

PREDIKSI TIPE KEPERIBADIAN MBTI ARTIS K-POP BERDASARKAN CAPTION INSTAGRAM MENGGUNAKAN WORD2VEC DAN LONG-SHORT TERM MEMORY (LSTM)

Alfian Hakim^{*1}, Satrio Hadi Wijoyo², Nanang Yudi Setiawan³

^{1,2,3}Universitas Brawijaya, Malang

Email : ¹alfianhkm@student.ub.ac.id, ²satriohadi@ub.ac.id, ³nanang@ub.ac.id

*Penulis Korespondensi

(Naskah masuk: 13 Februari 2023, diterima untuk diterbitkan: 26 September 2023)

Abstrak

Myers-Briggs Type Indicator (MBTI) adalah metode pengujian psikologi yang membedakan kepribadian seseorang. MBTI termasuk pembagian tipe kepribadian yang paling populer di dunia, termasuk di Korea Selatan. Tren MBTI di Korea Selatan juga dimanfaatkan oleh para artis K-Pop untuk berbagi tipe MBTI sehingga bisa mendekatkan hubungan antara penggemar dan idolanya. Salah satu media sosial yang umum digunakan oleh artis K-Pop adalah Instagram. Penelitian ini mencoba membuat model klasifikasi tipe kepribadian berdasarkan *caption* Instagram artis K-Pop menggunakan *Word2Vec* dan *Long-Short Term Memory* (LSTM). Terdapat 118.401 data *caption* yang sudah dibersihkan melalui serangkaian langkah *pre-processing* dari 458 artis. Distribusi tipe kepribadian menunjukkan bahwa target tidak seimbang sehingga perlu dilakukan penanganan yaitu penggeseran *threshold* yang dilakukan pasca pemodelan. Evaluasi kombinasi model menghasilkan nilai *macro f1* 0,65 pada data artis, dengan rincian model *Extroversion-Introversion*, *Sensing-Intuition*, *Thinking-Feeling* memiliki nilai *macro f1* yang sama yaitu 0,88, sedangkan model *Judging-Perceiving* memiliki nilai *macro f1* yang sedikit lebih baik yaitu 0,90. Model diimplementasikan dalam aplikasi web Streamlit agar penggemar K-Pop dapat menggunakannya untuk memprediksi tipe MBTI dengan masukan *caption* Instagram. Aplikasi web dievaluasi menggunakan kuesioner *System Usability Scale* (SUS) dan mendapatkan skor 84,55 sehingga sudah termasuk kategori *acceptable*.

Kata kunci: MBTI, K-Pop, *caption* Instagram, *Word2Vec*, LSTM

PREDICTING MBTI PERSONALITY TYPES OF K-POP ARTISTS USING WORD2VEC AND LONG-SHORT TERM MEMORY (LSTM)

Abstract

Myers-Briggs Type Indicator (MBTI) is a psychological test that distinguishes a person's personality. The MBTI is one of the most popular personality types in the world, including South Korea. The MBTI trend in South Korea is also used by K-Pop artists to share their MBTI types so they could be closer to their fans. One of the social media commonly used by K-Pop artists is Instagram. This study tries to develop a personality type classification model based on Instagram captions of K-Pop artists using *Word2Vec* and *Long-Short Term Memory* (LSTM). There are 118,401 caption data that have been cleaned through a series of *pre-processing* steps from 458 artists. The distribution of personality types shows that the target is not balanced, so it is necessary to handle imbalanced data, namely shifting the threshold after modeling. Evaluation of the combination model yields a *macro f1* value of 0.65 in the artist data, with the details of the *Extroversion-Introversion*, *Sensing-Intuition*, *Thinking-Feeling* models having the same *macro f1* value of 0.88, while the *Judging-Perceiving* model has a slightly better *macro f1* value of 0.90. The model is deployed in the Streamlit web application so that K-Pop fans can use it to predict the MBTI type by inputting Instagram captions. The web application is evaluated using the *System Usability Scale* (SUS) questionnaire and gets a score of 84.55 so it is considered *acceptable*.

Keywords: MBTI, K-Pop, Instagram captions, *Word2Vec*, LSTM

1. PENDAHULUAN

Myers-Briggs Type Indicator (MBTI) merupakan sebuah metode psikologi yang digunakan

untuk memetakan dan membedakan karakteristik kepribadian individu. MBTI disebut sebagai uji kepribadian paling populer di dunia, digunakan di

115 negara, tersedia dalam 29 bahasa, dan digunakan oleh 88 perusahaan yang termasuk Fortune 100 (Kerwin, 2018). Popularitas MBTI juga sampai ke Korea Selatan. Menurut dosen Psikologi Dankook University, Lim Myoung-ho, muda-mudi Korea Selatan menggunakan MBTI untuk mempercepat proses mencari kecocokan dengan calon pasangan, sehingga mereka gemar berbagi informasi tipe kepribadiannya satu sama lain. Pengelompokan kepribadian juga memberikan rasa kesatuan antar tiap tipe kepribadian (Yeung & Seo, 2022).

Artis K-Pop ikut menggunakan MBTI untuk memperdekat hubungan dengan penggemarnya. Banyak artis K-Pop mengungkapkan tipe kepribadiannya kepada penggemar melalui berbagai media sosial. Adapun situs web yang berisi profil artis K-Pop yang selalu mencantumkan informasi MBTI, situs web tersebut adalah kprofiles.com dan kpopmap.com. Penggemar biasanya mengikuti akun media sosial idolanya seperti Instagram sebagai bentuk pendekatan dan memahami perilaku idolanya.

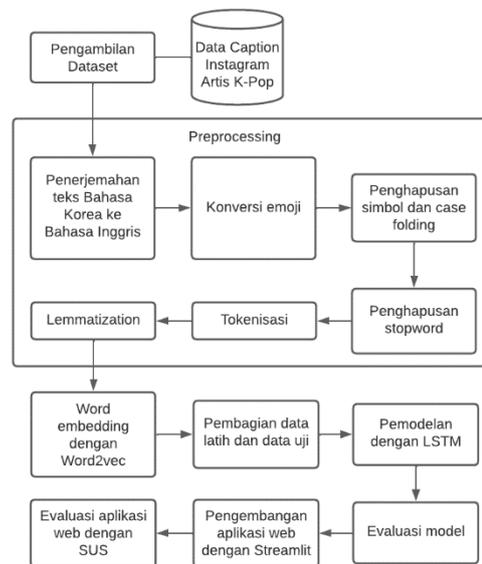
Berdasarkan penelitian sebelumnya, tipe kepribadian seseorang dapat memengaruhi perilakunya dalam berinternet, dan penulisan *caption* adalah salah satu bentuk perilaku berinternet (Choong & Varathan, 2021). Data teks berupa *caption* Instagram dapat dipelajari dengan pembelajaran mesin yang biasa disebut *Natural Language Processing* (NLP). NLP mampu mengenali aspek linguistik seperti kata benda, kata kerja, kata sifat dan bahkan struktur kalimat (Redhum dkk., 2018). NLP merupakan salah satu pendekatan dalam domain luas *data mining*. *Data mining* banyak digunakan sebagai solusi permasalahan dalam berbagai bidang, contohnya di bidang pendidikan seperti penelitian yang dilakukan Herlambang & Wijoyo (2018) yang mengklasifikasikan sumber belajar, Wicaksono dkk. (2019) yang mengklasifikasikan siswa *slow learner*, dan Herlambang dkk. (2019) yang mengklasifikasikan pencapaian siswa.

Salah satu metode yang bisa digunakan dalam NLP adalah pembobotan Word2Vec dipadukan dengan *Long-Short Term Memory* (LSTM). Penelitian yang dilakukan Muhammad dkk. (2021) dan Widayat (2021) dengan metode tersebut menghasilkan akurasi yang sangat baik yaitu di atas 85%. Penulis dalam penelitian ini mengajukan pengembangan model klasifikasi tipe kepribadian MBTI berdasarkan data teks *caption* akun Instagram artis-artis K-Pop menggunakan metode LSTM dan pembobotan Word2Vec. Terdapat empat model yang akan dikembangkan sesuai dengan aspek MBTI. Penelitian ini diharapkan dapat membantu penggemar K-Pop untuk mengetahui tipe kepribadian artis-artis idolanya yang belum diketahui tipe kepribadiannya, penulis mengembangkan sebuah aplikasi web sederhana

dalam penelitian ini demi memenuhi kebutuhan tersebut.

2. METODE PENELITIAN

Langkah-langkah yang dilakukan penulis dalam melakukan penelitian digambarkan dalam bagan berikut.



Gambar 1. Diagram Alir Penelitian

2.1. Pengambilan Dataset

Data *caption* Instagram diambil dengan memanfaatkan *library* *instaloader* yang tersedia pada bahasa pemrograman Python. Penulis membuat daftar nama artis K-Pop dengan nama penggunanya di Instagram terlebih dahulu sebelum mengambil data *caption*. Data MBTI dari para artis K-Pop diperoleh dengan cara manual melalui situs kprofiles.com dan kpopmap.com. Kriteria artis K-Pop yang menjadi subjek penelitian adalah sebagai berikut.

1. Memiliki laman profil yang dapat diakses di kprofiles.com atau kpopmap.com serta diketahui tipe kepribadiannya.
2. Memiliki akun Instagram pribadi, tidak satu akun dengan *member* lain dalam satu grup.

2.2. Pre-processing

Caption yang tersedia di profil artis K-Pop banyak berisi Bahasa Korea, agar mempermudah analisis maka teks tersebut diterjemahkan ke Bahasa Inggris terlebih dahulu memanfaatkan fitur dalam Google Sheet berupa fungsi `GOOGLETRANSLATE()`. Akurasi penerjemahan sepenuhnya terbatas pada algoritma Google Translate sehingga hasil terjemahan berpotensi tidak sesuai konteks atau kurang tepat penerjemahannya, penerjemahan menggunakan cara ini dipilih karena

jumlah data yang sangat banyak sehingga akan membutuhkan sumber daya berlebih jika menggunakan jasa penerjemah. Penggunaan cara penerjemahan lain sangat berpotensi mengubah hasil *pre-processing* dan hasil penelitian karena diksi yang digunakan sangat mungkin berbeda-beda. Data hasil terjemahan dikonversi menggunakan konversi emoji untuk mengubah simbol emoji menjadi teks. Data teks yang sudah dikonversi emoji dibersihkan nilai kosong dan duplikatnya serta simbol-simbol lalu huruf kapital diubah menjadi huruf kecil (*case folding*). Langkah selanjutnya adalah penghapusan *stopword*, *stopword* adalah daftar kata yang dianggap tidak mengandung informasi penting (Muhammad dkk., 2021). Langkah berikutnya adalah tokenisasi yaitu proses pemecahan kalimat menjadi daftar kata. Tahap terakhir pada *pre-processing* adalah *lemmatization*, berguna untuk mencari akar kata. Sampel *caption* yang diolah dalam *pre-processing* dapat dilihat pada Tabel 1.

Tabel 1. Sampel *Pre-processing*

Proses	Hasil
(belum diproses)	🍹카테일 안주가 친절한 사장님 맛집🍹 엄마 원픽 카테일 바🍹 너무 맘에 들어해서 다음 휴가 때 다시 가기로 했는데 엄마 언제와 🏠 집에 있다가도 롱티 생각나면 별떡 일어나서 감ㅋㅋㅋㅋ
Terjemah ke Bahasa Inggris	🍹Cocktail snacks are kind boss restaurant 🍹Mom One Pick Cocktail Bar 🍹I liked it so much that I decided to go back to the next vacation.
Konversi emoji	fire Cocktail snacks are kind boss restaurant fire Mom One Pick Cocktail Bar cocktail_glass I liked it so much that I decided to go back to the next vacation.
Penghapusan simbol dan <i>case folding</i>	fire cocktail snacks are kind boss restaurant fire mom one pick cocktail bar cocktail glass i liked it so much that i decided to go back to the next vacation
Penghapusan <i>stopword</i>	fire cocktail snack kind boss restaurant fire mom one pick cocktail bar cocktail glass liked much decided go back next vacation
Tokenisasi	['fire', 'cocktail', 'snack', 'kind', 'boss', 'restaurant', 'fire', 'mom', 'one', 'pick', 'cocktail', 'bar', 'cocktail', 'glass', 'liked', 'much', 'decided', 'go', 'back', 'next', 'vacation']
<i>Lemmatization</i>	['fire', 'cocktail', 'snack', 'kind', 'boss', 'restaurant', 'fire', 'mom', 'one', 'pick', 'cocktail', 'bar', 'cocktail', 'glass', 'like', 'much', 'decide', 'go', 'back', 'next', 'vacation']

2.3. Word Embedding dengan Word2Vec

Tahap ini dilakukan untuk mengkonversi data berupa daftar kata menjadi daftar bilangan atau biasa disebut vektor sehingga dapat diproses oleh komputer.

2.4. Pembagian Data Latih dan Data Uji

Tahap berikutnya adalah pembagian data menjadi data latih dan data uji. Skala pembagian yang digunakan adalah 80:20 dengan data latih lebih besar.

2.5. Pemodelan dengan LSTM

Penulis mengembangkan model berdasarkan data latih hasil konversi. Model yang dikembangkan tersusun dari beberapa *layer*, yaitu:

3. *Embedding layer*, berisi vektor-vektor yang sudah diproses sebelumnya.
4. *Dropout layer*, berfungsi untuk mencegah *overfitting* dengan cara mengabaikan neuron secara acak.
5. *Encoding layer* berdasarkan model LSTM.
6. *Dense layer*, lapisan di mana setiap neuron terkoneksi dengan semua neuron di lapisan setelahnya.

2.6. Evaluasi Model

Evaluasi dilakukan dengan dua skema yaitu dengan menguji model menggunakan data uji (*caption*) dan juga menguji pada data artis (kumpulan *caption*) untuk mengukur seberapa akurat model bekerja. Evaluasi dengan dua skema ini dilakukan untuk membandingkan performa model dalam memprediksi kepribadian berdasarkan satu *caption* dengan berdasarkan kumpulan *caption*, apakah prediksi berdasarkan satu *caption* lebih sulit daripada berdasarkan kumpulan *caption*. Selain dilakukan evaluasi pada tiap model, dilakukan pula evaluasi untuk kombinasi keempat model mengingat penggunaan tipe MBTI adalah dengan mengkombinasikan sifat dari empat domain. Metriks yang dihitung antara lain akurasi dan juga *macro f1* (nilai rata-rata *f1-score* kedua kelas).

2.7. Pengembangan Aplikasi Web

Model yang sudah dikembangkan akan diimplementasikan dalam sebuah aplikasi web yang bisa digunakan oleh penggemar K-Pop untuk memprediksi tipe kepribadian artis K-Pop. Aplikasi web dikembangkan menggunakan Streamlit. Aplikasi web akan memiliki halaman prediksi yang dapat digunakan oleh pengguna untuk memasukkan data *caption* sendiri kemudian keluar hasil prediksi tipe kepribadiannya. Halaman lain akan memuat data-data *caption* dari seorang artis K-Pop dan hasil prediksi tipe kepribadian berdasarkan data-data *caption* tersebut.

2.8. Evaluasi Aplikasi Web dengan SUS

Aplikasi web yang telah dibuat akan dievaluasi menggunakan pengujian *System Usability Scale* (SUS). Responden yang akan mengisi kuesioner pengujian SUS adalah penggemar K-Pop. Evaluasi

ini akan menghasilkan nilai yang menentukan apakah aplikasi sudah dapat diterima atau belum.

3. LANDASAN KEPUSTAKAAN

3.1. Myers-Briggs Type Indicator (MBTI)

Tipe kepribadian MBTI ditentukan dengan cara mengisi kuesioner yang berisi sejumlah pertanyaan singkat. MBTI mengadopsi dikotomi preferensi yang diajukan oleh Carl Jung yaitu *Extroversion-Introversion (E/I)*, *Sensing-Intuition (S/N)*, dan *Thinking-Feeling (T/F)*, Myers dan Briggs kemudian menambahkan preferensi keempat yaitu *Judging-Perceiving (J/P)* (Kerwin, 2018). Berdasarkan empat domain tersebut, sebagai contoh jika hasil evaluasi psikologis individu menunjukkan cenderung pada *Introversion (I)*, *Intuition (N)*, *Feeling (F)*, dan *Judging (J)*, maka individu tersebut akan diklasifikasikan sebagai tipe kepribadian INFJ. Dengan demikian, terdapat total 16 tipe kepribadian yang dihasilkan dari kombinasi empat domain ini. Karakteristik dikotomi preferensi ini dipaparkan lebih lanjut pada Tabel 1.

Tabel 2. Karakteristik Preferensi MBTI

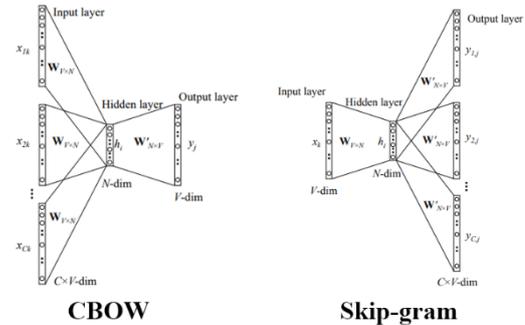
<i>Extroversion (E)</i>	<i>Introversion (I)</i>
Mendapat energi dari bertemu dengan orang lain, menikmati tugas-tugas yang bervariasi, bertempo cepat, dan baik dalam <i>multitasking</i> .	Seringkali lebih menyukai bekerja sendiri atau dalam kelompok kecil, bertempo semauanya, dan berfokus pada satu tugas dalam satu waktu.
<i>Sensing (S)</i>	<i>Intuition (N)</i>
Berpikir realistis dengan fokus pada fakta dan informasi terperinci, menerapkan pengetahuan umum dan pengalaman terdahulu untuk menyelesaikan masalah secara praktis.	Berpikir imajinatif dengan fokus pada kemungkinan-kemungkinan dan hal yang lebih besar, baik dalam mengenali pola dan nilai, serta mencari solusi kreatif terhadap sebuah masalah.
<i>Thinking (T)</i>	<i>Feeling (F)</i>
Menentukan keputusan berdasarkan analisis logis, objektif menilai kelebihan dan kekurangan, dan menghargai kejujuran, konsistensi, serta keadilan.	Menentukan keputusan berdasarkan perasaan, sensitif dan kooperatif, mempertimbangkan orang lain dalam pengambilan keputusan.
<i>Judging (J)</i>	<i>Perceiving (P)</i>
Terorganisir dan tertata, senang membuat rencana dan menjalaninya, serta tak keberatan mengikuti aturan.	Senang dengan pilihan terbuka, bertindak secara spontan, fleksibel dengan jadwal dan aturan.

Sumber: Tieger dkk. (2014) dalam Amirhosseini & Kazemian (2020)

3.2. Word2Vec

Word embedding adalah proses konversi sebuah kata menjadi nilai bilangan (vektor) sehingga dapat diolah oleh komputer. Word2Vec dapat menghitung kemiripan antar kata dan juga bekerja cepat pada data yang besar. Model arsitektur Word2Vec bekerja menggunakan *Neural Network (NN)*, mengkalkulasi *cosine similarity* antara vektor-vektor kata. Kata yang mirip akan memiliki nilai *cosine similarity* yang tinggi. Word2Vec pertama kali dicetuskan oleh Mikolov dkk. (2013). Terdapat dua model arsitektur pada Word2Vec, yaitu model *Continuous Bag of*

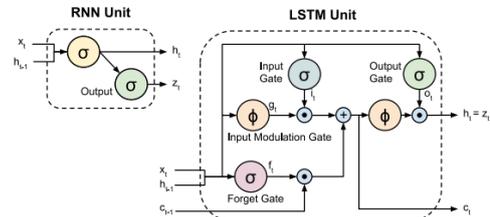
Words (CBOW) dan model *Skip-gram*. Cara kerja kedua model ini saling berkebalikan, masukan untuk CBOW adalah sejumlah kata dan luarannya berupa satu kata, sedangkan masukan untuk *Skip-gram* adalah satu kata dan luarannya berupa sejumlah kata (Sivakumar dkk., 2020). Ilustrasi perbedaan arsitektur CBOW dan *Skip-gram* dapat dilihat pada Gambar 2.



Gambar 2. Arsitektur CBOW dan *Skip-gram*
Sumber: Sivakumar dkk. (2020)

3.3. Long-Short Term Memory (LSTM)

LSTM adalah variasi *Recurrent Neural Network (RNN)* yang dikembangkan dengan tujuan menghindari masalah dependensi jangka panjang pada RNN, pertama kali dikenalkan pada tahun 1997 oleh Hochreiter & Schmidhuber. Dalam RNN, *network loop* hanya menggunakan satu *layer* yaitu tanh, sedangkan pada LSTM menggunakan satu *layer*. Perbedaan arsitektur RNN dan LSTM dapat dilihat pada Gambar 2. LSTM terdiri dari tiga gerbang yaitu *forget gate*, berguna untuk menentukan informasi mana yang akan dihapus dari sel; *input gate*, berguna untuk menentukan nilai masukan mana yang akan diperbarui pada *memory state*; dan *output gate*, berguna untuk menentukan luaran apa yang akan dihasilkan berdasarkan masukan dan memori yang ada dalam sel (Manalu dkk., 2020).



Gambar 3. Arsitektur RNN vs LSTM
Sumber: Donahue dkk. (2017)

LSTM memiliki gerbang input i_t , modulasi input g_t , *forget* f_t dan output o_t , ditambah lagi sel memori c_t dan unit tersembunyi h_t . Relasi komponen-komponen ini dirumuskan dalam Persamaan 1 – 6 dengan keterangan W adalah matriks bobot antara input atau *hidden state* dengan gerbang-gerbang dan b adalah bias. Setiap blok dalam jaringan memiliki banyak sel dan sel-sel

tersebut berbagi gerbang dalam satu blok (Donahue dkk., 2017).

$$\begin{aligned}
 i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) & (1) \\
 f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) & (2) \\
 o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) & (3) \\
 g_t &= \sigma(W_{xc}x_t + W_{hc}h_{t-1} + b_c) & (4) \\
 c_t &= f_t \odot c_{t-1} + i_t \odot g_t & (5) \\
 h_t &= o_t \odot \phi(c_t) & (6)
 \end{aligned}$$

Keterangan:

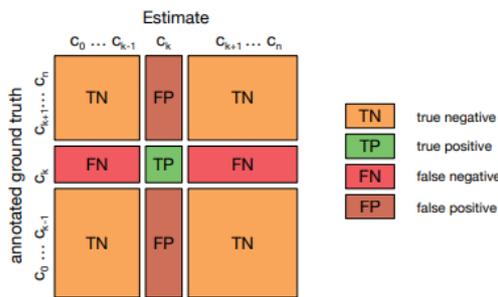
- i_t : input
- g_t : modulasi input
- f_t : forget
- o_t : output
- c_t : sel memori
- h_t : unit tersembunyi
- W : matriks bobot
- b : bias

3.4. Evaluasi Model

Model yang telah dikembangkan perlu dievaluasi agar dapat mengetahui kinerja yang dihasilkan, salah satu bentuk evaluasi model prediksi yang umum dilakukan adalah *confusion matrix*. *Confusion matrix* adalah sebuah matriks yang berguna untuk menilai performa model prediksi yang sudah dikembangkan. Pada klasifikasi biner, *confusion matrix* terdiri dari empat blok yang terdiri dari *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, *False Negative (FN)*. Akan tetapi, jumlah blok bisa bertambah banyak sesuai dengan jumlah kelas pada label targetnya seperti pada Gambar 3. Berdasarkan blok-blok tersebut, dapat dikalkulasikan nilai akurasi, *precision*, *recall*, dan *f1-score*. Tabel 2 merupakan bentuk dasar *confusion matrix*. Rumus untuk akurasi, *precision*, *recall*, dan *f1-score* dijabarkan pada Persamaan 7 – 10.

Tabel 3. *Confusion Matrix* pada Klasifikasi Biner

Kategori	Kelas Prediksi	
	Positive	Negative
Kelas True	TP	FN
Kelas False	FP	TN



Gambar 4 Ilustrasi *Multiclass Confusion Matrix*
Sumber: Krüger (2016)

$$accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (7)$$

$$precision = \frac{TP}{TP+FP} \quad (8)$$

$$recall = \frac{TP}{TP+FN} \quad (9)$$

$$f = 2 \times \frac{precision \times recall}{precision + recall} \quad (10)$$

3.5. Penggeseran *Threshold*

Penggeseran *threshold* adalah salah satu teknik penanganan masalah klasifikasi dengan data yang tidak seimbang. Teknik ini dilakukan setelah pemodelan selesai. Teknik ini disebut teknik yang mudah dilakukan dan jarang memberikan efek negatif pada klasifikasi data tidak seimbang (Zhou & Liu, 2006). Nilai *default threshold* dalam klasifikasi biner dengan kelas 0 dan 1 adalah 0,5. Penggeseran *threshold* bisa dilakukan untuk memperbaiki performa dengan cara menggesernya mendekati kelas mayoritas.

3.6. Penelitian Terkait

Terdapat banyak penelitian yang sudah terlebih dahulu dilakukan dengan topik prediksi tipe kepribadian berdasarkan data media sosial. Penulis melakukan studi literatur dan menghimpun ringkasan penelitian-penelitian tersebut dalam Tabel 3 sebagai acuan. Subjek penelitian pada penelitian yang sudah ada biasanya adalah responden yang diminta melakukan uji kepribadian lalu data media sosialnya digunakan sebagai bahan analisis, belum ada yang menggunakan data kepribadian tokoh publik seperti artis K-Pop.

Tabel 4. Penelitian Terdahulu

Peneliti	Teori Kepribadian	Sumber Data	Metode	Akurasi
(Choong & Varathan, 2021)	MBTI, khusus J/P	Kaggle (Personal ity Café)	LightGBM	81,68%
(Harahap & Muslim, 2020)	MBTI (dibagi menjadi 4 generalisasi)	Twitter	TF-IDF, <i>Random Forest Classifier</i>	70%
(Amirhossaini & Kazemian, 2020)	MBTI	Kaggle	TF-IDF, XGBoost	E/I: 78,17% , S/N: 86,06% , T/F: 71,78% , J/P: 65,70%
(Keh & I-Tsun, 2019)	MBTI	Personality Cafe	BERT	E/I: 75,83% , S/N: 74,41% , T/F: 75,75% , J/P: 71,90%

Tabel 5. Penelitian Terdahulu (lanjutan)

Peneliti	Teori Kepribadian	Sumber Data	Metode	Akurasi
(Celli & Lepri, 2018)	BFI & MBTI	Twitter	<i>Support Vector Machine</i>	E/I: 61,3% , S/N: 61,3% , T/F: 61,3%

Peneliti	Teori Kepribadian	Sumber Data	Metode	Akurasi
				J/P: 61,3%

4. HASIL DAN PEMBAHASAN

4.1. Hasil Persiapan Data

Data artis K-Pop yang dikumpulkan berjumlah 458 artis. Tipe MBTI yang paling banyak muncul pada artis K-Pop adalah ENFP dengan 67 artis atau 14,63%. Tipe MBTI yang paling sedikit muncul adalah ESFP dan ESTP dengan hanya 5 dan 6 artis.

Proses pengumpulan data *caption* (*scraping*) dilakukan dua kali, pada 5 Oktober 2022 untuk 148 artis, dan pada 20 November 2022 untuk 310 artis. Data *caption* yang berhasil dikumpulkan sejumlah 140.332 baris *caption*. Data tersebut dibersihkan dengan serangkaian langkah *pre-processing* menyisakan 118.401. Tipe MBTI dengan *caption* bersih paling banyak adalah ENFJ sejumlah 17.804 dan yang paling sedikit adalah ESTP sejumlah 1.832. Distribusi jumlah artis dan jumlah *caption* per tipe MBTI dapat dilihat pada Tabel 6 dan Tabel 7.

Tabel 6. Distribusi Tipe MBTI

Tipe	Jumlah Artis	Jumlah Caption Bersih
ENFP	67	17.271
INFP	50	13.687
ENFJ	43	19.148
INFJ	42	13.247
ISFP	42	9.845
ISFJ	35	7.622
ENTP	29	6.823
INTP	27	3.935
ESFJ	25	5.818
ENTJ	21	7.293
ESTJ	17	4.398
INTJ	17	4.934
ISTJ	16	3.875
ISTP	16	3.545
ESTP	6	1.908
ESFP	5	2.552
Total	458	118.401

Tabel 7. Distribusi Tipe MBTI per Aspek

		E	I
		Artis	213
E/I	Caption	61.175	57.226
		S	N
S/N	Artis	162	296
	Caption	37.459	80.942
		T	F
	Artis	149	309
T/F	Caption	34.625	83.776
		J	P
J/P	Artis	216	242
	Caption	62.348	56.053
Total	Jumlah Artis	458	
	Jumlah Caption Bersih		118.401

4.2. Hasil Pemodelan

Data *caption* bersih dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Jumlah data latih terdiri dari 94.720 *caption* dan data uji terdiri dari 23.681 *caption*. Kolom target (y_{train}

dan y_{test}) dibagi menjadi 4 sesuai aspek MBTI dan tiap sifat diberi nilai 0 atau 1. Konfigurasi Word2Vec yang digunakan adalah ukuran dimensi 50 dan arsitektur Skip-gram. Parameter ukuran *dropout* pada model LSTM adalah 0,2. Ukuran *dropout* ini dipilih karena nilai ini umum digunakan dalam pemodelan *neural network* untuk mencegah *overfitting*. Pelatihan model dilakukan 20 kali ($epoch=20$). Nilai 20 diambil karena nilai ini umum digunakan dan juga menyesuaikan ukuran dataset serta sumber daya waktu, karena *epoch* terlalu sedikit dapat menyebabkan model kurang menggambarkan pola yang ada dalam data, sedangkan *epoch* yang terlalu banyak dapat menyebabkan *overfitting* atau waktu pelatihan yang tidak efisien.

Terdapat dua skema evaluasi yang dilakukan, yaitu evaluasi terhadap data uji (*caption*) sehingga setiap *caption* memiliki label tipe kepribadian, dan evaluasi terhadap data artis. Tujuan perbandingan dua skema evaluasi ini adalah membuktikan bahwa memprediksi tipe kepribadian hanya berdasarkan 1 *caption* itu lebih sulit daripada berdasarkan kumpulan *caption*. Prediksi pada data artis didasarkan pada rata-rata prediksi terhadap kumpulan *caption*-nya, misalkan artis A memiliki 10 *caption*, dan rata-rata nilai probabilitas dari 10 *caption* tersebut adalah 0,6, nilai ini berada di atas *default threshold* yaitu 0,5 artinya dia akan masuk kelas 1 (contoh jika aspek E/I maka dia termasuk kelas E). Ilustrasi sederhana dapat dilihat pada Tabel 7 dan Tabel 8. Tabel 7 menggambarkan skema I yaitu setiap *caption* diprediksi labelnya, sedangkan Tabel 8 menggambarkan skema II yaitu dengan mengambil rata-rata prediksi dari kumpulan *caption* lalu mengelompokkannya berdasarkan artisnya. Contohnya, artis bernama Bros memiliki 3 *caption* yang diprediksi INTP, ISFJ, dan ENFJ, sedangkan tipe aslinya adalah INFJ, hasil per domain dirata-rata dan didapatkan Bros bertipe INFJ.

Tabel 8. Ilustrasi Evaluasi Skema I

Caption	Artis	MBTI Asli	MBTI Prediksi	Akurasi
Sunny beach	Mario	E-S-T-J	E-N-F-P	1-0-0-0
High hopes	Mario	E-S-T-J	E-N-F-P	1-0-0-0
Lay in my bed	Mario	E-S-T-J	E-S-T-P	1-1-1-0
Lost in sea	Bros	I-N-F-J	I-N-T-P	1-1-0-0
Sleeping bag	Bros	I-N-F-J	I-S-F-J	1-0-1-1
This town is amazing	Bros	I-N-F-J	E-N-F-J	0-1-1-1

Tabel 9. Ilustrasi Evaluasi Skema II

Artis	MBTI Asli	MBTI Prediksi Rata-rata	Akurasi
Mario	E-S-T-J	E-N-F-P	1-0-0-0
Bros	I-N-F-J	I-N-F-J	1-1-1-1

Evaluasi model pada data uji dan data artis menghasilkan tabel *confusion matrix* seperti pada Tabel 10, Tabel 11, Gambar 5, dan Gambar 6.

Tabel 10. *Confusion Matrix* per Model pada Data Uji

Kategori		Kelas Prediksi	
		E	I
Kelas	E	8.909	3.391
	I	6.913	4.468
Kategori		Kelas Prediksi	
		S	N
Kelas	S	854	6.619
	N	611	15.597
Kategori		Kelas Prediksi	
		T	F
Kelas	T	551	6.376
	F	290	16.464
Kategori		Kelas Prediksi	
		J	P
Kelas	J	8.782	3.715
	P	6.720	4.464

Tabel 11. *Confusion Matrix* per Model pada Data Artis

Kategori		Kelas Prediksi	
		E	I
Kelas	E	213	0
	I	159	86
Kategori		Kelas Prediksi	
		S	N
Kelas	S	5	157
	N	0	296
Kategori		Kelas Prediksi	
		T	F
Kelas	T	4	145
	F	0	309
Kategori		Kelas Prediksi	
		J	P
Kelas	J	215	1
	P	165	77

Gambar 5. *Confusion Matrix* Kombinasi 4 Model pada Data Uji*

Berdasarkan *confusion matrix* dapat dihitung akurasi dan *macro f1*, lebih jelasnya dapat dilihat pada Tabel 12. Metrik yang menjadi acuan utama dalam penelitian ini adalah *macro f1* karena metrik tersebut sangat mempertimbangkan hasil prediksi kelas minoritas pada data yang tidak seimbang.

Tabel 12. Performa per Model

Model	Data Uji		Data Artis	
	Akurasi	Macro f1	Akurasi	Macro f1
E/I	0,56	0,55	0,65	0,62
S/N	0,69	0,50	0,66	0,43
T/F	0,72	0,49	0,68	0,43
J/P	0,56	0,54	0,64	0,60
Kombinasi	0,19	0,12	0,21	0,10

Gambar 6. *Confusion Matrix* Kombinasi 4 Model pada Data Artis*

* *confusion matrix* untuk kombinasi 4 model dibuat visualisasi karena dalam bentuk tabel tidak cukup ruang

Akurasi pada model S/N dan T/F cenderung lebih tinggi tetapi *macro f1* cenderung jauh lebih rendah karena hasil prediksi masih sangat bias ke kelas mayoritas (N dan F). Model E/I dan J/P cenderung lebih seimbang tetapi juga masih bias ke kelas E dan J. Evaluasi kombinasi empat model menghasilkan akurasi dan *macro f1* yang sangat rendah pada data uji menunjukkan bahwa sulit untuk memprediksi tipe kepribadian MBTI dengan tepat pada keempat aspeknya hanya berdasarkan satu *caption*. Evaluasi kombinasi empat model pada data artis menghasilkan nilai *macro f1* yang lebih rendah menandakan bahwa bias ke kelas mayoritas masih tinggi, lebih tinggi daripada pada data uji.

Langkah *threshold moving* atau penggeseran *threshold* dilakukan pada tiap model untuk memperbaiki performa model. Nilai *threshold* awal pada klasifikasi biner adalah 0,5, dilakukan penggeseran melalui iterasi mencari *macro f1* terbaik. Iterasi dilakukan berdasarkan data artis mengingat tujuan utama penelitian ini adalah memprediksi tipe kepribadian artis. Proses ini berhasil menemukan *threshold* optimal untuk tiap model seperti yang tercantum pada Tabel 13.

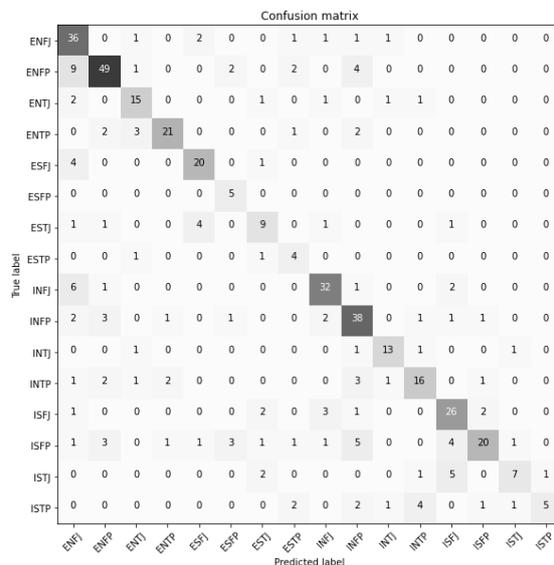
Tabel 13. Hasil Penggeseran *Threshold*

Model	Kelas 0	Kelas 1	<i>Threshold Optimal</i>
E/I	I	E	0,526
S/N	N	S	0,330
T/F	F	T	0,299
J/P	P	J	0,524

Evaluasi model dilakukan lagi setelah ditemukan *threshold* optimal, menghasilkan *confusion matrix* baru pada data artis yang dapat dilihat pada Tabel 14 dan Gambar 7.

Tabel 14. *Confusion Matrix* per Model pada Data Artis Setelah Penggeseran *Threshold*

Kategori		Kelas Prediksi	
		E	I
Kelas Aktual	E	199	14
	I	39	206
Kategori		Kelas Prediksi	
		S	N
Kelas Aktual	S	130	32
	N	16	280
Kategori		Kelas Prediksi	
		T	F
Kelas Aktual	T	118	31
	F	15	294
Kategori		Kelas Prediksi	
		J	P
Kelas Aktual	J	203	13
	P	34	208



Gambar 7. *Confusion Matrix* Kombinasi 4 Model pada Data Artis Setelah Penggeseran *Threshold*

* *confusion matrix* untuk kombinasi 4 model dibuat visualisasi karena dalam bentuk tabel tidak cukup ruang

Angka prediksi benar (*true positive*) tiap kelas pada *confusion matrix* menunjukkan bahwa model setelah menggunakan *threshold* optimal berhasil mengklasifikasikan kelas lebih baik dengan jumlah benar jauh lebih banyak daripada jumlah salah.

Tabel 15. Perbandingan Macro f1 Sebelum dan Sesudah Penggeseran *Threshold*

Model	Data Uji		Data Artis	
	Macro f1 sebelum PT	Macro f1 sesudah PT	Macro f1 sebelum PT	Macro f1 sesudah PT
E/I	0,55	0,56	0,62	0,88
S/N	0,50	0,56	0,43	0,88
T/F	0,49	0,55	0,43	0,88
J/P	0,54	0,56	0,60	0,90
Kombinasi	0,12	0,12	0,10	0,65

Keterangan:
PT = Penggeseran *Threshold*

Performa model dibandingkan dari sebelum dan sesudah penggeseran *threshold*. Perbandingan juga dilakukan pada data uji untuk membuktikan bahwa memprediksi tipe kepribadian hanya berdasarkan satu *caption* itu sangat sulit. Perbandingan dapat dilihat pada Tabel 15.

Terdapat peningkatan nilai *macro f1* yang drastis pada evaluasi dengan data artis baik tiap model maupun kombinasi model. Namun, pada data uji peningkatan tidak terlalu signifikan dan pada kombinasi tidak terdapat peningkatan sama sekali (0,12 ke 0,12), hal ini mempertegas bahwa memprediksi tipe kepribadian secara akurat hanya dari satu *caption* sangatlah sulit sedangkan memprediksi tipe kepribadian secara akurat dari kumpulan *caption* lebih memungkinkan. Kumpulan *caption* lebih mudah diprediksi daripada satu *caption* saja karena dengan lebih banyak data maka pola yang ada dalam data cenderung lebih terbentuk misalkan dengan munculnya kata atau frasa tertentu yang berulang-ulang di berbagai *caption*.

Apabila ditinjau dari jumlah *caption* tiap artis, ditemukan bahwa artis dengan jumlah *caption* lebih banyak cenderung lebih akurat hasil prediksinya. Artis yang benar diprediksi tipe kepribadian keempat aspek memiliki median jumlah *caption* 145, benar diprediksi pada tiga aspek (misal ENFP diprediksi ENFJ) memiliki median jumlah *caption* 75, benar diprediksi pada dua aspek (misal ENFP diprediksi ESFJ) memiliki median jumlah *caption* 52, dan benar diprediksi hanya pada satu aspek (misal ENFP diprediksi ESTJ) memiliki median jumlah *caption* 27.

4.3. Hasil Implementasi Model

Penulis mengembangkan aplikasi web Streamlit dengan tujuan agar penggemar K-Pop dapat menggunakan model untuk memprediksi tipe kepribadian berdasarkan masukan berupa *caption*. Aplikasi yang dikembangkan memiliki tiga halaman utama, meliputi:

1. Halaman Prediksi

Halaman ini menyediakan kolom masukan yang dapat diisi pengguna untuk memasukkan data *caption*. Halaman ini akan menampilkan hasil prediksi tipe kepribadian dan daftar artis yang paling mirip berdasarkan kedekatan nilai probabilitas tiap model setelah pengguna selesai memasukkan *caption* dan menekan tombol 'Submit'.

2. Halaman Dasbor

Halaman ini menampilkan persebaran tipe kepribadian artis K-Pop, artis dengan tipe kepribadian tertentu dan *caption* dari artis-artis tertentu, serta kemiripan antara satu artis dengan artis yang lain berdasarkan kedekatan nilai probabilitas tiap model.

3. Halaman Performa Model

Halaman ini menyediakan informasi terkait performa model yaitu *confusion matrix* dan juga *classification report*.

Prediksi MBTI Berdasarkan Caption Instagram Artis K-Pop

Aplikasi ini adalah aplikasi untuk memprediksi tipe MBTI berdasarkan kemiripan karakteristik penulisan caption Instagram, anda dapat menginputkan caption secara manual, upload file .xlsx atau mengambil data Twitter.

Gambar 8. Tampilan Masukan di Halaman Prediksi

Aplikasi web yang sudah dikembangkan dapat diakses melalui tautan <https://predict-kpop-mbti.streamlit.app>. Per tanggal 29 Desember 2022, aplikasi web tersebut sudah digunakan oleh 280 *unique user* setelah penulis menyebarkannya melalui media sosial Twitter.

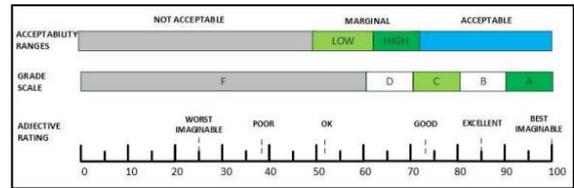
Evaluasi aplikasi web dilakukan menggunakan pengujian *System Usability Scale* (SUS). SUS merupakan pengujian *usability* yang melibatkan pengguna akhir dari responden (Ependi dkk., 2019). Instrumen pengujian pada SUS adalah 10 pernyataan yang menjadi tolak ukur pengujian. Hasil perhitungan instrumen akan dinilai dari tiga sudut pandang yaitu *acceptability*, *grade scale*, dan *adjective rating*. Pada sudut pandang *acceptability*, terdapat tiga tingkatan yaitu *not acceptable*, *marginal*, dan *acceptable*, sedangkan *grade scale* terdapat lima tingkatan yaitu A, B, C, D, F, dan *adjective rating* terdiri dari banyak tingkatan yaitu *worst imaginable*, *poor*, *ok*, *good*, *excellent*, dan *best imaginable*.

Tabel 16. Instrumen Pengujian SUS

No	Pernyataan
1	Saya berpikir akan menggunakan sistem ini lagi
2	Saya merasa sistem ini rumit untuk digunakan
3	Saya merasa sistem ini mudah digunakan
4	Saya membutuhkan bantuan dari orang lain atau teknisi dalam menggunakan sistem ini

Tabel 17. Instrumen Pengujian SUS (lanjutan)

No	Pernyataan
5	Saya merasa fitur-fitur sistem ini berjalan dengan semestinya
6	Saya merasa ada banyak hal yang tidak konsisten (tidak serasi) pada sistem ini
7	Saya merasa orang lain akan memahami cara menggunakan sistem ini dengan cepat
8	Saya merasa sistem ini membingungkan
9	Saya merasa tidak ada hambatan dalam menggunakan sistem ini
10	Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem ini



Gambar 9. Sudut Pandang Penilaian Pengujian SUS
Sumber: Sauro (2011) dalam Ependi dkk. (2019)

Responden pengujian SUS adalah 11 penggemar K-Pop berusia 17-22 tahun, dengan rincian 10 responden berstatus mahasiswa dan 1 responden sudah bekerja. Responden didapatkan secara acak dengan memanfaatkan kuesioner terbuka yang disebar melalui media sosial Twitter dan Instagram, lalu responden diminta untuk mencoba fungsionalitas aplikasi secara mandiri sehingga tidak ada intervensi dari penulis terkait pemilihan responden tertentu atau penilaian yang diberikan. Pengisian kuesioner dilakukan memanfaatkan Google Form. Jumlah responden dinilai sudah memadai karena tidak ada panduan baku terkait minimal jumlah responden dalam penilaian SUS. Beberapa kajian SUS terdahulu menunjukkan perbedaan jumlah responden yang variatif, mulai dari yang berjumlah ratusan sampai hanya lima responden (Ependi dkk., 2019). Skor SUS yang dihasilkan untuk aplikasi web ini adalah 84,55. Skor tersebut sudah termasuk kategori *acceptable*, termasuk *grade B*, dan termasuk *excellent* pada *adjective rating* mengacu pada Gambar 9. Evaluasi SUS yang dilakukan menunjukkan bahwa aplikasi web yang dikembangkan dapat diterima oleh responden.

5. KESIMPULAN DAN SARAN

Berdasarkan langkah-langkah penelitian yang sudah dilakukan dan hasil yang didapatkan dalam penelitian, didapatkan kesimpulan bahwa model LSTM dengan pembobotan Word2Vec dalam memprediksi 16 tipe kepribadian MBTI artis K-Pop memiliki nilai *macro f1* 0,65 atau 65%. Dengan perincian setiap model aspek MBTI yaitu model *Extroversion-Introversion*, *Sensing-Intuition*, *Thinking-Feeling* memiliki nilai *macro f1* yang sama yaitu 0,88, sedangkan model *Judging-Perceiving* memiliki nilai *macro f1* yang sedikit lebih baik yaitu 0,90. Model yang dikembangkan dinilai akurat dalam mengklasifikasikan dua sifat tiap aspek MBTI tetapi kurang akurat dalam mengklasifikasikan 16 tipe kepribadian. Model yang sudah dikembangkan diimplementasikan dalam aplikasi web menggunakan Streamlit untuk membantu para penggemar K-Pop dalam mengidentifikasi dan memahami tipe kepribadian idolanya. Aplikasi web yang dikembangkan sudah digunakan oleh 280 *unique user* dan dievaluasi menggunakan kuesioner SUS menghasilkan nilai 84,55 dan masuk kategori *acceptable* atau dapat diterima.

Saran yang dapat penulis berikan untuk penelitian selanjutnya dengan tema serupa adalah eksperimen menggunakan teknik *pre-processing* lain dalam mengolah data mentah misalnya *POS-tagging* atau LIWC, menggunakan teknik pembobotan dan pemodelan lain misalnya Fasttext, Glove, BERT, dan sebagainya, menggunakan data teks media sosial lain yang digunakan oleh artis K-Pop selain Instagram seperti Lysn, Weverse atau Bubble.

DAFTAR PUSTAKA

- AMIRHOSSEINI, M. H. & KAZEMIAN, H., 2020. Machine Learning Approach to Personality Type Prediction Based on the Myers-Briggs Type Indicator. *Multimodal Technologies and Interaction*, 4(9).
- CELLI, F. & LEPRI, B., 2018. Is Big Five better than MBTI? A personality computing challenge using Twitter data. *Computational Linguistics CLiC-it 2018*.
- CHOONG, E. J. & VARATHAN, K. D., 2021. Predicting judging-perceiving of Myers-Briggs Type Indicator (MBTI) in online social forum. *PeerJ*, 9(e11382), pp. 1-27.
- DONAHUE, J. dkk., 2017. Long-Term Recurrent Convolutional Networks for Visual Recognition and Description. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 39(4), pp. 677-691.
- EPENDI, U., KURNIAWAN, T. B. & PANJAITAN, F., 2019. System usability scale vs heuristic evaluation: a review. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, 10(1), pp. 65-74.
- HARAHAP, R. N. & MUSLIM, K., 2020. Peningkatan Akurasi Pada Prediksi Kepribadian MBTI Pengguna Twitter Menggunakan Augmentasi Data. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 7(4), pp. 815-822.
- HERLAMBANG, A. D. & WIJOYO, S. H., 2019. Algoritma Naive Bayes untuk Klasifikasi Sumber Belajar Berbasis Teks pada Mata Pelajaran Produktif di SMK Rumpun Teknologi Informasi dan Komunikasi. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 6(4), pp. 430-435.
- HERLAMBANG, A. D., WIJOYO, S. H. & RACHMADI, A., 2019. Intelligent Computing System to Predict Vocational High School Student Learning Achievement Using Naive Bayes Algorithm. *Journal of Information Technology and Computer Science*, 4(1), pp. 15-25.
- KEH, S. S. & I-TSUN, C., 2019. Myers-Briggs Personality Classification and Personality-Specific Language Generation Using Pre-trained Language Models. *arXiv preprint arXiv:1907.06333*.
- KERWIN, P. L., 2018. Creating clarity: addressing misconceptions about the MBTI assessment.
- KRÜGER, F., 2016. *Activity, Context, and Plan Recognition with Computational Causal Behaviour Models*. Rostock: University of Rostock.
- MANALU, B. U., TULUS & EFENDI, S., 2020. Deep Learning Performance In Sentiment Analysis. *2020 4th International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM)*, pp. 97-102.
- MUHAMMAD, P. F., KUSUMANINGRUM, R. & WIBOWO, A., 2021. Sentiment Analysis Using Word2vec And Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews. *Procedia Computer Science*, Volume 179, pp. 728-735.
- REDHUM, S., SRIVASTAVA, S., BANSAL, B. & GUPTA, G., 2018. Sentiment Analysis Using Text Mining: A Review. *International Journal on Data Science and Technology*, 4(2), pp. 49-53.
- SIVAKUMAR, S. dkk., 2020. Review on Word2Vec Word Embedding Neural Net. *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 282-290.
- WICAKSONO, A. H. dkk., 2022. Klasifikasi Siswa Slow Learner untuk Mendukung Sekolah dalam Meningkatkan Pemahaman Siswa Menggunakan Algoritma Naive Bayes. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 9(3), pp. 589-596.
- WIDAYAT, W., 2021. Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning. *Jurnal Media Informatika Budidarma*, 5(3), pp. 1018-1026.
- YEUNG, J. & SEO, Y., 2022. *MBTI: How South Koreans fell in love with an American World War II era personality test*. [Online] Available at: <https://edition.cnn.com/2022/07/22/asia/south-korea-mbti-personality-test-dating-briggs-myers-intl-hnk-dst/index.html> [Accessed 15 August 2022].
- ZHOU, Z.-H. & LIU, X.-Y., 2006. Training cost-sensitive neural networks with methods addressing the class imbalance problem. *IEEE Transactions on Knowledge and Data Engineering*, 18(1), pp. 63-77.