

Deteksi Teks Berita *Hoax* Bahan Bakar Minyak (BBM) Menggunakan Metode K-Nearest Neighbor

Nabilla Rizqi Amalia Nur Asri¹, Rendra Soekarta², Muhammad Yusuf³

Program Studi Teknik Informatika, Universitas Muhammadiyah Sorong

e-mail: nabillamalia123@gmail.com, rendrasoekarta@gmail.com, yusuf@um-sorong.ac.id,

Abstrak

Hoax merupakan informasi yang tidak benar dan berbahaya karena menipu orang untuk mempercayai sesuatu yang tidak benar. Hoax dapat membuat masyarakat resah karena informasi yang tidak diketahui kebenarannya. Karena semakin berkembangnya teknologi komunikasi dan informasi juga membuat hoax dapat beredar dengan cepat. Pada penelitian ini dibutuhkan suatu sistem yang bertujuan untuk meminimalisir keresahan yang akan terjadi dengan membedakan berita hoax dan bukan hoax. Text mining adalah jenis data mining yang mencari pola yang menarik dalam koleksi data teks. Metode yang digunakan dalam text mining yaitu k-nearest neighbor (KNN). Proses mendeteksi berita hoax dapat dilakukan dengan tahap pengumpulan data, preprocessing data membagi data dengan jumlah data train 80% dan data test 20% kemudian mengklasifikasi dengan k-nearest neighbor dengan nilai $k = 1$. Hasil akurasi yang didapatkan sebesar 70,83%.

Kata Kunci : *Klasifikasi, Hoax, Bahan Bakar Minyak, K-Nearest Neighbor*

1. PENDAHULUAN

Hoax merupakan informasi yang tidak benar dan berbahaya karena menipu orang untuk mempercayai sesuatu yang tidak benar [1] Hoax dapat mengganggu masyarakat karena mengandung informasi yang salah. Semakin berkembangnya teknologi dan komunikasi yang sangat pesat mengakibatkan menyebarnya informasi dengan cepat.

Sebagai contoh isu yang dibicarakan beberapa waktu lalu akibat berkembangnya media sosial yang ada di Indonesia yang berakibat buruk bagi masyarakat yaitu penyebaran informasi hoax yang menimbulkan keresahan. Karena hal tersebut perlu diingatkan agar masyarakat dapat menggunakan media sosial secara positif dan jangan dengan mudah untuk percaya berita-berita disekitar dengan tidak benar. Salah satu hoax yang beredar di masyarakat adalah hoax tentang bahan bakar minyak (BBM) yang meluas melalui media sosial. Informasi tersebut membuat panik masyarakat dan berdampak terhadap munculnya kekhawatiran massal.

Untuk mengetahui apakah sebuah berita mengandung informasi hoax atau tidak yaitu dengan melakukan teknik teks mining. Teks mining adalah salah satu metode yang dapat digunakan untuk mengklasifikasi data. Teks mining adalah jenis data mining yang mencari pola yang menarik dalam koleksi data teks. Masalah *clustering*, ekstraksi informasi dan retrieval informasi juga ditangani oleh teks mining. Metode yang dipakai untuk melakukan *text mining* salah satunya adalah *K-Nearest Neighbor* (KNN).

Penelitian terkait teks mining pernah dilakukan oleh dengan judul “Klasifikasi Berita Hoax Pilpers Menggunakan Metode *Modified K-Nearest Neighbor* Dan Pembobotan Menggunakan TF-IDF”[2] dimana didapatkan masalah terkait berita hoax, dikembangkan sebuah prototype untuk mengklasifikasi berita tersebut dan metode yang digunakan Metode K-Nearest Neighbor dengan hasil yang didapatkan nilai *precision* sebesar 93,75%, nilai *Recall* sebesar 90,90% dan nilai *Accuracy* sebesar 92.30%. Ketidakakurat pengujian dari penentuan hasil klasifikasi

dikarenakan kurangnya data latih yang diinputkan kedalam sistem sehingga menimbulkan rendahnya *prosentase* kemiripan antar dokumen.

Penelitian yang telah dilakukan terkait dengan judul “Klasifikasi Berita Clickbait Menggunakan *K-Nearest Neighbor* (KNN)” [3]. Penelitian ini membahas tentang klasifikasi pada berita clickbait dengan sistem yang akan dibuat menggunakan algoritma *k-nearest neighbor* dengan teknik *text mining*. Hasil penelitian yang dilakukan untuk pengklasifikasian berita clickbait, didapatkan hasil terbaik pada jumlah $k=11$ dengan menggunakan scenario 1 pada pembagian data dengan jumlah data sebanyak 800 data latih dan 200 data uji yang menghasilkan akurasi sebesar 71% precision 72% dan Recall 71%. Dapat disimpulkan bahwa jumlah data latih yang semakin banyak akan menambah ketepatan klasifikasi.

Penelitian yang dilakukan dengan judul “Analisa dan Deteksi Konten Hoax Pada Media Berita Indonesia Menggunakan *Machine Learning*” [4]. Penelitian ini membahas tentang menganalisa dan mendeteksi konten hoax, untuk percobaan yang dilakukan bertujuan memilih algoritma terbaik dalam membedakan berita hoax dan berita asli menggunakan metode *text mining* serta pendekatan dengan *machine learning* dan 150 artikel berbahasa Indonesia (50 artikel hoax dan 100 artikel asli) sebagai data yang akan digunakan. Hasil dari penelitian ini didapatkan kesimpulan bahwa algoritma *Random Forest* memiliki akurasi terbaik dalam mengklasifikasikan berita hoax dan berita asli dibandingkan dengan algoritma *Multilayer Perceptron*, *Naïve Bayes*, dan *Support Vector Machine* dengan nilai akurasi 75.37%

Penelitian yang dilakukan dengan judul “Perbandingan Kinerja Metode *Naïve Bayes* Dan *K-Nearest Neighbor* Untuk Klasifikasi Artikel Berbahasa Indonesia” [5] Penelitian ini membahas pengklasifikasi artikel berbahasa Indonesia yang bertujuan menentukan keterkaitan artikel dengan sebuah tema jurnal. Menggunakan banyak metode klasifikasi yaitu metode *Naïve Bayes* dan metode *K-Nearest Neighbor*. *Naïve Bayes*. Dataset yang digunakan sebanyak 40 jurnal. *Naivebayes accuracy* 70% *Recall* 70% *Precision* 70,9% *Error* 30% sedangkan KNN *Accuracy* 40% *Recall* 40% *Precision* 64,1% *Error* 60%. Dari masing-masing metode terbukti lebih unggul kinerja *Naïve Bayes* yang dapat mengklasifikasikan 28 dokumen sedangkan *K-Nearest Neighbor* hanya dapat mengklasifikasikan 16 dokumen.

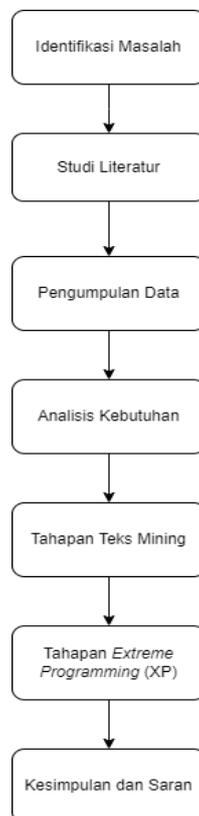
Penelitian yang dilakukan dengan judul “Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode *K-Nearest Neighbor*” [6] Penelitian ini membahas melakukan analisis sentimen pengguna twitter terhadap ketiga pasangan kandidat pada pilkada 2017. Dengan input berupa tweet dalam Bahasa Indonesia, yang akan dilakukan klasifikasi menggunakan algoritma *K-Nearest Neighbor* untuk menentukan sentiment positif atau negatif. Data yang dipakai dalam penelitian ini berjumlah 2000 data tweets dengan perincian 1500 sebagai data latih dan 500 sebagai data uji. Dari hasil yang didapatkan bahwa nilai akurasi terbesar adalah 67,2% ketika $k=5$, presisi tertinggi 56,94% ketika $k=5$, dan recall 78,24% dengan $k=15$.

Penelitian yang dilakukan dengan judul “Klasifikasi Berita Hoax Menggunakan Algoritma *Naïve Bayes* Berbasis PSO” [7] Penelitian ini membahas seberapa besar pengklasifikasi berita hoax menggunakan algoritma *naïve bayes* dan juga seberapa besar pengaruh PSO yang berfokus kepada ketepatan website *turbackhoax.id* dalam mengklasifikasi berita hoax dan non hoax yang kemudian di prosentase menggunakan algoritma *naïve bayes*. Dataset yang digunakan adalah berita dari media sosial yang diperoleh dari *turnbackhoax.id* yang menggunakan dataset 110 dataset dengan rincian 55 berita hoax dan 55 berita non hoax. Hasil terbaik dari *Naïve bayes* berbasis (PSO) adalah dengan nilai *Population size* sebesar 5 dan nilai *inertia Weight* sebesar 1.0, *Accuracy* yang dihasilkan adalah sebesar 91,82%.

2. METODE PENELITIAN

2.1 Alur Penelitian

Dalam melakukan penelitian ini dilakukan beberapa tahapan alur penelitian yaitu dengan identifikasi Masalah, Studi Literatur, Pengumpulan Data, Analisis Kebutuhan, *Data Training dan Data Testing*, Algoritma *K-Nearest Neighbor (KNN)*, Tahapan *Extreme Programming*, Kesimpulan dan Saran.:



Gambar 1 Alur Penelitian

2.2 Tahapan *Text Mining*

Pengolahan teks dilakukan dengan menggunakan model yang disebut text mining [8] untuk mengidentifikasi kata-kata yang dapat secara akurat menggambarkan isi dokumen sehingga antara hubungan dokumen dapat dipelajari, *text mining* didefinisikan sebagai proses ekstraksi data berbasis teks dari dokumen.

1. Scraping Data

Data scraping adalah metode untuk mengumpulkan, menganalisis, dan memproses informasi dari situs web.

2. *Preprocessing Data*

Preprocessing data bertujuan untuk meningkatkan akurasi dan mempercepat komputasi. Secara teknis, *preprocessing*, juga dikenal sebagai transformasi data, dapat mengurangi nilai data tanpa mengubah informasi yang mereka miliki.

3. TF-IDF

Tahapan ini mengubah teks menjadi vector TF-IDF. *Term frequency* (TF) akan menghitung frekuensi setiap kata dalam teks berita, kemudian *inverse document frequency* (IDF) menghitung kebalikan frekuensi setiap kata.

2. Data Train dan data test

Data ini akan dibagi menjadi dua yaitu data train dan data test. Data train atau data latih data yang akan digunakan untuk melatih dan mengembangkan model. Data test digunakan untuk testing model atau simulasi penggunaan model. Data train dan data test ini untuk memberikan hasil prediksi yang lebih akurat. Untuk data train 80% dan data test 20%.

3. *K-Nearest Neighbor*

Metode yang dipakai dalam klasifikasi data adalah metode *K Nearest Neighbor*. Metode tersebut diperlukan data latih dulu sebelum diproses. Tujuan dari algoritma ini mencari klasifikasi data baru berdasarkan tetangga terdekat berdasarkan data latih sebelumnya. Jumlah tetangga terdekat untuk metode ini ditunjukkan dengan variable $k=1$. variable $k=1$ tersebut untuk menentukan nilai prediksinya.

4. *Confusion Matrix* dan *Classification Report*

Pada tahap confusion matrix digunakan untuk menganalisis seberapa baik classifier mengenali tuple dari kelas yang berbeda.

- True Positive* (TP) : sebagai ketepatan prediksi, actual hoax terdeteksi sebagai hoax
- True Negative* (TN) : sebagai aktual bukan hoax terdeteksi sebagai bukan hoax
- False Positive* (FP) : Sebagai jumlah aktual hoax yang terdeteksi sebagai bukan hoax
- False Negative* (FN) : sebagai kesalahan prediksi, dimana aktual bukan hoax dideteksi sebagai hoax.

Berikut adalah nilai-nilai yang digunakan:

- Accuracy* : melakukan deteksi bukan hoax dengan benar sebagai presentase dari jumlah total prediksi yang benar.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

- Precession* : merupakan angka prediksi benar positif dibandingkan dengan nilai keseluruhan hasil yang diprediksi positif.

$$Precision = \frac{TP}{TP + FP}$$

- Recall* : melakukan deteksi sebagian besar bukan hoax tetapi akan ada disertai kesalahan dalam klasifikasi data sehingga ada beberapa hoax terdeteksi sebagai bukan hoax.

$$Recall = \frac{TP}{TP + FN}$$

- F1-Score* : merupakan perbandingan rata-rata dari *precision* dan *recall* yang dibobotkan.

$$F1 = \frac{2 \times Recall \times Precision}{Recall + Precision}$$

Classification Report hasil evaluasi yang didapatkan sulit untuk diukur seberapa akurat model klasifikasi yang telah dibuat. Jadi, data dari *confussion matrix* akan digunakan untuk menghitung nilai-nilai yang telah disimpan oleh *classification report*.

2.3 Tahapan *Extreme Programming* (XP)

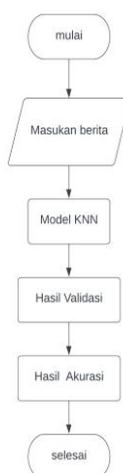
Model pengembangan system yang digunakan *extreme programming* memiliki 4 tahap sebagai berikut:

a. Planning

Tahap ini dilakukan analisa terhadap proses berjalan, kemudian menganalisa kebutuhan sebagai pemecahan masalah pada penelitian dan ini sudah dilakukan di bab sebelumnya.

b. Design

Setelah melakukan tahapan kebutuhan dan pengumpulan data langkah berikutnya yaitu pembuatan desain sistem aplikasi. Design sistem yang dibuat yaitu perancangan design *flowchart system*, *use case diagram system* dan *user interface*.



Gambar 2 Design Flowchart Sistem

c. Coding

Tahap pengkodean merupakan implementasi dari pemodelan sistem yang telah dibuat kedalam kode program agar menghasilkan sebuah sistem.

d. Testing

Tahap pengujian atau peninjauan sistem yang telah dibangun. Langkah ini dilakukan evaluasi memperbaiki kesalahan menguji fitur-fitur sistem secara menyeluruh.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan data

Pengumpulan data berita dilakukan dengan cara scrapping data menggunakan library dari *python* yaitu *newspaper3k* untuk mengakses dan mengambil data berita dari web kompas dan kominfo. Data berita hoax yang diambil dari *website* kominfo sebanyak 41 dan data berita bukan hoax dari kompas sebanyak 285 dengan memasukan bahan bakar minyak (BBM) sebagai kata kunci pencarian.

| Unnamed: 0 | Judul | Penulis | Teks | Rangkuman | Sumber |
|------------|--|--------------|---|---|---------------------------|
| 0 | [DISINFORMASI] KPU Nyatakan Ganjar Pranowo Gag... | Pdsi Kominfo | Penjelasan '\n\nBeredar sebuah unggahan video ... | Penjelasan :Beredar sebuah unggahan video di k... | https://www.kominfo.go.id |
| 1 | [HOAKS] Dirut Pertamina Sita Aset Ahok untuk G... | Pdsi Kominfo | Penjelasan '\n\nBeredar di platform YouTube se... | Penjelasan :Beredar di platform YouTube sebuah... | https://www.kominfo.go.id |
| 2 | [DISINFORMASI] Pertamina Berisi Pertalite | Pdsi Kominfo | Penjelasan '\n\nBeredar sebuah video melalui p... | Pesan berantai itu disertai narasi yang menyeb... | https://www.kominfo.go.id |
| 3 | [DISINFORMASI] Mobil Wakil Presiden Kehabisan ... | Pdsi Kominfo | Penjelasan '\n\nBeredar sebuah video di media ... | Penjelasan :Beredar sebuah video di media sosi... | https://www.kominfo.go.id |
| 4 | [DISINFORMASI] BBM Langka Akibat Depo Pertamina... | Pdsi Kominfo | Penjelasan '\n\nBeredar unggahan di media sosi... | Penjelasan :Beredar unggahan di media sosial F... | https://www.kominfo.go.id |

Gambar 3 Proses Pengumpulan Data

3.2 Preprocessing Data

Pada tahapan ini data yang sudah dikumpulkan akan diolah menjadi data yang siap digunakan dalam proses klasifikasi yang nantinya akan diuji dan dilatih terlebih dahulu.

| Unnamed: 0 | Judul | Penulis | Teks | Rangkuman | Sumber | label |
|------------|--|-----------------------------|---|---|---------------------------|------------|
| 0 | [DISINFORMASI] KPU Nyatakan Ganjar Pranowo Gag... | Pdsi Kominfo | Penjelasan '\n\nBeredar sebuah unggahan video ... | Penjelasan :Beredar sebuah unggahan video di k... | https://www.kominfo.go.id | hoax |
| 1 | [HOAKS] Dirut Pertamina Sita Aset Ahok untuk G... | Pdsi Kominfo | Penjelasan '\n\nBeredar di platform YouTube se... | Penjelasan :Beredar di platform YouTube sebuah... | https://www.kominfo.go.id | hoax |
| 2 | [DISINFORMASI] Pertamina Berisi Pertalite | Pdsi Kominfo | Penjelasan '\n\nBeredar sebuah video melalui p... | Pesan berantai itu disertai narasi yang menyeb... | https://www.kominfo.go.id | hoax |
| 3 | [DISINFORMASI] Mobil Wakil Presiden Kehabisan ... | Pdsi Kominfo | Penjelasan '\n\nBeredar sebuah video di media ... | Penjelasan :Beredar sebuah video di media sosi... | https://www.kominfo.go.id | hoax |
| 4 | [DISINFORMASI] BBM Langka Akibat Depo Pertamina... | Pdsi Kominfo | Penjelasan '\n\nBeredar unggahan di media sosi... | Penjelasan :Beredar unggahan di media sosial F... | https://www.kominfo.go.id | hoax |
| ... | ... | ... | ... | ... | ... | ... |
| 320 | Ini Efek jika Sering Gonta-ganti Jenis BBM | Erwin Setiawan | BANDUNG, KOMPAS.com - Mahalnya harga bahan bak... | Kadang pakai BBM berkualitas tinggi, kadang re... | https://www.google.com | bukan hoax |
| 321 | Ini Efek jika Sering Gonta-ganti Jenis BBM | Kompas Com Nur Jamal Sha'id | BANDUNG, KOMPAS.com - Mahalnya harga bahan bak... | Kadang pakai BBM berkualitas tinggi, kadang re... | https://www.google.com | bukan hoax |
| 322 | Ini Efek jika Sering Gonta-ganti Jenis BBM | Dok. Pertamina | BANDUNG, KOMPAS.com - Mahalnya harga bahan bak... | Kadang pakai BBM berkualitas tinggi, kadang re... | https://www.google.com | bukan hoax |
| 323 | Ini Efek jika Sering Gonta-ganti Jenis BBM | Kompas Adityo | BANDUNG, KOMPAS.com - Mahalnya harga bahan bak... | Kadang pakai BBM berkualitas tinggi, kadang re... | https://www.google.com | bukan hoax |

Gambar 4 Preprocessing Data

a. *Cleaning*

Pada tahap ini menghapus data yang terduplikat pada kolom judul dan menghapus data disinformasi dan hoax.

| | Judul | Sumber | label |
|---|---|---------|------------|
| 0 | Alasan Jokowi Tetap Naikkan Harga BBM meski Me... | KOMPAS | bukan hoax |
| 1 | Apa Bahan Bakar Pesawat? | KOMPAS | bukan hoax |
| 2 | Awas Disinformasi! Massa Bakar Pos Polisi saat... | KOMINFO | hoax |
| 3 | BPH Migas: Penyelundupan BBM Bersubsidi Capai ... | KOMPAS | bukan hoax |
| 4 | Bahan Bakar Air, Mungkinkah? | KOMPAS | bukan hoax |

Gambar 5 Proses Cleaning

b. *Case folding*

Pada tahap ini data berita dirubah menjadi huruf kecil hasilnya akan ditempatkan kembali kedalam kolom berita dalam objek dataframe.

| | Judul | Sumber | Label |
|-----|---|---------|------------|
| 0 | alasan jokowi tetap naikkan harga bbm meski me... | KOMPAS | bukan hoax |
| 1 | apa bahan bakar pesawat? | KOMPAS | bukan hoax |
| 2 | awas disinformasi! massa bakar pos polisi saat... | KOMINFO | hoax |
| 3 | bph migas: penyelundupan bbm bersubsidi capai ... | KOMPAS | bukan hoax |
| 4 | bahan bakar air, mungkinkah? | KOMPAS | bukan hoax |
| ... | ... | ... | ... |
| 115 | perbandingan pertalite lama dan pertalite bar... | KOMINFO | hoax |
| 116 | pertamina bagikan subsidi bbm rp700 ribu untu... | KOMINFO | hoax |
| 117 | segera isi bbm karna malam ini harganya akan ... | KOMINFO | hoax |
| 118 | selebaran kenaikan harga bbm 1 september 2022 | KOMINFO | hoax |
| 119 | tautan kuesioner hadiah subsidi bbm pertam... | KOMINFO | hoax |

120 rows x 3 columns

Gambar 6 Proses Case Folding

c. Stopword

Pada tahap ini akan dilakukan proses *stopword*, proses ini bertujuan menghilangkan kata-kata yang tidak relevan dari kolom judul berita. Untuk prosesnya digunakan modul *stopword* dari pustaka *Natural Language Toolkit* (NLTK) dalam *python* untuk mengambil *stopwords* dalam bahasa indonesia.

| | Judul | Sumber | Label |
|---|---|---------|------------|
| 0 | alasan jokowi naikkan harga bbm mengaku pilih... | KOMPAS | bukan hoax |
| 1 | bahan bakar pesawat? | KOMPAS | bukan hoax |
| 2 | awas disinformasi! massa bakar pos polisi demo... | KOMINFO | hoax |
| 3 | bph migas: penyelundupan bbm bersubsidi capai ... | KOMPAS | bukan hoax |
| 4 | bahan bakar air, mungkinkah? | KOMPAS | bukan hoax |

Gambar 7 Proses Stopword

d. Tokenizing

Tahap ini dilakukan proses memecah kalimat atau teks menjadi sebuah kata yang disebut dengan token. Untuk memecahkan kata-kata menjadi token digunakan modul *word_tokenize library* NLTK (*Natural Language Toolkit*) dengan memakai lambda function untuk menerima parameter x yang merupakan teks berita pada setiap baris dan kolom judul berita.

| | Judul | Sumber | Label |
|---|--|---------|------------|
| 0 | [alasan, jokowi, naikkan, harga, bbm, mengaku, ... | KOMPAS | bukan hoax |
| 1 | [bahan, bakar, pesawat, ?] | KOMPAS | bukan hoax |
| 2 | [awas, disinformasi, !, massa, bakar, pos, pol... | KOMINFO | hoax |
| 3 | [bph, migas, :, penyelundupan, bbm, bersubsidi... | KOMPAS | bukan hoax |
| 4 | [bahan, bakar, air, ,, mungkinkah, ?] | KOMPAS | bukan hoax |

Gambar 8 Proses Tokenizing

e. Stemming

Tahap ini dilakukan proses mengubah token yang berimbuhan menjadi kata kunci dengan menghilangkan kata awalan dan akhiran dari kata tersebut.

| | Judul | Sumber | label |
|---|---|---------|------------|
| 0 | [alas, jokowi, naik, harga, bbm, aku, pilih, s... | KOMPAS | bukan hoax |
| 1 | [bahan, bakar, pesawat,] | KOMPAS | bukan hoax |
| 2 | [awas, disinformasi, , massa, bakar, pos, poli... | KOMINFO | hoax |
| 3 | [bph, migas, , selundup, bbm, subsidi, capai, ... | KOMPAS | bukan hoax |
| 4 | [bahan, bakar, air, , mungkin,] | KOMPAS | bukan hoax |

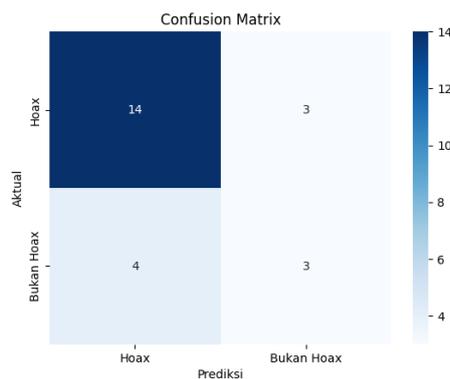
Gambar 9 Proses Stemming

3.3 K- Nearest Neighbor

Hasil akurasi yang didapatkan dari model klasifikasi menggunakan algoritma *K-Nearest Neighbor* dengan nilai $k=1$ adalah 70,83%.

3.4 Confusion Matrix

Hasil dari pengujian ini didapatkan nilai TP (True Positive) dengan nilai sebesar 14, TN (True Negative) dengan nilai sebesar 3, (False Positive) dengan nilai sebesar 3, dan FN (False Negative) dengan nilai sebesar 4.

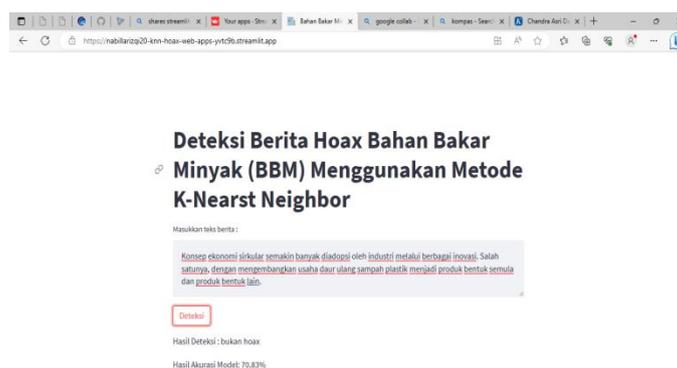


Gambar 10 Confusion Matrix

Berdasarkan perhitungan diatas, maka didapatkan nilai *accuracy* sebesar 71% *precision* 82%, *recall* 78%, dan *f1-score* 75%.

3.5 Halaman Web

Sistem yang dibangun sesuai dengan tahapan pengembangan sistem. Sistem yang akan diimplementasikan di website. Website yang akan dibangun menggunakan bahasa pemrograman *python* dan *streamlit* yang merupakan sebuah dari *library python* untuk membuat web app.



Gambar 11 Tampilan Web

4. KESIMPULAN

Berdasarkan hasil penelitian dan pengujian yang dilakukan peneliti maka dapat diambil kesimpulan dari klasifikasi berita hoax bahan bakar minyak yang peneliti menggunakan algoritma *K-Nearest Neighbor* dengan menggunakan nilai $k=1$. Maka didapatkan nilai akurasi sebesar 70,83%. Peneliti juga melakukan pengujian dengan memasukkan 2 judul berita yang diambil di *website* kompas dan kominfo untuk mendeteksi apakah berita tersebut merupakan berita hoax atau bukan hoax. Hasilnya diperoleh teks berita pertama terdeteksi sebagai berita bukan hoax dengan tingkat akurasi 70,83%, judul teks berita kedua terdeteksi sebagai berita hoax dengan tingkat akurasi 70,83%. Hal ini menunjukkan bahwa proses implementasi model pada deteksi berita hoax dan bukan hoax bahan bakar minyak (BBM) berhasil dilakukan. Dalam penelitian ini dapat disimpulkan klasifikasi berita hoax bahan bakar minyak penting dilakukan.

5. SARAN

Dari kesimpulan yang ada dapat diketahui bahwa penelitian yang dilakukan masih terdapat kekurangan. Oleh karena itu di perlukan saran-saran untuk kebaikan pengembangan sistem lebih lanjut diantaranya :

1. Dapat menambahkan beberapa klasifikasi yang lain agar lebih mempertajam hasil analisis
2. Perlu adanya penambahan dari media sosial lain seperti facebook dan lain lain.
3. Diperlukan lebih banyak lagi data berita yang jelas-jelas hoax untuk melihat tingkat akurasi

UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih kepada keluarga yang telah memberi “dukungan financial” terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] Afriza, A., & Adisantoso, J. (n.d.). Metode Klasifikasi Rocchio untuk Analisis Hoax. *Ilmu Komputer Agri-Informatika*, Volume 5 Nomor Halaman 1-10 .
- [2] Nur Rozi, F., & Harini Sulistyawati, D. (2019). Klasifikasi Berita Hoax Pilpres Menggunakan Metode Modified K-Nearest Neighbor Dan Pembobotan Menggunakan TF-IDF. *Volume 15, Nomor 1*.

- [3] Sagita, R., Enri, U., & Primajaya, A. (2020). Klasifikasi Berita Clickbait Menggunakan K-Nearest Neighbor. *Journal Of Information System*, 230-239
 - [4] Munirul, Ula, Alvanof, M. M., & Triandi, R. (n.d.). Analisa Dan Deteksi Konten Hoax Pada Media Berita Indonesia Menggunakan Machine Learning. 102-110.
 - [5] Wibawa, A. P. (2018). Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa Indonesia. *Jurnal Teknologi Informasi dan Ilmu Komputer*, hal 427-434.
 - [6] Deviyanto, A., & Didik R. Wahyudi, M. (2018). Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor. *Jurnal Informatika Susuna Kalijaga*, 1-13.
 - [7] Muhabatin, H., Prabowo, C., Ali, I., Rohmat, C. L., & Amalia, D. R. (2021). Klasifikasi Berita Hoax Menggunakan Algoritma Naive Bayes Berbasis PSO. *Informatics For Educator And Professionals*, 156-165.
 - [8] Firdaus, A., & Firdaus, W. I. (2021). Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan). *JUPITER*, Hal 66-78.
-