

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Universitas adalah institusi pendidikan tinggi yang menyediakan program-program studi sarjana (S1), magister (S2), dan doktor (S3) dalam berbagai bidang ilmu seperti teknologi, sains, seni, sosial, ekonomi, hukum, kedokteran, dan lain-lain. Salah satu Universitas yang ada di Indonesia adalah Universitas Komputer Indonesia (UNIKOM). UNIKOM adalah perguruan tinggi swasta yang fokus pada pendidikan ilmu komputer dan teknologi informasi, banyak orang membicarakan universitas tersebut di media sosial, salah satunya di platform *twitter*.

Twitter digunakan tidak hanya untuk bersosialisasi dan berkomunikasi namun juga digunakan untuk menyampaikan aspirasi dan mempresentasikan hal-hal yang sedang terjadi pada masyarakat [1]. Isi dari *tweet* bervariasi, bisa berupa hal yang positif dan juga negatif. Apabila dianalisis dan diolah dengan tepat, *tweet* memiliki potensi manfaat yang besar. Salah satu cara yang dapat digunakan untuk menganalisis dan mengolah teks *tweet* tersebut adalah analisis sentimen.

Analisis sentimen adalah proses memahami dan mengolah data tekstual secara otomatis untuk mendapatkan informasi. Analisis sentimen dilakukan untuk mendeteksi opini terhadap suatu subjek dan objek (misalnya individu, organisasi dll) dalam sebuah kumpulan data[1]. Analisis sentimen aspek adalah variasi dari analisis sentimen yang mengidentifikasi dan mengevaluasi opini atau perasaan yang terkait dengan aspek tertentu dari teks tersebut[2]. *Twitter* dapat digunakan untuk mengumpulkan data dari aspek tertentu mengenai UNIKOM, seperti kepuasan mahasiswa dengan akademik perguruan tinggi dan, aspek sosial, serta fasilitas kampus pada mahasiswa[3] dan juga kualitas pelayanan yang memiliki hubungan yang erat dengan kepuasan mahasiswa[4]. Aspek pada suatu universitas bermanfaat bagi manajemen UNIKOM untuk mengidentifikasi masalah dan mengambil tindakan untuk meningkatkan kualitas universitas. Saat ini, belum

terdapat implementasi yang memanfaatkan data dari media sosial *twitter* untuk menganalisis berbagai aspek di suatu universitas. Maka dari itu peneliti akan mencoba melakukan analisis sentimen berbasis aspek mengenai topik pada UNIKOM.

Tweet pada platform *twitter* memiliki karakteristik tersendiri. Dengan batas karakter yang ada, pengguna *twitter* cenderung menggunakan bahasa yang ringkas, singkat, dan padat informasi. Selain itu, slang, emoji, dan hashtag menjadi ciri khas komunikasi di platform ini. Karena keunikannya, variasi topik yang dibahas dapat membuat beberapa data tampak berbeda secara signifikan atau menjadi *outlier*. Oleh karena itu, diperlukan pendekatan khusus dalam melakukan analisis pada data *Twitter*. Untuk melakukan analisis sentimen berbasis aspek perlu menggunakan algoritma klasifikasi. Salah satu algoritma yang cukup terkenal digunakan dalam klasifikasi adalah Algoritma *Modified K-Nearest Neighbors (MKNN)*. Algoritma *MKNN* merupakan Algoritma penyempurnaan dari *K-Nearest Neighbors (KNN)*. *KNN* adalah metode klasifikasi yang memanfaatkan jarak atau kemiripan antara suatu data dengan data lainnya. Dalam melakukan klasifikasi, *KNN* mengidentifikasi sejumlah *k* tetangga terdekat dari suatu data dan menentukan jenis kelas data tersebut berdasarkan mayoritas kelas tetangga terdekat tersebut. *KNN* termasuk algoritma yang relatif mudah dibandingkan dengan algoritma lain karena tidak membangun model pada saat pembelajaran mesin, melainkan hanya mengandalkan memori [5]. Adapun kelebihan *MKNN* dari *KNN* adalah tingkat akurasi yang lebih tinggi, dimana dalam *MKNN* ditambahkan perhitungan nilai validitas yang berguna dalam mengatasi masalah *outlier* dalam perhitungan bobot pada *KNN* Tradisional [6].

Pada penelitian Okfalisa, disebutkan bahwa hasil pengujian menghasilkan bahwa metode *MKNN* lebih baik dari metode *KNN*, disebutkan bahwa hasil pengujian menghasilkan akurasi *KNN* tertinggi sebesar 94,95% dengan rata-rata akurasi saat tes sebesar 93,94% dan akurasi tertinggi *MKNN* sebesar 99,51% dengan rata-rata akurasi selama tes sebesar 99,20% [7]. Pada penelitian lainnya, Analisis Sentimen Terhadap Ulasan Pengguna *MRT* Jakarta Menggunakan

Modified K-Nearest Neighbor, metode ini menghasilkan akurasi 86% dengan parameter $k=3$ [8]. Hasil pengujian ini sudah cukup baik, namun belum diketahui bagaimana hasil penerapan *MKNN* untuk kasus analisis sentimen berbasis aspek.

Berdasarkan permasalahan yang telah dijelaskan, maka dari itu dalam penelitian ini akan mencoba melakukan analisis sentimen berbasis aspek dengan topik Universitas Komputer Indonesia pada media sosial *twitter* dengan menerapkan metode *MKNN*.

1.2 Rumusan Masalah

Berdasarkan penjelasan dari latar belakang di atas, maka didapatkan rumusan masalah yaitu bagaimana melakukan analisis sentimen berbasis aspek terhadap *tweet* mengenai Universitas Komputer Indonesia dengan menerapkan metode *Modified K-Nearest Neighbors (MKNN)*?

1.3 Maksud Penelitian

Maksud dari penelitian ini adalah mengklasifikasi aspek dan sentimen yang ada pada *tweet* mengenai Universitas Komputer Indonesia menggunakan metode *Modified K-Nearest Neighbors (MKNN)*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengukur akurasi metode *Modified K-Nearest Neighbors (MKNN)* dalam melakukan analisis sentimen berbasis aspek terhadap *tweet* mengenai Universitas Komputer Indonesia berdasarkan aspek – aspek yang terkait.

1.5 Batasan Masalah

Batasan masalah penelitian diperlukan agar ranah penelitian tidak terlalu meluas dan lebih terarah, adapun batasan masalahnya di antaranya:

1. Data yang digunakan hanya mengambil *tweet* dengan kata pencarian “unikom” dan hastag “#unikom” pada rentang waktu 1 Januari 2020 – 8 mei 2023 dari paltform *twitter*.
2. Menggunakan *tweet* dengan Bahasa Indonesia.
3. Proses pengumpulan data diperoleh dengan teknik *scraping* menggunakan *library snsscrape*.
4. Aspek yang digunakan meliputi aspek Fasilitas, Akademik, Pelayanan, Kegiatan, dan Lingkungan Kampus.
5. Sentimen terbagi atas Positif dan Negatif.

1.6 Metode Penelitian

Metode penelitian adalah pendekatan sistematis yang digunakan untuk merencanakan, melaksanakan, dan menganalisis suatu penelitian. Metode Penelitian yang digunakan pada penelitian ini adalah dengan metode penelitian deskriptif. Beberapa langkah yang dilakukan selama penelitian ini diantaranya sebagai berikut:



Gambar 1. 1 Metode Penelitian

1.6.1 Literature Review

Literature review (tinjauan literatur) adalah proses sistematis dalam penelitian yang melibatkan pengumpulan, penelaahan, dan sintesis literatur atau karya-karya tulis yang relevan dalam suatu bidang tertentu. Tujuan dari *literature review* adalah untuk menyajikan pemahaman yang komprehensif tentang penelitian yang telah dilakukan sebelumnya, pemikiran dan teori yang telah ada, dan perkembangan pengetahuan di bidang yang diteliti. Pada tahap ini dilakukan pencarian studi *literature review* melalui berbagai sumber seperti *Google scholar*, *Academia*, dan referensi file skripsi di Perpustakaan.

1.6.2 Pengumpulan dan Pengolahan Data

Pada tahap ini pengumpulan data dilakukan dengan teknik *web scraping* di *twitter* dengan menggunakan *library snsrape*. Ketika data sudah didapatkan maka dilakukan pelabelan secara manual, data *tweet* di labeli menurut aspek dan sentimen nya masing-masing.

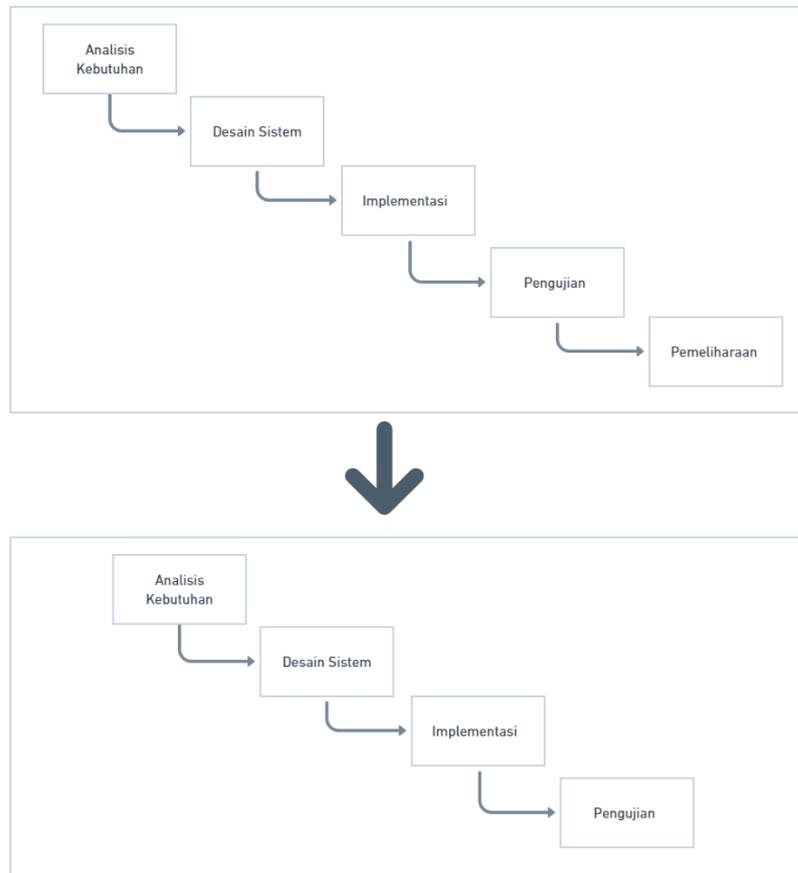
1.6.3 Analisis

Tahap ketiga yang dilakukan adalah analisis, langkah ini digunakan mulai dari analisis metode untuk *preprocessing*, ketidakseimbangan data, pembobotan kata *tf-idf*, dan tahap klasifikasi aspek dan sentimen menggunakan metode *MKNN*.

1.6.4 Pembangunan Sistem

Tahap keempat yang dilakukan adalah pembangunan sistem, yaitu proses pengembangan perangkat lunak atau aplikasi yang meliputi tahap analisis, desain, pengembangan, pengujian, implementasi, dan pemeliharaan. Dalam tahap pembangunan ini, pengembangan perangkat lunak akan dilakukan menggunakan metode *waterfall* secara berurutan.

Namun, terdapat satu pengurangan pada tahapan proses, yaitu tahapan *Maintenance* tidak akan dilakukan. Proses ini hanya akan mencakup sampai pengujian saja. Berikut gambaran dari metode *waterfall* yang akan diterapkan [9]:



Gambar 1. 2 Metode *Waterfall*

Tiap langkah harus dilakukan secara berurutan, berikut penjelasan dari tahapan pada metode *waterfall* yang akan dilakukan:

1. Analisis kebutuhan, tahap ini bertujuan untuk mengumpulkan informasi tentang kebutuhan dan menentukan fitur-fitur yang harus dimiliki oleh sistem. Ada 2 Fitur yang ada dalam sistem yang akan dibangun yaitu fitur pelatihan yang digunakan untuk

melatih data, serta fitur pengujian yang digunakan untuk menganalisis teks yang akan dimasukan user berdasarkan data yang sudah dilatih sebelumnya.

2. Desain sistem, tahap ini bertujuan untuk merancang sistem secara keseluruhan seperti struktur menu dan *wireframe* dari sistem yang akan dibangun.
3. Implementasi, tahap ini merupakan tahap pengembangan sistem, dimana program-program yang telah dirancang pada tahap sebelumnya diimplementasikan menjadi sebuah sistem yang lengkap. Sistem yang akan dibangun adalah berbasis *website* dengan menggunakan bahasa pemrograman *python*.
4. Pengujian, tahap ini bertujuan untuk menguji kehandalan dan kinerja sistem. Pengujian meliputi pengujian fungsional sistem bekerja dengan *blackbox testing* dan pengujian akurasi.

1.6.5 Kesimpulan

Berdasarkan penelitian yang telah dilakukan mulai dari awal hingga pengujian maka dapat ditarik kesimpulan mengenai sistem yang dibangun serta saran untuk penelitian selanjutnya.

1.7 Sistematika Penulisan

BAB 1 PENDAHULUAN

Bab ini menguraikan berbagai hal yang berkaitan dengan permasalahan dimana pada penelitian ini, menjelaskan latar belakang penelitian, rumusan masalah dari penelitian, tujuan dan manfaat penelitian, batasan – batasan penelitian serta sistematika penulisan dari tulisan penelitian ini.

BAB 2 LANDASAN TEORI

Bab ini menguraikan tinjauan umum dilakukannya penelitian. Mulai dari landasan teori yang berkaitan dengan topik penelitian berupa bahan kajian,

konsep dasar, dan teori dari para ahli, serta tinjauan terhadap penelitian sebelumnya.

BAB 3 ANALISIS DAN PERANCANGAN

Bab ini menguraikan mengenai analisis serta perancangan sistem yang akan dibangun, mulai dari awal mula hingga hasil akhir sistem untuk menghasilkan sebuah keputusan untuk mengklasifikasikan aspek dan sentimen dari data yang ada. Mencakup mulai dari data mentah, tahapan – tahapan *preprocessing* (*cleaning, case folding, tokenizing, normalization, stemming, dan stopword removal*), tahapan pembobotan kata menggunakan *tf-idf*, dan pembahasan yang mendalam mengenai algoritma yang digunakan beserta perhitungannya sehingga memunculkan hasil akhir.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini akan menguraikan tentang implementasi dari sistem yang telah dibangun. Selain itu, bab ini akan melakukan perhitungan akurasi dari hasil pengujian yang telah dilakukan.

BAB 5 KESIMPULAN DAN SARAN

Bab ini merupakan penutup dari sistematika penulisan penelitian, yang berisi ringkasan kesimpulan yang dapat diambil dari seluruh penelitian, serta saran-saran yang diperoleh setelah penelitian dilakukan. Diharapkan bahwa saran-saran ini dapat memberikan masukan yang bermanfaat dan membangun.