

# ANALISIS SENTIMEN KINERJA DEWAN PERWAKILAN RAKYAT (DPR) PADA TWITTER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER

Dianati Duei Putri<sup>1</sup>, Gigih Forda Nama<sup>2</sup>, Wahyu Eko Sulistiono<sup>3</sup>

<sup>1,2,3</sup> Teknik Informatika, Universitas Lampung; Jl. Prof. Sumantri Brojonegoro, Lampung

Received: 19 November 2021

Accepted: 14 Desember 2021

Published: 10 Januari 2022

**Keywords:**

Sentiment Analysis, DPR, Naive Bayes Classifier

**Correspondent Email:**

dianaputri.dp119@gmail.com

**How to cite this article:**

Dianati (2022). Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (Dpr) Pada Twitter Menggunakan Metode Naive Bayes Classifier . Jurnal Informatika dan Teknik Elektro Terapan, 10(1).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY NC)

**Abstrak.** Dalam penelitian ini akan dilakukan analisis sentimen masyarakat terhadap kinerja Dewan Perwakilan Rakyat (DPR) yang diungkapkan melalui media sosial twitter. Ada beberapa tahap untuk melakukan analisis sentimen , yaitu pengumpulan data (*crawling*), *preporcessing data* yang terdiri dari proses *cleaning data*, *tokenization*, *stop remova* dan *case folding*, *splitting data* dan klasifikasi data menggunakan metode *Naive Bayes Classifier*. Penelitian ini menggunakan sebanyak 1546 data *tweet*. Hasil dari penelitian ini didapatkan bahwa DPR mendapatkan 95 *tweet* positif dengan polaritas 0.75 atau 75% sentimen positif, 693 *tweet* netral dengan polaritas 0.79 atau 79% sentimen netral dan 758 *tweet* negatif dengan polaritas 0.82 atau 82% sentimen negatif dengan *accuracy score* 0.8 atau 80% berdasarkan data testing sebanyak 20%.

**Abstract.** In this study, an analysis of public sentiment will be carried out on the performance of the House of Representatives (DPR) which is expressed through Twitter social media. There are several stages to carry out sentiment analysis, namely data collection (*crawling*), data preprocessing consisting of data cleaning, tokenization, stop remova and case folding processes, data splitting and data classification using the Naive Bayes Classifier method. This research uses 1546 tweet data. The results of this study showed that the DPR received 95 positive tweets with a polarity of 0.75 or 75% positive sentiment, 693 neutral tweets with a polarity of 0.79 or 79% neutral sentiment and 758 negative tweets with a polarity of 0.82 or 82% negative sentiment with an accuracy score of 0.8 or 80%. based on testing data as much as 20%.

## I. PENDAHULUAN

Indonesia merupakan negara demokrasi, hal ini ditunjukkan dengan dilakukannya pemilihan umum untuk memilih kepala negara, kepala daerah serta badan legislatif(DPR). DPR merupakan suatu lembaga negara yang bergerak didalam lingkup politik hukum, dan

Undang-Undang sebagai manifestasi dari politik tersebut. DPR memiliki kekuasaan sebagai pembentuk Undang-Undang yang telah diatur dalam Undang-Undang Dasar Republik Indonesia Tahun 1945 pasal 20 ayat 1. DPR sebagai lembaga yang mewakilkan rakyat menyanggung tanggung jawab yang harusnya dipenuhi secara demokratis dan responsif untuk mengatasi permasalahan sosial yang

ada pada masyarakat serta tidak memprioritaskan kepentingan partai politik. Namun pada saat ini banyak lapisan masyarakat bahkan memberikan pendapat mengenai kinerja dari DPR yang dianggap tidak berpihak kepada rakyat. Banyak masyarakat yang memberikan opini mereka mengenai kinerja DPR melalui media sosial. Media sosial adalah sebuah layanan yang memfasilitasi pengguna untuk berkomunikasi dengan berbagai topik. Media sosial saat ini yang populer di masyarakat adalah Twitter.

## II. TINJAUAN PUSTAKA

### 2.1 Twitter

Twitter adalah layanan jejaring sosial yang membantu penggunanya mengirim dan membaca pesan berbasis teks hingga 140 karakter. Pada awal 2013, pengguna twitter mengirimkan lebih dari 500 juta kicauan per hari. Tingginya popularitas twitter menyebabkan layanan ini telah dimanfaatkan untuk berbagai keperluan dalam berbagai aspek, misalnya sebagai sarana protes, kampanye politik, sarana pembelajaran dan sebagai media komunikasi darurat[1]. API merupakan cara program komputer "berbicara" satu sama lain agar mereka dapat meminta dan menyajikan informasi. Ini dilakukan dengan mengizinkan aplikasi perangkat lunak memanggil apa yang disebut sebagai endpoint: alamat yang terkait dengan informasi jenis tertentu yang disediakan (endpoint umumnya unik seperti nomor telepon). Twitter mengizinkan akses ke bagian dari layanan melalui API untuk memungkinkan orang-orang membangun perangkat lunak yang terintegrasi dengan Twitter seperti solusi yang membantu sebuah perusahaan menjawab umpan balik pelanggan di Twitter. Data Twitter berbeda dari data yang dibagi oleh kebanyakan platform sosial lain karena data tersebut mencerminkan informasi yang dipilih pengguna untuk dibagikan ke publik. Platform API twitter menyediakan akses luas ke data Twitter publik yang telah dipilih pengguna untuk dibagikan ke dunia. Twitter juga mendukung API yang memungkinkan pengguna mengelola informasi Twitter pengguna yang non-publik dan memberikan informasi ini ke pengembang yang telah diizinkan pengguna untuk melakukannya [2].

### 2.2 Text Mining

*Text Mining* adalah ekstraksi informasi dari data sumber yang belum terstruktur yang mengacu pada teknik penambangan data untuk menganalisis dan memproses data. Pada *text*

*mining*, perlu pertama diperlukan pengambilan data kemudian data tersebut perlu di *preprocessing* sebelum proses klasifikasi. Proses *preprocessing* melalui beberapa tahap yaitu *case folding, tokenization, filtering, stop removal*[3].

### 2.3 Python

Python merupakan bahasa pemrograman tingkat tinggi yang mendukung pemrograman berorientasi objek. Python memiliki perbedaan dengan bahasa pemrograman lain yaitu dalam penulisan sintaks[4]. Dalam bahasa pemrograman python terdapat berbagai *library* dan *framework* yang digunakan untuk melakukan analisis data[5]. Dalam penelitian ini menggunakan *library* sebagai berikut :

#### 1. Tweepy

Tweepy adalah *library* python yang digunakan untuk mengakses API ilik twitter. Tweepy dapat mengakses API twitter sehingga dapat mengambil informasi dari twitter dengan skrip python[6]. Tweepy digunakan untuk menjembatani twitter dengan python, dalam penelitian ini tweepy digunakan untuk mengambil informasi mengenai DPR yang ada di twitter.

#### 2. NLTK (*Natural Language Tool Kit*)

*Library* NLTK ini digunakan pada tahapan *tokenization* yaitu memisahkan kalimat menjadi per kata yang didapat untuk diolah[7].

#### 3. Sastrawi

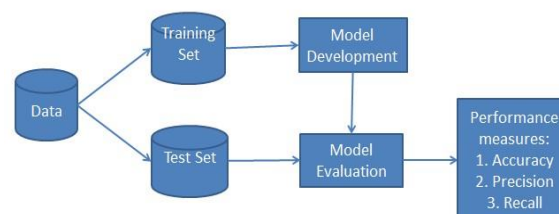
*Library* Sastrawi digunakan dalam tahapan *stop removal* untuk menghapus kata-kata yang dianggap tidak berhubungan.

#### 4. Scikit-learn

*Scikit-lear* merupakan *library* pada python yang digunakan untuk membangun sebuah program *machine learning*. *Library* ini juga digunakan untuk Algoritma *Naive Bayes* untuk keperluan *Data Science*.

### 2.4 Naive Bayes

*Naive Bayes* merupakan salah satu algoritma yang digunakan dalam *text mining*. Algoritma *naive bayes* memprediksi kejadian masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal dengan nama *Teorema Bayes*[8]



Gambar 2.1 Naive Bayes

Metode NBC menempuh dua tahap dalam proses klasifikasi teks, yaitu tahap pelatihan dan tahap klasifikasi. Pada tahap pelatihan dilakukan proses analisis terhadap sampel dokumen berupa pemilihan vocabulary, yaitu kata yang mungkin muncul dalam koleksi dokumen sampel yang sedapat mungkin dapat menjadi representasi dokumen. Selanjutnya adalah penentuan probabilitas prior bagi tiap kategori berdasarkan sampel dokumen. Pada tahap klasifikasi ditentukan nilai kategori dari suatu dokumen berdasarkan term yang muncul dalam dokumen yang diklasifikasi[9]. Dalam algoritma naïve bayes classifier setiap dokumen direpresentasikan dengan pasangan atribut “x1, x2, x3,...xn” dimana x1 adalah kata pertama, x2 adalah kata kedua dan seterusnya. Sedangkan V adalah himpunan kategori Tweet. Pada saat klasifikasi algoritma akan mencari probabilitas tertinggi dari semua kategori dokumen yang diujikan (VMAP), dimana persamaannya adalah sebagai berikut :

$$V_{MAP} = \underset{V_j \in V}{\operatorname{argmax}} \frac{P(X_1, X_2, X_3, \dots, X_n | V_j) P(V_j)}{P(X_1, X_2, X_3, \dots, X_n)}$$

Untuk P(x1, x2, x3,...xn) nilainya konstan untuk semua kategori (Vj)[10]. Persamaan diatas dapat disederhanakan menjadi sebagai berikut :

$$V_{MAP} = \underset{V_j \in V}{\operatorname{argmax}} \prod_{i=1}^n P(x_i | V_j) P(V_j)$$

Keterangan :

Vj = Kategori tweet j =1, 2, 3,...n. Dimana dalam penelitian ini

j1= kategori tweet sentimen negatif,

j2 = kategori tweet sentimen positif, dan

j3 = kategori tweet sentiment netral

P(xi|Vj) = Probabilitas xi pada kategori

VjP(Vj) = Probabilitas dari Vj

Untuk P(Vj) dan P(xi|Vj) dihitung pada saat pelatihan dimana persamaannya adalah sebagai berikut :

$$P(V_j) = \frac{|docs\ j|}{|contoh|}$$

$$P(X_i | V_j) = \frac{nk+1}{n+|kosakata|}$$

Keterangan :

|docs j| = jumlah dokumen setiap kategori j

|contoh| = jumlah dokumen dari semua kategori

nk = jumlah frekuensi kemunculan setiap kata

n = jumlah frekuensi kemunculan kata dari setiap kategori

|kosakata| = jumlah semua kata dari semua kategori

Pengujian pada performa algoritma Naïve Bayes Classifier dilakukan dengan menguji tingkat akurasi, presisi dan recall. Dalam melakukan penghitungan tingkat akurasi algoritma Naïve Bayes Classifier dengan rumus sebagai berikut:

$$Akurasi = \frac{\sum \text{data benar}}{n \text{ dokumen}} \times 100 \%$$

Rumus untuk melakukan penghitungan tingkat presisi sebagai berikut:

$$Presisi = \frac{\sum \text{data positif atau negatif}}{n \text{ dokumen positif atau negatif}} \times 100 \%$$

Rumus untuk melakukan penghitungan recall sebagai berikut:

$$recall = \frac{\sum \text{doc relevan dan terambil}}{\sum \text{seluruh dokumen relevan}} \times 100\%$$

Proses perhitungan prioritas bantuan dilakukan untuk menyamakan data dan melihat data mana yang lebih tinggi dalam suatu dataset sehingga dapat ditampilkan prioritas dari bantuan yang dibutuhkan. Perhitungan prioritas bantuan perkategori dilakukan dengan rumus sebagai berikut:

$$Jumlah = \frac{\sum \text{seluruh dataset}}{\sum \text{dataset perkategori}}$$

Rumus untuk menghitung rata-rata dari dataset perkategori :

$$JumlahKategori = \frac{\sum \text{dataset perkategori}}{\sum \text{dataset positif atau negatif}}$$

## 2.5 Jupyter

Perez dan Granger menciptakan aplikasi *Ipython Notebook* berbasis *web browser* yang telah digunakan oleh banyak bidang penelitian. Aplikasi ini kini menjadi *Jupyter Notebook App*, dengan menggunakan bahasa pemrograman utama yang lebih beragam, yaitu *Julia*, *Python* dan *R*[11]. *Jupyter* memiliki beberapa produk yang dapat digunakan yaitu :

### 1. Jupyter Notebook

*Jupyter Notebook* yang sebelumnya dengan nama *Ipython Notebook* merupakan *web-based* yang digunakan untuk membuat dokumen notebook jupyter[12]. Dokumen notebook jupyter disimpan dengan format “.ipynb”.

### 1. Jupyter Hub

*Jupyter Hub* merupakan server yang digunakan untuk menjalankan *jupyter notebook*.

2.6 Design Science Research (DSR)

*Design Science Research* (DSR) merupakan suatu metode penelitian yang dikembangkan oleh Ken Peffers, Tuure Tuunanen dan Charles E.Gengler. DSR memiliki banyak gagasan untuk melakukan penelitian dan memberikan konsep bagaimana kita dapat menyimpulkan hasil suatu proses penelitian yang dilakukan[13].

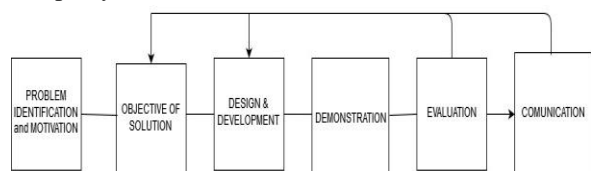
2.7 Sentiment Analysis

*Sentiment analysis* merupakan suatu proses dalam data mining yang digunakan untuk mengidentifikasi dan mengekstrak suatu informasi dari suatu teks yang bertujuan untuk memahami *social sentiment* pada teks tersebut[14]. *Sentiment analysis* juga digunakan untuk mendapatkan informasi mengenai sikap, pendapat dan emosi yang ada pada teks informasi tersebut. *Sentiment Analysis* difokuskan untuk review klasifikasi berdasarkan polaritas[15]. Berdasarkan klasifikasi, analisis sentimen dibagi menjadi dua yaitu klasifikasi subjektivitas dan klasifikasi ke dalam positif atau negatif[16].

III. METODOLOGI PENELITIAN

3.1 Metode Penelitian

Pada penelitian ini metode yang digunakan yaitu metode *Design Science Research* (DSR). Metode ini memiliki 6 tahapan yaitu :



Gambar 3.1 Metode DSR

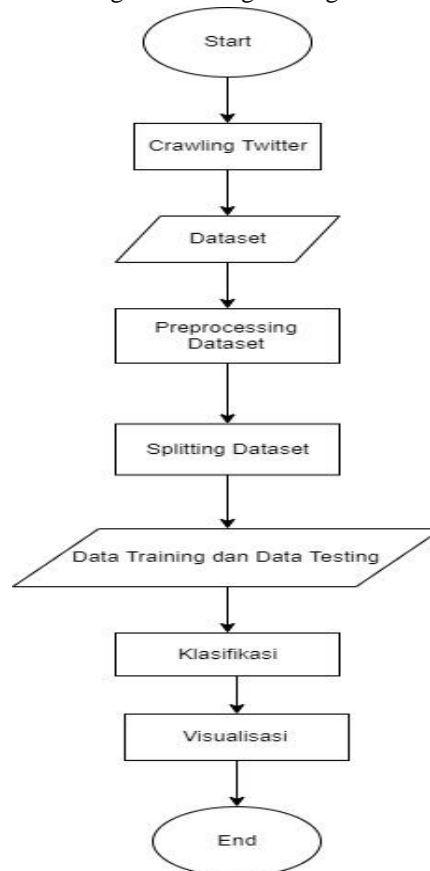
1. Identifikasi Masalah dan Motivasi

Proses pertama yang dilakukan pada penelitian ini adalah mencari permasalahan apa yang terjadi pada dunia politik pada saat ini terutama untuk permasalahan DPR.

2. Menentukan Objek Solusi

Pada tahapan ini dilakukan dengan analisis opini masyarakat yang akan dibutuhkan untuk membangun sistem. Pada twitter dilakukan *crawling* untuk mendapatkan data dan akan di analisis menggunakan algoritma *naive bayes*.

3. Perancangan dan Pengembangan



Gambar 3.2 Diagram Alur Sistem

4. Demonstrasi

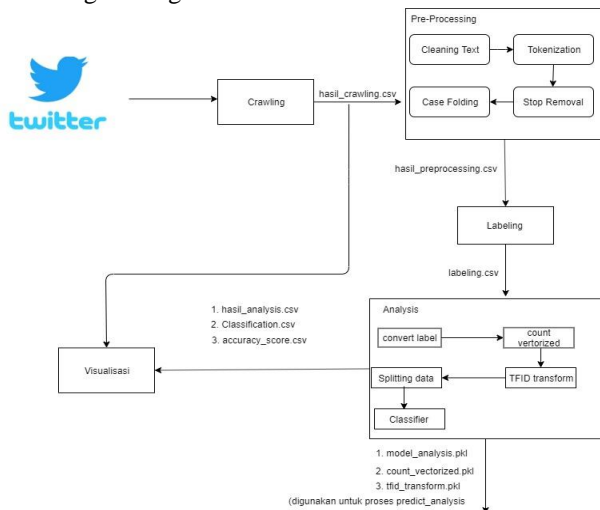
Tahapan selanjutnya adalah demonstrasi, pada tahapan ini akan dilakukan pengujian dari sistem yang dibuat. Pengujian kemampuan sistem ini dilakukan dengan tujuan mengetahui apakah sistem yang dibuat berjalan sesuai dengan rancangan sebelumnya.

5. Evaluasi

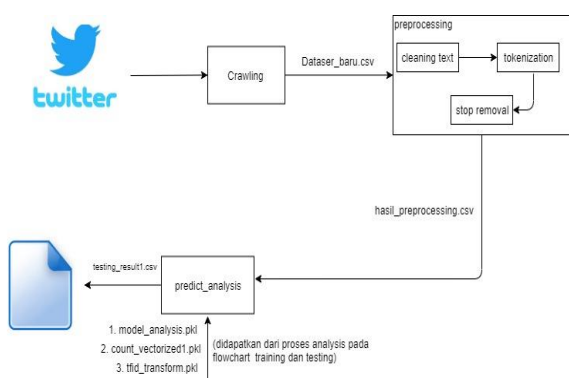
Tahapan evaluasi merupakan tahapan untuk memberikan ulasan dari hasil pengujian yang telah dilakukan pada sistem. Tahapan evaluasi ini penting untuk mengetahui suatu sistem *analyzer sentiment* dan algoritma *naive bayes* yang diimplementasikan pada program.

#### IV. HASIL DAN PEMBAHASAN

##### 4.1 Pengembangan



Gambar 4.1 Flowchart Training dan Testing



Gambar 4.2 Flowchart Model Predict

##### 1. Crawling

Pada penelitian ini dilakukan *crawling* twitter dengan *query* DPR untuk mendapatkan *tweet* dari pengguna twitter. Proses *crawling* menggunakan *library tweepy* dari python serta API yang tersedia dari twitter.

##### 2. Preprocessing data

Preprocessing merupakan tahapan yang dilakukan untuk cleaning data yang telah di dapatkan. Preprocessing data ini menghapus karakter yang sering ada pada tulisan tetapi tidak terlalu memiliki makna, selain itu proses ini juga mengubah huruf menjadi huruf kecil (case folding).

##### 3. Labeling

Proses selanjutnya yaitu labeling. Proses ini dilakukan untuk memberikan label pada isi komentar pengguna twitter yang kemudian akan disimpan kedalam dataset untuk dilakukan proses Splitting Data. Terdapat tiga pilihan dalam memberikan label ke suatu kalimat yaitu Positif, Netral dan Negatif, hal ini dimaksudkan agar pilihan label menjadi konsisten dan menghindari adanya salah input.

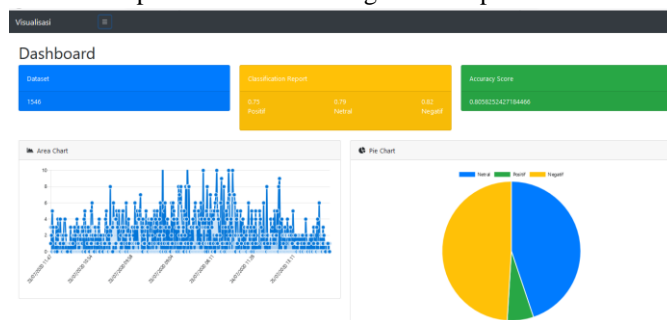
##### 4.2 Demonstrasi

###### 1. Analysis

Setelah melakukan splitting data, tahapan selanjutnya yaitu sentiment analysis. Tahapan ini digunakan untuk mengetahui classification report dan accuracy score berdasarkan data training. Pada proses ini menggunakan algoritma naive bayes untuk membuat klasifikasi

###### 2. Visualisasi

Setelah melakukan proses Analysis tahapan selanjutnya yang dilakukan yaitu visualisasi. Tahapan visualisasi ini merupakan tahapan akhir yang digunakan untuk menampilkan hasil kedalam grafik maupun tabel.



Gambar 4.3 Visualisasi

Tables

Dashboard / Tables

No	Tanggal	User	Author	isi	Stop Removal	Label	Polaritas
0	23/01/2020	melambilga	riqineng	mentar anggota wakil rakyat ngga beda jauh dengan omad yang soka ngabiz th	mentar anggota wakil rakyat ngga beda omad soka ngabiz th	negatif	-1
1	23/01/2020	hendia_id	thousuichotamah	isi baru dipr mpr dan pemerintah sikh sepatut hentikan pembahasan rru hp tapi malah apa ancam getar aka jidit	baru dipr pemerintah sikh sepatut hentikan pembahasan rru hp apa ancam getar aka jidit	positif	1
2	23/01/2020	herlaksana	haldiripakaga	ada benemnya juga bebua peran pelak senior isi dipr	benemnya peran pelak senior dipr	negatif	-1
3	23/01/2020	zalfahris	ernestprakasa	kalo mau gampang jangin jadi anggota dipr pak bikin akun report apa tinggal nyaring konten konting show sama stan	kalo gampang anggota dipr bikin akun report apa tinggal nyaring konten konting show stan	negatif	-1
4	23/01/2020	kariembah	fadison	pemerintah dan dipr benemngi tolak anekasi surat pual kementerian luar negeri republik indonesia bkaupdr dipr	pemerintah dipr benemngi tolak anekasi surat pual kementerian negeri republik indonesia bkaupdr dipr	netral	0
5	23/01/2020	stamou_h	wilanggersi_084	ruh isi dia nastr anggota dipr komisi vi dari traksi pual dimalak arangan tnd ditud mind sava pat	nastr anggota dipr komisi vi dari traksi pual dimalak arangan tnd ditud mind sava pat	netral	0
6	23/01/2020	halowdy	chitdyr_dpr	bagayannya mana rih boleh ngakak kalo bayar	bagayannya rih ngakak kalo bayar	negatif	-1
7	23/01/2020	apronemaka	janam_jip	peranan kritis bika hama sama hannya th memund saya pual bapag pialak maah pak jakes	peranan kritis pual bapag pialak maah pak jakes	positif	1

Gambar 4.4 Tabel Visualisasi

##### 4.3 Evaluasi

Sesuai dengan hasil yang didapatkan yaitu *accuracy score* dari 20% *dataset* yang digunakan sebagai *data testing* didapatkan nilai 0,8 atau 80%, *accuracy score* didapat dari

membandingkan value dari label yang terdapat pada dataset dengan value label yang diprediksi. Kemudian terdapat classification report, pada *classification report* terdapat berbagai macam score yaitu *precision*, *recall*, *f1-score* dan *support score*.

#### 1. Precision Score

Nilai rata-rata dari precision score untuk label positif adalah 0.75 atau 75% hal ini berarti sistem mampu mendapatkan prediksi secara akurat sebesar 75% dari total prediksi yang dilakukan oleh 95 data untuk label positif. Nilai rata-rata dari precision score untuk label netral adalah 0.79 atau 79% hal ini berarti sistem mampu mendapatkan prediksi secara akurat sebesar 79% dari total prediksi yang dilakukan oleh 693 data untuk label netral. Nilai rata-rata dari precision score untuk label negatif adalah 0.82 atau 82% hal ini berarti sistem mampu mendapatkan prediksi secara akurat sebesar 82% dari total prediksi yang dilakukan oleh 758 data untuk label negatif.

#### 2. Recall Score

Recall Score adalah tingkat keberhasilan sistem dalam menemukan kembali informasi. Nilai rata-rata dari recall score untuk label positif 0.29 atau 29% hal ini berarti sistem mampu mendapatkan prediksi secara akurat sebesar 29% dari total data yang diberi label positif yang terdapat pada data testing. Nilai rata-rata dari recall score untuk label netral adalah 0.67 atau 67% hal ini berarti sistem mampu mendapatkan prediksi secara akurat sebesar 67% dari total data yang diberi label netral yang terdapat pada data testing. Nilai rata-rata dari recall score untuk secara akurat sebesar 84% dari total data yang dilabel negatif yang terdapat pada data testing.

#### 3. F1-Score

F1-Score adalah perbandingan antara precision dan recall yang digunakan untuk menentukan akurasi data yang diuji. Nilai rata-rata dari f1-score untuk label positif adalah 0.43 atau 43% hal ini menunjukkan bahwa sistem mampu memprediksi data secara akurat sebesar 43% dari total data yang diberi label positif. Nilai f1-score untuk label netral 0.70 atau 70% hal ini berarti sistem mampu memprediksi data yang diberi label netral secara akurat sebesar 70% dari total data yang diberi label. Nilai f1-score untuk label negatif 0.77 atau 77% hal ini berarti sistem dapat memprediksi data yang diberi label negatif secara akurat sebesar 77% dari total data yang diberi label negatif.

#### 4. Model Predict

Model predict adalah proses untuk mengklasifikasi data yang baru dengan hasil predict analisis yang telah dibuat. Pada *model predict* ini melakukan beberapa tahapan yang sama dengan tahapan sebelumnya tetapi terdapat satu tahapan yang berbeda yaitu tahapan *predict analysis*. Tahapan ini akan membuat data baru yang akan langsung mengetahui polaritasnya tanpa harus melalui tahapan *labeling*.

## V. KESIMPULAN

### 1. Kesimpulan

Kesimpulan berdasarkan penelitian yang telah dilakukan sebagai berikut :

1. Berdasarkan penelitian ini sistem *sentiment analyzer* telah berhasil dibuat untuk mendapatkan *tweet* dari twitter mengenai kinerja DPR.
2. Data hasil analisis berupa jumlah klasifikasi, *accuracy score* berhasil disajikan ke dalam website.
3. Berdasarkan penelitian menggunakan algoritma *naive bayes* didapatkan *accuracy score* sebesar 0.8 atau 80% hal ini berarti sistem mampu memprediksi 80% secara akurat dari total data testing sebesar 20%.
4. Berdasarkan penelitian hasil analisis dapat digunakan untuk memprediksi dataset baru tanpa harus dilakukan *labeling* terlebih dahulu.
5. Berdasarkan hasil analisis sistem mendapatkan klasifikasi *tweet* dari twitter mengenai DPR sebanyak 95 positif, 693 netral dan 758 negatif dari data hasil *crawling* sebanyak 1546.

### 2. Saran

Saran yang diberikan untuk penelitian selanjutnya adalah sebagai berikut :

1. Sistem *sentiment analysis* ini sebaiknya menggunakan *database* agar mampu menyimpan data dalam jumlah besar.
2. Visualisasi hasil analisis sebaiknya menggunakan fitur filter pada bagian tabel berdasarkan tanggal *crawling* sehingga mempermudah untuk melihat *sentiment analysis*.

## DAFTAR PUSTAKA

- [1] B. M. Pintoko and K. M. L, "Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naïve Bayes Classifier," in *e-Proceeding of Engineering*, 2018, vol. 5, no. 3, pp. 8121–8130.
- [2] "Tentang API Twitter." [Online]. Available: <https://help.twitter.com/id/rules-and-policies/twitter->

- api. [Accessed: 24-Sep-2020].
- [3] A. T. Jaka, "Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining," 2015.
- [4] "index | TIOBE - The Software Quality Company." [Online]. Available: <https://www.tiobe.com/tiobe-index/>. [Accessed: 26-Aug-2020].
- [5] M. C. Kirana, N. P. Perkasa, M. Z. Lubis, and M. Fani, "Visualisasi Kualitas Penyebaran Informasi Gempa Bumi di Indonesia Menggunakan Twitter," *J. Appl. Informatics Comput.*, vol. 3, no. 1, pp. 23–32, 2019.
- [6] Marco Bonzanini, *Mastering Social Media Mining with Python*, vol. 53, no. 9, 2019.
- [7] C. D. Sherrill, *Python Programming Workshop*, no. September, 2009.
- [8] D. Heksaputra, Y. Azani, Z. Naimah, and L. Iswari, "Penentuan Pengaruh Iklim Terhadap Pertumbuhan Tanaman dengan Naïve Bayes," in *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, 2013, pp. 34–36.
- [9] I. Fahrur Rozi, A. Taufika Firdausi, and K. Islamiyah, "Analisis Sentimen Pada Twitter Mengenai Pasca Bencana Menggunakan Metode Naïve Bayes Dengan Fitur N-Gram," *J. Inform. Polinema*, vol. 6, no. 2, pp. 33–39, 2020.
- [10] Y. Cahyono, "Analisis Sentiment pada Sosial Media Twitter Menggunakan Naïve Bayes Classifier dengan Feature Selection Particle Swarm Optimization dan Term Frequency," *J. Inform. Univ. Pamulang*, vol. 2, no. 1, p. 14, 2017.
- [11] "Project Jupyter | Home." [Online]. Available: <https://jupyter.org/>. [Accessed: 26-Aug-2020].
- [12] K. Sahoo, A. K. Samal, J. Pramanik, and S. K. Pani, "Exploratory data analysis using python," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 4727–4735, 2019.
- [13] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007.
- [14] F. Nurhuda, S. Widya Sihwi, and A. Doewes, "Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier," *J. Teknol. Inf. ITSmart*, vol. 2, no. 2, p. 35, 2016.
- [15] E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Ureña-López, and A. R. Montejó-Ráez, "Sentiment analysis in Twitter," *Nat. Lang. Eng.*, vol. 20, no. 1, pp. 1–28, 2014.
- [16] J. V. Praneeth Sai and B. Balachander, "Sentimental analysis of twitter data using tweepy and textblob," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 3, pp. 6537–6544, 2020.