



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Daya saing merupakan salah satu bentuk implementasi dari peningkatan kualitas suatu perguruan tinggi (Kurniawan, 2018). Hal ini dapat dilakukan dengan mengoptimalkan potensi perguruan tinggi, salah satunya dengan mengaplikasikan teknologi informasi dalam proses penyelenggaraan pelayanan (Kurniawan, 2018). Penelitian Sulistyawati dan Seminari (2015) menjelaskan bahwa kualitas pelayanan merupakan faktor yang penting dalam memberikan kepuasan bagi pelanggan. Hal tersebut juga terbukti pada penelitian yang dilakukan oleh Suratno, dkk. (2016) bahwa tingginya kualitas pelayanan akan meningkatkan kepuasan terhadap pelanggan. Survei dapat dijadikan sebagai salah satu cara untuk mendapatkan umpan balik langsung dari mahasiswa dalam mengukur kualitas pelayanan pada perguruan tinggi. Pertanyaan terbuka dapat diberikan dalam survei untuk mengumpulkan data saran dan masukan.

Di Universitas Multimedia Nusantara, unit kampus yang bertugas untuk mengumpulkan data dan membuat laporan hasil survei kepuasan alumni adalah divisi Career Development Centre (External Student Affairs) (Reni, 2019). Hasil wawancara dapat dilihat pada Lampiran A. Berdasarkan wawancara dengan Reni (2019) selaku petugas Career Development Centre di Universitas Multimedia Nusantara, survei kepuasan alumni wajib diisi oleh lulusan untuk mengetahui tingkat kepuasan mahasiswa selama menempuh pendidikan di Universitas Multimedia Nusantara dari tahun ke tahun. Reni (2019) juga menjelaskan bahwa pemrosesan hasil survei dilakukan oleh bagian Alumni Relations setiap akhir tahun

berjalan. Proses perekapan data sampai kategorisasi masih berjalan secara manual untuk diketahui poin apa saja yang harus diperbaiki ataupun yang sudah baik.

Proses kategorisasi dilakukan dengan cara melihat beberapa kata kunci yang berhubungan dengan bagian-bagian terkait yang dijelaskan sebelumnya, contohnya kata-kata seperti “kurikulum” dan “koleksi perpustakaan” akan dikategorikan sebagai “Akademik”, kata-kata seperti “jaringan internet”, “parkir”, dan “kantin” akan dikategorikan sebagai “Fasilitas”, dan seterusnya. Hasil survei yang telah dikategorikan kemudian diproses dan diteruskan ke bagian terkait yang membutuhkan perbaikan atau yang mendapat masukan (Reni, 2019). Adapun bagian terkait yang mendapatkan saran dan masukan mengacu pada angket yang ditanyakan, seperti bagian keamanan kampus, Student Services, Program Studi, Perpustakaan, dan lain-lain.

Proses yang dilakukan dalam pembuatan laporan hasil survei secara manual sangat menguras tenaga dan waktu. Terlebih lagi kesalahan bisa timbul saat proses kategorisasi manual ini yang membuat laporan menjadi kurang akurat. Hal tersebut dibuktikan dari lamanya waktu yang dibutuhkan dalam memproses pembuatan laporan yang dapat mencapai berhari-hari dan jumlah data yang banyak terpaksa harus diperiksa satu per satu dikarenakan data yang diperoleh hanya berbentuk *file* Microsoft Office Excel yang di dalamnya sama sekali belum diolah (Reni, 2019). Melihat dari permasalahan yang ada, aplikasi yang dapat secara otomatis memberikan hasil data yang sudah terkategori tentunya akan sangat membantu proses pembuatan laporan yang lebih akurat, efektif, dan efisien.

Aplikasi yang mampu melakukan kategorisasi secara otomatis terhadap saran dan masukan dari hasil survei tersebut tentunya memerlukan algoritma untuk melakukan klasifikasi terhadap teks. Berdasarkan penelitian yang dilakukan oleh Imron (2019), algoritma Naive Bayes digunakan untuk klasifikasi sentimen dari data teks komentar di situs ulasan dan media sosial terhadap tempat wisata di Kabupaten Rembang dan menghasilkan nilai akurasi di atas 80%. Algoritma ini terbukti mampu untuk melakukan klasifikasi teks dan juga memiliki keuntungan, seperti mampu melakukan klasifikasi teks dengan data latih yang sedikit (Goldwater, 2018).

Hal tersebut dibuktikan dalam penelitian yang dilakukan oleh Fanissa, dkk. (2018), algoritma Naive Bayes mampu menganalisis sentimen pariwisata di Kota Malang dengan data latih sebanyak 200 buah dan menghasilkan akurasi melebihi 85%. Algoritma Naive Bayes juga memiliki performa yang lebih tinggi dibandingkan algoritma lain untuk melakukan klasifikasi teks. Penelitian Wawre dan Deshmukh (2016) memaparkan bahwa hasil akurasi algoritma Naive Bayes menghasilkan akurasi lebih tinggi dari algoritma Support Vector Machine (SVM). Penelitian lain oleh Dey, dkk. (2016) juga menyimpulkan bahwa algoritma Naive Bayes memiliki akurasi lebih tinggi dibandingkan dengan algoritma k-Nearest Neighbors (KNN) pada kasus klasifikasi analisis sentimen.

Selain itu, dijelaskan juga oleh Nugroho (2014) bahwa algoritma Naive Bayes memiliki kelebihan berkaitan dengan akurasi dan waktu komputasi yang lebih hemat. Penelitian yang dilakukan Wu, dkk. (2018) juga menggunakan algoritma Naive Bayes dengan alasan memiliki kesederhanaan dalam komputasinya. Dari beberapa penelitian tersebut dapat disimpulkan bahwa algoritma Naive Bayes

memiliki kelebihan dan performa yang baik dibandingkan algoritma-algoritma klasifikasi lainnya dalam klasifikasi teks.

Dalam implementasinya, tidak jarang algoritma Naive Bayes dipadukan dengan N-gram. Metode N-gram bekerja dengan membagi kata menjadi beberapa urutan berdasarkan nilai N-nya (Jurafsky dan Martin, 2014). N-gram berguna dalam proses kategorisasi, salah satunya untuk melihat kemiripan dari beberapa kata, sehingga kesalahan penulisan yang ada pada data masih dapat diproses dengan baik. Penelitian yang dilakukan oleh Mathew dan Bai (2016) menjelaskan bahwa tingkat efektivitas algoritma Naive Bayes meningkat dengan dipadukan dengan metode N-gram.

Berdasarkan hasil wawancara yang dilakukan dan analisis dari beberapa penelitian sebelumnya, dibutuhkan perancangan dan pembangunan aplikasi web untuk membuat *survey report* kepuasan alumni Universitas Multimedia Nusantara dengan menggunakan algoritma Naive Bayes dan N-gram untuk melakukan kategorisasi saran dan masukan.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang sudah dijelaskan, beberapa rumusan masalah yang akan dibahas adalah sebagai berikut.

1. Bagaimana cara merancang dan membangun aplikasi web *survey report* dengan algoritma Naive Bayes dan N-gram untuk kategorisasi saran?
2. Bagaimana nilai *accuracy*, *precision*, *recall*, dan *f-measure* dari algoritma Naive Bayes dan N-gram dalam mengategorikan saran dari survei kepuasan alumni Universitas Multimedia Nusantara?

1.3 Batasan Masalah

Batasan-batasan masalah yang ada dalam penelitian ini adalah sebagai berikut.

1. Pengguna akhir aplikasi web adalah petugas Alumni Relations dan petugas Career Development Centre Universitas Multimedia Nusantara.
2. Kategori yang digunakan untuk saran dan masukan dibagi menjadi lima buah yang mengacu pada bagian-bagian pertanyaan survei kepuasan alumni Universitas Multimedia Nusantara, yaitu Akademik, Fasilitas, Kemahasiswaan, Lainnya, dan Bukan Saran.
3. Saran dan masukan menggunakan Bahasa Indonesia.
4. Model N-gram yang digunakan bernilai 1 hingga 5 pada level karakter.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang sudah dijelaskan, tujuan dilakukan penelitian ini adalah sebagai berikut.

1. Merancang dan membangun aplikasi web *survey report* dengan algoritma Naive Bayes dan N-gram untuk kategorisasi saran.
2. Mengukur dan mengetahui nilai *accuracy*, *precision*, *recall*, dan *f-measure* dari algoritma Naive Bayes dan N-gram dalam mengategorikan saran dari survei kepuasan alumni Universitas Multimedia Nusantara.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dalam penelitian ini bagi peneliti adalah peneliti dapat memahami bagaimana algoritma Naive Bayes dan N-gram bekerja untuk mengklasifikasikan data yang tidak terstruktur, dalam penelitian ini yaitu teks. Manfaat yang diharapkan untuk pengguna adalah dapat memudahkan petugas

Alumni Relations dan Career Development Centre untuk membuat laporan hasil survei kepuasan alumni Universitas Multimedia Nusantara menjadi lebih efektif dan efisien. Nilai *accuracy*, *precision*, *recall*, dan *f-measure* dari algoritma Naive Bayes dan N-gram dalam mengategorikan saran dari survei kepuasan alumni Universitas Multimedia Nusantara ini juga dapat dimanfaatkan untuk pengembangan penelitian selanjutnya.

1.6 Sistematika Penulisan

Sistematika penulisan yang digunakan untuk penyajian laporan skripsi ini terdiri dari BAB I, BAB II, BAB III, BAB IV, dan BAB V, masing-masing akan diuraikan sebagai berikut.

BAB I PENDAHULUAN

Pada bab ini terdapat enam bagian, yaitu latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan laporan.

BAB II LANDASAN TEORI

Bab ini berisi teori-teori yang digunakan untuk penelitian. Teori-teori yang mendukung penelitian ini adalah klasifikasi teks, algoritma Naive Bayes, *Laplace smoothing*, N-gram, metrik evaluasi, *Technology Acceptance Model*, dan skala Likert.

BAB III METODOLOGI PENELITIAN DAN PERANCANGAN SISTEM

Pada bab ini berisi gambaran beserta penjelasan tahapan metode yang dilakukan dan rancangan aplikasi yang dibuat dengan bantuan diagram, tabel, dan gambar.

BAB IV IMPLEMENTASI DAN ANALISIS

Pada bab ini berisi pembahasan mengenai implementasi dan analisis dari algoritma Naive Bayes dan N-gram untuk aplikasi yang dibuat. Bab ini juga memberikan nilai *accuracy*, *precision*, *recall*, dan *f-measure* dari algoritma yang digunakan, serta hasil uji penerimaan aplikasi.

BAB V SIMPULAN DAN SARAN

Bab ini berisi simpulan dari hasil penelitian dan analisis yang dilakukan dalam penelitian, serta saran yang dapat digunakan untuk pengembangan aplikasi maupun penelitian lebih lanjut.