

PERBANDINGAN KINERJA KLASIFIKASI SENTIMEN ULASAN PRODUK PEMBELIAN BERAS DI MARKETPLACE SHOPEE

Dedy Sugiarto¹, Syandra Sari¹, Anung B. Ariwibowo¹, Fitria Nabilah Putri², Dimmas Mulya², Tasya Aulia¹, Arviandri Naufal Zaki²

¹ Program Studi Sistem Informasi, Fakultas Teknologi Industri, Universitas Trisakti, Jakarta, Indonesia

² Program Studi Teknik Informatika, Fakultas Teknologi Industri, Universitas Trisakti, Jakarta, Indonesia

dedy@trisakti.ac.id, syandra_sari@trisakti.ac.id, anung@trisakti.ac.id, fitria064001900002@std.trisakti.ac.id, dimmas.mulya@trisakti.ac.id, tasya065001800022@std.trisakti.ac.id, arviandri064001800035@std.trisakti.ac.id

ARTICLE INFO

ABSTRACT

Keywords

Rice
Product review
Sentiment analysis
Support vector machine
Naïve bayes
Logistic regression
k-nearest neighbor
TF-IDF
Bag of word

This study aims to compare the performance of product purchase sentiment classification in market place shopee using four classification algorithms, namely support vector machine (SVM), naïve bayes (NB), logistic regression (LR), k-nearest neighbor (KNN) and associated with the feature extraction model used, namely term frequency - inverse document. frequency (TF-IDF) and bag of word (BOW). Data collection was carried out by extracting rice product review data through the Shopee website using a web scraping technique which was then saved in the form of a file with CSV format. The number of product reviews obtained is 3531 reviews and after pre-processing through the elimination of duplicate reviews, there are 464 reviews with details 16.17% having a negative label (rating 1 or 2), 15.52% having a neutral label (rating 3), and 68.32% have a positive label (rating 4 or 5). The composition of the rankings shows that the data is not balanced. The experimental results show that the combination of LR with TF-IDF shows the best performance with an accuracy of 80%.

1. Pendahuluan

Pembeli dapat memberikan ulasan setelah menerima barang dan sekaligus memberikan rating penilaian ketika belanja aneka produk secara online di sebuah marketplace. Semakin bagus rating yang dimiliki, maka semakin baik pula reputasi yang dimiliki produk tersebut. Penelitian yang dilakukan [1][2] menunjukkan bahwa ulasan konsumen dan penilaiannya memiliki pengaruh positif terhadap keputusan pembelian melalui marketplace Shopee. Ulasan yang positif dari pelanggan dapat meningkatkan persepsi calon pembeli dan meningkatkan minat calon pembeli untuk membeli produk. Begitu pula sebaliknya, ulasan negatif dari pelanggan menjadi dapat membatalkan keputusan calon pembeli untuk melakukan pembelian.

Beberapa penelitian telah dilakukan untuk menganalisis sentiment ulasan produk dan membuat prediksi kelas sentiment ulasan produk berdasarkan kelas positif, netral atau negative. Berdasarkan percobaan yang telah dilakukan, didapatkan bahwa dari dua metode yang digunakan yaitu Naive Bayes (NB) dan Support Vector Machine (SVM), akurasi metode SVM lebih tinggi dari metode NB dengan nilai akurasi berturut-turut sebesar 85% dan 79% dan waktu eksekusi SVM lebih pendek dari NB [3]. Dalam kasus lain model SVM dan model NB identik dalam hal kinerja klasifikasinya [4].

Penelitian analisis sentiment yang lebih khusus terkait ulasan produk pada marketplace Shopee antara dilakukan oleh [5] dimana didapatkan nilai akurasi sebesar 85% dengan menggunakan algoritma NB. Ulasan produk yang diteliti adalah *Xiaomi Redmi Note 9*. Beberapa penelitian yang melakukan evaluasi sentiment ulasan produk pada marketplace Shopee pada umumnya menggunakan algoritma NB sebagai mesin klasifikasinya serta model term frequency – inverse document frequency (TF-IDF) sebagai ekstraksi fitur yang mengubah data teks menjadi numerik [6][7] serta algoritma NB yang integrasikan dengan algoritma K-Means sebelum kelas sentiment diprediksi [8]. Hal ini sejalan dengan pernyataan

yang disampaikan [9] dalam konteks penelitian analisis sentiment dalam sektor pertanian bahwa algoritma NB merupakan algoritma yang paling banyak digunakan.

Hasil penyelidikan oleh [10] menunjukkan bahwa SVM dan LR adalah model yang paling umum digunakan kasus prediksi penyakit kronis yang diderita seseorang. Penelitian lain menunjukkan bahwa SVM paling sering diterapkan disusul oleh algoritma NB [11]. KNN juga disebutkan dalam sebuah penelitian merupakan algoritma yang paling banyak digunakan dalam kasus pembelajaran mesin [12]

Berdasarkan uraian latar belakang di atas, dirasakan perlu untuk membandingkan lebih banyak mesin klasifikasi yang digunakan untuk prediksi kelas sentiment dari ulasan produk di marketplace shopee. penelitian ini bertujuan untuk membandingkan akurasi klasifikasi sentiment pembelian produk di market place shopee menggunakan empat mesin klasifikasi yaitu SVM, NB, regresi logistik (LR) dan *k-nearest neighbor* (KNN) serta dikaitkan dengan model ekstraksi fitur yang digunakan yaitu TF-IDF dan *bag of word* (BOW).

2. Metodologi Penelitian

Pengumpulan data dilakukan dengan mengekstrak data produk melalui website Shopee menggunakan teknik web scraping yang kemudian disimpan dalam bentuk file dengan format CSV. Produk yang akan dianalisis yaitu produk beras. Dataset terdiri atas 4 *field* yaitu *url*, *title*, *comment* serta rating seperti dapat dilihat pada Gambar 1.

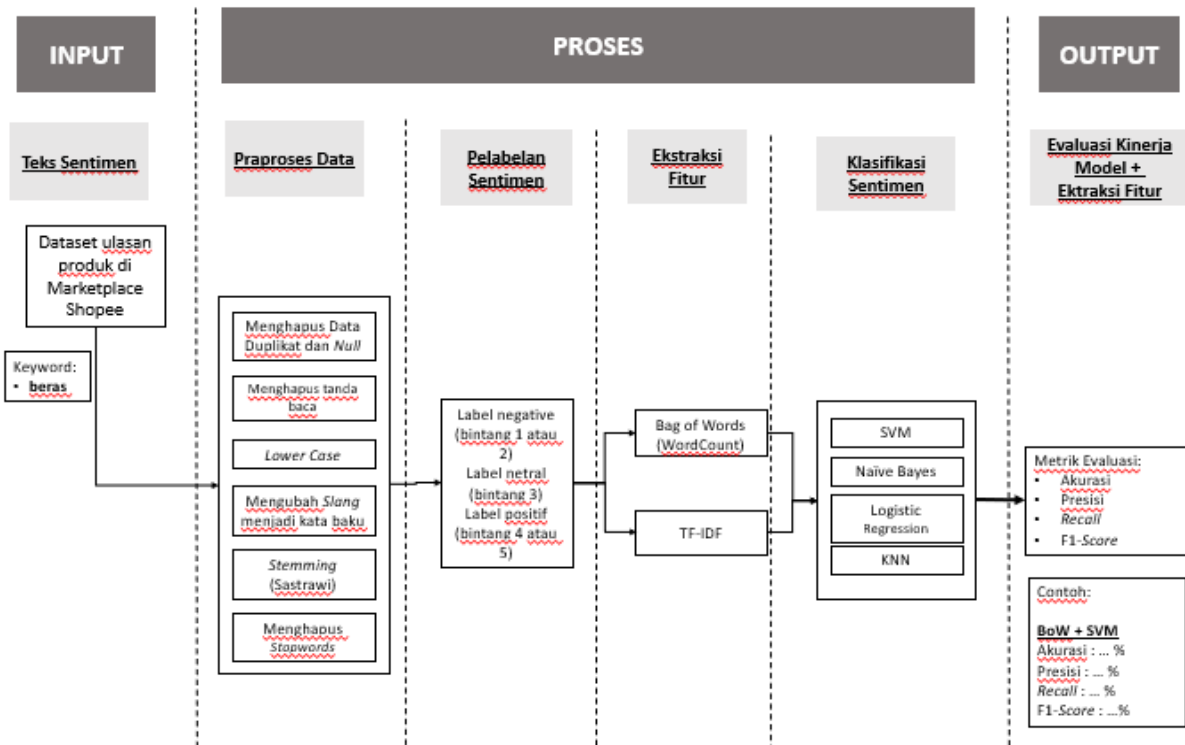
url	title	comment	rating
https://shopee.co.id/beras-online-beras-idola-...	Beras Online - Beras Idola 5kg/10kg Beras Pule...	Terimakasih atas paket produknya sudah sampai ...	5 Bintang (44)
https://shopee.co.id/beras-pera-cap-bung-i.137...	Beras Pera Cap Bunga 5KG	Bubble wrapnya tebal banget, beli ini untuk ke...	5 Bintang (9)
https://shopee.co.id/asagri-beras-organik-1-kg...	Asagri Beras Organik 1 Kg Premium Pulen Putih ...	mau beli yg 500gr buat coba2 dulu, tapi dikiri...	3 Bintang (4)
https://shopee.co.id/asagri-beras-merah-organi...	Asagri Beras Merah Organik 1 Kg Untuk Diet Ren...	Rasa: blm tau\nHarga: oke\nKualitas: oke\nBaru...	5 Bintang (2,6RB)
https://shopee.co.id/asagri-beras-merah-organi...	Asagri Beras Merah Organik 1 Kg Untuk Diet Ren...	Rasa: blm dicoba\nHarga: agak sedikit mahal\nK...	4 Bintang (174)

Gambar 1. Sample dataset hasil scraping

Jumlah ulasan produk yang diperoleh adalah 3531 ulasan dan setelah dilakukan *pre-processing* melalui penghapusan duplikat review didapatkan 464 ulasan dengan rincian 36 ulasan memiliki rating 1 (7,76%), 39 review dengan rating 2 (8,41%), 72 ulasan dengan rating 3 (15,52%), 141 ulasan dengan peringkat 4 (30,39%) dan 176 ulasan dengan peringkat 5 (37,93%). Komposisi peringkat menunjukkan bahwa data tidak seimbang. Proses selanjutnya dalam pra-pemrosesan adalah menghapus tanda baca, ubah menjadi huruf kecil, menghapus angka, mengubah kata yang tidak baku (*slang*), menghapus kata yang tidak penting (*stopword*) dan mengubah menjadi kata dasar (*stemming*) melalui library *pysastrawi*.

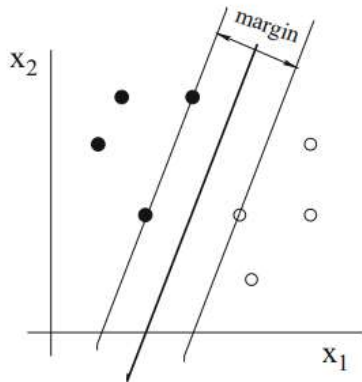
Setelah data dinilai bersih maka proses selanjutnya adalah mengubah ulasan menjadi matriks fitur yang bersifat numerik dengan dua model yaitu BOW dan TF_IDF untuk kemudian matriks tersebut

dipasok ke dalam 4 mesin klasifikasi yaitu SVM, LR, NB dan KNN. Tahapan penelitian dapat dilihat pada Gambar 2.

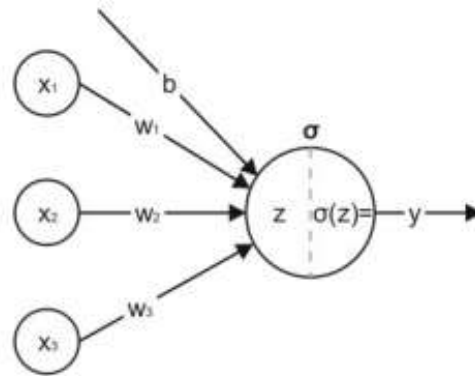


Gambar 2. Tahapan penelitian

SVM, LR, NB dan KNN merupakan beberapa algoritma pembelajaran mesin yang termasuk dalam kelompok metode klasifikasi [13][14]. Metode SVM bekerja melalui pencarian *hyperplane* dengan margin maksimal (*maximal margin hyperplane*) [13][15]. Dalam kasus klasifikasi biner, *hyperplane* adalah batas keputusan yang membedakan dua kelas di SVM. Sedangkan vektor-vektor yang terdekat dengan *hyperplane* dengan margin maksimal disebut *Support Vector*. LR adalah metode regresi dengan variable target bersifat kategorik dan merupakan gambar sederhana dari jaringan syaraf tiruan dimana terdapat dua bagian perhitungan yaitu penjumlahan terboboti dan fungsi aktivasi sigmoid atau logistic seperti dapat dilihat pada Gambar 2 [14].



Gambar 3. Hyperplane yang Memiliki Margin Maksimum dalam SVM [15]



Gambar 4. Model regresi logistik sebagai contoh jaringan syaraf tiruan sederhana [14]

Metode NB merupakan sebuah algoritma pembelajaran paling sederhana terkait pengklasifikasian yang memanfaatkan Teorema Bayes [14] dengan formula seperti dapat dilihat pada persamaan (1).

$$p(t | f) = \frac{p(f|t)p(t)}{p(f)} \dots\dots\dots(1)$$

dimana $p(t)$ adalah probabilitas prior dari suatu nilai target (label kelas), $p(f)$ adalah probabilitas dari sebuah fitur atau predictor, $p(f|t)$ adalah probabilitas dari fitur f yang diberikan target t , dan $p(t|f)$ adalah probabilitas dari target t bila diberikan fitur f tertentu.

Metode KNN merupakan metode klasifikasi sederhana yang didasarkan pada mekanisme untuk mengevaluasi kesamaan atau kedekatan antara vektor atribut. Faktanya bahwa setiap contoh dapat diwakili oleh titik dalam ruang n -dimensi yang memungkinkan untuk menghitung jarak geometris antara pasangan contoh misalnya menggunakan jarak Euclidean [14].

3. Hasil dan Pembahasan

Contoh data bersih yang telah didapatkan dari tahapan pra-pemrosesan data dapat dilihat pada Gambar 5. Terlihat bahwa data yang awalnya tertulis sebagian dengan huruf besar telah berubah menjadi semuanya huruf kecil. Kemudian angka seperti 500 juga sudah dihilangkan, juga beberapa tanda baca. Tampilan sampel dari teks yang telah diubah menjadi angka dapat dilihat pada Gambar 6 dan 7. Gambar 6 menampilkan jumlah kata unik dari kata ‘murah’, ‘cepat’, ‘enak’, ‘bagus’ dan ‘kualitas’ yang terdapat pada setiap dokumen yang dikenal dengan istilah *term frequency*. Sedangkan tampilan nilai TF-IDF dapat dilihat pada Gambar 7 yang didapatkan dari nilai *term frequency* yang dikalikan dengan nilai dari *inverse document frequency*.

comment	ratings	comment_processed	Label
Terimakasih atas paket produknya sudah sampai ...	5	terimakasih paket produk	Positif
Bubble wrapnya tebal banget beli ini untuk ke...	5	bubble wrapnya tebal banget beli butuh jual na...	Positif
mau beli yg 500gr buat coba2 dulu tapi dikiri...	3	beli gede coba kirim kg bayar ribet cocok mpas...	Netral
Rasa blm tau\nHarga oke\nKualitas oke\nBaru...	5	harga oke kualitas oke nyoba makan hitam moga ...	Positif
Rasa blm dicoba\nHarga agak sedikit mahal\nK...	4	coba harga mahal kualitas coba packing rapih k...	Positif
...

Gambar 5. Sampel data hasil pra-pemrosesan

```
count_vect_train_df[['murah', 'cepat', 'enak', 'bagus', 'kualitas']]
```

	murah	cepat	enak	bagus	kualitas
0	0	0	0	0	0
1	1	0	0	2	1
2	0	0	0	1	0
3	0	0	0	0	1
4	0	0	0	0	0
...
409	0	0	1	0	1
410	0	0	0	0	0
411	0	1	1	0	1
412	1	1	1	0	2
413	0	0	1	1	1

414 rows x 5 columns

Gambar 6. Sampel vektor fitur dari model BOW

```
[22] tfidf_train_df[['murah', 'cepat', 'enak', 'bagus', 'kualitas']]
```

	murah	cepat	enak	bagus	kualitas
0	0.000000	0.000000	0.000000	0.000000	0.000000
1	0.114454	0.000000	0.000000	0.195654	0.059088
2	0.000000	0.000000	0.000000	0.112589	0.000000
3	0.000000	0.000000	0.000000	0.000000	0.109294
4	0.000000	0.000000	0.000000	0.000000	0.000000
...
409	0.000000	0.000000	0.256873	0.000000	0.148611
410	0.000000	0.000000	0.000000	0.000000	0.000000
411	0.000000	0.183511	0.213269	0.000000	0.123384
412	0.143418	0.110122	0.127980	0.000000	0.148082
413	0.000000	0.000000	0.224793	0.215314	0.130051

414 rows × 5 columns

Gambar 7. Sampel vektor fitur dari model TF-IDF

▼ Bag of Words

```
[ ] SVM = svm.SVC(C=1.0, kernel='linear', degree=3, gamma='auto')
SVM.fit(countword_train,y_train)
y_pred = SVM.predict(countword_test)
```

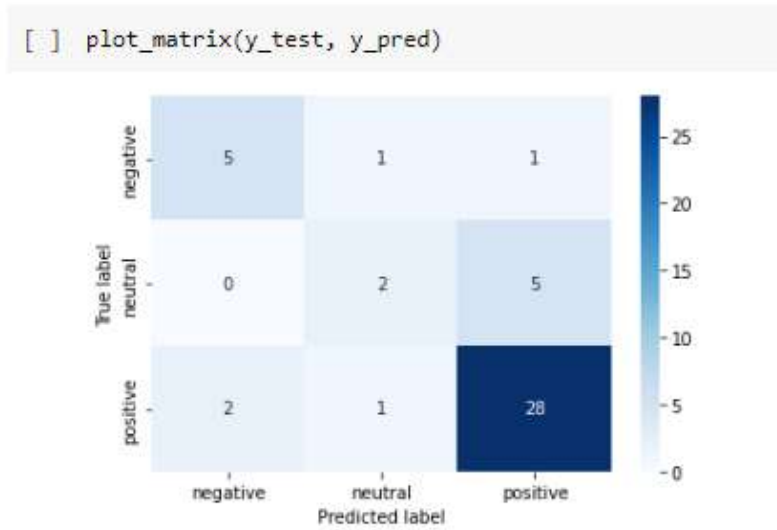
```
[ ] print(classification_report(y_test, y_pred, target_names=target_names))
```

	precision	recall	f1-score	support
negative	0.71	0.71	0.71	7
neutral	0.50	0.29	0.36	7
positive	0.82	0.90	0.86	31
accuracy			0.78	45
macro avg	0.68	0.63	0.65	45
weighted avg	0.76	0.78	0.76	45

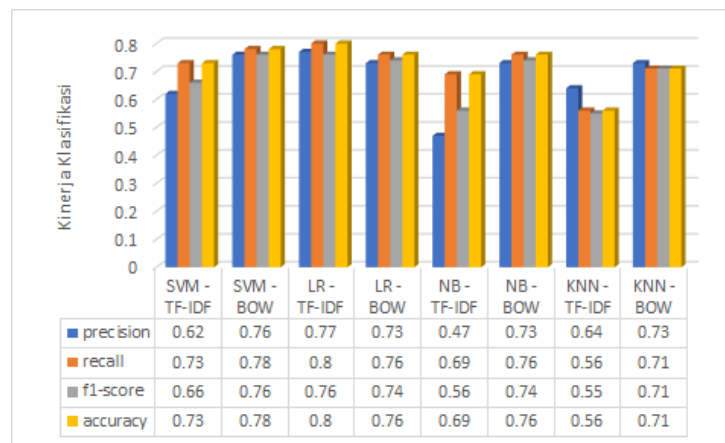
Gambar 8. Output klasifikasi dari data uji dengan metode SVM-BOW

Gambar 8 memperlihatkan hasil pengukuran kinerja dari data uji dengan menggunakan SVM dan ekstraksi fitur BOW. Terlihat bahwa F1-score tertinggi didapatkan pada data dengan label positif yaitu sebesar 0,86 dan terendah pada label netral yaitu 0,36. Confusion matrix dapat dilihat pada Gambar 9 yang merupakan hasil klasifikasi dari data uji dengan komposisi 10% dari dataset yang sudah bersih dengan menggunakan model SVM-BOW. Terlihat bahwa pada kategori netral, dari 7 ulasan yang bersifat netral hanya 2 yang dapat diprediksi dengan benar bahwa ulasan tersebut memang netral, sisanya diprediksi sebagai positif.

Secara keseluruhan perbandingan kombinasi mesin klasifikasi dan ekstraksi fitur dapat dilihat pada Gambar 10. Kinerja tertinggi didapatkan dari kombinasi LR dengan TF-IDF dan terlihat terendah adalah NB-TF-IDF serta KNN-TF-IDF. Secara umum LR dan SVM menunjukkan hasil yang lebih baik dibandingkan KNN dan NB. Hasil ini mengkonfirmasi penelitian sebelumnya yang menunjukkan SVM dan LR lebih baik dibandingkan metode shallow learning lainnya [16].



Gambar 9. Confusion matrix dari SVM-BOW



Gambar 10. Perbandingan kinerja dari seluruh mesin klasifikasi dan ekstraksi fitur yang digunakan

4. Kesimpulan

Perbandingan akurasi klasifikasi sentiment pembelian produk di marketplace Shopee yang menggunakan empat mesin klasifikasi yaitu SVM, NB, LR dan KNN serta dikaitkan dengan model TF-IDF dan BOW menunjukkan bahwa LR-TF-IDF menunjukkan kinerja terbaik dengan nilai akurasi sebesar 80%. Penelitian lanjutan perlu dilakukan dengan melakukan proses *hyperparameter tuning* pada keempat mesin klasifikasi yang digunakan serta perbandingan komposisi dari data latih dan data uji.

• Ucapan Terimakasih

Ucapan terima kasih disampaikan kepada Universitas Trisakti dan Fakultas Teknologi Industri atas pendanaan yang diberikan sehingga penelitian ini dapat diselesaikan.

Daftar Pustaka

- [1] A. N. Ardianti and M. A. Widiartanto, "Pengaruh Online Customer Review dan Online Customer Rating terhadap Keputusan Pembelian melalui Marketplace Shopee .," *J. Ilmu Adm. Bisnis*, pp. 1–11, 2019.
- [2] R. S. Damayanti, "Pengaruh Online Costumer Review and Rating, E-Service Quality dan Price Terhadap Minat Beli pada Online Marketplace (Studi Empiris Pada Mahasiswa Universitas Muhammadiyah Magelang)," *Pros. 2nd Bus. Econ. Conf. Util. Mod.*, pp. 684–693, 2019, [Online]. Available: <http://journal.ummg1.ac.id/index.php/conference/article/download/3559/1731>.
- [3] Y. Basani, H. V. Sibuea, S. Ida Patona Sianipar, and J. Presly Samosir, "Application of Sentiment Analysis on Product Review E-Commerce," *J. Phys. Conf. Ser.*, vol. 1175, no. 1, 2019, doi: 10.1088/1742-6596/1175/1/012103.
- [4] X. Fang and J. Zhan, "Sentiment analysis using product review data," *J. Big Data*, vol. 2, no. 1, 2015, doi: 10.1186/s40537-015-0015-2.
- [5] L. O. Sihombing, H. Hannie, and B. A. Dermawan, "Sentimen Analisis Customer Review Produk Shopee Indonesia Menggunakan Algoritma Naive Bayes Classifier," *Edumatic J. Pendidik. Inform.*, vol. 5, no. 2, pp. 233–242, 2021, doi: 10.29408/edumatic.v5i2.4089.
- [6] Y. R. Saputri and H. Februariyanti, "Jurnal Mantik SENTIMENT ANALYSIS ON SHOPEE E-COMMERCE USING THE," vol. 6, no. 36, pp. 1349–1357, 2022.
- [7] R. Kosasih and A. Alberto, "Sentiment analysis of game product on shopee using the TF-IDF method and naive bayes classifier," *Ilk. J. Ilm.*, vol. 13, no. 2, pp. 101–109, 2021, doi: 10.33096/ilkom.v13i2.721.101-109.
- [8] T. Hariguna, W. M. Baihaqi, and A. Nurwanti, "Sentiment Analysis of Product Reviews as A Customer Recommendation Using the Naive Bayes Classifier Algorithm," *IJIS Int. J. Informatics Inf. Syst.*, vol. 2, no. 2, pp. 48–55, 2019, doi: 10.47738/ijis.v2i2.13.
- [9] J. Novák, P. Benda, E. Šilerová, J. Vaněk, and E. Kánská, "Sentiment Analysis in Agriculture," *Agris On-line Pap. Econ. Informatics*, vol. 13, no. 1, pp. 121–130, 2021, doi: 10.7160/aol.2021.130109.
- [10] J. Sen *et al.*, *Machine Learning: Algorithms, Models, and Applications*, no. January. 2022.
- [11] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," *BMC Med. Inform. Decis. Mak.*, vol. 19, no. 1, pp. 1–16, 2019, doi: 10.1186/s12911-019-1004-8.
- [12] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, "Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction," *Sci. Rep.*, vol. 12, no. 1, pp. 1–11, 2022, doi: 10.1038/s41598-022-10358-x.
- [13] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Boston: Pearson Education Inc., 2006.
- [14] S. Skansi, *Introduction to deep learning - From Logical Calculus to Artificial Intelligence*. Cham: Springer, 2020.
- [15] M. Kubat, *An Introduction to Machine Learning*. 2017.
- [16] T. Sabri, O. El Beggar, and M. Kissi, "Comparative study of Arabic text classification using feature vectorization methods," *Procedia Comput. Sci.*, vol. 198, no. 2021, pp. 269–275, 2021, doi: 10.1016/j.procs.2021.12.239.