

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Dalam perkembangan teknologi di masa sekarang dapat mempermudah dalam memperoleh informasi. Mudah-mudahan perolehan informasi diakibatkan oleh penggunaan internet dan kecepatan akses internet yang semakin cepat. Banyak informasi yang bisa didapatkan ketika menggunakan internet salah satunya adalah informasi tentang suatu kejadian atau berita. Saat ini sudah banyak situs yang menyediakan informasi berita yaitu Kompas.com, detik.com, Liputan6.com, dan masih banyak lagi. Setiap suatu artikel berita memiliki topik utama yang dibahas dalam berita tersebut, contohnya berita dengan topik olahraga, fashion, politik, dan yang lainnya. Banyaknya suatu artikel berita dan juga topik yang dibahas didalamnya terkadang mengakibatkan sulitnya untuk mendapatkan suatu berita dengan topik pembahasan yang diinginkan. Maka dari itu dibutuhkanlah suatu sistem untuk mengklasifikasikan berita tersebut sesuai dengan topik yang dibahas didalamnya.

Ada beberapa penelitian yang sudah membahas mengenai klasifikasi dokumen berita, seperti pada penelitian Andi Ahmad Irfa dan kawan-kawan [1] yang menggunakan metode *k-Nearest neighbor* dalam melakukan klasifikasi dokumen artikel berita [1]. Pada penelitian oleh Andi Ahmad Irfa dan teman-temannya, didapat hasil pada penelitian yang dilakukan menghasilkan nilai *f-measure* sebesar 69,9% [1]. Penelitian mengenai klasifikasi dokumen menggunakan metode *k-Nearest neighbor* juga pernah dilakukan oleh Claudio Fresta Suharno dan teman-temannya [2]. Pada penelitian yang dilakukan Claudio dijelaskan perbandingan klasifikasi dokumen menggunakan *k-Nearest neighbor* dengan seleksi fitur dan tanpa seleksi fitur [2]. Metode seleksi fitur yang digunakan oleh Claudio adalah metode seleksi fitur *Chi-square*, pada penelitian yang dilakukan dijelaskan bahwa penggunaan seleksi fitur dalam melakukan klasifikasi dokumen dengan menggunakan metode *k-Nearest neighbor* mendapat hasil lebih baik dibandingkan dengan tanpa menggunakan seleksi fitur. [2]

Seleksi fitur merupakan tahap untuk memilih fitur yang paling penting dalam suatu data atau dokumen [2]. Tujuan dari seleksi fitur adalah untuk meningkatkan performa dari klasifikasi dokumen dengan menghilangkan fitur yang dianggap tidak relevan dalam klasifikasi untuk mengurangi dimensi dari himpunan fitur [2]. Seleksi fitur merupakan tahapan yang sangat penting dalam mengoptimalkan kinerja dari metode klasifikasi [3].

Particle swarm optimization merupakan salah satu algoritma evolusioner yang mirip dengan algoritma genetika dan menggunakan fungsi *fitness* untuk mengevaluasi kualitas dari solusi [4]. *Particle swarm optimization* banyak digunakan untuk memecahkan serta sebagai masalah seleksi fitur [5].

Mehdi Hosseinzadeh Aghdam dan Setareh Heidari pernah melakukan penelitian mengenai penggunaan metode *Particle Swarm Optimization* sebagai seleksi fitur pada *Text Categorization* [6]. Pada penelitian yang dilakukan oleh Mehdi dijelaskan seleksi fitur dengan menggunakan metode *Particle swarm optimization* mendapatkan hasil yang cukup baik dengan menghasilkan nilai *precision* sebesar 77.4619. Pada Penelitian ini Mehdi juga melakukan perbandingan hasil seleksi fitur menggunakan *Particle Swarm Optimization* dengan beberapa metode lain yaitu *Information Gain*, *Chi-Square*, dan *Genetic Algorithm*. Hasil nilai *precision* pada penelitian Mehdi berturut-turut adalah 70,3791, 72,204, dan 76,6511. Berdasarkan hal itu dalam penelitian ini akan digunakan metode *K-Nearest Neighbor* berbasis *Particle swarm optimization* untuk klasifikasi topik berita serta mengetahui nilai akurasi pada metode yang digunakan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka dapat dirumuskan masalah yang timbul adalah bagaimana tingkat Akurasi klasifikasi berita dengan menggunakan metode *K-Nearest neighbor* berbasis *Particle swarm optimization*

1.3 Maksud dan Tujuan

Berdasarkan permasalahan yang ada, maka maksud dari penelitian yang dilakukan adalah menerapkan metode *K-Nearest neighbor* berbasis *Particle swarm*

optimization untuk mengklasifikasikan topik Artikel Berita. Tujuan yang ingin dicapai adalah mengukur tingkat akurasi dari metode *K-Nearest neighbor* berbasis *Particle swarm optimization* untuk mengklasifikasikan topik Artikel Berita

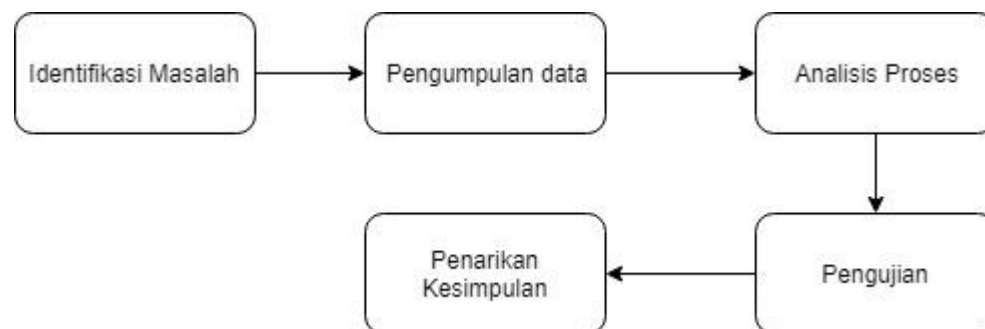
1.4 Batasan Masalah

Agar penelitian yang dilakukan lebih terarah, maka diperlukan suatu pembatasan masalah atau ruang lingkup kajian sebagai berikut:

1. Data yang digunakan hanya berupa artikel berita berbahasa Indonesia
2. Data diambil dari portal berita online
3. Data diklasifikasikan ke dalam 5 Kategori yaitu travel, olahraga, kesehatan, otomotif dan teknologi Informasi.
4. Banyak Data yang digunakan adalah 250 data yang akan dibagi menjadi data latih dan data uji

1.5 Metodologi Penelitian

Tahapan yang dilakukan dalam penelitian ini terdapat pada Gambar 1.1 sebagai berikut:



Gambar 1.1 Gambar Tahapan Penelitian

Berikut ini merupakan penjelasan dari setiap tahapan yang dilakukan pada penelitian ini :

1. Identifikasi Masalah

Pada tahap identifikasi masalah dilakukan analisis mengenai masalah yang dihadapi pada penelitian yang dilakukan.

2. Pengumpulan data

Pada tahapan pengumpulan data yaitu mengumpulkan data literatur dan data Artikel Berita. Literatur yang dibutuhkan adalah mengenai metode K-Nearest Neighbor dan Particle Swarm Optimization. Sedangkan data Artikel Berita diambil dari berbagai situs seperti <https://kompas.com>, <https://detik.com/>, <http://liputan6.com> dan situs lainnya.

3. Analisis Proses

Pada tahap analisis proses dilakukan analisis mengenai proses klasifikasi topik. Analisis yang dilakukan meliputi analisis masalah, analisis preprocessing, analisis *particle swarm optimization* dan analisis metode K-Nearest Neighbor.

4. Pengujian

Pada tahap ini dilakukan pengujian terhadap hasil klasifikasi dengan menggunakan *confusion matrix*.

5. Penarikan Kesimpulan

Pada tahapan ini yaitu menentukan kesimpulan berdasarkan hasil pengujian yang telah dilakukan.

1.6 Sistematika Penulisan

BAB 1 PENDAHULUAN

Bab 1 menguraikan tentang latar belakang permasalahan, perumusan masalah, maksud dan tujuan, batasan masalah, metodologi penelitian, dan sistematika penulisan.

BAB 2 TINJAUAN PUSTAKA

Dalam bab ini penulis menjelaskan tentang landasan teori penunjang yang dapat mendukung pokok pembahasan dari pembuatan simulasi untuk klasifikasi topik berita menggunakan metode *k-Nearest neighbor* dan menggabungkannya dengan *Particle swarm optimization*.

BAB 3 ANALISIS DAN PERANCANGAN

Dalam bab ini penulis akan menjelaskan tentang analisis serta perancangan. Analisis yang akan dibahas berupa analisis sistem, analisis algoritma, analisis kebutuhan fungsional, analisis kebutuhan non-fungsional. Pada perancangan yang akan dibahas adalah perancangan antarmuka.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Dalam bab ini penulis akan membahas tentang implementasi dan pengujian dari perangkat lunak yang dibangun sesuai dengan analisis dan perancangan yang telah dibuat. Implementasi yang akan dibahas adalah implementasi antarmuka. Pada bagian pengujian yang akan dibahas adalah pengujian sistem yang telah dibuat.

BAB 5 KESIMPULAN DAN SARAN

Dalam bab ini penulis akan menyimpulkan hasil dan nilai akurasi yang didapat serta memberikan saran kepada penulis selanjutnya yang ingin melakukan pengembangan terhadap metode yang digunakan dalam klasifikasi topik berita.

