# **BAB 1**

# **PENDAHULUAN**

## **Latar Belakang**

Pemilu terjadi setiap 5 tahun sekali dengan jumlah pemilih yang semakin bertambah setiap periodenya, tercatat penambahan pemilih tetap sejak tahun 2004 sampai 2019 sebanyak 28,90% [1]. Media sosial seperti Facebook, Twitter, LinkedIn, YouTube sudah menjadi bagian dari kehidupan manusia, dibuktikan dengan cara berinteraksi kebanyakan dapat dilakukan dengan sosia media, sampai menjadi identitas diri ke masyarakat umum [2].

Pada sosial media terutama *Twitter*, banyak penyebaran informasi, hiburan, bahkan opini politik, hal ini tidak luput dari terjadinya pemilu yang merupakan kegiatan politik yang akan dilaksanakan pada tahun 2024, disebut sebagai pesta demokrasi, hal tersebut pastinya ramai diperbincangkan di sosial media. Salah satu cara menyampaikan opini dan lewat media sosial adalah dengan menggunakan tanda pagar tertentu pada *tweet* yang dipost di *Twitter*. *Twitter* sendiri digunakan karena kemudahan dalam membalas *tweet* dengan *like*, komentar bahkan *reply*, sehingga topik yang sedang ramai diperbincangkan dapat dengan mudah tersebar [3].

Oleh karena itu, maka diperlukannya analisa lebih lanjut untuk melihat seberapa besar kesetujuan dan ketidaksetujuan pengguna *Twitter* terhadap pemilihan calon presiden 2024, lalu data akan diperoleh dari *tweet* pengguna *Twitter* yang berkaitan dengan topik ini yaitu pemilihan umum 2024. *Sentiment Analysis* atau dapat disebut *Opinion Mining* merupakan studi yang mempelajari opini, sentimen, evaluasi, tingkah laku dan emosi terhadap suatu entitas [2] [3]. *Sentiment Analysis* ini digunakan untuk mengetahui *tweet* yang diposting mengandung kata positif, negatif atau netral [4].

Pada penelitian ini, akan menggunakan metode Support Vector Machine (SVM) yaitu salah satu metode klasifikasi pada *machine learning* (*supervised learning*) yang memprediksi kelas berdasarkan pola dari hasil training. Klasifikasi dilakukan dengan membuat garis pemisah (*hyperlane*) antara kelas positif dan negatif [2]. Sehingga hasil dari penelitian ini adalah sentiment pengguna *Twitter* terhadap pemilihan calon presiden pada tahun 2024.

## **Rumusan Masalah**

Berdasarkan latar belakang yang telah disebutkan, maka penulis mendapati rumusan masalah:

1. Bagaimana penerapan *Sentiment Analysis* menggunakan algoritma *Support Vector Machine* (SVM) diimplementasikan pada topik pemilihan calon presiden 2024?
2. Bagaimana akurasi *Sentiment Analysis* untuk tweet pemilihan calon presiden 2024 dengan menggunakan algoritma *10-Fold Cross Validation* dan *Confusion Matrix*?
3. Apakah ada keterkaitan antara hasil analisa dengan survey elektabilitas portal berita?

## **Tujuan dan Manfaat**

### **Tujuan Penelitian**

Berdasarkan masalah yang dinyatakan diatas, maka tujuan dari penelitian ini adalah:

1. Mengembangkan aplikasi web sederhana sebagai bentuk implementasi *Sentiment Analysis* menggunakan algoritma *Support Vector Machine* (SVM) dengan topik pemilu 2024
2. Menguji akurasi *Sentiment Analysis* dengan algoritma *Support Vector Machine* (SVM) untuk *tweet* pemilu 2024
3. Melakukan perbandingan hasil prediksi dengan website berita terpercaya seperti detik.com

### **Manfaat Penelitian**

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Manfaat bagi penulis:
   1. Hasil penelitian dapat memberikan manfaat bagi penulis tentang penggunaan *Text Mining* dalam memprediksi hasil pemilu 2024 berdasarkan data *Twitter*
   2. Mengimplementasi ilmu dan teori yang penulis dapatkan selama perkuliahan dalam bidang *Machine Learning*
2. Manfaat bagi akademis:
   1. Penulis berharap dengan riset yang dilakukan dapat menjadi referensi bacaan dengan topik *Sentiment Analysis* dan *Text Mining*
3. Manfaat bagi masyarakat:
   1. Sebagai prediksi pemilihan calon presiden 2024 berdasarkan data Twitter
   2. Memberikan referensi dan bacaan bagi masyarakat yang berkaitan dengan tema politik

## **Ruang Lingkup**

Pada penelitian ini, ruang lingkup yang dibatasi oleh penulis adalah sebagai berikut:

1. Perancangan aplikasi ini dijalankan berbasis website
2. Penerapan *Sentiment Analysis* menggunakan algoritma *Support Vector Machine* (SVM)
3. Tema pemilu yang diambil hanya seputar pemilihan calon presiden, bukan calon lainnya.
4. Pengujian algoritma akan dilakukan dengan metode *10-fold* *cross validation*
5. Bahasa pemrogaman yang digunakan dalam perancangan aplikasi ini adalah bahasa pemrograman python.
6. Sentimen yang diteliti penulis dibagi menjadi 3 yaitu positif, netral, dan negatif
7. Data diambil dari *Twitter* dengan batas 1 Januari 2023 sampai dengan 31 Maret 2023

## **Metodologi Penelitian**

Metode yang akan dipakai pada penelitian ini adalah algoritma *Support Vector Machine* (SVM) yang akan diterapkan pada data *tweet* dari media sosial *Twitter* dengan tema pemilihan capres 2024, lalu data tersebut akan dilakukan beberapa proses diantaranya:

1. *Data Labeling* melakukan pelabelan setiap data menjadi 3 jenis, positif, netral, dan negatif.
2. *Preprocessing* yaitu pembersihan data dari noise sehingga siap digunakan pada tahap *Transformation*
3. *Transformation* menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) sehingga data diubah dari kalimat atau kata menjadi sebuah nilai
4. *Klasifikasi* dilakukan pemrosesan data menggunakan algoritma *Support Vector Machine* (SVM).
5. *Evaluasi* akan dilakukan untuk mengetahui nilai akurasi, presisi, *recall*, dan F1-*score* menggunakan *10*-*Fold Cross Validation* dan *Confusion Matrix* dengan membagi dataset kedalam data *training* dan data *testing* secara bertahap. Lalu dilakukan validasi dengan dibandingkan dengan label data sebenarnya.

## **Sistematika Penulisan**

Sistematika penulisan yang akan digunakan adalah sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini menjelaskan tentang latar belakang masalah yang diangkat, rumusan masalah yang didapatkan, batasan masalah yang dibuat, manfaat dan tujuan penelitian bagi berbagai pihak serta sistematika penulisan yang akan digunakan di penelitian ini.

BAB 2 LANDASAN TEORI

Bab ini menjelaskan tentang teori yang berkaitan tentang S*entiment Analysis* dan algoritma *Support Vector Machine* yang diambil dari berbagai sumber sebagai landasan dalam pembuatan penelitian.

BAB 3 ANALISIS DAN PERANCANGAN

Pada bagian ini berisi tahapan rancangan aplikasi *Sentiment Analysis* dengan algoritma *Support Vector Machine* berbasis website sederhana.

BAB 4 IMPLEMENTASI DAN EVALUASI

Bab ini berisikan hasil pengembangan yang diimplementasi dalam website dan uji akurasi yang diukur dengan metode 10-*Fold Cross Validation*

BAB 5 SIMPULAN DAN SARAN

Bab ini berisikan pemaparan kesimpulan dari hasil uji coba dan pendapat penulis tentang riset yang telah dilakukan.