

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Al-Qur'an Indonesia

Al-Qur'an Indonesia merupakan salah satu aplikasi Al-Qur'an digital yang tersedia di *Google Play Store*. Aplikasi ini dilengkapi dengan terjemahan bahasa Indonesia serta menyediakan audio mp3 *murottal* dengan pilihan 8 *qori*, semua surat dalam aplikasi ini dapat dibaca secara *offline* (Ari Hartawan et al., 2022). Selain itu, aplikasi ini juga memiliki fitur lain seperti jadwal adzan dan penunjuk arah kiblat.

2.1.2 Analisis Sentimen

Analisis sentimen merupakan suatu metode komputasi yang melibatkan pemahaman, ekstraksi, dan pengolahan data teks secara otomatis untuk menghasilkan informasi mengenai sentimen yang terkandung dalam suatu kalimat pendapat, perilaku, dan emosi individu terhadap sebuah entitas (mencakup individu, peristiwa atau topik), dengan tujuan untuk menyajikan informasi yang berharga dari sebuah data yang bersifat tidak terstruktur (Herlinawati et al., 2020). Dengan analisis sentimen, teks yang terdapat dalam sebuah kalimat atau dokumen akan dihitung nilai polaritasnya untuk mengidentifikasi apakah pendapat tersebut bersifat positif, negatif atau netral (Amrullah et al., 2020).

Proses-proses yang umumnya dilakukan pada analisis sentimen adalah pengumpulan data, *data preprocessing* (*case folding, filtering, tokenizing* dan

stemming), *data labeling*, pembobotan kata dan pembagian data latih dan data uji sebelum diolah dengan algoritma klasifikasi dan diuji dengan menggunakan *confusion matrix*. Keunggulan dari analisis sentimen ialah penghematan waktu, tenaga dan biaya karena dapat dilakukan secara konsisten dan efektif (Dewi, 2022).

2.1.3 Web Scraping

Web Scraping merupakan salah satu cara yang bisa digunakan untuk mengambil data dari sebuah halaman web. Proses pengambilan data dengan *web scraping* dilakukan secara otomatis dengan menggunakan aplikasi atau kode pemrograman (Rudini et al., 2023). Data yang sudah dikumpulkan melalui proses *web scraping* biasanya disimpan di sistem file untuk diproses dan dianalisis pada tahap selanjutnya (Hafiz & Sudarmilah, 2023).

2.1.4 Data Preprocessing

Data Preprocessing adalah tahapan untuk mengubah data mentah menjadi data yang memiliki format yang lebih terstruktur dan mudah dipahami, proses ini menjadi penting karena data mentah seringkali memiliki format yang tidak teratur (Khaerul Mukhlisin, 2021). Berikut ini merupakan tahapan-tahapan yang dilakukan pada *data preprocessing*:

a. Case Folding

Case folding merupakan proses yang dilakukan untuk menyeragamkan karakter pada data. *Case folding* adalah proses dimana seluruh huruf dalam teks diubah menjadi huruf kecil, sedangkan karakter selain huruf 'a' sampai 'z' seperti angka dan tanda baca akan

dihapus karena dianggap sebagai *delimiter* atau karakter pemisah (Jumeilah, 2017).

b. *Tokenizing*

Tokenizing merupakan proses membagi kalimat menjadi kata-kata. pada tahap *tokenizing*, strategi yang umum digunakan adalah memisahkan kata-kata berdasarkan *white space* atau spasi (Handayani et al., 2022).

c. *Spelling Correction*

Spelling correction merupakan proses pengecekan ejaan dalam data. Karena banyak *user* yang memiliki kebiasaan menulis sebuah kata dengan cara disingkat, *typo*, hingga penggunaan bahasa gaul, sehingga bisa mempengaruhi hasil analisis akhirnya, maka *spelling correction* menjadi solusi untuk mengatasi masalah ini (Utomo et al., 2021).

d. *Filtering*

Filtering atau biasa juga disebut *stopword removal* adalah langkah yang digunakan untuk menghapus kata-kata yang sering muncul tapi dianggap tidak memiliki makna dalam suatu konteks (Mahendra et al., 2023).

e. *Stemming*

Stemming merupakan proses yang dilakukan untuk mendapatkan kata dasar dari setiap kata dengan cara menghilangkan *affixes* atau imbuhan yang terdapat dalam kata tersebut baik itu berupa sisipan,

awalan atau kombinasi antara awalan dan akhiran pada kata turunan (Sinaga et al., 2023).

2.1.5 Valence Aware Dictionary and Sentiment Reasoner (VADER)

Valence Aware Dictionary and Sentiment Reasoner (VADER) merupakan metode analisis yang menganalisis teks berdasarkan *lexicon* (*lexicon based*). VADER menghasilkan kelas sentiment berupa positif, negatif dan netral dengan tambahan skor total atau *compound score* (Sumitro et al., 2021). VADER melakukan pelabelan data menggunakan kamus leksikon untuk menghitung nilai polaritas kata. Nilai polaritas kata digunakan untuk menghitung *compound score*. *Compound score* merupakan skor yang digunakan untuk menghitung jumlah nilai semua kosakata (Effendi and Noviana, 2021)

$$Compound\ Score = \frac{x}{\sqrt{x^2 + \alpha}} \quad (2.1)$$

(Anggraini & Suroyo, 2019)

Keterangan:

x = jumlah dari sentimen skor unsur kata kalimat

a = parameter normalisasi yaitu a = 15

Proses pelabelan bergantung pada nilai *compound score*. Ketentuan nilai *compound score* pada pelabelan VADER dapat dilihat pada tabel 2.1

Tabel 2.1 Ketentuan pelabelan VADER

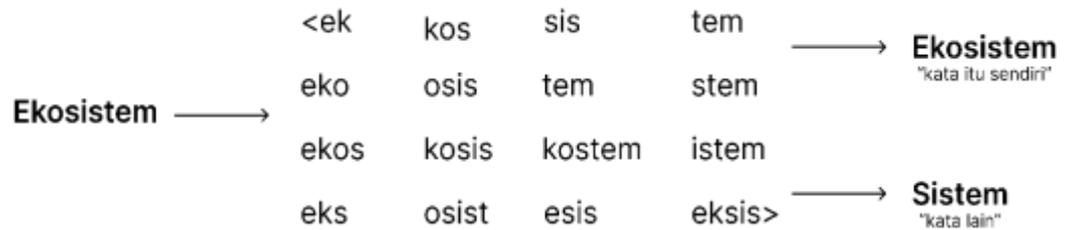
Sentimen	<i>Compound Score</i>
Positif	$CS \geq 0.05$

Netral	$-0.05 > CS < 0.05$
Negatif	$CS \leq -0.05$

Berdasarkan tabel 2.1 proses pelabelan dengan VADER didasarkan pada *compound score*. Jika nilai *compound score* lebih dari sama dengan 0.05 maka ulasan tersebut masuk pada sentimen positif. Jika nilai *compound score* antara -0.05 dan 0.05 maka ulasan tersebut masuk pada sentimen netral. Jika nilai *compound score* kurang dari sama dengan -0.05 maka ulasan tersebut akan masuk pada sentimen negatif.

2.1.6 *FastText*

FastText merupakan salah satu metode NLP yang dikembangkan oleh *Facebook AI Research (FAIR)*. Metode *FastText* memperoleh representasi kata dengan mempertimbangkan informasi *subword*, dimana setiap kata akan diurai menjadi sekelompok *n-gram*, ini membuat *FastText* memungkinkan untuk menangkap makna kata-kata dan memahami sufiks serta prefiks kata (Nurdin et al., 2020). Nilai vektor kata dalam *FastText* direpresentasikan dengan melakukan penjumlahan nilai setiap *n-gram*, sehingga kata-kata yang tidak ditemui dalam korpus tetap dapat direpresentasikan dengan baik karena kemungkinan besar *n-gram* pembentuk kata tersebut muncul dalam *n-grams* yang ada dalam korpus (Alfariqi et al., 2020). Ilustrasi cara kerja *n-gram* pada *FastText* dapat dilihat pada gambar 2.1



Gambar 2.1 Ilustrasi *n-grams* pada *FastText*

Kemampuan *FastText* untuk memecahkan kata menjadi *n-grams* ini memungkinkan *FastText* untuk mengatasi permasalahan *Out of Vocabulary* (OOV), yang berarti *FastText* dapat memberikan nilai vektor untuk kata yang bahkan asing atau baru muncul, sehingga meningkatkan kemampuan model untuk memahami dan merepresentasikan kata-kata yang belum pernah dilihat sebelumnya (Mujahidin et al., 2022).

2.1.7 *Naïve Bayes Classifier*

Naïve Bayes Classifier merupakan metode klasifikasi yang didasarkan pada *Teorema Bayes*. *Naïve Bayes Classifier* bukanlah suatu algoritma tunggal, tetapi serangkaian algoritma klasifikasi berdasarkan *Teorema Bayes* yang menggambarkan probabilitas suatu peristiwa berdasarkan pengetahuan sebelumnya atau probabilitas lain yang diketahui dari peristiwa tersebut (Ernianti Hasibuan & Elmo Allistair Heriyanto, 2022). Secara umum rumus *Naïve Bayes Classifier* dapat dilihat pada persamaan 2.2

$$P(H|X) = \frac{P(X | H) P(H)}{P(X)} \quad (2.2)$$

(Anjasmoros et al., 2020)

Keterangan:

X : data dengan kelas yang belum diketahui.

H : hipotesis data yang merupakan suatu kelas spesifik

$P(H | X)$: probabilitas hipotesis H berdasar kondisi X (posterior)

$P(H)$: probabilitas hipotesis H . (prior probabilitas)

$P(X | H)$: probabilitas X berdasar kondisi pada hipotesis H

$P(X)$: probabilitas X

Naïve Bayes Classifier menggunakan asumsi independensi antarfitur untuk memprediksi probabilitas keanggotaan suatu kelas (Tanggraeni & Sitokdana, 2022). Penggabungan posterior probabilitas dan prior probabilitas dalam rumus *Naïve Bayes* digunakan untuk menghitung probabilitas dari setiap kemungkinan klasifikasi (Indarso et al., 2023).

2.1.8 *Convolutional Neural Network*

Convolutional Neural Network (CNN) adalah jenis jaringan saraf tiruan yang menggunakan umpan balik untuk mempertahankan struktur hierarki dengan mempelajari representasi fitur internal, CNN mampu menggeneralisasi fitur-fitur dalam konteks masalah termasuk pengenalan objek dan *computer vision* (Hermanto et al., 2021). Walaupun pertama kali dirancang untuk memproses gambar, seiring berjalannya waktu, *Convolutional Neural Network* (CNN) telah berkembang menjadi alat yang efisien dan cepat dalam tugas-tugas pemrosesan bahasa alami (Ahmad et al., 2023).

Convolutional Neural Network memiliki 3 layer yaitu, *convolutional layer*, *pooling layer*, dan *fully-connected layer*. *Convolutional layer* berisi

kumpulan filter dengan ukuran tetap yang digunakan untuk melakukan konvolusi pada data input dengan output berupa *featur maps* sementara *Pooling layer* berfungsi untuk mengurangi dimensi pada *feature maps* dengan cara menerapkan beberapa operasi *linear* dan *non linear* (Hidayat & Handayani, 2023). Sedangkan layer terakhir yaitu *fully-connected layer* berfungsi untuk memahami pola yang telah dihasilkan oleh layer sebelumnya (Ayu Shafirra N, 2020).

Dalam metode *Convolutional Neural Network*, fungsi aktivasi diterapkan ketika berada diantara layer *convolutional* dan *pooling*, aktivasi tersebut menggunakan fungsi aktivasi ReLU, sementara fungsi aktivasi *output* menggunakan *sigmoid* (Ayu Shafirra N, 2020). Persamaan fungsi ReLU dapat dilihat pada persamaan 2.3.

$$f_h(0, FMa, b) = \max(0, FMa, b) = \begin{cases} FMa, b, & \text{jika } FMa, b \geq 0 \\ 0, & \text{jika } FMa, b < 0 \end{cases} \quad (2.3)$$

(Ayu Shafirra N, 2020)

Keterangan:

FMa, b : *feature map* ke a, b

a, b : panjang *fixel* pada *feature map*

Fungsi *sigmoid* yang dilakukan pada pada *layer output* hanya dikhususkan pada klasifikasi dengan dua kelas karena hanya memiliki jarak nilai 0 sampai 1 (Ayu Shafirra N, 2020). Persamaan fungsi *aktivasi sigmoid* dapat dilihat pada persamaan 2.4.

$$F(x) = \frac{1}{1 + e^{-x}} \quad (2.4)$$

(Ayu Shafirra N, 2020)

Keterangan:

$F(x)$: output dari sigmoid

e : konstanta Euler, merupakan angka logaritma natural

Setelah menghitung fungsi *sigmoid* pada layer *output*, langkah selanjutnya adalah meminimalkan nilai *loss* dengan menggunakan *Adam optimizer*. *Adam optimizer* merupakan teknik optimasi stokastik yang digunakan dalam model CNN untuk memperbarui parameter jaringan dengan tujuan untuk mengoptimalkan fungsi (Ayu Shafirra N, 2020).

2.1.9 Confusion Matrix

Confusion Matrix merupakan cara untuk evaluasi kinerja hasil yang umumnya digunakan dalam pemodelan klasifikasi apakah model tersebut memiliki kinerja yang baik atau tidak. Tabel *confusion matrix* dapat dilihat pada tabel 2.2.

Tabel 2.2 *Confusion matrix*

		Kelas Prediksi		
		Negatif	Netral	Positif
Kelas Aktual	Negatif	TN	FNeu	FP
	Netral	FN	TNeu	FP
	Positif	FN	FNeu	TP

Confusion Matrix memiliki beberapa *output* yang dapat dihitung untuk mengukur performa algoritma klasifikasi, diantaranya:

- a. Akurasi

Akurasi pada *confusion matrix* menunjukkan seberapa akurat model klasifikasi atau seberapa dekat prediksi dengan nilai sebenarnya. Rumus untuk menghitung akurasi dapat dilihat pada persamaan 2.5.

$$Akurasi = \frac{TP+TN}{TP++TN+FP++FN} \quad (2.5)$$

(Firdaus et al., 2022)

b. Presisi

Presisi merupakan nilai yang dihasilkan dari perbandingan nilai prediksi yang benar dan hasil dari prediksi secara keseluruhan. Rumus untuk menghitung nilai presisi dapat dilihat pada persamaan 2.6

$$Presisi = \frac{TP}{TP+FP} \quad (2.6)$$

(Firdaus et al., 2022)

c. Recall

Recall merupakan nilai yang dihasilkan dari perbandingan nilai prediksi yang benar dan keseluruhan hasil yang benar. Rumus yang digunakan untuk menghitung nilai *recall* dapat dilihat pada persamaan 2.7

$$Recall = \frac{TP}{TP+FN+FN_{neu}} \quad (2.7)$$

(Firdaus et al., 2022)

d. F1-Score

F1-score merupakan nilai yang dihasilkan dari nilai rata-rata harmonik dari presisi dan *recall*. Rumus yang digunakan untuk menghitung nilai *f1-score* dapat dilihat pada persamaan 2.8

$$F1 - score = \frac{2 * Presisi * Recall}{Presisi + Recall} \quad (2.8)$$

(Firdaus et al., 2022)

Keterangan:

True Negative (TN) : banyaknya nilai data negatif yang tergolong negatif

False Negative (FN) : banyaknya nilai data negatif yang tergolong positif atau netral

True Neutral (Tneu) : banyaknya nilai data netral yang tergolong netral

False Neutral (Fneu) : banyaknya nilai data netral yang tergolong positif atau negatif

True Positive (TP) : banyaknya nilai data positif yang tergolong positif

False Positive (FP) : banyaknya nilai data positif yang tergolong negatif atau netral

2.1.10 Word Cloud

Word Cloud merupakan gambaran visual yang digunakan untuk memvisualisasikan kemunculan kata-kata dalam suatu teks, semakin besar ukuran huruf maka semakin besar frekuensi kemunculan sebuah kata dan semakin kecil huruf semakin kecil frekuensi kemunculan kata tersebut (Suryati et al., 2023)

2.2 Penelitian Terkait dan Kebaruan Penelitian

2.2.1 State of The Art (SOTA)

Perbandingan penelitian yang berhubungan dengan permasalahan penelitian dan hasil penelitian yang dihasilkan dapat dilihat pada tabel 2.3.

Tabel 2.3 SOTA

No	Konten	Deskripsi
1.	22 Juni 2023	
	Paper Ke-1	
	Judul Paper	Analisis Sentimen <i>Zoom Cloud Meetings</i> di <i>Play Store</i> Menggunakan <i>Naïve Bayes</i> dan <i>Support Vector Machine</i>
	Penulis	Nuraeni Herlinawati, Yuri Yuliani, Siti Faizah, Windu Gata, Samudi
	Jurnal/Konferensi	CESS (Journal of Computer Engineering System and Science)
	URL	https://jurnal.unimed.ac.id/2012/index.php/cess/article/view/18186
Permasalahan	Penggunaan aplikasi <i>Zoom Cloud Meetings</i> mulai <i>booming</i> karena adanya wabah virus corona sehingga pemerintah menghimbau untuk tidak keluar rumah, para pekerja diminta untuk WFH begitupun siswa dan mahasiswa diminta belajar dari rumah. Pada <i>Google Play Store</i> terdapat fitur	

		rating dan ulasan dari pengguna. Ulasan penting karena sering digunakan sebagai tolak ukur apakah suatu aplikasi tersebut baik atau tidak.
	Hasil utama	Pada penelitian ini jumlah data setelah dilakukan <i>preprocessing</i> menjadi 1007. Data hampir seimbang dengan label positif sebanyak 546 dan label negatif 461 ulasan. Evaluasi model menggunakan <i>10fold cross validation</i> diperoleh nilai akurasi dan nilai AUC dari masing-masing algoritma yaitu Naïve Bayes nilai akurasi 74,37% dan nilai AUC 0,659. Sedangkan untuk algoritma SVM nilai akurasi 81,22% dan nilai AUC 0,886. Dalam penelitian ini dapat diketahui bahwa algoritma SVM lebih unggul 6,85% dari algoritma Naïve Bayes.
	Batasan	Analisis Sentimen hanya menggunakan algoritma <i>Naïve Bayes Classifier</i> dan <i>Support Vector Machine</i>
2.	22 Juni 2023	
	Paper Ke-2	
	Judul Paper	Analisis Sentimen Aplikasi Novel <i>Online</i> di <i>Google Play Store</i> Menggunakan Algoritma <i>Support Vector Machine</i> (SVM)
	Penulis	Selva Indah Nurhafida, Falentino Sembiring
	Jurnal/Konferensi	Jurnal Sains Komputer & Informatika (J-SAKTI)
	URL	http://tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/view/447

	Permasalahan	Novel saat ini sudah tersedia melalui platform online android, terutama <i>Google Play Store</i> . Ulasan pada <i>Play Store</i> dapat digunakan untuk menentukan apakah suatu aplikasi baik atau buruk, serta untuk mencari masalah yang dimiliki orang-orang dengan aplikasi tersebut
	Hasil utama	Pada penelitian ini, dari dua aplikasi yang diteliti yaitu <i>Wattpad</i> dan <i>Dreame</i> . Terdapat 2356 ulasan positif dan 1058 ulasan negatif untuk aplikasi <i>Wattpad</i> , sementara ada 1865 ulasan positif dan 761 ulasan negatif untuk <i>Dreame</i> . Berdasarkan hasil pengujian akurasi kernel SVM, kernel <i>sigmoid</i> memiliki akurasi terbaik untuk memprediksi data review <i>Wattpad</i> dengan akurasi sebesar 88,60%. Sementara untuk aplikasi <i>Dreame</i> memiliki akurasi 87,45% dengan menggunakan kernel <i>Linear</i>
	Batasan	Analisis Sentimen hanya menggunakan algoritma <i>Support Vector Machine</i>
3.	22 Juni 2023	
	Paper Ke-3	
	Judul Paper	Analisis Sentimen Aplikasi <i>E-Government</i> Pada <i>Google Play</i> Menggunakan Algoritma <i>Naïve Bayes</i>
	Penulis	Artanti Inez Tanggraeni, Melkior N. N. Sitokdana
	Jurnal/Konferensi	JATISI (Jurnal Teknik Informatika dan Sistem Informasi)
	URL	https://jurnal.mdp.ac.id/index.php/jatisi/article/view/1835
	Permasalahan	Aplikasi Sentuh Tanahku sebagai wujud dari inovasi digital pemerintah dalam bidang pertahanan memiliki banyak ulasan dari pengguna. Hasil ulasan tersebut memiliki pengaruh bagi pengembangan

		aplikasi. Banyaknya data ulasan akan sulit jika diolah secara manual, maka diperlukan metode untuk melihat kecenderungan pengguna terhadap aplikasi.
	Hasil Utama	Pada penelitian ini diambil data dari <i>Play Store</i> pada rentang waktu september 2021 sampai oktober 2022 dengan jumlah 674 ulasan. Setelah dilakukan penghapusan data tersisa sebanyak 642 ulasan. Dengan perbandingan 407 ulasan positif dan 235 ulasan negatif dengan menggunakan algoritma Naïve Bayes dengan pembobotan TF-IDF menghasilkan akurasi sebesar 89%, <i>precision</i> sebesar 83% dan <i>recall</i> sebesar 87%
	Batasan	Data ulasan yang digunakan sedikit dan tidak menggunakan penanganan negasi sehingga tidak meningkatkan akurasi dari metode yang digunakan.
4.	22 Juni 2023	
	Paper Ke-4	
	Judul Paper	Analisis Sentimen <i>Review</i> Halodoc Menggunakan <i>Naïve Bayes Classifier</i>
	Penulis	Asep Hendra, Fitriyani
	Jurnal/Konferensi	JISKA (Jurnal Informatika Sunan Kalijaga)
	URL	https://ejournal.uin-suka.ac.id/saintek/JISKA/article/view/2076
	Permasalahan	Halodoc merupakan aplikasi layanan kesehatan yang bisa digunakan oleh masyarakat secara gratis atau berbayar, pengguna dapat memberikan ulasan pada <i>Play Store</i> untuk menjadi bahan evaluasi

		Halodoc. Karena banyaknya ulasan yang ada, oleh karena itu diperlukan analisis sentimen terhadap ulasan Halodoc.
	Hasil Utama	Pada penilitan ini digunakan data sebanyak 950 data dari ulasan <i>Play Store</i> dengan pengujian data acak 250 data dan 650 data. Algoritma <i>Naïve Bayes Classifier</i> yang telah diterapkan untuk analisis sentimen Halodoc dari 475 sentimen positif terdapat 387 sentimen yang tepat dikategorikan positif dan sisanya masuk kategori negatif, kemudian dari 475 sentimen negatif, 389 sentimen tepat dikategorikan negatif dan sisanya positif dengan nilai akurasi 81,68%
	Batasan	Klasifikasi teks pada penelitian ini hanya menggunakan algoritma <i>Naïve Bayes Classifier</i> saja dan tidak dilakukan seleksi fitur optimasi seperti <i>prune methode</i> untuk meningkatkan hasil akurasi
5.	22 Juni 2023	
	Paper Ke-5	
	Judul Paper	Analisis Sentimen Pada <i>Review</i> Aplikasi Grab di <i>Google Play Store</i> Menggunakan Support Vector Machine
	Penulis	Rizki Wahyudi, Gilang Kusumawardhana
	Jurnal/Konferensi	JURNAL INFORMATIKA
	URL	https://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/9681
	Permasalahan	Pada <i>Google Play Store</i> pengguna sering membaca <i>review</i> pengguna lain dan reputasi aplikasi sebelum mengunduh sebuah aplikasi. Analisis <i>review</i> dari pengguna sangat menarik bagi pemilik

		aplikasi untuk mengambil keputusan di masa depan. Penelitian ini bertujuan menganalisis <i>review</i> aplikasi Grab pada <i>Play Store</i>
	Hasil Utama	Pada penelitian ini digunakan data sebanyak 1000 data ulasan dari <i>Play Store</i> . Dengan hasil, kata bernilai positif yang sering diulas oleh pengguna adalah “ovo” yang muncul sebanyak 266 kali. Yang paling sering ulas oleh pengguna adalah tentang “driver” yang muncul sebanyak 585 kali. Hasil dari analisis sentimen terhadap 900 ulasan data testing dengan menggunakan nilai pengujian fold=5 pada metode SVM diperoleh akurasi sebesar 85,54% yang menghasilkan prediksi pada data testing yaitu 59 review positif dan 675 review negatif
	Batasan	Data yang digunakan pada penelitian dibatasi untuk <i>review</i> yang berkategori paling relevan, dan algoritma yang digunakan hanya <i>Support Vector Machine</i> .
6.	22 Juni 2023	
	Paper Ke-6	
	Judul Paper	Analisis Sentimen Ulasan Aplikas Mola pada <i>Google Play Store</i> Menggunakan Algoritma <i>Support Vector Machine</i>
	Penulis	Muhammad Diki Hendriyanto, Azhad Ali Ridha, Ultach Enri
	Jurnal/Konferensi	Journal of Information Technology and Computer Science (INTECOMS)
	URL	https://journal.ipm2kpe.or.id/index.php/INTECOM/article/view/3708

	Permasalahan	Aplikasi Mola adalah aplikasi <i>streaming</i> video di <i>Google Play Store</i> yang sudah di unduh sebanyak 5 juta kali namun hanya memiliki rating 3,2 pada <i>Google Play Store</i> . Pemberian rating aplikasi diikuti dengan ulasan dari para pengguna. Ulasan yang ada cukup banyak dan tidak terstruktur serta mengandung opini pengguna tentang kepuasan terhadap aplikasi tersebut. Berdasarkan hal tersebut dilakukan analisis sentimen menggunakan algoritma SVM untuk mengetahui bagaimana sentimen para pengguna terhadap aplikasi MOLA.
	Hasil Utama	Pada penilitan ini digunakan data sebanyak 520 data dari hasil scraping pada tanggal 1 desember 2021 sampai 31 januari 2022 menggunakan <i>library google play store</i> . Setelah dilakukan pelabelan data tersebut menghasilkan 312 ulasan positif dan 208 ulasan negatif. Hasil terbaik didapatkan pada skenario 1 dengan rasio perbandingan data <i>training</i> 90% dan data <i>testing</i> 10% dengan hasil <i>accuracy</i> 92,31%, <i>precision</i> 96,3%, <i>recall</i> 89,66% dan <i>f1-score</i> 92,86%.
	Batasan	Algoritma yang digunakan pada penelitian hanya <i>Support Vector Machine</i> , dan dataset yang digunakan sedikit
7.	23 Juni 2023	
	Paper Ke-7	
	Judul Paper	Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit dan Bareksa Dengan Algoritma KNN
	Penulis	Aluisius Dwiki Adhi Putra, Safitri Juanita
	Jurnal/Konferensi	JATISI (Jurnal Teknik Informatika dan Sistem Informasi)

	URL	https://jurnal.mdp.ac.id/index.php/jatiasi/article/view/962
	Permasalahan	Terdapat banyak aplikasi investasi online yang sudah banyak diunduh masyarakat diantaranya adalah Bibit dan Bareksa. Bareksa memiliki pengguna sebanyak 11 ribu dan bibit sebesar 85 ribu. Tujuan penelitian ini adalah untuk memberikan rekomendasi apakah aplikasi Bibit dan Bareksa yang lebih unggul dan aman untuk investasi online bagi masyarakat dengan cara analisis sentimen
	Hasil Utama	Pada penelitian ini digunakan data sebanyak 998 yang terdiri dari 484 sentimen positif dan 514 sentimen negatif untuk aplikasi Bareksa sedangkan untuk aplikasi Bibit menggunakan 1063 data yang terdiri dari 541 sentimen positif dan 522 sentimen negatif. Data tersebut melewati tahapan <i>preprocessing</i> dan <i>modelling</i> menggunakan model CRISP-DM dan algoritma yang digunakan adalah KNN. Dengan perbandingan 60:40 untuk data <i>training</i> dan data <i>testing</i> maka nilai akurasi, <i>precision</i> dan <i>recall</i> yang dihasilkan dari tiap aplikasi yaitu untuk Bibit 85,14%, 91,91% dan 76,44% sedangkan untuk Bareksa yaitu 81,70%, 87,15% dan 75,73%
	Batasan	Algoritma yang digunakan hanya terbatas pada algoritma KNN, dan <i>dataset</i> yang digunakan tidak banyak.
8.	23 Juni 2023	
	Paper Ke-8	
	Judul Paper	Analisis Sentimen Aplikasi Gojek Menggunakan Metode SVM dan NBC (Studi Kasus: Komentar Pada Play Store)

Penulis	Mochamad Tri Anajsmoro, Istiadi, Fitri Marisa
Jurnal/Konferensi	Conference on Innovation and Application of Science and Technology (CIASTECH)
URL	http://publishing-widyagama.ac.id/ejournal-v2/index.php/ciastech/article/view/1905
Permasalahan	Go-Jek adalah aplikasi transportasi online yang sudah menguasai masyarakat Indonesia. Ulasan pada <i>Google Play Store</i> dapat digunakan untuk evaluasi dan meningkatkan layanan aplikasi. Untuk mengetahui kualitas layanan, analisis sentimen dapat digunakan untuk mengklasifikasikan ulasan kedalam sentimen positif dan negatif
Hasil Utama	Pada penelitian ini digunakan data komentar aplikasi Go-jek diantara tanggal 9 juli 2020 sampai 1 september 2020. Diambil 1000 komentar positif dan 1000 komentar negatif. Pada metode SVM dalam pengujian ini menggunakan 4 kernel yaitu kernel linear, kernel polynomial, kernel RBF dan kernel Sigmoid. Pada kernel linear <i>score</i> paling tinggi <i>accuracy</i> = 0.8612 dengan komposisi <i>testing</i> 30% dan <i>training</i> 70%, <i>recall</i> = 0.9922 dengan komposisi <i>testing</i> 90% dan <i>training</i> 10%, <i>precision</i> = 0.7967 dengan komposisi <i>testing</i> 30% dan <i>training</i> 70%. Pada kernel polynomial <i>score</i> paling tinggi <i>accuracy</i> = 0.5008 dengan komposisi <i>testing</i> 60% dan <i>training</i> 40%, pada kernel RBF <i>score</i> paling tinggi <i>accuracy</i> = 0.5401 dengan komposisi data <i>testing</i> 60% dan <i>training</i> 40%. Pada kernel sigmoid <i>score</i> paling tinggi <i>accuracy</i> = 0.6244 dengan komposisi <i>testing</i> 60% dan <i>training</i> 40%. Sementara hasil pengujian metode NBC <i>score</i> paling tinggi <i>accuracy</i> = 0.6445 dengan komposisi <i>testing</i> 10% dan <i>training</i> 90% dan <i>precision</i> = 0.7636 dengan komposisi <i>testing</i> 10% dan <i>training</i> 90% dan <i>recall</i> = 0.4242 dengan komposisi <i>testing</i> 10% dan <i>training</i> 90%.

	Batasan	Data yang digunakan hanya dibatasi pada jumlah 1000 data ulasan negatif dan 1000 data ulasan positif
9.	23 Juni 2023	
	Paper Ke-9	
	Judul Paper	Analisis Sentimen Aplikasi WETV di <i>Google Play Store</i> Menggunakan Algoritma Support Vector Machine
	Penulis	Ummi Kulsum, Mohamad Jajuli, Nina Sulistiyowati
	Jurnal/Konferensi	Journal of Applied Information and Computing (JAIC)
	URL	https://jurnal.polibatam.ac.id/index.php/JAIC/article/view/4802
	Permasalahan	WeTv merupakan aplikasi <i>streaming online</i> yang sudah diunduh lebih dari 50 juta kali di situs resmi <i>Google Play Store</i> . Banyaknya jumlah pengguna membuat ulasan juga berjumlah banyak yang membuat sulit untuk dibaca secara manual, analisis sentimen diperlukan untuk mengklasifikasikan ulasan kedalam kelas positif dan negatif.
Hasil Utama	Pada penelitian ini didapatkan data sebanyak 5378 data yang diambil dari ulasan aplikasi WeTv di <i>Google Play Store</i> dari tanggal 11 desember 2020 sampai 15 januari 2021. Selanjutnya data dibagi menjadi 3 kelas yaitu positif, negatif dan netral. Diperoleh 3737 data positif, 258 data netral dan 1383 data negatif. Karena penelitian ini hanya menggunakan sentimen positif dan negatif, maka ulasan netral dihapus sehingga tersisa 5120 data. Selanjutnya dilakukan lagi penghapusan untuk ulasan membahas asing sehingga tersisa 3996 data. Setelah dilakukan penghapusan simbol, angka dan ulasan	

		duplikat data akhir berjumlah 3121 data. Hasil tertinggi didapatkan oleh 3 skenario. Yang pertama pada pembagian 90% data latih dan 10% data uji dengan <i>accuracy</i> sebesar 85%, <i>precision</i> 86%, <i>recall</i> 90% dan <i>f1-score</i> 88%. Yang kedua pada skenario pembagian 80% data latih dan 20% data uji menghasilkan <i>accuracy</i> sebesar 85%, <i>precision</i> 85%, <i>recall</i> 90%, <i>f1-score</i> 88%. Dan skenario dengan pembagian 70% data latih dan 30% data uji dengan hasil <i>accuracy</i> 85%, <i>precision</i> 86%, <i>recall</i> 89%, <i>f1-score</i> 87%.
	Batasan	Algoritma yang digunakan dalam proses <i>data mining</i> hanya <i>Support Vector Machine</i> dan kernel yang digunakan adalah kernel linear, sehingga tidak dapat melihat perbandingan antar linear dan algoritma.
10.	23 Juni 2023	
	Paper Ke-10	
	Judul Paper	Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naïve Bayes, Random Forest dan Support Vector Machine
	Penulis	Evita Fitri, Yuri Yuliani, Susy Rosyida, Windu Gata
	Jurnal/Konferensi	JURNAL TRANSFORMATIKA
	URL	http://journals.usm.ac.id/index.php/transformatika/article/view/2317#
	Permasalahan	Saat ini sudah banyak tersedia aplikasi belajar <i>online</i> atau situs web belajar <i>online</i> yang dapat diakses. salah satu aplikasi tersebut adalah Ruangguru, aplikasi ini sudah diunduh sebanyak 15 juta lebih. Namun banyaknya aplikasi belajar <i>online</i> yang tersedia membuat pengguna lebih selektif dalam

		menentukan aplikasi apa yang cocok atau yang lebih baik untuk digunakan. Salah satu yang dapat mempengaruhi pengguna untuk memilih aplikasi adalah ulasan yang ada pada layanan pengunduhan aplikasi, apabila ulasan-ulasan ini dikumpulkan lalu diolah maka hasil tersebut dapat dijadikan sebuah kesimpulan aplikasi belajar <i>online</i> mana yang memiliki sentimen baik
	Hasil Utama	Pada penelitian ini dataset didapat dari <i>review</i> aplikasi Ruangguru di <i>Google Play Store</i> . Diambil dengan menggunakan <i>web scrapping</i> dari <i>website Google Play Store</i> . Data diambil sebanyak 1629 dengan bentuk <i>text review</i> dalam rentang waktu 1 maret sampai 30 maret 2020. Hasil nilai akurasi tertinggi yaitu pada pengujian dengan menggunakan algoritma Random Forest dengan jumlah akurasi sebesar 97,16% serta nilai AUC 0.996, kemudian algoritma Support Vector Machine yang menghasilkan akurasi sebesar 96,01% dengan nilai AUC 0,543 dan pengujian dengan algoritma Naïve Bayes dengan nilai akurasi sebesar 94,16% serta nilai AUC sebesar 0,999.
	Batasan	Data yang digunakan pada penelitian ini dibatasi dalam rentang waktu 1 maret sampai 30 maret 2020, sehingga data yang digunakan tidak banyak
11.	24 Juni 2023	
	Paper Ke-11	
	Judul Paper	Analisis Sentimen Pada Ulasan Aplikasi Jakarta Terkini (JAKI) di <i>Google Play Store</i> Menggunakan Metode <i>Support Vector Machine</i>
	Penulis	Daniel Dwi Eryanto Manurung, Nur Hafifah Matondang, Desta Sandya Prasvita

	Jurnal/Konferensi	Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)
	URL	https://conference.upnvj.ac.id/index.php/senamika/article/view/2149
	Permasalahan	Aplikasi Jakarta Terkini (JAKI) terus berkembang dengan baik namun beberapa pengguna mengeluh terhadap layanan aplikasi. Beberapa keluhan seperti sertifikat vaksin yang tidak tersedia, kuota vaksin selalu penuh dan sebagainya. Analisis sentimen dilakukan pada aplikasi JAKI melalui ulasan <i>Google Play store</i> untuk memberikan informasi kepada masyarakat terkait kinerja aplikasi.
	Hasil Utama	Pada penelitian ini digunakan data sebanyak 1000 data yang diambil dngan menggunakan id aplikasi 'id.go.jakarta.smartcity.jaki' pada tanggal 26 november 2019 sampai 15 maret 2022. Selanjutnya dilakukan pelabelan secara manual dan didapatkan 453 positif dan 547 negatif. Kemudian dilakukan seleksi fitur dengan <i>chi-square</i> . Tahapan selanjutnya dilakukan pembagian data sebesar 80% 20% dan diklasifikasikan dengan metode SVM. Hasilnya adalah fitur terseleksi sebanyak 66 fitur dan pemodelan menggunakan kernel RBF, C=40 dan gamma=0.1 dari SVM, diperoleh 120 data uji mendapatkan <i>accuracy</i> 97%, <i>precision</i> 100%, <i>recall</i> 93%, dan <i>specificity</i> 100%
	Batasan	Data yang digunakan pada penelitian ini dibatasi hanya dengan 1000 data ulasan. Dan algoritma yang digunakan dibatasi hanya dengan algoritma <i>Support Vector Machine</i> .
12.	24 Juni 2023	
	Paper Ke-12	

Judul Paper	Sentiment Analysis on Twitter Tweets About Covid-19 Vaccines Using NLP and Supervised KNN Classification Algorithm
Penulis	F. M. Javed Mehedi Shamrat, Sovon Chakraborty, M. M. Imran, et al
Jurnal/Konferensi	Indonesian Journal of Electrical Engineering and Computer Science
URL	https://www.researchgate.net/publication/353182015
Permasalahan	Para ilmuwan telah menemukan beberapa vaksin untuk virus <i>covid-19</i> . Masyarakat umum telah mengungkapkan pendapat mereka tentang keamanan dan keefektifan vaksin di media sosial seperti Twitter. Dilakukan pengklasifikasian sentimen dari <i>tweet</i> untuk mengetahui sentimen masyarakat terkait vaksin <i>covid-19</i> .
Hasil Utama	Pada penelitian ini data diambil dari <i>tweet</i> pada Twitter menggunakan <i>library Tweepy</i> dengan jumlah 10.000 dari tagar #Pfizer, 10.000 dari tagar #Moderna, dan 10.000 dari tagar #AstraZeneca. Data yang sudah diambil selanjutnya diproses untuk menghilangkan karakter khusus, <i>hyperlink</i> , <i>reetwet</i> , emoji dan stiker. Selanjutnya dilakukan penghitungan <i>polarity</i> dan <i>subjectivity</i> . <i>Subjectivity</i> berisi fakta, opini dan keinginan, sedangkan <i>polarity</i> berisi perasaan dan emosi. Klasifikasi KNN dilakukan terhadap skor <i>polarity</i> . Jika skor <i>polarity</i> lebih dari 0 maka tweet tersebut positif, jika kurang dari 0 negatif dan jika sama dengan 0 maka netral. Hasilnya vaksin Pfizer memiliki sentimen positif sebesar 47,29, Moderna 46,16, dan AstraZeneca 40,08.

	Batasan	Data yang digunakan pada penelitian ini dibatasi hanya dengan data yang diambil dari twitter dan algoritma yang digunakan dibatas hanya dengan menggunakan algoritma KNN.
13.	24 Juni 2023	
	Paper Ke-13	
	Judul Paper	Study Analysis of the Effectiveness of Using the Digital Al-Quran among College Students in the Dimensions of Globalization Flow
	Penulis	Fahri Husaeni, Zaenal Arifin, Ajat Rukajat, Fadil Santosa
	Jurnal/Konferensi	ATTHULAB: Islamic Religin Teaching and Learning Journal
	URL	http://journal.uinsgd.ac.id/index.php/atthulab/article/view/12921#
	Permasalahan	Globalisasi yang berdampak signifikan terhadap nilai spiritual, pola pikir dan sikap. Perlu adanya upaya yang dilakukan untuk membendung dampak negatifnya. Salah satunya adalah dengan mendorong mahasiswa untuk lebih meningkatkan interaksinya dengan Al-Qur'an
Hasil Utama	Hasil dari penelitian ini menunjukkan, total skor dari 71 responden adalah 851 skor. Skor maksimal adalah 1349. Dalam membahas hasil data pada penelitian ini menggunakan PAP (<i>Benchmark Assesment</i>) untuk mengkategorikan data. Berdasarkan kategori skor, keefektifan Al-Quran berada pada kategori "efektif". Adapun pola skor menggunakan kerangka skala Guttman menunjukkan hasil adanya peningkatan efektifitas sebesar 53,5% dalam membaca Al-Qur'an, 92,9% sepakat bahwa Al-Qur'an digital sangat memudahkan mahasiswa untuk lebih intens dalam berinteraksi dengan Al-Quran dan	

		57,7% mahasiswa menyatakan Al-Quran digital menjadi salah satu motivasi utama dalam memperbaiki ibadah mereka.
	Batasan	Data yang digunakan pada penelitian ini dibatasi hanya pada 71 responden mahasiswa bimbingan konseling tingkat 3.
14.	27 Juni 2023	
	Paper Ke-14	
	Judul Paper	Analisis Sentimen Review Aplikasi PeduliLindungi pada <i>Google Play Store</i> Menggunakan NBC
	Penulis	George Leneth Locarso
	Jurnal/Konferensi	JTIK (Jurnal Teknik Informatika Kaputama)
	URL	https://jurnal-backup.kaputama.ac.id/index.php/JTIK/article/view/800
	Permasalahan	Sentimen dari masyarakat merupakan faktor yang penting dalam pemerintah melakukan suatu kebijakan. Dengan memperhatikan <i>feedback</i> dari masyarakat pemerintah dapat mengetahui apa saja yang dapat diubah atau ditingkatkan pada sebuah aplikasi. Untuk itu dilakukan analisis sentimen agar ulasan dari masyarakat lebih mudah disaring.
	Hasil Utama	Data yang digunakan pada penelitian ini merupakan data ulasan yang didapatkan dari <i>review</i> aplikasi PeduliLindungi pada situs <i>Google Play Store</i> dengan jumlah 1179 data. Data komentar kemudian dibagi menjadi 70% data latih dan 30% data uji. Hasil uji menunjukkan akurasi sebesar 83,3%, rata-rata presisi 65% dan rata-rata <i>recall</i> 63%.

	Batasan	Data yang digunakan pada penelitian ini sedikit dan tidak ada algoritma lain sebagai pembanding
15.	22 Juni 2023	
	Paper Ke-15	
	Judul Paper	Analisis Sentimen Pilkada di Tengah Pandemi Covid-19 Menggunakan <i>Convolution Neural Network</i> (CNN)
	Penulis	Sukma Nindi Listyarini, Dimas Aryo Anggoro
	Jurnal/Konferensi	Jurnal Pendidikan dan Teknologi Indonesia (JPTI)
	URL	http://jpti.journals.id/index.php/jpti/article/view/60
	Permasalahan	Pemilihan kepala daerah 2020 menjadi kontroversi karen dilaksanakan ditengah pandemi covid-19. Komentar muncul diberbagai sosial media seperti twitter. Ada masyarakat yang setuju Pilkada dilaksanakan dan ada yang berpendapat untuk menunda sampai pandemi berakhir. Untuk itu dilakukan analisis senimen untuk mengetahui persepsi umu dari masyarakat terkait pelaksanaan Pilkada 2020.
Hasil Utama	Data yang digunakan pada penelitian ini sebanyak 500 <i>tweet</i> . Data akan dikelompokkan kedalam sentimen negatif dan positif. <i>Dataset</i> penelitian dibagi menjadi tiga bagian, 70% <i>data training</i> , 20% <i>data validation</i> dan 10 <i>data testing</i> . Hasil akurasi tertinggi sebesar 90% dengan 4-layer <i>concolutional</i> dan 100 <i>epoch</i>	

	Batasan	<i>Dataset</i> yang digunakan hanya dibatasi pada jumlah 500. Dan hanya menggunakan algoritma CNN dalam melakukan proses klasifikasi
16.	22 Juni 2023	
	Paper Ke-16	
	Judul Paper	Komparasi Algoritma Naïve Bayes dan Support Vektor Machine (SVM) pada Analisis Sentimen Spotify
	Penulis	Ayu Sri Rahayu, Ahmad Fauzi, Rahmat
	Jurnal/Konferensi	Jurnal Sistem Komputer dan Informatika (JSON
	URL	https://ejurnal.stmik-budidarma.ac.id/index.php/JSON/article/view/5398
	Permasalahan	Spotify merupakan aplikasi penyedia musik. Musik menjadi subjek yang menarik untuk diteliti karena bisa membuat perbedaan opini besar dalam opini dan sentimen pendengarnya.
	Hasil Utama	Jumlah data yang digunakan pada penelitian ini sebanyak 1.500 ulasan, setelah dilakukan <i>data preprocessing</i> dan klasifikasi data tebagi menjadi 1159 ulasan positif dan 341 ulasan negatif. Naïve Bayes memperoleh akurasi yang lebih baik dibanding SVM dengan akurasi 86,4%, presisi 69% dan <i>recall</i> 75% sementara SVM akurasi 84%, presisi 71% dan <i>recall</i> 51%
	Batasan	Ulasan yang digunakan hanya berjumlah 1500
17.	14 Januari 2024	
	Paper Ke-17	

	Judul Paper	Analisis Sentimen Aplikasi E-Learning Selama Pandemi Covid-19 dengan Menggunakan Metode Support Vector Machine dan Convolutional Neural Network
	Penulis	Nina Ismaya Pangaribuan, Alek Sander Simbolon, Neni Mona Aruan
	Jurnal/Konferensi	SEMINASTIKA
	URL	https://journal.universitasmulia.ac.id/index.php/seminastika/article/view/236
	Permasalahan	Aplikasi <i>E-learning</i> dibutuhkan masyarakat dalam meningkatkan pendidikan. <i>E-learning</i> yang diteliti adalah Ruangguru dan Zenius karena memiliki banyak pengguna dan memiliki tingkat kepuasan yang tinggi
	Hasil Utama	Jumlah data yang digunakan pada penelitian ini sebanyak 37.438 untuk Ruangguru dan 17.350 untuk Zenius. Data dipisah menjadi data latih dan data uji dengan split data 80:20. Klasifikasi data Ruangguru dengan menggunakan SVM menghasilkan akurasi 83% dan untuk Zenius menghasilkan akurasi 76%. Sementara CNN menghasilkan akurasi 86% untuk aplikasi Ruangguru dan menghasilkan 84% untuk aplikasi Zenius
	Batasan	Tidak menggunakan algoritma pendeteksi <i>typo</i> yang membuat <i>dataset</i> lebih baik
18.	15 Januari 2024	
	Paper Ke-18	
	Judul Paper	Analisis Sentimen Pengguna Twitter terhadap Pembayaran <i>Cashless</i> menggunakan Shopeepay dengan algoritma Random Forest

	Penulis	Thifal Fadiyah Basar, Dian Eka Ratnawati, Issa Arwani
	Jurnal/Konferensi	Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer
	URL	https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/10830
	Permasalahan	Banyaknya pengguna shopeepay di Indonesia menimbulkan banyaknya opini yang diberikan diberbagai platform salah satunya adalah twitter
	Hasil Utama	Data yang digunakan pada penelitian ini berasal dari media sosial twitter. Data tersebut diambil dengan menggunakan pemrograman python dan library tweepy. Data yang sudah terkumpul selanjutnya dibersihkan dan diberikan bobot dengan menggunakan TF-IDF. Selanjutnya data diklasifikasi dengan menggunakan metode Random Forest dengan hasil terbaik yaitu akurasi 95%, presisi 95%, <i>recall</i> 94% dan <i>f1-score</i> 95%
	Batasan	Tidak ada perbandingan dengan algoritma lain karena hanya menggunakan metode klasifikasi Random Forest dan hanya menggunakan satu skenario pembagian data latih dan data uji.
19.	15 Januari 2024	
	Paper Ke-19	
	Judul Paper	Algoritma LSTM-CNN untuk Sentimen Klasifikasi dengan Word2Vec pada Media Online
	Penulis	Dedi Tri Hermanto, Arief Setyanto, Emha Taufiq Luthfi
	Jurnal/Konferensi	Creative Information Technology Journal (CITEC JOURNAL)
	URL	https://citec.amikom.ac.id/main/index.php/citec/article/view/264

	Permasalahan	Media online menghasilkan berbagai macam berita seperti ekonomi, politik, kesehatan dan olahraga. Topik ekonomi adalah salah satu bahasan yang paling menarik karena memiliki dampak langsung kepada masyarakat.
	Hasil Utama	Jumlah data yang digunakan pada penelitian ini adalah sebanyak 1.200 data. Data tersebut dibagi menjadi data training dan data testing dengan masing-masing jumlah 900 dan 300 data. Setelah dilakukan klasifikasi nilai akurasi pada model CNN-LSTM lebih tinggi dari model LSTM dan LSTM-CNN yaitu sebesar 62%, recall 76% dan presisi 66%.
	Batasan	Jumlah data yang digunakan pada penelitian ini hanya 1.200 dan kombinasi klasifikasi yang digunakan hanya CNN dan LSTM
20.	15 Januari 2024	
	Paper Ke-20	
	Judul Paper	Klasifikasi Sentimen Ulasan Film Indonesia dengan Konversi Spech-to-Text (STT) Menggunakan Metode Convolutional Neural Network (CNN)
	Penulis	Nadhifa Ayu Shafirra dan Irhamah
	Jurnal/Konferensi	Jurnal Sains dan Seni ITS
	URL	https://ejurnal.its.ac.id/index.php/sains_seni/article/view/51825
	Permasalahan	Ulasan sebuah film merupakan opini yang bersifat subjektif, ulasan tersebut dapat diolah dengan analisis sentimen untuk menentukan opini tersebut kedalam sentimen tertentu.

	Hasil Utama	Data pada penelitian ini berasal dari audio berupa suara orang yang dikonversi menjadi teks. Data tersebut diproses dan didapatkan 224 potongan ulasan. Film yang dipilih adalah 5 film indonesia yang tayang pada tahun 2019. Film dengan persentase sentimen positif tertinggi adalah film Bumi Manusia sebesar 74% dan film dengan sentimen positif terendah adalah Dilan sebesar 50%
	Batasan	Metode klasifikasi yang digunakan pada penelitian ini hanya CNN
21.	15 Januari 2024	
	Paper Ke-21	
	Judul Paper	Klasifikasi Sentimen pada Twitter dalam Membantu Pemilihan Kandidat Karyawan dengan Menggunakan <i>Convolutional Neural Network</i> dan <i>FastText Embeddings</i>
	Penulis	Farhan Alfarifqi, Warih Maharani, Jati Hiliamsyah Husen
	Jurnal/Konferensi	eProceedings of Engineering
	URL	https://openlibrarypublications.telkomuniversity.ac.id/13085
	Permasalahan	Proses penerimaan karyawan menjadi tahapan awal yang berperan penting untuk meningkatkan perusahaan. Proses penerimaan yang buruk akan berdampak pada tingginya angka resign karyawan. Masalah ini bisa diatasi dengan mengamati karyawan dari sosial media.
	Hasil Utama	Data pada penelitian ini berasal dari twitter dengan jumlah 3582 tweet. Data tersebut dibagi menjadi 2900 data latih dan 682 data uji. Akurasi tertinggi yang didapat yaitu sebesar 84.01%
	Batasan	Kurangnya jumlah data yang digunakan dan metode klasifikasi yang digunakan hanya CNN

22.	15 Januari 2024	
	Paper Ke-22	
Judul Paper	Implementasi Analisis Sentimen Masyarakat Mengenai Kenaikan Harga BBM Pada Komentar Youtube dengan Metode Gaussian Naïve Bayes	
Penulis	Syamsul Mujahidin, Bagus Prasetio, Muchammad Chandra Cahyo Utomo	
Jurnal/Konferensi	Jurnal Vocational Teknok Elektronika dan Informatika	
URL	https://ejournal.unp.ac.id/index.php/voteknika/article/view/118299	
Permasalahan	Banyaknya jumlah pengguna Youtube membuat banyaknya pendapat yang diberikan masyarakat terkait beberapa hal salah satunya adalah tentang kenaikan harga BBM.	
Hasil Utama	Data yang digunakan berasal dari komentar pada salah satu video youtube yang membahas tentang kenaikan BBM. Hasil akurasi terbaik yang didapat pada penelitian ini ada pada akurasi sebesar 74%.	
Batasan	Hanya mengambil data dari satu video yang membahas tentang kenaikan BBM	
23.	15 Januari 2024	
	Paper Ke-23	
Judul Paper	Perbandingan Kinerja <i>Word Embeddings Word2Vec, GloVe</i> , dan <i>FastText</i> pada Klasifikasi Teks	
Penulis	Ariyanti Nurdin, Bernadus Anggo Seno Aji, Anugrayani Bustamin, Zaenal Abidin.	
Jurnal/Konferensi	JURNAL TEKNOKOMPAK	
URL	https://ejurnal.teknokrat.ac.id/index.php/teknokompak/article/732	

	Permasalahan	Karakteristik teks yang tidak terstruktur menjadi tantangan dalam ekstraksi fitur pada bidang pemrosesan teks. Penelitian ini bertujuan untuk membandingkan kinerja metode <i>word embedding</i> yaitu <i>word2vec</i> , <i>glove</i> dan <i>fasttext</i>
	Hasil Utama	<i>FastText</i> menjadi metode word embedding yang menghasilkan akurasi terbaik setelah dilakukan klasifikasi menggunakan CNN. <i>Fasttext</i> mendapatkan akurasi sebesar 97.9% sementara <i>word2vec</i> mendapatkan akurasi 92.5% dan <i>glove</i> sebesar 95.8%. sedangkan pada dataset <i>Reuteers News fasttext</i> memiliki akurasi 71.5%, <i>word2vec</i> 69.4% dan <i>glove</i> 68.8%. <i>FastText</i> mampu merepresentasikan vektor dari kata yang tidak ada dalam korpus
	Batasan	Klasifikasi yang digunakan pada penelitian ini hanya menggunakan CNN
24.	17 Januari 2024	
	Paper Ke-24	
	Judul Paper	Efektivitas Aplikasi Islami Terhadap Muslim Generasi Z dalam Meningkatkan Kemampuan Membaca Al-Qur'an
	Penulis	Ari Hartawan, Imamul Arifin, Yogi Dwi Prasetyo
	Jurnal/Konferensi	Jurnal Studi Al-Qur'an
	URL	https://journal.unj.ac.id/unj/index.php/jsq/article/view/23727
	Permasalahan	Generasi Z melihat <i>smartphone</i> sebagai gaya hidup, banyaknya inovasi yang ditawarkan memudahkan segala aspek kehidupan, salah satunya dalam beribadah dengan membaca Al-Qur'an.

	Untuk itu, perlu diketahui efektivitas penggunaan aplikasi islami dalam meningkatkan kemampuan membaca Al-Qur'an.
Hasil Utama	Dari total 85 responden, 72 diantaranya menggunakan aplikasi Islami di <i>smartphone</i> . 70 diantaranya menggunakan aplikasi Islami untuk membaca Al-Qur'an. Karena dianggap memberikan fleksibilitas untuk membaca Al-Qur'an dimanapun, 70% responden menganggap aplikasi Islami dapat meningkatkan kemampuan mereka untuk membaca Al-Qur'an
Batasan	Jumlah data hanya berasal dari 85 responden

Berdasarkan tabel 2.3 penelitian yang telah dilakukan memiliki perbedaan dan juga persamaan pada latar belakang masalah atau hasil penelitian yang didapat. Faktor seperti latar belakang masalah, algoritma yang digunakan, jumlah data, serta faktor-faktor lain dapat mempengaruhi hasil penelitian. Berdasarkan hal tersebut, pada penelitian yang akan dilakukan juga memiliki faktor-faktor pembeda yang dapat mempengaruhi hasil penelitian, seperti metode yang digunakan, jumlah data, ruang lingkup dan lain-lain.

Penelitian terkait dengan judul “Efektivitas Aplikasi Islami Terhadap Muslim Generasi Z dalam Meningkatkan Kemampuan Membaca Al-Qur'an” yang dilakukan oleh Ari Hartawan, Imamul Arifin, dan Yogi Dwi Prasetyo. Masalah pada penelitian ini yaitu generasi Z saat ini melihat *smarthphone* sudah seperti gaya hidup, hal ini membuat generasi Z mudah terbawa

oleh arus negatif globalisasi. Adanya aplikasi Islami diharapkan dapat menjadi salah satu solusi agar intensitas interaksi muslim generasi Z dengan Al-Qur'an tetap terjaga. Interaksi dengan Al-Qur'an diharapkan dapat menjadi salah satu upaya untuk menjaga moralitas dan akhlak. Objek dari penelitian ini adalah 85 responden muslim generasi Z, dimana 72 diantaranya menggunakan aplikasi Islami dan 70 orang menggunakannya untuk membaca Al-Qur'an. Pada penelitian ini menunjukkan aplikasi yang paling sering digunakan adalah aplikasi Al-Qur'an Indonesia. Hasil dari penelitian ini adalah 70% responden menganggap aplikasi Islami bisa meningkatkan kemampuan membaca Al-Qur'an mereka.

Penelitian terkait selanjutnya dengan judul “Komparasi Algoritma *Naïve Bayes Classifier* dan *Support Vector Machine* (SVM) pada Analisis Sentimen Spotify” yang dilakukan oleh Ayu Sri Rahayu, Ahmad Fauzi dan Rahmat. Masalah pada penelitian ini yaitu aplikasi Spotify sebagai salah satu aplikasi musik yang paling banyak digunakan, mendapatkan banyak ulasan dari pengguna. Ulasan yang diberikan akan berpengaruh untuk pengembangan aplikasi, namun dengan banyaknya ulasan tentunya dibutuhkan metode untuk mengetahui kecenderungan pendapat pengguna untuk aplikasi. Hasil dari penelitian ini menunjukkan metode *Naïve Bayes Classifier* lebih unggul dibanding dengan metode *Support Vector Machine*. NBC memiliki akurasi 86,4%, presisi 69% dan *recall* 75%, sementara SVM memiliki akurasi 84%, presisi 71% dan *recall* 51%.

Penelitian terkait selanjutnya dengan judul “Analisis Sentimen Aplikasi *E-Learning* Selama Pandemi Covid-19 dengan Menggunakan Metode *Support Vector Machine* dan *Convolutional Neural Network*” yang dilakukan oleh Nina Ismaya Pangaribuan, Alek Sander Simbolon dan Nenni Mona Aruan. Masalah pada penelitian ini adalah meningkatnya pengguna aplikasi *e-learning* selama pandemi covid-19, Ruangguru dan Zenius sebagai 2 aplikasi yang banyak digunakan, analisis sentimen dilakukan untuk mengetahui hal apa saja yang bisa dikembangkan oleh aplikasi berdasarkan ulasan yang diberikan oleh penggunanya. Hasil penelitian ini menunjukkan metode *Convolutional Neural Network* lebih unggul dibanding dengan metode *Support Vector*. Pada aplikasi Ruangguru, CNN memperoleh akurasi 86%, sementara SVM memperoleh akurasi 83%. Sedangkan pada aplikasi Zenius CNN memperoleh akurasi 84%, sementara SVM 76%.

Penelitian terkait selanjutnya dengan judul “Perbandingan Kinerja *Word Embeddings Word2Vec, Glove, dan FastText* pada Klasifikais Teks” yang dilakukan oleh Ariyanti Nurdin, Bernadus Anggo Seno Aji, Anugrayani Bustamin dan Zaenal Abidin. Masalah pada penelitian ini adalah karakteristik teks yang tidak terstruktur menjadu tantangan dalam ekstrasi fitur pada bidang pemrosesan teks. Untuk itu diperlukan metode terbaik untuk melakukan proses *word embeddings*. Hasil dari penelitian ini menunjukkan hasil setelah dilakukan klasifikasi menggunakan *Convolutional Neural Network, FastText* mendapatkan akurasi sebesar 97,9%, *Word2Vec* mendapatkan akurasi sebesar 92,5% dan *Glove* memiliki akurasi 95,8%. Pada *dataset* lain yang diuji,

FastText memiliki akurasi sebesar 71,5%, *Word2Vec* 69,4% dan *Glove* 68,8%. Penelitian ini menunjukkan bahwa metode *word embeddings FastText* dapat digunakan dengan baik pada proses pengolahan data teks.

Banyaknya pengguna aplikasi Al-Qur'an Indonesia dan beragamnya fitur yang dimiliki berbanding lurus dengan banyak dan beragamnya ulasan yang diberikan pengguna pada *Google Play Store*. Karena ulasan yang diberikan bersifat tidak terstruktur, oleh karena itu diperlukan analisis sentimen untuk mengetahui kecenderungan pengguna terhadap aplikasi. Pada penelitian-penelitian sebelumnya, metode klasifikasi *Naïve Bayes Classifier* dan *Convolutional Neural Network* masing-masing menghasilkan nilai akurasi yang lebih baik dibandingkan dengan metode lain pada proses analisis sentimen. Oleh karena itu, pada penelitian kali ini akan melakukan analisis sentimen terhadap ulasan aplikasi Al-Qur'an Indonesia dengan menggunakan dua metode yang memiliki akurasi tinggi pada penelitian sebelumnya yaitu *Naïve Bayes Classifier* dan *Convolutional Neural Network*. Selain itu, pada penelitian ini akan digunakan metode pembobotan kata *FastText* yang pada penelitian sebelumnya mempunyai performa yang lebih baik dibanding metode lain yaitu *Word2Vec* dan *Glove*.

2.2.2 Matriks Penelitian

Matriks penelitian berisi informasi seputar judul dan algoritma atau metode yang digunakan dalam penelitian. Matriks penelitian dapat dilihat pada tabel 2.4

Tabel 2.4 Matriks penelitian

No	Judul	Metode / Algoritma						
		SVM	Random Forest	KNN	NBC	CNN	LSTM	RNN
1	<i>Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine</i>	√			√			
2	<i>Analisis Sentimen Aplikasi Novel Online di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)</i>	√						
3	<i>Analisis Sentimen Aplikasi E-Government Pada Google Play Menggunakan Algoritma Naïve Bayes</i>				√			

No	Judul	Metode / Algoritma						
		SVM	Random Forest	KNN	NBC	CNN	LSTM	RNN
4	Analisis Sentimen <i>Review</i> Halodoc Menggunakan <i>Naïve Bayes Classifier</i>				√			
5	Analisis Sentimen Pada <i>Review</i> Aplikasi Grab di <i>Google Play Store</i> Menggunakan Support Vector Machine	√						
6	Analisis Sentimen Ulasan Aplikasi Mola pada <i>Google Play Store</i> Menggunakan Algoritma <i>Support Vector Machine</i>	√						
7	Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit dan Bareksa Dengan Algoritma KNN			√				
8	Analisis Sentimen Aplikasi Gojek Menggunakan Metode SVM dan NBC (Studi Kasus: Komentar Pada Play Store)	√			√			

No	Judul	Metode / Algoritma						
		SVM	Random Forest	KNN	NBC	CNN	LSTM	RNN
9	Analisis Sentimen Aplikasi WETV di <i>Google Play Store</i> Menggunakan Algoritma Support Vector Machine	√						
10	Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naïve Bayes, Random Forest dan Support Vector Machine	√	√		√			
11	Analisis Sentimen Pada Ulasan Aplikasi Jakarta Terkini (JAKI) di <i>Google Play Store</i> Menggunakan Metode <i>Support Vector Machine</i>	√						
12	Sentiment Analysis on Twitter Tweets About Covid-19 Vaccines Using NLP and Supervised KNN Classification Algorithm			√				

No	Judul	Metode / Algoritma						
		SVM	Random Forest	KNN	NBC	CNN	LSTM	RNN
13	Study Analysis of the Effectiveness of Using the Digital Al-Quran among College Students in the Dimensions of Globalization Flow							
14	Analisis Sentimen Review Aplikasi PeduliLindungi pada <i>Google Play Store</i> Menggunakan NBC				√			
15	Analisis Sentimen Pilkada di Tengah Pandemi Covid-19 Menggunakan <i>Convolution Neural Network (CNN)</i>					√		
16	Komparasi Algoritma Naïve Bayes dan Support Vektor Machine (SVM) pada Analisis Sentimen Spotify	√			√			
17	Analisis Sentimen Aplikasi E-Learning Selama Pandemi Covid-19 dengan Menggunakan Metode SVM dan CNN	√				√		

No	Judul	Metode / Algoritma						
		SVM	Random Forest	KNN	NBC	CNN	LSTM	RNN
18	Analisis Sentimen Pengguna Twitter terhadap Pembayaran <i>Cashless</i> menggunakan Shopeepay dengan algoritma Random Forest		√					
19	Algoritma LSTM-CNN untuk Sentimen Klasifikasi dengan Word2Vec pada Media Online					√	√	
20	Klasifikasi Sentimen Ulasan Film Indonesia dengan Konversi Spech-to-Text (STT) Menggunakan Metode Convolutional Neural Network (CNN)					√		
21	Klasifikasi Sentimen pada Twitter dalam Membantu Pemilihan Kandidat Karyawan dengan Menggunakan Convolutional Neural Network dan <i>FastText Embeddings</i>					√		

No	Judul	Metode / Algoritma						
		SVM	Random Forest	KNN	NBC	CNN	LSTM	RNN
22	Implementasi Analisis Sentimen Masyarakat Mengenai Kenaikan Harga BBM Pada Komentar Youtube dengan Metode Gaussian Naïve Bayes				√			
23	Perbandingan Kinerja Word Embeddings <i>Word2Vec</i> , <i>Glove</i> , dan <i>FastText</i> pada Klasifikasi Teks					√		
24	Efektivitas Aplikasi Islami Terhadap Muslim Generasi Z dalam Meningkatkan Kemampuan Membaca Al-Quran							

2.2.3 Relevansi Penelitian

Penelitian yang dilakukan memiliki relevansi atau keterkaitan dengan penelitian sebelumnya. Perbandingan terkait relevansi penelitian dapat dilihat pada tabel 2.5

Tabel 2.5 Relevansi Penelitian

Peneliti	(Ari Hartawan et al., 2022).	(Alif Alparisi, 2023)
Judul	Efektivitas Aplikasi Islami Terhadap Muslim Generasi Z dalam Meningkatkan Kemampuan Membaca Al-Qur'an	Analisis Sentimen Ulasan Aplikasi Al-Qur'an Indonesia menggunakan metode <i>Convolutioanl Neural Network</i> dan <i>Naïve Bayes Classifier</i>
Masalah Penelitian	Generasi Z melihat <i>smartphone</i> sebagai gaya hidup, banyaknya inovasi yang ditawarkan memudahkan segala aspek kehidupan, salah satunya dalam beribadah dengan membaca Al-Qur'an. Untuk itu, perlu diketahui efektivitas penggunaan aplikasi islami dalam meningkatkan kemampuan membaca Al-Qur'an.	Banyaknya pengguna aplikasi Al-Qur'an Indonesia berbanding lurus dengan banyaknya ulasan yang diberikan. Sehingga diperlukan suatu metode untuk mengetahui kecenderungan pengguna terhadap aplikasi Al-Qur'an Indonesia.

Peneliti	(Ari Hartawan et al., 2022).	(Alif Alparisi, 2023)
Objek Penelitian	85 responden muslim generasi Z	Ulasan aplikasi Al-Qur'an Indonesia di <i>Google Play Store</i>
Algoritma/Metode	Pengkategorian pendapat responden	<i>Convolutional Neural Network, Naïve Bayes Classifier, FastText</i>
Dataset	72 responden generasi Z yang memiliki aplikasi islami	Dataset yang digunakan sebanyak 13.436 ulasan aplikasi Al-Qur'an Indonesia pada rentang waktu 1 januari sampai 9 agustus 2023..

Berdasarkan tabel 2.5 terdapat satu penelitian yang memiliki relevansi dengan penelitian yang dilakukan. Penelitian sebelumnya meneliti terkait efektivitas penggunaan aplikasi Al-Qur'an digital. Pada penelitian sebelumnya menunjukkan bahwa aplikasi Al-Qur'an yang paling banyak digunakan adalah aplikasi Al-Qur'an Indonesia. Banyaknya pengguna akan berbanding lurus dengan banyaknya ulasan yang diberikan. Oleh karena itu, perlu dikembangkan dengan meneliti ulasan yang diberikan pengguna untuk mengetahui kecenderungan pendapat pengguna terkait aplikasi Al-Qur'an Indonesia.