

**ANALISIS SENTIMEN TERHADAP ISU KHILAFAH PADA MEDIA
SOSIAL *TWITTER* MENGGUNAKAN ALGORITME *K-NEAREST
NEIGHBORS* DAN *SUPPORT VECTOR MACHINE***

Disusun Sebagai Salah Satu Persyaratan dalam Memperoleh Gelar Sarjana (S-1)
pada Program Studi Teknik Informatika



Diajukan oleh:

FARES ANALIS SYAHAD

17106050033

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA

2023



PENGESAHAN TUGAS AKHIR

Nomor : B-1133/Un.02/DST/PP.00.9/05/2023

Tugas Akhir dengan judul : ANALISIS SENTIMEN TERHADAP ISU KHILAFAH PADA MEDIA SOSIAL TWITTER BERBASIS TEXT MINING MENGGUNAKAN ALGORITME K-NEAREST NEIGHBORS DAN SUPPORT VECTOR MACHINE

yang dipersiapkan dan disusun oleh:

Nama : FARES ANALIS SYAHAD
Nomor Induk Mahasiswa : 17106050033
Telah diujikan pada : Senin, 10 April 2023
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Dr. Ir. Bambang Sugiantoro, S.Si., M.T., IPM.
SIGNED

Valid ID: 645a6c5933af



Penguji I

Agus Mulyanto, S.Si., M.Kom.
SIGNED

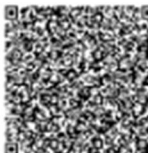
Valid ID: 6459b3e925dc



Penguji II

Dwi Otik Kurniawati, M.Eng.
SIGNED

Valid ID: 64587f1674fc



Yogyakarta, 10 April 2023

UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 645b12899cf64



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir Lamp :

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu 'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Fares Analisis Syahad
NIM : 17106050033
Judul Skripsi : ANALISIS SENTIMEN TERHADAP ISU KHILAFAH PADA
MEDIA SOSIAL TWITTER BERBASIS *TEXT MINIG* MENGGUNAKAN
ALGORITME *K-NEAREST NEIGHBOR* DAN *SUPPORT VECTOR MACHINE*

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Teknik Informatika

Dengan ini saya mengharap agar skripsi/tugas akhir tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya saya ucapkan terima kasih.

Wassalamu 'alaikum wr. wb.

Yogyakarta, 28 Maret 2023

Pembimbing,

Dr. Ir. Bambang Sugiantoro, S.Si., M.T.

NIP. 19751024 200 912 1 002

SURAT PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini:

Nama : Fares Analis Syahad

NIM : 17106050033

Jurusan : Teknik Informatika

Fakultas : Sains dan Teknologi

Menyatakan bahwa:

1. Skripsi yang berjudul: ANALISIS SENTIMEN TERHADAP ISU KHLIFAH PADA MEDIA SOSIAL TWITTER BERBASIS TEXT MINING MENGGUNAKAN ALGORITME K-NEAREST NEIGHBORS DAN SUPPORT VECTOR MACHINE Merupakan hasil karya asli saya yang diajukan untuk memenuhi salah satu persyaratan memperoleh gelar sarjana strata satu (S1) di jurusan Teknik Informatika, Fakultas Sains dan Teknologi, UIN Sunan Kalijaga, Yogyakarta.
2. Semua sumber yang saya gunakan dalam penulisan skripsi ini telah saya cantumkan sesuai dengan ketentuan yang berlaku. Jika di kemudian hari terbukti bahwa karya ini bukan hasil karya saya atau hasil plagiat dari orang lain, maka saya bersedia menerima sanksi yang berlaku.

Yogyakarta, 03 Maret 2023

Yang menyatakan,



Fares Analis Syahad

NIM. 17106050033

INTISARI

Pada bulan Juni 2022 di Indonesia terjadi sebuah penangkapan ketua organisasi masyarakat Khilafatul Muslimin oleh polisi dengan alasan kegiatan kelompok Khilafatul muslimin tidak sesuai dengan nilai-nilai Pancasila dan mengancam keutuhan NKRI. Hal ini mengakibatkan isu khilafah menjadi sebuah topik yang hangat diperbincangkan pada media sosial khususnya pada sosial media twitter. Para pengguna twitter banyak memperbincangkan isu ini dengan berbagai opini mereka baik opini yang memiliki sentimen positif, netral maupun negatif. Penelitian ini bertujuan untuk menganalisis sentimen netizen twitter tentang isu tersebut dengan text mining. Pada penelitian ini algoritma Machine Learning yang digunakan adalah K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM) dan data yang digunakan mengandung kata “khilafah” didapat dengan menggunakan metode crawling. Data tersebut adalah 3630 data tweet dengan label dan 397 data tweet tanpa label. Analisis data menggunakan confusion matriks dan k-fold cross validation. Hasil dari penelitian ini dapat disimpulkan bahwa metode SVM lebih baik daripada KNN dengan nilai akurasi menggunakan confusion matriks adalah 87% untuk SVM dan 56% untuk KNN dan nilai rata-rata akurasi menggunakan k-fold cross validation adalah 60% untuk SVM dan 58% untuk KNN. Hasil dari implementasi model SVM pada 397 data tanpa label adalah 241 tweet (60.71%) sentimen positif, 107 tweet (26.95%) sentimen netral dan 49 tweet (12.34%) sentimen negatif. Sedangkan implementasi menggunakan KNN menghasilkan 220 tweet (55.42%) sentimen positif, 94 tweet (23.68%) sentimen netral dan 83 tweet (20.91%) sentimen negatif. Dari statistik tersebut dapat ditarik kesimpulan bahwa sentimen netizen twitter terhadap isu khilafah cenderung positif.

Kata Kunci: Analisis Sentimen, Text Mining, Machine Learning, Data Mining.

ABSTRACT

In June 2022 in Indonesia there was an arrest by the police of the head of the Khilafatul Muslimin community organization on the grounds that the activities of the Khilafatul Muslimin group were not in accordance with the value of Pancasila and threatened the integrity of the Unitary State of the Republic of Indonesia. This resulted in the khilafah issue becoming a hot topic of discussion on social media, especially on social media Twitter. Twitter users discuss this issue a lot with their various opinions, both opinions that have positive, neutral and negative sentiments. This research aims to analyze the sentiment of twitter netizens about the issue with text mining. In this research, the Machine Learning algorithms used are K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) and the data used contains the word "khilafah" obtained using the crawling method. The data is 3630 tweet data with labels and 397 tweet data without labels. Data analysis uses confusion matrix and k-fold cross validation. The results of this study can be concluded that the SVM method is better than KNN with the accuracy value using confusion matrix is 87% for SVM and 56% for KNN and the average accuracy value using k-fold cross validation is 60% for SVM and 58% for KNN. The results of the SVM model implementation on 397 unlabeled data are 241 tweets (60.71%) positive sentiment, 107 tweets (26.95%) neutral sentiment and 49 tweets (12.34%) negative sentiment. While the implementation using KNN resulted in 220 tweets (55.42%) positive sentiment, 94 tweets (23.68%) neutral sentiment and 83 tweets (20.91%) negative sentiment. From these statistics it can be concluded that the sentiment of twitter netizens on the issue of khilafah tends to be positive.

Keyword: Sentiment analysis, Text Mining, Machine Learning, Data Mining.

KATA PENGANTAR



Assalamu'alaikum Wr. Wb.

Alhamdulillahirobbil'alamin segala puji dan syukur kehadirat Allah SWT yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penelitian. Shalawat serta salam tidak lupa senantiasa penulis sanjungkan kepada Nabi Muhammad SAW yang akan sangat penulis harapkan syafaatnya di *yaumulqiyamah* nanti, aamiin.

Penelitian ini disusun dengan judul “Analisis Sentimen Terhadap Isu Khilafah Pada Media Sosial *Twitter* Menggunakan Algoritme *K-Nearest Neighbors* Dan *Support Vector Machine*” sebagai syarat untuk menyelesaikan tugas akademik mahasiswa Jurusan Teknik Informatika di Universitas Islam Negeri Sunan Kalijaga Yogyakarta.

Dalam penyusunan penelitian ini banyak hambatan dan rintangan yang penulis hadapi, namun pada akhirnya dapat terselesaikan berkat adanya bimbingan dan bantuan dari berbagai pihak yang senantiasa membantu secara moral maupun spiritual. Untuk itu penulis tidak lupa menyampaikan ucapan terimakasih kepada:

1. Kedua Orang Tua yang selalu memberikan dukungan, semangat, inspirasi dan motivasi dalam bentuk moral, material, maupun spiritual, sehingga penulis dapat menyelesaikan penelitian ini dengan keadaan sehat.
2. Prof. Dr.Phil Al Makin, S.Ag., M.A., selaku Rektor UIN Sunan Kalijaga Yogyakarta.
3. Dr. Dra. Hj. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
4. Maria Ulfa Siregar, S.Kom., MIT., Ph.D., selaku Ketua Jurusan Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta dan Dosen Pembimbing Akademik.
5. Dr. Ir. Bambang Sugiantoro, S.Si., M.T., IPM. selaku Dosen Pembimbing Skripsi.

6. Seluruh Dosen, Staf dan Karyawan di Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta, atas segala bantuan dan kerjasamanya selama ini.
7. Teman-teman seperjuangan mahasiswa-mahasiswi Jurusan Teknik Informatika UIN Sunan Kalijaga Yogyakarta yang telah memberikan dukungan serta do'a sehingga penelitian ini terselesaikan.
8. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah membantu dalam penelitian ini.

Penulis menyadari penyusunan penelitian ini jauh dari kesempurnaan. Oleh sebab ini, penulis memohon kepada pembaca atas kritik dan saran yang membangun guna melengkapi dan memperbaiki kekurangan dalam penyusunan penelitiannya. Penulis berharap penelitian ini dapat bermanfaat bagi pembaca, dan khususnya bagi penulis.

Yogyakarta, 29 Maret 2023

Penulis

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA
FARES ANALIS SYAHAD
NIM. 17106050033

HALAMAN PERSEMBAHAN

“ Pertama, penulis persembahkan skripsi ini kepada Allah SWT, Tuhan Yang Maha Esa. Kedua, penulis persembahkan kepada kedua orang tua tercinta yang telah banyak memberikan dukungan dan bimbingan dalam segala aspek kehidupan. Dan yang terakhir, penulis persembahkan untuk almamater Universitas Islam Negeri Sunan Kalijaga Yogyakarta dan teman-teman Jurusan Teknik Informatika angkatan 2017. “



MOTTO

“Ilmu itu lebih baik daripada harta. Ilmu menjaga engkau sedangkan engkau menjaga harta. Ilmu itu penghukum sedangkan harta terhukum. Harta itu kurang apabila dibelanjakan, tapi ilmu bertambah bila dibelanjakan”. – Ali bin abi thalib



DAFTAR ISI

PENGESAHAN TUGAS AKHIR	i
SURAT PERSETUJUAN SKRIPSI.....	ii
SURAT PERNYATAAN KEASLIAN SKRIPSI	iii
INTISARI.....	iv
ABSTRACT	v
KATA PENGANTAR.....	vi
HALAMAN PERSEMBAHAN	viii
MOTTO	ix
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan	2
1.4 Batasan Masalah	3
1.5 Manfaat	3
1.6 Keaslian Penelitian	4
1.7 Sistematika Penelitian.....	4
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	6
2.1 TINJAUAN PUSTAKA	6
2.2 LANDASAN TEORI.....	8
2.2.1 Khilafah	8
2.2.2 <i>Data mining</i>	9
2.2.3 <i>Text Mining</i>	9
2.2.4 Natural Language Processing.....	9
2.2.5 Analisis Sentimen	10
2.2.6 <i>Twitter</i>	10

2.2.7	Crawling Twiter API	11
2.2.8	Persiapan data	11
2.2.9	<i>Feature Selection</i>	12
2.2.10	Penilaian Sentimen	13
2.2.11	Pembobotan Kata.....	13
2.2.12	Python	15
2.2.13	Confusion Matrix.....	16
2.2.14	Cross Validation	17
2.2.15	<i>Support Vector Machine</i>	18
2.2.16	<i>K-Nearest Neighbor</i>	20
BAB III METODE PENELITIAN.....		23
3.1	Metode Penelitian	23
3.2	Obyek Penelitian.....	23
3.3	Perangkat Penelitian	23
3.3.1	Perangkat Keras (<i>Hardware</i>).....	23
3.3.2	Perangkat Lunak (<i>Software</i>)	23
3.4	Tahapan Penelitian.....	24
3.4.1	Pengumpulan Data.....	24
3.4.2	Seleksi dan Penilaian Sentimen Data.....	25
3.4.3	Persiapan Data	26
3.4.4	Pelatihan Model	32
BAB IV HASIL DAN PEMBAHASAN.....		48
4.1	Analisis Hasil dan Pembahasan	48
4.2	Klasifikasi Kelas Sentimen.....	54
BAB V PENUTUP		58
5.1	Kesimpulan	58
5.2	Saran	59

DAFTAR PUSTAKA.....	60
LAMPIRAN	65
CURRICULUM VITAE	70



DAFTAR TABEL

Tabel 2.1 Kriteria Confusion Matrix	16
Tabel 2.2 Persamaan Setiap Kernel SVM	20
Tabel 3.1 Contoh Data Tweet.....	25
Tabel 3.2 Contoh Penilaian Sentimen Data.....	26
Tabel 3.3 Contoh hasil Cleansing dan Case Folding.....	27
Tabel 3.4 Contoh hasil Tokenisasi	28
Tabel 3.5 Contoh hasil Konversi Bahasa Gaul.....	29
Tabel 3.6 Contoh hasil Stopword Removal.....	30
Tabel 3.7 Contoh hasil Stemming	31
Tabel 3.8 Perbandingan sebelum dan sesudah persiapan data	32
Tabel 3.9 Contoh Data Tweet untuk Pelatihan Model	33
Tabel 3.10 Contoh nilai TF dan IDF masing-masing kata	34
Tabel 3.11 Contoh nilai <i>TF-IDF</i>	36
Tabel 3.12 Pendekatan One-vs-all (OvA)	38
Tabel 3.13 Contoh data latih	38
Tabel 3.14 Contoh perhitungan Linear Kernel.....	39
Tabel 3.15 Contoh perhitungan nilai Y	41
Tabel 3.16 Contoh Nilai <i>Support Vector</i>	42
Tabel 3.17 Menghitung <i>Cosine Similarity</i> Tahap 1.....	44
Tabel 4.1 Perbandingan Akurasi	54
Tabel 4.2 Perbandingan Hasil Prediksi Klasifikasi sentimen.....	56

DAFTAR GAMBAR

Gambar 2.1 Diagram Alir <i>Cross Validation</i>	18
Gambar 2.2 Hyperplane Terbaik dan Margin Maksimum	19
Gambar 2.3 Memperoleh Hyperplane Terbaik dengan Menghitung Tiga Support Vector	19
Gambar 2.4 Ilustrasi <i>K-Nearest Neighbor</i>	21
Gambar 3.1 Diagram Alir Tahapan Penelitian	24
Gambar 3.2 Jumlah tweet masing-masing kelas sentimen.....	33
Gambar 4.1 Ilustrasi K-Fold Cross Validation.....	48
Gambar 4.2 Akurasi 10-Fold <i>Support Vector Machine</i>	49
Gambar 4.3 Akurasi 10-Fold <i>K-Nearest Neighbor</i>	50
Gambar 4.4 Perbandingan akurasi antar nilai K yang berbeda	51
Gambar 4.5 Grafik Confusion matrix dari model SVM.....	52
Gambar 4.6 Grafik confusion matrix dari model KNN.....	53
Gambar 4.7 Diagram Data Hasil Prediksi <i>Support Vector Machine</i>	55
Gambar 4.8 Diagram Data Hasil Prediksi <i>K-Nearest Neighbour</i>	56

BAB I

PENDAHULUAN

1.1 Latar Belakang

Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) telah melakukan survei penetrasi dan perilaku pengguna internet pada tahun 2022. Hasil survei mengatakan ada sekitar 210 juta pengguna internet di Indonesia (naik secara signifikan dari tahun 2018 sebanyak 40 juta pengguna). Dari survei tersebut dengan total sebanyak 7568 responden yang tersebar di 34 provinsi menyatakan alasan utama menggunakan internet adalah untuk mengakses sosial media. Seperti *Facebook, Whatsapp, Telegram, Line, Instagram, Youtube, dan Twitter*.

Sosial media *Twitter* memiliki 18 juta pengguna di Indonesia pada awal tahun 2022 (Datareportal, 2022). Jumlah yang cukup besar tersebut menjadikan *tweet* para penggunanya bervariasi dalam banyak hal yakni Pendidikan, hiburan, pekerjaan, politik, juga agama (Deviyanto & Wahyudi, 2018). Salah satu isu politik juga agama yang menjadi *trending topic* di *Twitter* pada tahun 2022 adalah tentang khilafah. Isu ini didasari oleh penangkapan pemimpin organisasi masyarakat Khilafatul Muslimin oleh polisi pada tanggal 7 Juni 2022 (CNN Indonesia, 2022). Polisi menganggap kegiatan kelompok Khilafatul Muslimin tidak sesuai dengan nilai-nilai Pancasila dan mengancam keutuhan Negara Kesatuan Republik Indonesia (Kompas.com, 2022).

Melalui isu tersebut pengguna *Twitter* menuliskan opini mereka melalui sosial media *Twitter* berdasarkan isu khilafah, hal ini menjadikan *Twitter* sebagai sumber data yang mempunyai potensi untuk dianalisis dan diolah untuk menghasilkan informasi yang bermanfaat mengenai isu politik dan agama di Indonesia. Namun, data yang diperoleh dari *Twitter* bukanlah data yang terstruktur, sehingga membutuhkan proses *Text Mining*.

Dalam penelitian ini, peneliti mengumpulkan data dengan topik khilafah yang diambil dari media sosial *Twitter* berupa opini para pengguna dalam bentuk *tweet* di *Twitter*. Data tersebut kemudian diolah dan dianalisis melalui proses *text mining* sehingga mendapatkan kesimpulan dari para pengguna *Twitter* tentang khilafah.

Dalam penelitian-penelitian sebelumnya banyak yang telah menggunakan algoritme *KNN* dan *Support Vector Machine* untuk membandingkan kedua algoritme tersebut terutama untuk analisis sentimen pada *Twitter*. Dalam penelitian ini peneliti mencoba untuk membandingkan kedua algoritme tersebut dengan rencana menentukan mana yang memiliki akurasi terbaik dari algoritme *KNN* dan algoritme *Support Vector Machine*.

Berdasarkan uraian latar belakang tersebut, maka dirancanglah sebuah program untuk analisis sentimen dengan python. Penelitian ini berjudul “Analisis Sentimen Terhadap Isu Khilafah Menggunakan Algoritme *K-Nearest Neighbors* Dan *Support Vector Machine*” sebagai objek dalam menyelesaikan Tugas Akhir.

1.2 Rumusan Masalah

Berdasarkan pada latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana cara mengumpulkan data yang tepat untuk analisis sentimen?
2. Apa metode yang tepat untuk melakukan persiapan data pada analisis sentimen ?
3. Bagaimana melakukan klasifikasi sentimen pada data yang sudah melewati proses persiapan ?
4. Bagaimana melakukan evaluasi kinerja model pada analisis sentimen?

1.3 Tujuan

Tujuan yang ingin dicapai dari penelitian ini yaitu mengimplementasikan algoritme *K-Nearest Neighbors* dan *Support Vector Machine* dalam mengklasifikasikan sentimen terhadap isu khilafah pada *Twitter*, mengetahui tingkat perbandingan akurasi dari algoritme *K-Nearest Neighbors* dan *Support Vector Machine* dalam mengklasifikasikan sentimen terhadap isu khilafah pada *Twitter*, dan mengetahui hasil opini publik tentang isu khilafah setelah melakukan analisis sentimen menggunakan algoritme *K-Nearest Neighbors* dan algoritme *Support Vector Machine*

1.4 Batasan Masalah

Diperlukan Batasan masalah dalam penelitian ini sehingga dapat diperoleh hasil yang sesuai dengan tujuan pembuatan, Adapun pembatasan masalah pada penelitian ini adalah sebagai berikut:

1. Objek penelitian dilakukan terhadap topik khilafah pada *Twitter*.
2. Data yang dianalisis merupakan data yang diambil dari *Twitter* pada bulan Juni dan Agustus 2022.
3. Data *Tweet* yang diambil yaitu *tweet* yang mengandung topik khilafah.
4. Jumlah data yang digunakan untuk dilabeli sebanyak 3630 data.
5. Rasio pembagian data berlabel yang digunakan adalah 9 banding 1.
6. Jumlah data uji yang tidak dilabeli sebanyak 397 data.
7. Data *tweet* yang diambil berbahasa Indonesia.
8. Data diklasifikasikan menjadi 3 kelas yaitu Positif, Negatif dan Netral.
9. Kata kunci yang digunakan Ketika pengambilan data adalah “khilafah”.

1.5 Manfaat

Hasil penelitian ini diharapkan memiliki manfaat penelitian sebagai berikut:

1. Memperkaya khazanah salah satu bidang ilmu pengetahuan yaitu *Machine learning*, khususnya penerapan *Text Mining* menggunakan algoritme *K-Nearest Neighbor* dan *Support Vector Machine* untuk mengklasifikasikan sentimen pada topik khilafah dengan *platform* media sosial *Twitter*.
2. Mengetahui algoritme mana yang lebih baik antara algoritme *K-Nearest Neighbor* dan *Support Vector Machine* untuk mengklasifikasikan sentimen pada topik khilafah dengan melihat nilai akurasi.
3. Dengan penerapan analisis sentimen tentang isu khilafah dapat digunakan sebagai alat bantu untuk melihat respons masyarakat terhadap topik tersebut, sehingga dapat menjadi informasi yang bermanfaat bagi netizen maupun pemerintah.

1.6 Keaslian Penelitian

Berdasarkan studi Pustaka yang telah dilakukan oleh penulis, belum ditemukan adanya penelitian tentang ANALISIS SENTIMEN TERHADAP ISU KHILAFAH PADA MEDIA SOSIAL *TWITTER* MENGGUNAKAN ALGORITME *K-NEAREST NEIGHBORS* DAN *SUPPORT VECTOR MACHINE*.

1.7 Sistematika Penelitian

Sistematika Penulisan laporan tugas akhir ini disusun dalam beberapa bab yang masing-masing bab menguraikan beberapa pokok pembahasan, adapun sistematika penulisan ini yaitu sebagai berikut:

BAB I PENDAHULUAN

Bab ini meliputi latar belakang, rumusan masalah, batasan masalah, manfaat penelitian, tujuan penelitian, keaslian penelitian, dan sistematika penelitian.

BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI

Bab ini membahas mengenai dasar teori dan pandangan umum tentang konsep/teori apa saja yang berkaitan dengan topik yang diangkat, yang telah dibuat berdasarkan hasil penelitian dan hal-hal yang berguna dalam proses penulisan tugas akhir ini.

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan metode dan tahapan – tahapan yang dilakukan ketika melakukan penelitian , mulai dari cara mengumpulkan data, mempersiapkan data, menganalisis data, dan penerapan software yang dibuat untuk memproses data.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memuat hasil dari penelitian, terdiri dari melakukan pengujian pada metode beberapa kali dan pengujian pada data baru. Hasil pengujian kemudian dianalisis berdasarkan pada teori dan perbandingan dari standar yang sudah ada.

BAB V PENUTUP

Berisi tentang pernyataan singkat berupa kesimpulan dari pembahasan *software* yang dibuat secara keseluruhan dan saran untuk mengembangkan sistem yang lebih baik.



BAB V

PENUTUP

5.1 Kesimpulan

Setelah menyelesaikan serangkaian tahapan analisa sentimen pada data teks *Twitter* dengan objek khilafah maka dapat diambil kesimpulan:

1. Penerapan model klasifikasi dengan menggunakan algoritme *Support Vector Machine* pada data uji berjumlah 397 data menghasilkan 241 label positif, 107 label netral dan 49 label negatif. Sedangkan menggunakan algoritme *K-Nearest Neighbor* diperoleh 220 label positif, 94 label netral dan 83 label negatif.
2. Pengujian analisa sentimen pada isu khilafah lebih baik menggunakan algoritme *Support Vector Machine* daripada algoritme *K-Nearest Neighbor* berdasarkan dari nilai akurasi *Confusion Matrix* sebesar 87%. Sedangkan akurasi *K-Nearest Neighbor* sebesar 56%.
3. Netizen Twitter cenderung memiliki sentimen yang positif terhadap isu khilafah disusul oleh yang memiliki sentimen netral diikuti dengan netizen Twitter yang memiliki sentimen negatif.

5.2 Saran

Pada penelitian selanjutnya disarankan beberapa hal yaitu:

1. Menambah *keyword* lebih dari satu selain *keyword* “khilafah” agar data lebih kaya akan informasi namun, *keyword* tersebut masih harus memiliki hubungan.
2. Jumlah data yang digunakan diharapkan lebih banyak dari sebelumnya untuk meningkatkan pemberian informasi terhadap isu khilafah dan agar dapat lebih meningkatkan keakuratan dan analisis yang dilakukan.
3. Keseimbangan jumlah data pada data latih di setiap kelas harus diperhatikan agar menghindari penurunan hasil akurasi karena data yang tidak seimbang.
4. Menambahkan metode *classifier* lainnya sehingga dapat mengetahui berbagai macam metode dan perbandingan model lainnya.
5. Besar kemungkinan hasil akurasi dapat meningkat apabila dilakukan beberapa optimalisasi seperti dapat mengatasi *emoticon* agar dapat melabeli data dengan lebih sesuai.

DAFTAR PUSTAKA

- 3.1. *Cross-validation: evaluating estimator performance — scikit-learn 1.2.0 documentation.* (t.t.). Diambil 20 Januari 2023, dari https://scikit-learn.org/stable/modules/cross_validation.html
- About Twitter | Our company purpose, principles, leadership.* (t.t.). Diambil 7 Juni 2022, dari <https://about.twitter.com/en/who-we-are/our-company>
- Anam, M. K., Pikir, B. N., & Firdaus, M. B. (2021). Penerapan Naïve Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk Menganalisis Sentimen pada Interaksi Netizen dan Pemerintah. *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 21(1), 139–150. <https://doi.org/10.30812/matrik.v21i1.1092>
- Baita, A., & Cahyono, N. (2021). *ANALISIS SENTIMEN MENGENAI VAKSIN SINOVAC MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) DAN K-NEAREST NEIGHBOR (KNN).*
- De, D., Miranda, S., & Duncan, B. (t.t.). *Python: a programming language for software integration and development Related papers Integration Computing and Visualization for Biomolecular Analysis: An Example Using Python... Bruce Duncan Integration Biomolecular Analysis and Visualization Programming: Flexibility and Interactivity in the Design of... Daniel Stoffer Biomolecular visualization using AVS.*
<http://www.python.org/doc/Comparisons.html>
- Deviyanto, A., & Wahyudi, M. D. R. (2018). Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 3(1), 1. <https://doi.org/10.14421/jiska.2018.31-01>
- Dikiyanti, T. D., Rukmi, A. M., & Irawan, M. I. (2021). Sentiment analysis and topic modeling of BPJS Kesehatan based on twitter crawling data using Indonesian Sentiment Lexicon and Latent Dirichlet Allocation algorithm. *Journal of Physics: Conference Series*, 1821(1). <https://doi.org/10.1088/1742-6596/1821/1/012054>

- Dwiki, A., Putra, A., & Juanita, S. (2021). *Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit Dan Bareksa Dengan Algoritma KNN*. 8(2). <http://jurnal.mdp.ac.id>
- Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. *Procedia Computer Science*, 161, 765–772. <https://doi.org/10.1016/j.procs.2019.11.181>
- Fitriyah, N., Warsito, B., Asih, D., & Maruddani, I. (2020). ANALISIS SENTIMEN GOJEK PADA MEDIA SOSIAL TWITTER DENGAN KLASIFIKASI SUPPORT VECTOR MACHINE (SVM). *JURNAL GAUSSIAN*, 9(3), 376–390. <https://ejournal3.undip.ac.id/index.php/gaussian/>
- Giovani, A. P., Ardiansyah, A., Haryanti, T., Kurniawati, L., & Gata, W. (2020). ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGGUNAKAN ALGORITMA KLASIFIKASI. *Jurnal Teknoinfo*, 14(2), 115. <https://doi.org/10.33365/jti.v14i2.679>
- Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Teknika*, 10(1), 18–26. <https://doi.org/10.34148/teknika.v10i1.311>
- Iskandar, J. W., & Nataliani, Y. (2021). Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(6), 1120–1126. <https://doi.org/10.29207/resti.v5i6.3588>
- Lengkong, N. C., Safitri, O., Machsus, S., Putra, Y. R., Syahadati, A., & Nooraeni, R. (2021). ANALISIS SENTIMEN PENERAPAN PSBB DI DKI JAKARTA DAN DAMPAKNYA TERHADAP PERGERAKAN IHSG. *Jurnal Teknoinfo*, 15(1), 20. <https://doi.org/10.33365/jti.v15i1.866>
- Mardiana, T., Syahreva, H., & Tuslaela, T. (2019). KOMPARASI METODE KLASIFIKASI PADA ANALISIS SENTIMEN USAHA WARALABA BERDASARKAN DATA TWITTER. *Jurnal Pilar Nusa Mandiri*, 15(2), 267–274. <https://doi.org/10.33480/pilar.v15i2.752>

- Muttaqin, M. N., & Kharisudin, I. (2021). Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor. *UNNES Journal of Mathematics*, 10(2), 22–27.
<http://journal.unnes.ac.id/sju/index.php/ujm>
- Novantirani, A., Kania Sabariah, M. S., & Effendy, V. (t.t.). *Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota dengan Metode Support Vector Machine*.
- Nurjanah, W. E., Setya Perdana, R., & Fauzi, M. A. (2017). *Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet*. 1(12), 1750–1757. <http://j-ptiik.ub.ac.id>
- Pamuji, A. (2021). Performance of the K-Nearest Neighbors Method on Analysis of Social Media Sentiment. *JUISI*, 07(01).
- Pradana, H. Y., Slamet, I., & Zukhronah, E. (2022). Analisis Sentimen Kinerja Pemerintahan Menggunakan Algoritma NBC, KNN, dan SVM. Dalam *Universitas Muhammadiyah Tangerang (Vol. 4)*.
<http://jurnal.umt.ac.id/index.php/senamu/index>
- Puspita, R., & Widodo, A. (2021). Perbandingan Metode KNN, Decision Tree, dan Naïve Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS. *Jurnal Informatika Universitas Pamulang*, 5(4), 646.
<https://doi.org/10.32493/informatika.v5i4.7622>
- Rahman Isnain, A., Indra Sakti, A., Alita, D., & Satya Marga, N. (2021). Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jdmsi*, 2(1), 31–37. <https://t.co/NfhmfMjtXw>
- Rangga, M., Nasution, A., & Hayaty, M. (2019). Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter. *JURNAL INFORMATIKA*, 6(2), 212–218.
<http://ejournal.bsi.ac.id/ejurnal/index.php/ji>
- Resti Wardani, N., Saepudin, S., & Warman, C. (2022). Sentimen Analisis Kegiatan Trading Pada Ap-likasi Twitter dengan Algoritma SVM, KNN Dan

- Random Forrest. Dalam *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 6, Nomor 2).
- Reza Noviansyah, M., Rismawan, T., Marisa Midyanti, D., Sistem Komputer, J., & MIPA Universitas Tanjungpura Jl Hadari Nawawi, F. H. (2018). PENERAPAN DATA MINING MENGGUNAKAN METODE K-NEAREST NEIGHBOR UNTUK KLASIFIKASI INDEKS CUACA KEBAKARAN BERDASARKAN DATA AWS (AUTOMATIC WEATHER STATION) (STUDI KASUS: KABUPATEN KUBU RAYA). Dalam *Jurnal Coding, Sistem Komputer Untan* (Vol. 06, Nomor 2).
- Rustiana Program Studi Sistem Komputer Perguruan Tinggi Raharja, D., & Rahayu Magister Teknologi Informatika Perguruan Tinggi Raharja, N. (2017). ANALISIS SENTIMEN PASAR OTOMOTIF MOBIL: TWEET TWITTER MENGGUNAKAN NAÏVE BAYES. *Jurnal SIMETRIS*, 8.
- Sabrila, T. S., Sari, V. R., & Minarno, A. E. (2021). Analisis Sentimen Pada Tweet Tentang Penanganan Covid-19 Menggunakan Word Embedding Pada Algoritma Support Vector Machine Dan K-Nearest Neighbor. *Fountain of Informatics Journal*, 6(2), 69. <https://doi.org/10.21111/fij.v6i2.5536>
- Syarifuddin, M. (2020). ANALISIS SENTIMEN OPINI PUBLIK TERHADAP EFEK PSBB PADA TWITTER DENGAN ALGORITMA DECISION TREE-KNN-NAÏVE BAYES. *INTI NUSA MANDIRI*, 15(1). <https://doi.org/10.33480/inti.v15i1.1433>
- Syarifuddin, M. (2020). ANALISIS SENTIMEN OPINI PUBLIK TERHADAP EFEK PSBB PADA TWITTER DENGAN ALGORITMA DECISION TREE, KNN, DAN NAÏVE BAYES. *INTI Nusa Mandiri*, 15(1), 87–94. <https://doi.org/10.33480/inti.v15i1.1433>
- Tanjung, I. U. (2018). *Studi Komparative Pendirian Negara Khilafah Di Indonesia*. <https://doi.org/http://dx.doi.org/10.58836/jpma.v0i0.3967>
- Wardani, N. S., Prahutama, A., & Kartikasari, P. (2020). ANALISIS SENTIMEN PEMINDAHAN IBU KOTA NEGARA DENGAN KLASIFIKASI NAÏVE BAYES UNTUK MODEL BERNOULLI DAN MULTINOMIAL. *Jurnal Gaussian*, 9(3), 237–246. <https://ejournal3.undip.ac.id/index.php/gaussian/>

Wiratama Putra, T., & Triayudi, A. (2022). Analisis Sentimen Pembelajaran Daring menggunakan Metode Naïve Bayes, KNN, dan Decision Tree. *Jurnal Teknologi Informasi dan Komunikasi*, 6(1), 2022.
<https://doi.org/10.35870/jti>

