, *recall*, *precision*, nilai *error*, dan waktu komputasinya sehingga ditemukan algoitme terbaik. Adapun tujuan dari penelitian ini diharapkan dapat mengklasifikasi tingkat pemahaman mahasiswa yang mengikuti pembelajaran daring berdasarkan pada faktor yang dapat mempengaruhi pemahaman tersebut. Dengan begitu pengajar dapat terbantu dalam identifikasi pemahaman mahasiswanya dan kampus dapat melakukan pengoptimalan terhadap proses belajar mengajar.

1.2 Rumusan Masalah

Penelitian ini memiliki rumusan masalah yaitu:

1. Sistem pendidikan di Indonesia menerapkan pembelajaran daring sejak Pandemi Covid-19 tahun 2020. Pada sistem pendidikan ini pengajar mengalami kesulitan dalam mengetahui tingkat pemahaman mahasiswa terhadap materi yang telah diajarkan.
2. Penentuan tingkat pemahaman dapat diproses menggunakan klasifikasi. Ada banyak algoritme yang dapat menangani klasifikasi sehingga dibutuhkan perbandingan performa untuk menentukan algoritme terbaik.

1.3 Pertanyaan Penelitian

Pertanyaan penelitian yang ada yaitu:

1. Apa hasil klasifikasi tingkat pemahaman mahasiswa dengan menggunakan Algoritme Naïve Bayes dan C4.5?
2. Apa hasil analisis performa dari perbandingan algoritme klasifikasi Naive Bayes dan C4.5?

1.4 Batasan Masalah

Penelitian ini memiliki batasan masalah yaitu:

1. Data yang digunakan adalah hasil dari jawaban kuesioner terkait faktor yang mempengaruhi tingkat pemahaman mahasiswa saat mengikuti pembelajaran daring.
2. Responden meliputi mahasiswa di Institut Teknologi Telkom Purwokerto yang pernah mengikuti pembelajaran daring pada tahun 2021.
3. Pengambilan hasil survei kuesioner berdasarkan perhitungan menggunakan Rumus Slovin (e=5%) diperoleh jumlah sampel minimalnya 365 dari 4200 populasi mahasiswa Institut Teknologi Telkom Purwokerto.
4. Data tersebut diambil pada tanggal 12 April 2022 sampai 21 Juni 2022 dengan jumlah 419 *record data*.

5. Performa yang dibandingkan dari Algoritme Naïve Bayes dan C4.5 yaitu akurasi, *recall, precision,* RMSE, dan waktu komputasi.
6. Indikator yang dipakai untuk menentukan kemampuan mahasiswa dalam memahami materi yaitu kondisi kesehatan, motivasi, dan cara pengajaran yang didapat.

1.5 Tujuan Penelitian

Penelitian ini memiliki tujuan yaitu:

1. Mengetahui hasil klasifikasi tingkat pemahaman mahasiswa terhadap materi yang diajarkan saat pembelajaran daring diterapkan secara menyeluruh.
2. Mengetahui hasil analisis dan validasi performa terbaik dalam perbandingan algoritme klasifikasi Naïve Bayes dan C4.5.

1.6 Manfaat Penelitian

Penelitian ini memiliki manfaat yaitu:

1. Bagi Penulis

   Penelitian ini dapat melatih diri untuk mampu menyelesaikan masalah yang ada yaitu dalam penerapan Algoritme Naïve Bayes dan C4.5 untuk klasifikasi. Dapat menguji ilmu yang diperoleh dari perkuliahan dan dari luar perkuliahan. Selain itu, juga dapat menambah pengetahuan ilmiah dalam bidang pendidikan dan teknologi di Indonesia.
2. Bagi Pengajar

   Penelitian ini dapat memberikan rekomendasi algoritme terbaik untuk diterapkan dalam klasifikasi tingkat pemahaman mahasiswa. Dengan adanya klasifikasi berbasis teknologi ini, pengajar bisa lebih mudah dalam mendapatkan informasi mana mahasiswa yang sudah paham atau belum dengan materi yang telah diajarkan. Selain itu, dapat diketahui seberapa efektif pembelajaran daring yang telah diterapkan dengan mengacu pada tingkat pemahaman tersebut.