1. PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi informasi pada saat ini meningkat dengan cepat. Jumlah data terus bertambah karena semakin banyak orang yang memanfaatkan kemudahan teknologi sebagai tempat untuk mencari infomasi, salah satunya adalah berita. Berita merupakan informasi mengenai suatu peristiwa terbaru dan penting mengenai fakta yang sedang terjadi untuk disampaikan kepada masyarakat [1].

Pembaruan artikel berita yang diproduksi setiap harinya menyebabkan aliran berita terus bertambah, sehingga pengklasifikasian artikel berita diperlukan agar tidak menyebabkan kelebihan informasi [2]. Klasifikasi teks dapat digunakan untuk berbagai hal, salah satunya adalah klasifikasi topik berita [3]. Klasifikasi teks merupakan bagian ilmu dari *text mining* yang dapat memberikan solusi dalam klasifikasi berita. Klasifikasi merupakan metode pengelompokan data yang dimiliki oleh *data mining* dan *text mining*, istilah ini merujuk kepada sebuah metode untuk menyusun data secara sistematis.

Namun, dalam pengklasifikasian berita untuk saat ini masih dilakukan secara manual, dimana dalam mengunggah berita, pengunggah harus terlebih dahulu mengetahui isi dari berita secara keseluruhan untuk selanjutnya dimasukkan ke dalam kategori yang tepat [4]. Sehingga diperlukan pengklasifikasian berita secara otomatis. Dalam proses klasifikasi, metode menjadi suatu hal yang berpengaruh dalam kinerja klasifikasi. Ada banyak metode yang dapat digunakan dalam pengklasifikasian topik berita, namun dalam tugas akhir ini diharapkan sebuah metode pengklasifikasian yang sederhana namun efektif agar pelabelan dapat dilakukan dengan mudah dan cepat serta memiliki akurasi yang baik. Salah satu metode dalam proses klasfikasi adalah *k-Nearest Neighbor* (k-NN)

K-Nearest Neighbor (k-NN) merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut, metode k-NN dapat diaplikasikan pada pengklasifikasian topik berita berbentuk teks [5]. Metode tersebut dipilih karena k-NN merupakan salah satu algoritma *machine learning* yang bersifat nonparametrik,

garis keputusan kelas nonparametik menjadi sangat fleksibel dan nonlinear dibandingkan dengan metode *logistic regression* [6].

Namun pada k-NN biasa, bobot setiap data pada k data terdekat besarnya sama, yaitu 1/k. Jika bobotnya tidak diatur (tidak 1/k) maka *classifier* dapat memiliki performansi yang lebih bagus [7]. Maka dari itu pada tugas akhir ini metode yang akan digunakan adalah pengembangan dari metode k-NN yang sudah ada sebelumnya, yaitu *weighted k-Nearest Neighbor* yang diharapkan dapat memberikan performasi yang lebih baik daripada metode k-NN biasa dalam klasifikasi topik berita.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, rumusan masalah penelitian ini adalah sebagai berikut:

1. Berapa parameter nilai k pada *weighted* k-NN yang tepat untuk klasifikasi topik berita?
2. Bagaimana perfromasi metode *weighted* k-NN dalam klasifikasi topik berita?

1.3. Tujuan

Sesuai dengan latar belakang dan rumusan masalah yang telah dijelaskan, tujuan dari penelitian ini adalah sebagai berikut:

1. Mengetahui parameter nilai k pada *weighted* k-NN yang tepat untuk klasifikasi topik berita.
2. Mengetahui performansi metode *weighted* k-NN untuk klasifikasi topik berita.

1.4. Batasan Masalah

Adapun batasan masalah pada penelitian ini yaitu:

1. Kategori berita yang digunakan dalam pembuatan tugas akhir ini adalah *World, Business, Sports,* dan *SciTech.*
2. Dataset yang digunakan diambil dari situs *Dataturks* [8].

1.5. Rencana Kegiatan

Berikut merupakan rencana kegiatan yang akan dilakukan:

1. Studi Literatur

   Dalam tahapan studi literatur akan dilakukan pemahaman permasalahan dan pencarian informasi yang berkaitan dengan permasalahan yang dapat menjadi acuan dalam melakukan penelitian untuk penyelesaian.

2. Analisis sistem

   Dalam analisis sistem akan dilakukan analisis dari permasalahan dan mencari metode-metode yang akan digunakan.

3. Perancangan sistem

   Dalam tahap perancangan sistem dilakukan dengan mendesain sistem sesuai dengan kebutuhan berdasarkan permasalahan yang telah diidentifikasi.

4. Implementasi sistem

   Dalam tahapan implementasi dilakukan proses pembangunan sistem berdasarkan rancangan yang telah dilakukan pada tahapan selanjutnya.

5. Penyusunan laporan

   Dalam tahapan pembuatan laporan akhir berisikan hasil analisis sistem yang telah dibangun berupa hasil klasifikasi dari nilai akurasi dari sistem. Hasil analisis sistem akan menjadi kesimpulan yang diharapkan menjadi solusi dari permasalahan.

1.6. Jadwal Kegiatan

Berikut jadwal pelaksanaan kegiatan berdasarkan rencana kegiatan:

Tabel 1.1 Rencana Kegiatan

Kegiatan	**2019**					
Studi Literatur						
Analisis Sistem						
Perancangan Sistem						
Implementasi Sistem						
Penyusunan Laporan						