

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Media sosial Instagram merupakan salah satu media sosial yang populer dan sangat diminati oleh masyarakat global terutama di kalangan dewasa muda, tercatat oleh Hootsuite (We are Social), pada tahun 2021 pengguna instagram di indonesia mencapai 85 juta pengguna dan menjadi media sosial populer di urutan ke tiga, jumlah pengguna instagram ini terjadi peningkatan sebanyak 22 juta dari tahun sebelumnya yaitu 63 juta [1]. Dengan meningkat nya jumlah pengguna tersebut maka meningkat pula jumlah unggahan setiap harinya.

Didalam Instagram ini berisikan kumpulan foto dan caption berupa text. Dimana caption ini digunakan untuk mengungkapkan sebuah kejadian, ungkapan diri mereka sendiri dan hal yang sedang dirasakannya atau menggambarkan tentang diri mereka sendiri, maka dari data yang ada dapat dimanfaatkan untuk bidang psikologi untuk menganalisis kepribadian seseorang secara digital. Dan untuk menganalisis kepribadian ini dapat memanfaatkan sosial media dimana memiliki banyak data dan memfokuskan dari masing-masing individunya. Data media sosial di Instagram berupa caption-caption berupa text digunakan sebagai sumber data penelitian untuk menganalisis profil kepribadian seseorang[2].

Dengan adanya data berupa text caption maka dapat digunakan sebagai data untuk menganalisa dari kepribadian seseorang dengan penerapan ilmu data mining yang dinilai dalam kepribadian DISC. model DISC dibagi menjadi empat yaitu : Dominance (Dominasi), influence (Pengaruh), Steadiness (Stabilitas), Compliance (Patuh).

Penelitian tentang analisis kepribadian DISC pernah dilakukan oleh peneliti sebelumnya dengan menggunakan metode multinomial naïve bayes classifier didapatkan akurasi sebesar 48 %. Maka dari itu dalam hal ini Penelitian ini didasari atas saran dari peneliti sebelumnya dimana untuk mendapatkan suatu hasil akurasi yang lebih tinggi perlu dilakukan uji dengan metode yang berbeda [3].

Pada Penelitian ini akan melakukan uji performa dari metode Support Vector machine dengan membandingkan performa dari beberapa metode lain, yaitu Multinomial Naïve Bayes (MNB), K-Nearest Neighbor (KNN), Decision Tree. Kemudian dalam melakukan Klasifikasi untuk menganalisis kepribadian DISC akan menggunakan algoritma Support Vector Machine dimana algoritma ini memiliki kemampuan untuk menggeneralisasi permasalahan dengan baik sehingga menghasilkan akurasi yang tinggi dan tingkat kesalahan yang lebih kecil. Prinsip dasar yang dimiliki oleh metode Support Vector Machine (SVM) adalah pengklasifikasi linier, dimana SVM dapat mengklasifikasikan data yang dapat dipisahkan secara linier, tetapi SVM telah dikembangkan untuk dapat mengatasi masalah non-linier dengan menerapkan konsep trik kernel di ruang kerja. Tujuan dari SVM adalah untuk mendapatkan hyperplane terbaik yang memisahkan dua kelas dalam SVM linier dan beberapa kelas dalam SVM non-linier [4]

Berdasarkan uraian, penelitian ini bertujuan untuk membangun sebuah data mining dengan algoritma Support Vector Machine yang dapat menganalisa kepribadian DISC pengguna pada instagram serta mengukur kinerja akurasi yang didapatkan dari algoritma Support Vector Machine.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka permasalahan yang dapat dirumuskan adalah sebagai berikut :

1. Apakah Metode Support Vector Machine dapat diimplementasikan untuk menganalisa Kepribadian DISC pengguna instagram berdasarkan caption bahasa indonesia
2. Berapa akurasi yang dihasilkan dari metode Support Vector Machine dalam menganalisa kepribadian DISC pengguna instagram berdasarkan caption instagram bahasa indonesia
3. Berapa Tingkat Performa metode SVM dibandingkan Metode KNN, MNB, dan Decision Tree

## 1.3. Batasan Masalah

Berdasarkan latar belakang dan rumusan masalah yang telah dijabarkan, Peneliti membatasi luasnya cakupan ruang lingkup yang dibahas pada skripsi ini agar tidak melebar ke topik lain dan memudahkan penulis dalam pengerjaannya.

Adapun batasan-batasan masalah sebagai berikut:

1. Penelitian ini lebih menekankan pada pengujian metode
2. Analisis yang digunakan hanya kepribadian DISC

3. Metode yang digunakan dalam mengklasifikasi adalah Support Vector Machine
4. Dataset yang digunakan berupa text caption bahasa indonesia

#### 1.4. Maksud dan Tujuan

Adapun maksud dan tujuan dari penelitian dan penulisan ini adalah

:

1. Mengetahui cara kerja metode Support Vector Machine dalam menganalisis kepribadian DISC pengguna instagram berdasarkan caption bahasa indonesia
2. Mengetahui kepribadian DISC pengguna instagram berdasarkan caption bahasa indonesia
3. Mengetahui berapa tingkat akurasi metode Support Vector Machine dalam menganalisa kepribadian DISC
4. Mengetahui performa metode SVM dibandingkan metode KNN, MNB, dan Decision Tree

#### 1.5. Manfaat Penelitian

Beberapa manfaat dari penelitian ini adalah :

1. Membantu dalam mengetahui Kepribadian DISC pengguna instagram berdasarkan caption bahasa indonesia
2. Diharapkan dapat menambah dan mengembangkan wawasan, informasi dan pengetahuan kepada pihak yang berkepentingan
3. Diharapkan setelah adanya penelitian ini mampu sebagai referensi untuk penelitian selanjutnya di masa yang akan datang.

## **1.6. Metode Penelitian**

### **1.6.1. Metode Studi Literatur**

Mengumpulkan beberapa referensi terkait Klasifikasi, Personality, Kepribadian DISC, dan metode Support Vector Machine dari beberapa jurnal, prosiding, buku, artikel, dan beberapa referensi lainnya

### **1.6.2. Metode Pengumpulan Data**

Pengumpulan data merupakan proses dalam penelitian yang bertujuan untuk mengumpulkan data untuk pengujian penelitian yang akan dilakukan. Pengumpulan data yang digunakan dalam penelitian ini yaitu menggunakan dataset yang sudah diberi label DISC oleh psikolog

### **1.6.3. Metode Analisis**

Analisis masalah didasarkan pada informasi yang diperoleh di tahap sebelumnya agar diperoleh metode yang tepat untuk memecahkan masalah dalam penelitian ini

### **1.6.4. Implementasi**

Pada tahap ini, implementasi dilakukan berdasarkan analisis yang telah dilakukan telah dilakukan dalam bentuk program (notebook) sesuai dengan desain.

### **1.6.5. Testing**

Pada metode testing atau pengujian ini dilakukan untuk menguji performa dari kinerja dari suatu metode yang dibangun, beberapa teknik yang digunakan adalah Confusion matrix.

### 1.7. Sistematika Penulisan

Secara keseluruhan laporan skripsi ini terdiri dari lima (5) bab sesuai dengan petunjuk laporan skripsi yang berlaku di Universitas Amikom Yogyakarta, sistematika penulisan laporan ini adalah sebagai berikut :

#### **BAB I PENDAHULUAN**

Bab ini berisikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penelitian.

#### **BAB II LANDASAN TEORI**

Bab II berisikan tentang pembahasan penelitian sebelumnya yang sudah dilakukan oleh peneliti lain sebagai referensi penelitian ini dijalankan dan juga berisikan teori-teori yang digunakan dalam penelitian.

#### **BAB III METODE PENELITIAN**

Bab ini menguraikan tentang alur penelitian dalam mendapatkan hasil dan juga pengujian yang dilakukan

#### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini berisi tentang tahapan yang dilakukan dalam penerapan alur yang telah dirancang yang meliputi rancangan implementasi didalam code untuk menghasilkan klasifikasi terbaik

#### **BAB V PENUTUP**

Bab ini berisi kesimpulan yang diperoleh dari perumusan masalah yang telah disampaikan, serta saran dalam pengembangan aplikasi supaya sistem dapat lebih baik lagi