

# BAB 1

## PENDAHULUAN

---

### 1.1 Latar Belakang

Aplikasi media sosial Instagram adalah salah satu aplikasi media sosial terpopuler dan memiliki akses pengguna terbanyak. Terdapat 111.187.100 pengguna Instagram di Indonesia pada September 2023, yang merupakan 39,6% dari seluruh populasi Indonesia [1]. Hal ini dikarenakan Instagram menjadikan aplikasi mereka sebagai wadah untuk masyarakat bebas bertukar pikiran dan beropini dengan menyediakan fitur *comments* dan *captions* dengan maksimal jumlah yaitu 2200 karakter.

Dengan banyaknya pengguna, maka semakin banyak data yang ditampung oleh Instagram. Instagram menjadi salah satu sumber pemanfaatan untuk mengetahui respon dan opini masyarakat terhadap sesuatu. Melihat peluang yang besar, maka banyak pelaku bisnis yang menjalankan usahanya di Instagram. Akan tetapi, tidak sedikit pelaku bisnis atau UMKM pengguna sosial media yang masih memiliki kesulitan dalam menjalankan bisnisnya secara *online*. Beberapa UMKM masih banyak yang memiliki kendala atau bahkan tidak tahu cara bersaing dalam dunia bisnis menggunakan teknologi sosial media, seperti apa saja faktor yang diperlukan agar UMKM tersebut dapat mengembangkan popularitas dan kualitas dari akun sosial media usahanya agar meraih kesuksesan dalam berbisnis.

Di sisi lain, selain meraih kesuksesan dan mendapatkan keuntungan dari menjalankan bisnis di Instagram, pelaku bisnis dapat melakukan suatu evaluasi perbaikan guna meningkatkan popularitas dan kesuksesan usahanya dengan memanfaatkan akun sosial medianya melalui analisis opini konsumen yang disampaikan pada komentar Instagram.

Opini masyarakat yang terdapat pada Instagram akan menjadi suatu data yang dapat menentukan pandangan masyarakat terhadap kualitas UMKM tersebut, apakah pandangannya positif, negatif atau netral. Dengan tersebar luasnya pengguna Instagram di seluruh Indonesia, maka opini atau data yang didapat akan menjadi

sangat banyak dan sulit untuk dianalisis secara manual, maka dari itu dibutuhkan suatu metode yang dapat menganalisis secara otomatis yaitu metode analisis sentimen.

Analisis sentimen atau dapat dikenal juga dengan *Opinion Mining* adalah suatu analisis terhadap dokumen tertentu yang berisikan opini mengenai suatu objek yang dapat dimanfaatkan untuk mengetahui respon atau pendapat pengguna media sosial terhadap isu atau permasalahan yang sedang dialami [2]. Proses berjalannya Analisis Sentimen ini perlu didukung dengan adanya klasifikasi dari sebuah model pembelajaran mesin. Seperti pada [3] melakukan klasifikasi data dengan menggunakan algoritma *Support Vector Machine (SVM)* dan *Decision Tree (DT)* dalam proses analisis sentiment pada komentar Instagram mengenai kepuasan kinerja PSSI. Penelitian [3] melakukan klasifikasi dengan cara membandingkan metode *Support Vector Machine (SVM)* dan *Decision Tree (DT)*. Penelitian ini menggunakan 5.500 data komentar dengan tingkat akurasi algoritma *Support Vector Machine (SVM)* sebesar 94% dengan pembagian *data training* dan *data testing* sebanyak 80:20 dan algoritma *Decision Tree (DT)* memiliki tingkat akurasi sebesar 86.67% dengan pembagian *data training* dan *data testing* sebanyak 80:20. Dari hasil perbandingan yang didapat, maka penelitian ini menghasilkan 20.13% komentar positif dan 79.87% komentar negatif dari data yang didapat menggunakan *Support Vector Machine (SVM)* sebagai algoritma dengan tingkat akurasi terbaik.

Adapula Metode *Naïve Bayes* dan *Random Forest* [4] sebagai algoritma yang digunakan dalam proses analisis sentimental mendeteksi komentar sarkasme di Instagram politikus Indonesia. Tahapan dari analisis sentimen yang dilakukan pada penelitian ini yaitu *Scrapping Data*, *Labelling Data*, *Preprocessing*, *Feature Extraction (TF-IDF)*, *Cross Validation*, *Sentimental Analysis Model Learning*, *Sarcasm Detection*, dan Evaluasi menggunakan *Confusion Matrix*. *Feature Extraction* dengan algoritma TF-IDF dilakukan untuk menganalisa hubungan antara setiap kalimat dengan dokumen [3]. TF-IDF juga digunakan untuk menghitung kemunculan dari setiap kata pada dokumen tersebut [4]. Penelitian yang menggunakan 3.140 data ini memperoleh hasil akurasi dengan algoritma *Random Forest* sebesar 71% dan

algoritma *Naïve Bayes* sebesar 60% pembagian *data training* dan *data testing* sebesar 80:20.

*Support Vector Machine (SVM)* tergolong sebagai salah satu algoritma terbaik untuk digunakan dalam proses klasifikasi Analisis Sentimen karena cenderung memiliki nilai akurasi yang lebih tinggi dibanding algoritma lainnya. Seperti pada penelitian [5] *Support Vector Machine (SVM)* berhasil memperoleh akurasi tertinggi sebesar 69,15% dibandingkan metode lainnya yaitu *K-Nearest Neighbors (KNN)*, *Naïve Bayes*, dan *Logistic Regression*. Penelitian ini menggunakan 2.765 data yang digunakan dalam Analisis Sentimen menggunakan data dari komentar yang berada pada komentar di aplikasi Twitter. Lalu pada penelitian [6] akurasi tertinggi juga didapatkan dari penggunaan metode *Support Vector Machine (SVM)* Radial RBF yaitu sebesar 79.67%, dengan menggunakan 1.483 data yang diperoleh dari komentar Instagram Kemdikbud dalam penelitian “Analisis Sentiment Pengguna Instagram terhadap Kebijakan Kemdikbud Mengenai Bantuan Kuota Internet dengan Metode SVM”. Selain metode *Support Vector Machine (SVM)*, *Naïve Bayes*, *Random Forest*, *K-Nearest Neighbors (KNN)*, *Naïve Bayes*, dan *Logistic Regression*. Adapula metode *Convolutional Neural Network (CNN)* yang digunakan [2] dalam penelitian terhadap pemberitaan Virus Corona-19 di Instagram. CNN termasuk dalam golongan *Deep Neural Network* yang memiliki kedalaman jaringan tinggi karena memiliki Layer Konvolusi atau *Convolution Layer*. CNN dapat mengolah data yang memiliki dimensi tinggi karena mengandung neuron yang saling terhubung di dalamnya. Pada [2] CNN berhasil memperoleh tingkat akurasi dengan tinggi yaitu sebesar 96%, dari 1.031 data yang digunakan dalam penelitian [2] yang diperoleh dari komentar Instagram @kumparancom, @detikcom, @cnnindonesia.

Pada Proyek Akhir ini, akan dilakukan penelitian untuk membantu para pelaku bisnis UMKM dalam meningkatkan popularitas dan kualitas usahanya agar menjadi lebih baik sesuai dengan keinginan dan kegemaran yang dirasakan oleh konsumen, serta membantu pelaku bisnis untuk mengetahui faktor apa saja yang dapat mempengaruhi popularitas dan kualitas dari akun sosial media UMKM tersebut. Penelitian ini akan dilakukan dengan cara pengumpulan data yang berupa komentar

dari akun Instagram sebuah UMKM untuk mengetahui bagaimana ulasan-ulasan dari konsumen.

Berdasarkan penelitian yang telah dilakukan sebelumnya, maka pada Proyek Akhir ini akan dilakukan pengembangan Aplikasi Klasifikasi dan Opini Pada Teks Komentar di Akun Sosial Media UMKM Menggunakan Model Mesin Pembelajaran dan Analisis Sentimen. Proyek akhir ini akan menggunakan dataset berupa teks komentar yang bersumber dari akun Instagram UMKM @by.neeth yang merupakan sebuah UMKM lokal yang bergerak di bidang *fashion* yang berdomisili di Bandung.

## **1.2 Rumusan Masalah**

Rumusan masalah pada Proyek Akhir ini adalah:

1. Bagaimana cara menentukan faktor - faktor yang dibutuhkan untuk meningkatkan popularitas dan kualitas UMKM dalam sosial media Instagram?
2. Bagaimana cara mengidentifikasi sentimen pada data komentar di sosial media Instagram UMKM termasuk ke dalam kelas sentimen positif, negatif, atau netral?

## **1.3 Tujuan**

Tujuan dari Proyek Akhir ini adalah:

1. Dapat mengklasifikasikan setiap komentar atau ulasan dari konsumen pada sosial media Instagram termasuk dalam kelas komentar positif, negatif, atau netral.
2. Membuat hasil visualisasi proyek akhir berbentuk *dashboard* untuk memberikan *insights* mengenai hasil analisis sentimen yang dilakukan.

## **1.4 Batasan Masalah**

Batasan masalah dari Proyek Akhir ini adalah:

1. Menggunakan data dari sosial media Instagram yang dimiliki oleh UMKM yaitu UMKM By.Neeth.

2. Analisis sentimental ini difokuskan pada komentar yang ditulis dalam postingan media sosial UMKM dalam Bahasa Indonesia.
3. Komentar atau postingan yang mengandung kata-kata kasar, tidak sopan, dan tidak relevan akan dihilangkan dari analisis sentimen ini.

## 1.5 Metode Pengerjaan

Adapun metode pengerjaan pada Proyek Akhir ini, sebagai berikut :

1. Studi Literatur  
Studi Literatur adalah proses pencarian data guna memperluas materi terkait melalui referensi dari sumber yang didapat.
2. Pengumpulan Data  
Pengumpulan data dilakukan untuk mendapatkan data sebagai masukan sistem, yaitu dataset. Perolehan dataset yang digunakan didapatkan dari media sosial Instagram UMKM berupa komentar konsumen. Dataset yang digunakan diperoleh dengan tools Apify melalui metode *scrapping data* atau pengambilan data.
3. Perancangan Model *Machine Learning*  
Tahap ini adalah analisis terhadap data yang diperoleh dari Studi Literatur kemudian dilakukan perancangan sistem menggunakan algoritma *Support Vector Machine (SVM)* dan *Random Forest*.
4. Pengembangan Dashboard  
Tahap ini merupakan tahap yang dilakukan untuk mengimplementasikan dan melakukan visualisasi terhadap hasil dari program analisis sentiment menggunakan model SVM dan *Random Forest* yang sudah dikembangkan. Dashboard ini dapat berfungsi sebagai antarmuka sarana informasi mengenai pandangan atau sentimen konsumen terhadap akun sosial media UMKM di Instagram.
5. Pengujian Sistem  
Tahap ini dilakukan dengan pengujian performa sistem yang telah dibangun.
6. Penyusunan Laporan

Penyusunan laporan berbentuk buku tugas akhir disusun sebagai wujud dokumentasi dari konsep, teori, perancangan.

### 1.6 Jadwal Pengerjaan

Berikut merupakan rencana jadwal pengerjaan Proyek Akhir Aplikasi Klasifikasi dan Opini pada Teks Komentar di Akun Sosial Media UMKM Menggunakan Model Mesin Pembelajaran dan Analisis Sentimen, dalam kurun waktu pada tahun 2023 – 2024.

Tabel 1. 1. Jadwal Pengerjaan Proyek Akhir

No	Deskripsi Kerja	Tahun 2023				Tahun 2024						
		Sep	Okt	Nov	Des	Jan	Feb	Mar	Apr	Mei	Jun	Jul
1	Diskusi Awal											
2	Literature Review											
3	Rumusan Masalah											
4	Tujuan Penelitian											
5	Pengumpulan Data											
6	Pengerjaan Proyek Akhir											
7	Pembuatan Laporan											
8	Kesimpulan											