

BAB 2

TINJAUAN PUSTAKA DAN LANDASAN TEORI

2.1 TINJAUAN PUSTAKA

Tinjauan pustaka ini membuat analisis sentimen pengguna jasa ekspedisi menggunakan metode NBC dan KNN yang diharapkan dapat membantu pedagang dalam mencari partner kerja untuk menyalurkan produk yang dijual melalui transaksi online, berikut ini landasan teori yang akan digunakan dalam pengerjaan penelitian berasal dari beberapa sumber jurnal yang hampir sama dengan penelitian ini yaitu.

Penelitian dengan judul “Analisis Sentimen Berdasarkan Opini Pada Twitter Menggunakan Metode Naive Bayes”. Penelitian ini membahas tentang analisis sentimen pada Twitter mengenai kasus pelayanan yang disediakan jasa ekspedisi barang untuk mengelompokkan respon seseorang terhadap pelayanan pengiriman, menggunakan metode Naive Bayes Classifier dengan 100 data yang diuji. Rata-rata tingkat akurasi mencapai 84% (Lorosae & Prakoso, 2018).

Penelitian dengan judul “Analisis Sentimen Data Computer Sosial Media Facebook dengan K-Nearest Neighbor (studi kasus akun jasa ekspedisi barang J&T ekspres Indonesia)”. Penelitian ini membahas tentang analisis sentimen pengguna facebook mengenai perusahaan J&T. Metode yang digunakan adalah NBC, pembobotan kata menggunakan *term frequency-invers document frequency* dan perhitungan similaritas menggunakan *cosine similarity* dengan tingkat tertinggi akurasi 79,21% dan 70,3% terendah dari 101 data uji (Salam et al., 2018).

Penelitian dengan judul “Analisis Sentimen Kepuasan Mahasiswa Terhadap Pembelajaran Online Selama Pandemic Covid-19 Pada Media Sosial Twitter Menggunakan Perbandingan Algoritma Klasifikasi”. Penelitian ini membahas tentang analisis sentimen pengguna Twitter mengenai keberlangsungan pembelajaran secara online selama pandemi, menggunakan Google Collaboration untuk perbandingan klasifikasi Naive Bayes, KNN dan C4.5. Bahasa pemrograman

Python data yang diambil menggunakan API Twitter yang diperoleh sebanyak 8856 (Ramadhanta et al., 2021).

Penelitian dengan judul “Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM Pada Media Sosial Twitter Menggunakan Naive Bayes Classifiers”. Penelitian ini membahas tentang analisis sentimen pengguna Twitter mengenai kebijakan PPKM yang dilakukan oleh pemerintah. Klasifikasi menggunakan Naive Bayes Classifier metode *crawling* menggunakan *software* RStudio, data yang diperoleh 1000 dengan hasil akurasi 99% positif dan 1% negative (Taofik Krisdiyanto, 2021).

Penelitian dengan judul “Analisis Sentimen Pada Media Sosial Twitter Menggunakan Naive Bayes Classifier Dengan Ekstraksi Fitur N-Gram”. Penelitian ini membahas tentang analisis sentimen pengguna Twitter mengenai persepsi masyarakat terhadap isu kenaikan tarif dasar listrik. klasifikasi dengan Naive Bayes, pengujian menggunakan *cross validation* dan *confusion matrix* kemudian menggunakan fitur N-Gram tingkat akurasi 92.00% (Nugroho, 2018).

Persamaan penelitian penulis secara umum dengan penelitian sebelumnya adalah data yang diambil dari *platform*, bahasa pemrograman yang digunakan yaitu Python, metode yang dipakai NBC dan KNN, hasil yang ditampilkan yaitu sentimen positif, negatif. Perbedaan penelitian penulis dengan penelitian sebelumnya adalah penelitian ini menggunakan bahasa pemrograman Python dengan *framework* Flask, pengambilan data menggunakan Library *snsrape* dengan bahasa Python, menampilkan hasil data yang telah diolah ke dalam sebuah sistem dashboard, data akan diolah menggunakan aplikasi Google Collab. Hasil akhir dari penggunaan metode untuk membandingkan tingkat akurasi data.

2.2 LANDASAN TEORI

2.2.1 Twitter

Twitter adalah media sosial yang dilambangkan oleh burung berwarna biru, Twitter diartikan sebagai pesan singkat antar pengguna yang sangat populer dikalangan masyarakat sekarang dimana unggahan yang dilakukan oleh pengguna disebut tweet. Tweet atau pesan yang dibagikan biasanya berisikan tentang

informasi-informasi yang sedang hangat diperbincangkan dan sering disebut juga trending topik di Twitter (Nugroho, 2018).

Twitter berdiri di San Francisco, California dan sekarang sudah memiliki lebih dari 25 kantor di seluruh dunia. Twitter dapat diakses dengan telepon genggam melalui situs webnya Twitter. Twitter merupakan salah satu jaringan sosial media yang banyak digunakan, sebagian orang menggunakan Twitter untuk memberikan opini, perbedaan pendapat akan memicu terjadinya sekelompok orang perang *tweet* karena tidak sepaham (Andrianto & Adinugroho, 2018).

2.2.2 Ekspedisi

Diartikan dari Kamus Besar Bahasa Indonesia (KBBI) ekspedisi adalah pengiriman surat dan barang atau perusahaan yang menyediakan jasa pengangkutan barang. Cara kerja ekspedisi dibagi beberapa tahapan yaitu barang yang akan dikirim, barang di kemas, dokumen dan administrasi pembiayaan (Rohimah et al., 2018).

Banyaknya perusahaan ekspedisi yang berdiri di Indonesia diantaranya yaitu J&T, JNE, Pos Indonesia, Sicepat, TIKI, dan Wahana. Setiap perusahaan memiliki kebijakan dan keunggulan tersendiri, layanan yang diberikan pun berbeda-beda antara lain ekonomi, regular, dan express. Keamanan dalam pengiriman barang menjadi salah satu tolok ukur konsumen.

2.2.3 Analisis Sentimen

Analisis sentimen adalah hasil perpaduan antara data mining dan teks mining, analisis sentimen biasanya digunakan untuk menganalisis opini-opini masyarakat yang bersumber dari data media sosial atau jaringan internet lainnya. Analisis sentimen dibagi menjadi 3 jenis yaitu sentimen positif, sentimen negatif dan sentimen netral (Salam et al., 2018).

Kegiatan yang melibatkan individu atau kelompok untuk melakukan analisis melalui opini seseorang terhadap pandangan sebuah produk, topik atau masalah sehingga bisa didapati persoalan tersebut masuk kedalam sentimen negatif, sentimen positif atau sentimen netral. Analisis sentimen dibagi menjadi 2 kelompok yaitu *coarse-grained sentiment* dan *fine grained sentiment analysis*.

Analisis sentimen digunakan untuk mendeteksi teks opini pada dokumen yang dikategorikan negatif, positif dan netral. Adapun tahapan dalam analisis sentimen adalah :

1. Pengumpulan data

Data yang diambil dari *tweet* masyarakat dalam media sosial Twitter, data diambil menggunakan Twitter API dengan Bahasa Python.

2. Pra pengolahan

Data yang telah didapatkan kemudian dibersihkan dan diolah untuk menormalisasi.

3. Pengelolaan

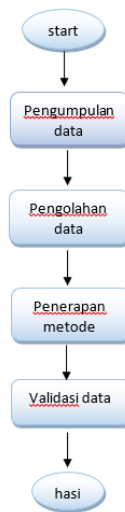
Tahapan ini untuk mengklasifikasikan data berdasarkan positif, negatif dan netral.

4. Validasi

Pada tahapan ini yaitu memvisualisasi hasil analisis sentimen ke dalam histogram

2.2.4 Naïve Bayes Classifier((NBC)

Naive Bayes Classifier atau yang sering disingkat NBC adalah metode yang pertama kali dikenalkan oleh seorang ilmuwan Inggris yang bernama Thomas bayes. NBC ini dikelompokkan berlandaskan metode probabilitas dan metode statistik, metode ini diklasifikasikan oleh teori Bayes untuk memperhitungkan berdasarkan ada dan tidaknya dari suatu golongan (Lorosae & Prakoso, 2018). Model penelitian ini dilakukan sesuai dengan Gambar 2.1



Gambar 2.1 Skema NBC

Proses perhitungan klasifikasi pada metode Naive Bayes Classifier sebagai berikut :

$$P(c_j|W_i) = P(c_j) \prod_{i=1}^k P(w_i|C_j) \quad (1)$$

Keterangan :

$P(c_j|W_i)$ = Belakang probabilitas digolongkan j apabila memperoleh kata i

$P(W_i|c_j)$ = Syarat probabilitas kata i ke golongan kata j

$P(c_j)$ = Prior probabilitas munculnya golongan j

$P(W_i)$ = Probabilitas kemunculan kata

i = Indeks suatu kata yang digolongkan mulai dari kata i hingga kata k

j = Indeks suatu kata yang digolongkan mulai dari kata i hingga kata ke-n.

2.2.5 K-Nearest Neighbor (KNN)

K-Nearest Neighbor merupakan algoritma yang digunakan untuk mengelompokkan objek pelatihan berdasarkan contoh pelatihan terdekat di ruang fitur. Pada penggunaan K-Nearest Neighbor dilakukan dengan mencari kelompok objek dalam data *training* dengan *testing*. Salah satu cara untuk menghitung dekat atau jauhnya kelompok tetangga menggunakan *Euclidean distance* dan *Manhattan distance*. Adapun persamaan seperti berikut:

$$d(x,y) = \sqrt{\sum_i^n (a_i - b_i)^2} \quad (2)$$

Keterangan :

$a = a_1, a_2, \dots, \dots, \dots, a_n$

$b = b_1, b_2, \dots, \dots, \dots, b_n$, mewakili n nilai atribut dari dua record.

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANI
YOGYAKARTA