

# BAB I PENDAHULUAN

## I.1 Latar Belakang

Pendidikan merupakan salah satu aspek penting dalam membentuk nilai dan norma masyarakat sehingga menghasilkan pribadi yang dapat berpikir *high order thinking* melalui kegiatan belajar dan mengajar (Pantiwati dan Permana, 2017). Seiring dengan berkembangnya IPTEK, metode pembelajaran berkembang menjadi *e-Learning*. *e-Learning* merupakan salah satu jenis kegiatan belajar mengajar yang menggunakan aplikasi elektronik dengan media internet, intranet, ataupun jaringan komputer (Wahano, 2003). *e-Learning* dapat digunakan untuk membantu meningkatkan pembelajaran serta pemahaman pelajar (Hanum, 2013) melalui soal-soal yang dikumpulkan menjadi bank soal (Harahap, 2015).

Bank soal *e-Learning* digunakan untuk menguji pengetahuan pelajar tentang materi yang bersangkutan. Oleh karena itu, diperlukan soal yang berkualitas tinggi sehingga soal-soal tersebut dapat mengukur tingkat intelektual pelajar serta mencakup keterampilan berpikir mulai dari paling sederhana hingga paling kompleks (Pantiwati dan Permana, 2017). Bank soal *e-Learning* juga perlu dievaluasi kembali. Hal ini karena banyaknya soal-soal yang masuk sehingga perlu dikelompokkan kembali untuk mempermudah pengajar mendapatkan soal yang sesuai dengan capaian kompetensinya. Salah satu standar identifikasi pendidikan yang sering digunakan untuk meningkatkan kualitas soal adalah *Bloom's Taxonomy* (Mohammed dan Omar, 2018).

Umumnya klasifikasi soal berdasarkan *Bloom's Taxonomy* dilakukan secara manual, namun cara tersebut dinilai relatif lama untuk mengolah data dalam skala besar (Aninditya dkk., 2019). Oleh karena itu, penulis memberikan solusi untuk mengklasifikasi soal secara otomatis dengan menggunakan metode *machine learning*. Metode *machine learning* yang dapat digunakan untuk memproses soal ujian adalah menggunakan NLP (Kusuma dkk., 2016). Terdapat banyak algoritma yang dapat digunakan untuk melakukan klasifikasi soal sehingga penulis ingin membandingkan dua algoritma untuk mengetahui perbandingan performa kedua algoritma.

Terdapat beberapa penelitian yang telah melakukan klasifikasi soal berdasarkan tingkat *Bloom's Taxonomy* menggunakan *machine learning*. Pada penelitian (Osman dan Yahya, 2016) yang menguji menggunakan beberapa algoritma seperti NB, SVM, *Logistic Regression*, dan *Decision Trees* menjelaskan bahwa klasifikasi pertanyaan menggunakan SVM dan *Logistic Regression* memberikan hasil yang lebih baik. Selain itu, pada penelitian (Aninditya dkk., 2019) menjelaskan bahwa klasifikasi pertanyaan menggunakan NB memiliki tingkat akurasi yang terbilang tinggi yaitu 85%. Adapun penelitian (Sangodiah dkk., 2014) yang menjelaskan bahwa algoritma SVM mendukung fitur BOW yang dapat meningkatkan performa dalam klasifikasi pertanyaan. Pada penelitian (Osadi dkk., 2017) yang membandingkan beberapa algoritma dengan *ensemble classifier* menjelaskan bahwa *ensemble classifier* menduduki posisi pertama dalam tingkat akurasi, dilanjutkan dengan algoritma SVM dan NB di urutan kedua dan urutan ketiga.

Salah satu universitas yang telah menerapkan *e-Learning* adalah Universitas Telkom. Universitas Telkom memiliki beberapa program studi, salah satunya adalah S1 Sistem Informasi. Salah satu mata kuliah wajib di program studi ini adalah DWBI. Mata kuliah ini berawal dari mata kuliah peminatan yang berubah menjadi mata kuliah wajib. Penelitian ini dilakukan sebagai solusi berupa model yang akan membantu pengajar untuk mendapatkan soal pada *e-Learning* yang sesuai dengan standar kompetensi sesuai dengan RBT.

Metode yang akan digunakan adalah SVM dan NB. SVM mengklasifikasi berdasarkan ruang hipotesis seperti fungsi linier dalam sebuah fitur berdimensi tinggi (*hyperplane*) (Sugara dan Subekti, 2019). NB mengklasifikasikan variabel tertentu menggunakan probabilitas dan statistik (Kurniawan, 2018). Jenis algoritma NB yang digunakan dalam penelitian ini adalah *Multinomial Naïve Bayes*. Algoritma ini banyak digunakan untuk mengklasifikasikan teks atau dokumen dimana perhitungannya menggunakan jumlah *term* dalam dokumen (Aninditya dkk., 2019). SVM dinilai dapat menjadi *classifier* karena dapat

mengklasifikasikan soal dari area yang bervariasi dengan akurasi yang tinggi serta mendukung fitur BOW yang menjadi salah satu indikator dalam klasifikasi teks seperti soal-soal (Sangodiah dkk., 2014). Sementara algoritma NB dinilai dapat dijadikan sebagai *classifier* karena memiliki hasil yang akurat, dinilai baik dalam mengklasifikasikan data *multi-class* serta memberi hasil akurasi yang baik dan memiliki kecepatan tinggi untuk mengolah data dalam jumlah besar seperti bank soal (Aninditya dkk., 2019). Hasil dari penelitian tugas akhir ini adalah akan menghasilkan hasil klasifikasi berdasarkan tingkat kognitif RBT, tingkat akurasi, dan prediksi untuk menentukan klasifikasi soal mata kuliah DWBI.

## **I.2 Perumusan Masalah**

Berdasarkan latar belakang di atas, maka rumusan masalah yang akan diteliti adalah sebagai berikut:

1. Bagaimana perbandingan implementasi algoritma SVM dan NB dalam mengklasifikasikan soal-soal mata kuliah DWBI pada program studi S1 Sistem Informasi Universitas Telkom berdasarkan tingkat kognitif RBT?
2. Bagaimana perbandingan tingkat akurasi dari klasifikasi menggunakan algoritma SVM dan NB dalam mengklasifikasikan soal-soal mata kuliah DWBI pada program studi S1 Sistem Informasi Universitas Telkom berdasarkan tingkat kognitif RBT?
3. Bagaimana cara menangani *imbalance data* pada *dataset* soal-soal mata kuliah DWBI pada program studi S1 Sistem Informasi Universitas Telkom berdasarkan tingkat kognitif RBT?
4. Bagaimana perbandingan hasil akurasi antara *dataset* yang menggunakan metode *oversampling* dengan *dataset* tanpa menggunakan metode *oversampling*?

## **I.3 Tujuan Penelitian**

Penelitian ini bertujuan untuk:

1. Membandingkan hasil klasifikasi antara algoritma SVM dan NB pada soal kuis mata kuliah DWBI berdasarkan tingkat kognitif RBT,

2. Membandingkan hasil akurasi antara algoritma SVM dan NB terhadap klasifikasi soal kuis mata kuliah DWBI berdasarkan tingkat kognitif RBT,
3. Menjelaskan secara singkat cara menangani *imbalance data* pada *dataset* soal-soal mata kuliah DWBI pada program studi S1 Sistem Informasi Universitas Telkom berdasarkan tingkat kognitif RBT,
4. Membandingkan hasil akurasi antara *dataset* yang menggunakan metode *oversampling* dengan *dataset* tanpa menggunakan metode *oversampling*.

#### **I.4 Batasan Penelitian**

Adapun batasan dalam melakukan penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah soal kuis mata kuliah DWBI angkatan 2018 program studi S1 Sistem Informasi Universitas Telkom yang didapatkan dari pengajar,
2. Jumlah soal adalah sebanyak 161 butir,
3. Berfokus pada tipe soal yang diberikan untuk mahasiswa yaitu soal berbahasa Indonesia pilihan ganda berbentuk teks,
4. Soal tidak mengandung *equation* atau persamaan rumus,
5. Soal diklasifikasi berdasarkan tingkat kognitif RBT (2001),
6. Algoritma yang digunakan adalah SVM dan NB,
7. Metode yang digunakan untuk menangani *imbalance data* adalah SMOTE,
8. *Tools* yang digunakan adalah *Python*.

#### **I.5 Manfaat Penelitian**

Manfaat penelitian tugas akhir ini dapat dibagi menjadi dua yaitu manfaat teoritis dan manfaat praktis. Penelitian tugas akhir ini memiliki manfaat teoritis sebagai kontribusi dalam referensi mengenai klasifikasi pertanyaan khususnya pada soal kuis mata kuliah DWBI berdasarkan RBT menggunakan dua algoritma, yaitu SVM dan NB. Penelitian ini juga dapat menjadi kontribusi dalam referensi mengenai cara menangani *imbalance data* dengan SMOTE. Selain itu, penelitian ini juga dapat menjadi referensi bagi penelitian selanjutnya yang ingin menggunakan atau mengembangkan topik penelitian ini.

Manfaat praktis penelitian tugas akhir ini memiliki bagi berbagai pihak yang dapat diimplementasikan dalam kehidupan mereka, di antaranya:

1. Bagi Pengajar

Penelitian tugas akhir ini diharapkan mampu membantu pengajar untuk mengevaluasi kembali soal kuis sesuai tingkat RBT menggunakan hasil penelitian menggunakan model algoritma SVM maupun NB.

2. Bagi Mahasiswa

Penelitian tugas akhir ini diharapkan mampu membantu mahasiswa untuk mendapatkan soal-soal yang sesuai dengan standar tingkat RBT sehingga dapat meningkatkan efisiensi pembelajaran.

3. Bagi Penulis

Penelitian tugas akhir ini diharapkan mampu membantu penulis untuk dapat menerapkan dan membandingkan performa antara algoritma SVM dan NB untuk mengklasifikasi soal-soal mata kuliah DWBI di program studi S1 Sistem Informasi Universitas Telkom, serta membandingkan tingkat akurasi tingkat kognitif soal berdasarkan RBT.

## **I.6 Sistematika Penulisan**

Penelitian ini diuraikan dengan sistematika penulisan dimulai dari bab I sampai dengan bab VI. Pengerjaan laporan Tugas Akhir terbagi menjadi dua bagian yaitu bagian proposal yang terdiri dari bab I hingga bab III, sementara bagian isi penelitian dimulai dari bab IV hingga bab VI.

Bab I atau Pendahuluan berisi uraian mengenai konteks permasalahan, latar belakang penelitian, perumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian, dan sistematika penulisan. Bab II atau Tinjauan Pustaka berisi literatur yang relevan dengan permasalahan yang diteliti dan dibahas pula hasil-hasil penelitian terdahulu. Bab ini juga berisi uraian alasan pemilihan kerangka kerja penelitian. Bab III ataupun Metodologi Penelitian berisi uraian mengenai kerangka konseptual masalah, sistematika penyelesaian masalah, serta prosedur pengerjaan penelitian.

Bab IV atau Analisis dan Perancangan berisi uraian hasil pengerjaan dari pembahasan studi kasus dalam penelitian. Pada bab ini dijelaskan proses pengumpulan data, *data preprocessing*, dan pembobotan TF-IDF. Bab V atau Implementasi dan Pengujian berisi uraian tahap implementasi dari metode yang telah ditentukan yaitu penerapan algoritma SVM dan NB. Pada bab ini juga dijelaskan evaluasi performansi dari implementasi penelitian. Bab VI atau Kesimpulan dan Saran berisi kesimpulan dari penelitian yang dilakukan serta jawaban dari pertanyaan penelitian yang disajikan di pendahuluan. Bab ini juga berisi saran yang dapat membangun penelitian selanjutnya.