

# Identifikasi Berita Hoax dengan Menerapkan Algoritma Text Mining

Nitha Kumala Dewi

Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: kumalanita7@gmail.com

Email Penulis Korespondensi: kumalanita7@gmail.com

**Abstrak**—Berita merupakan suatu informasi baru atau informasi mengenai sesuatu yang sedang terjadi, disajikan lewat bentuk media cetak, siaran, media sosial atau dari mulut ke mulut kepada orang ketiga atau banyak. Penyebaran berita tersebut kini sangat banyak terjadi baik dari media sosial atau message. Namun yang menjadi masalah adalah kebenaran dari berita yang diterima tersebut tidak ada yang dapat menjaminyanya. Text mining merupakan penerapan konsep dan teknik data mining untuk mencari pola dalam teks, proses penganalisaan teks guna menemukan informasi yang bermanfaat untuk tujuan tertentu. Defenisi lain yang berkaitan dengan text mining dikatakan bahwa text mining merupakan penambangan data yang berupa teks dimana sumber data biasanya didapatkan dari dokumen dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilaku kan analisa keterhubungan antara dokumen. Untuk memudahkan mengidentifikasi berita hoax membutuhkan algoritma text mining, dengan menggunakan algoritma text mining dapat berguna untuk mendapatkan berita yang sebenarnya. Berdasarkan perhitungan algoritma TF-IDF menyimpulkan bahwa kesamaan makna berita yang belum diketahui kebenarannya paling mendekati dengan berita merahputih.com dengan tingkat kepercayaan sebesar 23.11%.

**Kata Kunci:** Identifikasi; Hoax; Text Mining

**Abstract**—News is new information or information about something that is currently happening, presented in the form of print media, broadcasts, social media or word of mouth to third parties or multiple people. The spread of this news is now very much happening both from social media or messages. However, the problem is that no one can guarantee the truth of the news received. Text mining is the application of data mining concepts and techniques to look for patterns in text, the process of analyzing text to find useful information for a particular purpose. Another definition related to text mining is that text mining is data mining in the form of text where the data source is usually obtained from documents and the goal is to find words that can represent the contents of the document so that an analysis of the connectivity between documents can be carried out. To make it easier to identify news hoax requires a text mining algorithm, using a text mining algorithm can be useful to get real news. Based on the calculations of the TF-IDF algorithm, it can be concluded that the similarity of the meaning of the news whose truth is not yet known is closest to the news of merahputih.com with a confidence level of 23.11%.

**Keywords:** Identification; Hoax; Text Mining

## 1. PENDAHULUAN

Berita merupakan suatu informasi baru atau informasi mengenai sesuatu yang sedang terjadi, disajikan lewat bentuk media cetak, siaran, media sosial atau dari mulut ke mulut kepada orang ketiga atau banyak. Pada zaman sekarang ini penyebaran berita sangat berkembang pesat, dimana penyebaran berita pada mulanya dilakukan melalui media cetak kini penyebaran berita dominan dilakukan melalui media elektronik. Selain perkembangan media penyebaran berita juga mengalami perkembangan seperti penyebaran berita secara *broadcast* atau masal yaitu penyebaran berita yang tidak diketahui siapa saja penerimanya [1]-[2].

Berita merupakan suatu hal yang dianggap penting karena berita merupakan salah satu kebutuhan manusia untuk mengetahui informasi terbaru dari berbagai aspek kehidupan seperti ekonomi, cuaca, sosial, politik, agama dan lain sebagainya. Penyebaran berita tersebut kini sangat banyak terjadi baik dari media sosial atau *message*. Namun yang menjadi masalah adalah kebenaran dari berita yang diterima tersebut tidak ada yang dapat menjaminyanya. Informasi dari suatu berita dapat saja diputar balik dari kenyataan yang sebenarnya demi tercapainya tujuan dari suatu individu atau kelompok. Berita *Hoax* adalah pemberitaan palsu. Usaha untuk menipu atau mengakali pembaca/pendengarnya agar mempercayai sesuatu padahal sang pembuat berita palsu tersebut tahu bahwa berita tersebut palsu. Berita *hoax* dari hari ke hari semakin bertambah. Selain dalam bentuk tulisan, *hoax* juga bisa ditemukan dalam bentuk tayangan gambar, beberapa rekayasa gambar di internet yang dibuat dengan teknologi digital.

*Text mining* merupakan variasi dari data *mining* yang berusaha menemukan pola yang menarik dari sejumlah data tekstual yang berjumlah besar. *Text mining* merupakan penerapan konsep dan teknik data mining untuk mencari pola dalam teks, proses penganalisaan teks guna menemukan informasi yang bermanfaat untuk tujuan tertentu. Defenisi lain yang berkaitan dengan *text mining* dikatakan bahwa *text mining* merupakan penambangan data yang berupa teks dimana sumber data biasanya didapatkan dari dokumen dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilaku kan analisa keterhubungan antara dokumen [3]-[4].

berdasarkan penelitian terdahulu yang dilakukan oleh Danu Nur Irwanto pada tahun 2021 dengan judul “Identifikasi Berita Hoax dengan menggunakan kombinasi metode K-nearest dan TF-IDF Berbasis Web” menyimpulkan Sistem dapat mengidentifikasi berita hoax. Selama kategori berita ada dalam dataset [2].

Selanjutnya penelitian terdahulu yang dilakukan oleh Edward dengan judul “Identifikasi berita hoax berbasis web menggunakan algoritma c4.5” menyimpulkan bahwa “Algoritma C4.5 dapat menghasilkan nilai akurasi yang sangat baik dengan tingkat akurasi pada data pengujian untuk data uji sebanyak 100 sebesar 99.7%, dan data uji sebanyak 1.525 sebesar 99.56% [4].

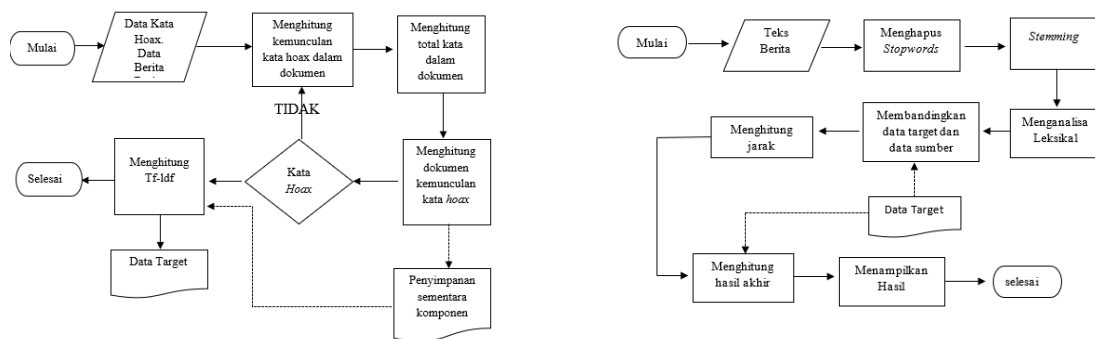
Selanjutnya penelitian terdahulu pada tahun oleh Hamdhan anshari Dkk yang berjudul "Perbandingan Kinerja Algoritma Multinomial Naïve Bayes (Mnb), Multivariate Bernoulli Dan Rocchio Algorithm Dalam Klasifikasi Konten Berita Hoax berbahasa Indonesia pada media sosial. Pada penelitian tersebut menyimpulkan bahwa hasil recall yang paling tinggi pada klasifikasi konten berita hoax berbahasa Indonesia diantara algoritma Multinomial Naïve Bayes (MNB), Multivariate Bernoulli dan Rocchio adalah pada algoritma Bernoulli yaitu sebesar 100% [5].

Berikutnya berdasarkan penelitian terdahulu pada tahun 2022 yang dilakukan oleh Rani Kurnia Putri Dkk yang membahas tentang "Identifikasi Berita Hoax Terkait Virus Corona Menggunakan Long Short-Term Memory. Pada penelitian tersebut menyimpulkan bahwa Hasil dari penelitian menunjukkan bahwa identifikasi berita hoax terkait corona virus dengan LST Mendapatkan hasil yang cukup baik dengan nilai rata-rata yang didapat adalah 51,09% untuk nilai Presisi 51,00% untuk nilai Recall sama dengan perhitungan hasil Akurasi dan 50,41% untuk nilai F-Measure, hasil ini mengindikasikan walaupun secara nilai hasilnya masih dikatakan kurang baik, namun secara konsistensi maka hasil identifikasi ini bisa dikatakan sangat baik jika dilihat dari hasil keseluruhan uji coba yang mana margin antara nilai setiap uji coba tidak lebih dari 5 digit [6].

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Berita merupakan suatu hal yang dianggap penting karena berita merupakan salah satu kebutuhan manusia untuk mengetahui informasi. Penyebaran berita sangat banyak terjadi baik dari media sosial atau *message*. Namun yang menjadi masalah adalah kebenaran dari berita yang diterima tersebut tidak ada yang dapat menjamainya. Adapun Tahapan Penelitian dengan menerapkan Text Mining ialah sebagai berikut :



Gambar 1. Skema penerapan algoritma *Text Mining*

### 2.2 Algoritma

Dalam matematika dan ilmu komputer, algoritma adalah prosedur langkah-langkah yang sistematis, logis, dan lengkap untuk penyelesaian masalah [7]-[8]. Algoritma digunakan untuk penghitungan, pemrosesan data, dan penalaran otomatis. Algoritma adalah metode efektif diekspresikan sebagai rangkaian terbatas dari intruksi-intruksi yang telah didefinisikan dengan baik untuk menghitung sebuah fungsi. Dimulai dari sebuah kondisi awal dan input awal, intruksi-intruksi tersebut menjelaskan sebuah komputasi yang bila dieksekusi diproses sejumlah urutan kondisi terbatas yang terdefinisi dengan baik yang pada akhirnya menghasilkan keluaran dan berhenti dikondisi akhir [9]-[7].

### 2.3 Text Mining

*Text Mining* merupakan penerapan konsep dan teknik data mining untuk mencari pola dalam teks, Proses penganalisaan teks guna menemukan informasi yang bermanfaat untuk tujuan tertentu. Definisi lain yang berkaitan dengan text mining dikatakan bahwa text mining merupakan penambangan data yang berupa teks dimana sumber data biasanya didapatkan dari dokumen dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen [10].

#### 2.2.1 Text Preprocessing

*Text Preprocessing* merupakan tahapan dari proses awal terhadap teks untuk mempersiapkan teks menjadi data yang akan diolah lebih lanjut. Sebuah teks yang harus dipisahkan, hal ini dapat dilakukan dalam beberapa tingkatan yang berbeda. Suatu dokumen dapat dipecahkan menjadi bab, sub bab, paragraf, kalimat dan pada akhirnya menjadi potongan kata. Selain itu pada tahapan ini keberadaan digit angka, huruf kapital atau karakter-karakter yang lainnya dihilangkan dan dirubah. Berikut ini proses penjelasan tahapan *text preprocessing*:

##### 1. Case Folding

Pada penjelasan diatas, telah dijelaskan bahwa pada tahapan *text preprocessing* terdapat beberapa hal yang dirubah, semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf 'a' sampai dengan 'z' yang diterima. Karakter selain huruf dihilangkan dan dianggap delimiter.

##### 2. Tokenizing

Tahap tokenizing adalah tahap pemotongan string input berdasarkan pada tiap kata yang menyusunnya.

### 3. Filtering

Tahap *filtering* adalah tahap mengambil kata-kata penting dari hasil *tokenizing*. Proses *filtering* dapat menggunakan algoritma *stoplist* (membuang kata yang kurang penting) atau *wordlist* (menyimpan kata penting). *Stoplist / stopword* adalah kata-kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-words*. Contoh *stopword* adalah “yang”, “dan”, “di”, “dari” dan lain-lain.

### 4. Stemming

*Stemming* merupakan suatu proses yang terdapat dalam sistem IR yang mentransformasi kata-kata yang terdapat dalam suatu dokumen ke kata-kata akarnya (*root word*) dengan menggunakan aturan-aturan tertentu. *Stemming* kebanyakan digunakan pada teks berbahasa Inggris dikarenakan teks berbahasa Inggris memiliki struktur imbuhan yang tepat dan mudah untuk diolah. Sementara *stemming* untuk proses bahasa Indonesia memiliki struktur imbuhan yang rumit / kompleks sehingga susah diolah.

### 5. Spelling Normalization

Merupakan perbaikan dan substitusi kata-kata yang salah eja ataupun disingkat dengan bentuk tertentu. Substitusi kata yang dilakukan untuk menghindari jumlah perhitungan dimensi kata yang melebar. Perhitungan dimensi kata akan melebar jika kata yang salah eja atau disingkat tidak diubah karena kata tersebut sebenarnya memiliki kontribusi dalam merepresentasikan dokumen tetapi akan dianggap sebagai entitas yang berbeda proses penyusunan matriks.

### 6. Analyzing

Merupakan tahap penentuan seberapa jauh keterhubungan antar kata-kata dokumen yang ada.

## 2.4 Term Frequency – Inverse Document Frequency (TF-IDF)

Algoritma *term Frequency - Inverse Document Frequency* (TF-IDF) merupakan algoritma yang berasal dari bidang *information retrieval*, namun saat ini semakin banyak digunakan dalam perbandingan dokumen [11].

## 2.5 Identifikasi

Identifikasi adalah kegiatan yang mencari, menemukan, mengumpulkan, meneliti, mendaftarkan, mencatat data dan informasi dari kebutuhan lapangan. Identifikasi masalah berisi semua variabel yang diteliti serta kedudukan variabel diantara semua variabel [12]. Fungsi dan tujuan identifikasi berita *hoax* untuk mengetahui berbagai masalah

## 2.6 Berita

Berita adalah informasi baru atau informasi mengenai sesuatu yang sedang terjadi, disajikan lewat bentuk cetak, siaran, internet, atau dari mulut ke mulut kepada orang ketiga atau orang banyak. Peliputan berita adalah proses pengumpulan data dan informasi di lapangan yang dilakukan wartawan atau jurnalis [13].

## 2.7 Hoax

*Hoax* adalah informasi yang sesungguhnya tidak benar, tetapi dibuat seolah-olah benar adanya. Pengertian lain dari *hoax* sebuah pemberitaan palsu atau usaha untuk menipu atau mengakali pembaca atau pendengarnya untuk mempercayai sesuatu padahal sang pencipta berita palsu tersebut tahu bahwa berita tersebut adalah palsu [14]-[15].

## 3. HASIL DAN PEMBAHASAN

Untuk mengidentifikasi berita *hoax*, penulis melakukan perbandingan kesamaan berita dengan berita yang sebenarnya. Untuk memudahkan dalam melakukan analisa, maka penulis telah menentukan judul berita “Bupati Malang Rendra Kresna Terima Suap Rp 3,45 Miliar Untuk Bayar Utang Kampaye” yang diambil dari 9 situs berita terpercaya dengan menggunakan teknik web data extraction, adapun berita tersebut berasal dari Tribunnews.com, Merahputih.com, Merdeka.com, Detiknews.com, Liputan6.com, Kompas.com, Kumparan.com, Suara.com, Tempo.com, Isi berita dari setiap website tersebut adalah sebagai berikut : Tribunnews.com :KPK Sebut Bupati Malang Rendra Kresna Terima Suap Rp 3,45 Miliar untuk Bayar Utang Kampaye

Bupati Malang Rendra Kresna disangka menerima uang hasil korupsi Rp 3,45 miliar. Hal itu disampaikan Wakil Ketua KPK Saut Situmorang dalam jumpa pers di Gedung KPK, Kuningan, Jakarta Selatan, Kamis (11/10/2018). Rendra disebut menerima uang dari Ali Murtopo, seorang pihak swasta. Rendra disuap oleh Ali terkait penunjang sarana peningkatan mutu pendidikan pada Dinas Pendidikan Kabupaten Malang Tahun Anggaran 2011. Selain kasus itu, KPK juga menyebut ada tindak pidana korupsi berupa gratifikasi. Saut mengatakan, Rendra diduga menerima suap dari Ali sebesar Rp 3,45 miliar. Uang diduga akan digunakan untuk membayar utang dana kampanye yang dikeluarkan untuk pemenangan Rendra saat maju menjadi Bupati Malang periode 2010-2015. "KPK menyesalkan masih terjadinya praktik korupsi yang dilakukan kepala daerah, terutama dalam kasus ini (Rendra Kresna) yang digunakan untuk membayar hutang saat kampanye pilkada," tulis KPK dalam rilisnya. Setelah Rendra menjabat, dikumpulkan fee proyek di Kabupaten Malang. Nantinya, uang dari sejumlah proyek itu digunakan untuk membayar hutang. Salah satu yang menjadi perhatian Rendra adalah Dana Alokasi Khusus (DAK) pada Dinas Pendidikan. Rendra menarget DAK Dinas Pendidikan pada 2010 sampai 2013. Khususnya proyek pengadaan buku tingkat SD dan SMP. "Dalam melakukan perbuatannya, Rendra diduga bersama-sama dengan mantan tim sukses saat Pilkada tahun 2010 dan berupaya mengatur

proyek pengadaan barang dan jasa," papar Saut. Sebagai pihak penerima, Rendra disangka melanggar Pasal 12 a atau Pasal 12 b atau Pasal 11 UU Nomor 31 Tahun 1999 sebagaimana diubah dengan UU Nomor 20 Tahun 2001 tentang Pemberantasan Tipikor. Sedangkan sebagai pihak pemberi, Ali disangka melanggar Pasal 5 Ayat (1) a atau b atau Pasal 13 UU Tipikor jo Pasal 55 Ayat (1) ke-1 KUHP.

### 3.1 Case Folding

Adapun tahapan yang dilakukan pada algoritma *Text Mining* diawali dengan melakukan proses *Case Folding*, yaitu melakukan perubahan string menjadi huruf kecil. Hasil dari proses case folding dapat dilihat pada paragraf dibawah ini :

Bupati malang rendra kresna disangka menerima uang hasil korupsi rp 345 miliar hal itu disampaikan wakil ketua kpk saut situmorang dalam jumpa pers di gedung kpk kuningan jakarta selatan Kamis 11/10/2018 rendra disebut menerima uang dari ali murtopo seorang pihak swasta rendra disuap oleh ali terkait penunjang sarana peningkatan mutu pendidikan pada dinas pendidikan kabupaten malang tahun anggaran 2011 selain kasus itu kpk juga menyebut ada tindak pidana korupsi berupa gratifikasi saut mengatakan rendra diduga menerima suap dari ali sebesar rp 3,45 miliar uang diduga akan digunakan untuk membayar utang dana kampanye yang dikeluarkan untuk pemenangan rendra saat maju menjadi bupati malang periode 2010-2015 kpk menyesalkan masih terjadinya praktik korupsi yang dilakukan kepala daerah terutama dalam kasus ini rendra kresna yang digunakan untuk membayar hutang saat kampanye pilkada tulis kpk dalam rilisnya setelah rendra menjabat dikumpulkan fee proyek di kabupaten malang nantinya uang dari sejumlah proyek itu digunakan untuk membayar hutang salah satu yang menjadi perhatian rendra adalah dana alokasi khusus dak pada dinas pendidikan rendra menarget dak dinas pendidikan pada 2010 sampai 2013 khususnya proyek pengadaan buku tingkat sd dan smp dalam melakukan perbuatannya, rendra diduga bersama-sama dengan mantan tim sukses saat pilkada tahun 2010 dan berupaya mengatur proyek pengadaan barang dan jasa papar saut sebagai pihak penerima rendra disangka melanggar pasal 12 a atau pasal 12 b atau pasal 11 uu nomor 31 tahun 1999 sebagaimana diubah dengan uu nomor 20 tahun 2001 tentang pemberantasan tipikor sedangkan sebagai pihak pemberi, ali disangka melanggar pasal 5 ayat 1 a atau b atau pasal 13 uu tipikor jo pasal 55 ayat 1 ke1 kuhp.

### 3.2 Tokenizing

Tahap tokenizing adalah tahap pemotongan string menjadi potongan kata kemudian disusun menjadi baris. Tahap tersebut dapat dilihat dibawah ini :

**malang** bupati malang rendra kresna disangka menerima uang hasil korupsi rp 345 miliar hal itu disampaikan wakil ketua kpk saut situmorang dalam jumpa pers di gedung kpk kuningan jakarta selatan Kamis 11/10/2018 rendra disebut menerima uang dari ali murtopo seorang pihak swasta rendra disuap oleh ali terkait penunjang sarana peningkatan mutu pendidikan pada dinas pendidikan kabupaten malang tahun anggaran 2011 selain kasus itu kpk juga menyebut ada tindak pidana korupsi berupa gratifikasi saut mengatakan rendra diduga menerima suap dari ali sebesar rp 3,45 miliar uang diduga akan digunakan untuk membayar utang dana kampanye yang dikeluarkan untuk pemenangan rendra saat maju menjadi bupati malang periode 2010-2015 kpk menyesalkan masih terjadinya praktik korupsi yang dilakukan kepala daerah terutama dalam kasus ini rendra kresna yang digunakan untuk membayar hutang saat kampanye pilkada tulis kpk dalam rilisnya setelah rendra menjabat dikumpulkan fee proyek di kabupaten malang nantinya uang dari sejumlah proyek itu digunakan untuk membayar hutang salah satu yang menjadi perhatian rendra adalah dana alokasi khusus dak pada dinas pendidikan rendra menarget dak dinas pendidikan pada 2010 sampai 2013 khususnya proyek pengadaan buku tingkat sd dan smp dalam melakukan perbuatannya, rendra diduga bersama-sama dengan mantan tim sukses saat pilkada tahun 2010 dan berupaya mengatur proyek pengadaan barang dan jasa papar saut sebagai pihak penerima rendra disangka melanggar pasal 12 a atau pasal 12 b atau pasal 11 uu nomor 31 tahun 1999 sebagaimana diubah dengan uu nomor 20 tahun 2001 tentang pemberantasan tipikor sedangkan sebagai pihak pemberi, ali disangka melanggar pasal 5 ayat 1 a atau b atau pasal 13 uu tipikor jo pasal 55 ayat 1 ke1 kuhp

### 3.3 Filtering

Pada tahapan ini akan dilakukan penghapusan kata yang tidak penting atau disebut dengan *stopwords*, dalam hal ini penulis memanfaatkan library tala, beberapa daftar kata *stopwords* sebagai berikut : *ada, adalah, adanya, adapun, akan, akankah, akhir, akhiri, akhirnya, aku, akulah, amat, amatlah, anda, andalah, antar, antara, dll.* Hasil dari menghilangkan kata *stopwords* dari berita dapat dilihat pada paragraf dibawah ini :

bupati malang rendra kresna disangka menerima uang hasil korupsi rp 345 miliar wakil ketua kpk saut situmorang jumpa pers gedung kpk kuningan jakarta selatan Kamis 11/10/2018 rendra menerima uang ali murtopo seorang swasta rendra disuap ali terkait penunjang sarana peningkatan mutu pendidikan dinas pendidikan kabupaten malang anggaran 2011 kasus kpk menyebut tindak pidana korupsi gratifikasi saut rendra diduga menerima suap ali rp 3,45 miliar uang diduga membayar utang dana kampanye dikeluarkan pemenangan rendra maju bupati malang periode 2010/2015 kpk menyesalkan praktik korupsi kepala daerah rendra kresna membayar hutang kampanye pilkada tulis kpk rilisnya rendra menjabat dikumpulkan fee proyek kabupaten malang uang proyek digunakan membayar hutang salah perhatian rendra dana alokasi khusus dak dinas pendidikan rendra menarget dak dinas pendidikan 2010/2013 proyek pengadaan buku tingkat sd dan smp perbuatannya rendra diduga mantan tim sukses pilkada 2010 berupaya mengatur proyek pengadaan barang jasa papar saut penerima rendra disangka melanggar pasal 12 a pasal 12 b atau pasal 11 uu nomor 31/1999

diubah uu nomor 20 2001 pemberantasan tipikor pemberi ali disangka melanggar pasal 5 ayat 1 a b atau pasal 13 uu tipikor jo pasal 55 ayat 1 ke1 kuhp.

### 3.4 Stemming

Pada tahapan ini dilakukan proses mentransformasikan kata-kata yang menggunakan aturan tertentu seperti menghilangkan awalan dan akhiran untuk mendapatkan kata-kata akarnya atau disebut dengan *root*. Hasil dari proses *stemming* dapat dilihat pada paragraf dibawah ini :

bupati malang rendra kresna sangka nerima uang hasil korupsi rp 345 miliar wakil ketua kpk saut situmorang jumpa pers gedung kpk kuningan jakarta selatan Kamis 11/10/2018 rendra nerima uang ali murtopo swasta rendra suap oleh ali terkait nunjang sarana peningkat mutu pendidik dinas pendidik kabupaten malang anggaran 2011 kasus kpk nyebut tindak pidana korupsi gratifikasi saut rendra duga nerima suap ali rp 345 miliar uang duga bayar utang dana kampanye keluar pemenang rendra maju bupati malang periode 2010-2015 kpk menyesal praktik korupsi kepala daerah rendra kresna bayar hutang kampanye pilkada tulis kpk rilis rendra jabat kumpul fee proyek kabupaten malang uang proyek guna bayar hutang salah perhatian rendra dana alokasi khusus dak dinas pendidik rendra narget dak dinas pendidik 2010 2013 proyek pengadaan buku tingkat sd smp perbuat rendra duga mantan tim sukses pilkada 2010 berupaya ngatur proyek pengadaan barang jasa papas saut nerima rendra sangka melanggar pasal 12 a pasal 12 b pasal 11 uu nomor 31 1999 ubah uu nomor 20 2001 berantas tipikor beri ali sangka melanggar pasal 5 ayat 1 a b pasal 13 uu tipikor jo pasal 55 ayat 1 ke1 kuhp

### 3.5 Spelling Normalization

Tahapan *spelling normalization* adalah perbaikan dan substitusi kata-kata yang salah eja ataupun disingkat dengan bentuk tertentu. Hasil dari tahapan *spelling normalization* dapat dilihat pada paragraf dibawah ini :

bupati malang rendra kresna sangka nerima uang hasil korupsi rp 345 miliar wakil ketua kpk saut situmorang jumpa pers gedung kpk kuningan jakarta selatan Kamis 11/10/2018 rendra nerima uang ali murtopo swasta rendra suap oleh ali terkait nunjang sarana peningkat mutu pendidik dinas pendidik kabupaten malang anggaran 2011 kasus kpk nyebut tindak pidana korupsi gratifikasi saut rendra duga nerima suap ali rp 345 miliar uang duga bayar utang dana kampanye keluar pemenang rendra maju bupati malang periode 2010-2015 kpk menyesal praktik korupsi kepala daerah rendra kresna bayar hutang kampanye pilkada tulis kpk rilis rendra jabat kumpul fee proyek kabupaten malang uang proyek guna bayar hutang salah perhatian rendra dana alokasi khusus dak dinas pendidik rendra narget dak dinas pendidik 2010 2013 proyek pengadaan buku tingkat sd smp perbuat rendra duga mantan tim sukses pilkada 2010 berupaya ngatur proyek pengadaan barang jasa papas saut nerima rendra sangka melanggar pasal 12 a pasal 12 b pasal 11 uu nomor 31 1999 ubah uu nomor 20 2001 berantas tipikor beri ali sangka melanggar pasal 5 ayat 1 a b pasal 13 uu tipikor jo pasal 55 ayat 1 ke1 kuhp

### 3.6 Analyzing

*Analyzing* merupakan tahap penentuan seberapa jauh keterhubungan antar kata-kata dokumen yang ada. Untuk memudahkan analisa keterhubungan antar kata-kata pada setiap dokumen maka penulis menerapkan algoritma *term frequency-inverse document frequency* (TF-IDF). Algoritma tersebut akan membantu penulis dalam menganalisa hubungan antara masing-masing frase/kalimat dari setiap berita dengan suatu berita yang belum diketahui kebenarannya. Untuk memudahkan penerapan algoritma TF-IDF maka penulis melakukan pengolahan teks yang sama terhadap berita yang belum diketahui kebenarannya, melalui tahapan *Case Folding*, *Tokenizing*, *Filtering*, *Stemming*, *Spelling Normalization* dan *Analyzing*. Berita yang belum diketahui kebenarannya dapat dilihat pada paragraf dibawah ini :

Bupati Malang Rendra Kresna disangka menerima uang hasil korupsi Rp 1,5 triliun. Hal itu disampaikan Wakil Ketua KPK Saut Situmorang dalam jumpa pers di Gedung KPK, Kuningan, Jakarta Selatan, Selasa (11/10/2018). Rendra disebut menerima uang dari Ali Murtopo, seorang pihak swasta. Rendra disuap oleh Ali terkait penunjang sarana pembangunan jalan raya Kabupaten Malang Tahun Anggaran 2011. Selain kasus itu, KPK juga menyebut ada tindak pidana korupsi berupa gratifikasi. Saut mengatakan, Rendra diduga menerima suap dari Ali sebesar Rp 1,5 Triliun. Uang diduga akan digunakan untuk membayar utang dana kampanye yang dikeluarkan untuk pemenangan Rendra saat maju menjadi Bupati Malang periode 2010-2015. "KPK menyesalkan masih terjadinya praktik korupsi yang dilakukan kepala daerah, terutama dalam kasus ini (Rendra Kresna) yang digunakan untuk membayar hutang saat kampanye pilkada," tulis KPK dalam rilisnya. Setelah Rendra menjabat, dikumpulkan fee proyek di Kabupaten Malang. Nantinya, uang dari sejumlah proyek itu digunakan untuk membayar hutang. Salah satu yang menjadi perhatian Rendra adalah Dana Alokasi Khusus (DAK) pada Pembangunan Jalan Raya. Rendra menarget Pembangunan Infrastruktur pada Tahun 2010 sampai 2013. Khususnya proyek Pembangunan Jalan Raya. Rendra diduga bersama-sama dengan mantan tim sukses saat Pilkada tahun 2010 dan berupaya mengatur proyek pembangunan jalan raya papas Saut. Sebagai pihak penerima, Rendra disangka melanggar Pasal 12 a atau Pasal 12 b atau Pasal 11 UU Nomor 31 Tahun 1999 sebagaimana diubah dengan UU Nomor 20 Tahun 2001 tentang Pemberantasan Tipikor. Sedangkan sebagai pihak pemberi, Ali disangka melanggar Pasal 5 Ayat (1) a atau b atau Pasal 13 UU Tipikor jo Pasal 55 Ayat (1) ke-1 KUHP.

Hasil dari proses pengolahan teks tersebut dapat dilihat dibawah ini :

bupati malang rendra kresna sangka nerima uang hasil korupsi rp 15 triliun wakil ketua kpk saut situmorang jumpa pers gedung kpk kuningan jakarta selatan selasa 11102018 rendra nerima uang ali murtopo swasta rendra suap ali terkait nunjang sarana bangun jalan raya kabupaten malang anggaran 2011 kpk nyebut tindak pidana korupsi gratifikasi saut rendra duga nerima suap ali rp 15 triliun uang duga bayar utang dana kampanye keluar pemenangan rendra maju bupati malang periode 20102015 kpk nyesal praktik korupsi kepala daerah, kasus rendra kresna bayar hutang kampanye pilkada tulis kpk rilis rendra jabat kumpul fee proyek kabupaten malang uang proyek bayar hutang salah perhatian rendra dana alokasi khusus dak bangun jalan raya rendra narget bangun insfrastruktur 2010 2013 proyek bangun jalan raya rendra duga mantan tim sukses pilkada 2010 berupaya ngatur proyek bangun jalan raya papar saut nerima rendra sangka langgar pasal 12 a pasal 12 b a pasal 11 uu nomor 31 1999 ubah uu nomor 20 2001 berantas tipikor beri ali sangka langgar pasal 5 ayat 1 a b pasal 13 uu tipikor jo pasal 55 ayat 1 ke 1 kuhp

Hasil dari keterkaitan frase/kalimat berita yang belum diketahui keberannya dengan berita terpercaya menerapkan algoritma TF-IDF dapat dilihat pada tabel 1 berikut.

Tabel 1. Term Frekuensi Masing-Masing Kata

Q	tf										df	$\frac{D}{df}$	idf	idf+1
	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10				
Malang	5	7	3	3	4	9	8	6	6	9	1	0	1	
Bupati	2	6	4	3	3	9	4	3	5	9	1	0	1	
Rendra	11	13	8	8	8	10	9	9	7	9	1	0	1	
Kresna	2	1	1	1	0	2	2	6	1	9	1	0	1	
Sangka	3	2	2	1	2	2	2	7	4	9	1	0	1	
Nerima	4	7	6	3	3	7	1	2	4	9	1	0	1	
Uang	4	1	1	0	1	0	2	0	1	6	1.5	0.176	1.176	
Hasil	1	0	0	0	0	0	0	0	0	1	9	0	1	
Korupsi	3	4	1	0	1	1	2	2	1	8	1.125	0.051	1.051	
rp	2	2	2	2	2	2	1	0	1	8	1.125	0.051	1.051	
15	0	0	0	0	0	0	0	0	0	0	0	0	1	
Triliun	0	0	0	0	0	0	0	0	0	0	0	0	1	
Wakil	1	2	1	1	1	1	1	1	1	9	1	0	1	
Ketua	1	2	1	1	1	2	1	1	1	9	1	0	1	
Kpk	5	3	5	3	5	4	3	4	4	9	1	0	1	
Saut	3	3	5	2	5	1	5	5	8	9	1	0	1	
situmorang	1	1	1	1	1	1	1	1	1	9	1	0	1	
Jumpa	1	1	1	0	0	0	0	0	0	3	3	0.477	1.477	
Pers	1	1	1	0	1	0	1	0	0	4	2.25	0.532	1.532	
Gedung	1	1	1	0	1	0	1	1	1	7	1.285	0.109	1.109	
Kuningan	1	1	1	1	1	0	1	1	0	7	1.285	0.109	1.109	
Jakarta	1	1	1	2	2	0	1	1	2	8	1.125	0.511	1.511	
selatan	1	0	1	1	1	0	0	1	0	5	1.8	0.255	1.255	
Selasa	0	0	0	0	0	0	0	0	0	0	0	0	1	
11102018	1	0	0	0	1	1	0	1	0	4	2.25	0.352	1.352	
Ali	4	3	2	2	2	1	4	4	2	9	1	0		
Murtopo	1	1	1	1	1	1	3	4	2	9	1	0	1	
Swasta	1	1	2	2	2	2	0	1	1	8	1.125	0.051	1.051	
Suap	2	7	4	3	4	4	4	4	4	9	1	0	1	
Terkait	1	2	1	1	2	1	0	2	1	8	1.125	0.051	1.051	
Nunjang	1	1	1	1	1	0	1	1	0	7	1.285	0.109	1.109	
Sarana	1	1	1	1	1	2	1	2	0	8	1.125	0.051	1.051	
Bangun	0	1	0	0	0	0	0	0	0	1	9	0.954	1.954	
Jalan	0	0	0	1	0	1	0	0	0	2	4.5	0.653	1.633	
Raya	0	0	0	0	0	0	0	0	0	0	0	0	1	
Kabupaten	2	5	2	2	3	5	6	3	2	9	1	0	1	
Anggaran	1	0	0	0	0	3	0	1	0	3	3	0.477	1.477	
2011	1	1	1	0	1	1	1	2	1	8	1.125	0.051	1.051	
Nyebut	1	0	0	2	0	0	0	0	0	2	4.5	0.653	1.653	
Tindak	1	3	0	0	0	0	2	1	0	4	2.25	0.352	1.352	
Pidana	1	3	0	0	0	0	2	2	0	4	2.25	0.352	1.352	
Korupsi	3	4	1	0	1	1	2	2	1	8	1.125	0.051	1.051	
Gratifikasi	1	5	3	2	3	3	0	1	2	8	1.125	0.051	1.051	
Duga	3	9	4	1	4	2	4	2	5	9	1	0	1	

Q	tf									D	idf	idf+1	
Bayar	3	2	1	2	1	0	2	1	2	8	1.125	0.051	1.051
Utang	3	2	2	2	2	0	2	1	2	8	1.125	0.051	1.051
Dana	2	2	3	1	3	1	5	4	4	9	1	0	1
Keluar	1	1	1	1	1	0	1	1	1	8	1.125	0.051	1.051
Pemenang	1	0	0	0	0	0	0	0	0	1	9	0.954	1.954
Maju	1	2	1	0	1	0	0	0	1	5	1.8	0.255	1.255
Periode	1	3	1	1	1	2	1	1	3	9	1	0	1
20102015	2	1	1	0	1	1	0	1	2	7	1.285	0.109	1.109
Menyesal	1	0	0	0	0	0	1	0	0	2	4.5	0.653	1.653
Praktik	1	0	0	0	0	0	0	0	0	1	9	0.954	1.954
Kepala	1	0	0	0	0	2	0	0	0	2	4.5	0.653	1.653
Daerah	1	0	0	0	0	3	0	0	0	2	4.5	0.653	1.653
Hutang	2	0	0	0	0	0	0	0	0	1	9	0.954	1.954
kampanye	2	2	2	2	2	0	4	2	3	8	1.125	0.051	1.051
Pilkada	2	3	0	0	0	0	2	1	1	5	1.8	0.255	1.255
Tulis	1	0	0	0	0	0	0	0	0	1	9	0.954	1.954
Rilis	1	0	0	0	0	0	0	0	0	1	9	0.954	1.954
Jabat	1	4	1	0	1	1	0	1	1	7	1.285	0.109	1.109
Kumpul	1	0	0	0	0	0	0	0	0	1	9	0.954	1.954
Fee	1	1	1	1	1	0	1	1	1	8	1.125	0.051	1.051
Proyek	4	4	4	2	4	1	6	3	4	9	1	0	1
Salah	1	0	1	0	1	0	0	0	1	4	2.25	0.352	1.352
Perhatian	1	1	1	0	1	0	1	0	1	6	1.5	0.176	1.176
Alokasi	1	1	1	0	1	1	2	2	1	8	1.125	0.051	1.051
Khusus	1	1	1	0	1	1	2	3	1	8	1.125	0.511	1.511
Dak	2	0	1	0	1	1	2	2	0	6	1.5	0.176	1.176
Narget	1	0	0	0	0	0	0	0	0	1	9	0.954	1.954
Infrastruktur	0	0	0	0	0	0	0	0	0	0	0	0	1
2010	2	4	1	0	1	1	3	2	2	8	1.125	0.051	1.051
2013	1	1	1	0	1	0	1	1	1	7	1.285	0.109	1.109
Mantan	1	1	0	1	0	1	1	1	1	7	1.285	0.109	1.109
Tim	1	0	1	0	1	0	2	2	2	6	1.5	0.176	1.176
Sukses	1	0	1	0	1	0	1	2	0	5	1.8	0.255	1.255
Pilkada	2	3	0	0	0	0	1	1	1	5	1.8	0.255	1.255
Berupaya	1	0	0	1	0	0	0	1	1	4	2.25	0.352	1.352
Ngatur	1	0	0	1	0	0	1	1	1	5	1.8	0.255	1.255
Papar	1	0	0	0	0	0	0	0	0	1	9	0.954	1.954
Langgar	2	0	0	0	0	0	1	1	0	3	3	0.477	1.477
Pasal	5	9	0	0	0	0	7	7	0	4	2.25	0.352	1.352
12	2	3	0	0	0	0	2	1	0	4	2.25	0.352	1.352
A	2	2	0	0	0	0	2	2	0	4	2.25	0.352	1.352
B	2	3	0	0	0	0	2	2	0	4	2.25	0.352	1.352
11	1	1	0	0	0	0	1	1	0	4	2.25	0.352	1.352
Uu	3	0	0	0	0	0	0	0	0	1	9	0.954	1.954
Nomor	2	6	0	0	0	1	4	1	0	5	1.8	0.255	1.255
31	1	3	0	0	0	0	2	0	0	3	3	0.477	1.477
1999	1	3	0	0	0	0	2	0	0	3	3	0.477	1.477
Ubah	1	3	0	0	0	0	2	0	0	3	3	0.477	1.477
20	1	3	0	0	0	0	2	1	0	4	2.25	0.352	1.352
2001	1	3	0	0	0	0	2	1	0	4	2.25	0.352	1.352
Berantas	1	0	1	0	1	1	2	2	1	7	1.285	0.109	1.352
Tipikor	2	0	0	0	0	0	0	1	0	2	4.5	0.653	1.653
Beri	1	1	0	0	0	1	0	2	0	4	2.25	0.352	1.352
5	1	2	0	0	0	0	2	2	0	4	2.25	0.352	1.352
Ayat	2	4	0	0	0	0	3	3	0	4	2.25	0.352	1.352
1	3	4	0	0	0	0	3	2	0	4	2.25	0.352	1.352
13	1	1	0	0	0	0	1	1	0	4	2.25	0.352	1.352
55	1	2	0	0	0	0	1	1	0	4	2.25	0.352	1.352
Kuhp	1	2	0	0	0	0	1	0	0	3	2.25	0.477	1.477

Tabel 2. Tabel Penghitungan Bobot

W= tf*(IDF+1)								
D1	D2	D3	D4	D5	D6	D7	D8	D9
5	7	3	3	4	9	8	6	6
2	6	4	3	3	9	4	3	5
11	13	8	8	8	10	9	9	7
2	1	1	1	0	2	2	6	1
3	2	2	1	2	2	2	7	4
4	7	6	3	3	7	1	2	4
4.704	1.176	1.176	1.176	1.176	0	2.352	0	1.176
1	0	0	0	0	0	0	0	0
3.153	4.204	1.051	0	1.051	1.051	2.102	2.102	1.051
2.102	2.102	2.102	2.102	2.102	2.102	1.051	0	1.051
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
1	2	1	1	1	1	1	1	1
1	2	1	1	1	2	1	1	1
5	3	5	3	5	4	3	4	4
3	3	5	2	5	1	5	5	8
1	1	1	1	1	1	1	1	1
1.477	1.477	1.477	0	0	0	0	0	0
1.532	1.532	1.532	0	1.532	0	1.532	0	0
1.109	1.109	1.109	0	1.109	0	1.109	1.109	1.109
1.109	1.109	1.109	1.109	1.109	0	1.109	1.109	0
1.511	1.511	1.511	3.022	3.022	0	1.511	1.511	3.022
1.255	0	1.255	1.255	1.255	0	0	1.255	0
0	0	0	0	0	0	0	0	0
1.352	0	0	0	1.352	1.352	0	1.352	0
4	3	2	2	2	1	4	4	2
1	1	1	1	1	1	3	4	2
2	7	4	3	4	4	4	4	4
1.051	2.102	1.051	1.051	2.102	1.051	0	2.102	1.051
1.109	1.109	1.109	1.109	1.109	0	1.109	1.109	0
1.051	1.051	1.051	1.051	1.051	2.102	1.051	2.102	0
0	1.954	0	0	0	0	0	0	0
0	0	0	1.633	0	1.633	0	0	0
0	0	0	0	0	0	0	0	0
2	5	2	2	3	5	6	3	2
1.477	0	0	0	0	4.431	0	1.477	0
1.051	1.051	1.051	0	1.051	1.051	1.051	2.102	1.051
1.653	0	0	3.306	0	0	0	0	0
1.352	4.056	0	0	0	0	2.704	1.352	0
1.352	4.056	0	0	0	0	2.704	2.704	0
3.153	4.204	1.051	0	1.051	1.051	2.102	2.102	1.051
1.051	5.255	3.153	2.102	3.153	3.153	0	1.051	2.102
3	9	4	1	4	2	4	2	5
3.153	2.102	1.051	2.102	1.051	0	2.102	1.051	2.102
3.153	2.102	2.102	2.102	2.102	0	2.102	1.051	2.102
2	2	3	1	3	1	5	4	4
1.051	1.051	1.051	1.051	1.051	0	1.051	1.051	1.051
1.954	0	0	0	0	0	0	0	0
1.255	2.51	1.255	0	1.255	0	0	0	1.255
1	3	1	1	1	2	1	1	3
2.218	1.109	1.109	0	1.109	1.109	0	1.109	2.218
1.653	0	0	0	0	0	1.653	0	0
1.954	0	0	0	0	0	0	0	0
1.653	0	0	0	0	3.306	0	0	0
1.653	0	0	0	0	4.959	0	0	0
3.908	0	0	0	0	0	0	0	0
2.102	2.102	2.102	2.102	2.102	0	4.204	2.102	3.153

W= tf*(IDF+1)								
D1	D2	D3	D4	D5	D6	D7	D8	D9
2.51	3.765	0	0	0	0	2.51	1.255	1.255
1.954	0	0	0	0	0	0	0	0
1.954	0	0	0	0	0	0	0	0
1.954	7.816	1.954	0	1.954	1.954	0	1.954	1.954
1.954	0	0	0	0	0	0	0	0
1.051	1.051	1.051	1.051	1.051	0	1.051	1.051	1.051
4	4	4	2	4	1	6	3	4
1.352	0	1.352	0	1.352	0	0	0	1.352
1.176	1.176	1.176	0	1.176	0	1.176	0	1.176
1.051	1.051	1.051	0	1.051	1.051	2.102	2.102	1.051
1.511	1.511	1.511	0	1.511	1.511	3.022	4.533	1.511
2.352	0	1.176	0	1.176	1.176	2.352	2.352	0
1.954	0	0	0	0	0	0	0	0
0.954	0.954	0.954	0.954	0.954	0.954	0.954	0.954	0.954
2.102	4.204	1.051	0	1.051	1.051	3.153	2.102	2.102
1.109	1.109	1.109	0	1.109	0	1.109	1.109	1.109
1.109	1.109	0	1.109	0	1.109	1.109	1.109	1.109
1.176	0	1.176	0	1.176	0	2.352	2.352	2.352
1.255	0	1.255	0	1.255	0	1.255	2.51	0
2.51	3.765	0	0	0	0	1.255	1.255	1.255
1.352	0	0	1.352	0	0	0	1.352	1.352
1.255	0	0	1.255	0	0	1.255	1.255	1.255
1.954	0	0	0	0	0	0	0	0
2.954	0	0	0	0	0	1.477	1.477	0
6.76	12.168	0	0	0	0	9.464	9.464	0
2.704	4.056	0	0	0	0	2.704	1.352	0
2.704	2.704	0	0	0	0	2.704	2.704	0
2.704	4.056	0	0	0	0	2.704	2.704	0
1.352	1.352	0	0	0	0	1.352	1.352	0
5.862	0	0	0	0	0	0	0	0
2.51	7.53	0	0	0	1.255	5.02	1.255	0
1.477	4.431	0	0	0	0	2.954	0	0
1.477	4.431	0	0	0	0	2.954	0	0
1.477	4.431	0	0	0	0	2.954	0	0
1.352	4.056	0	0	0	0	2.704	1.352	0
1.352	4.056	0	0	0	0	2.704	1.352	0
1.352	0	1.352	0	1.352	1.352	2.704	2.704	1.352
3.306	0	0	0	0	0	0	1.653	0
1.352	1.352	0	0	0	1.352	0	2.704	0
1.352	2.704	0	0	0	0	2.704	2.704	0
2.704	5.408	0	0	0	0	4.056	4.056	0
4.056	5.408	0	0	0	0	4.056	2.704	0
1.352	1.352	0	0	0	0	1.352	1.352	0
1.352	2.704	0	0	0	0	1.352	1.352	0
1.477	2.954	0	0	0	0	1.477	0	0
200.848	233.698	103.626	63.994	103.063	106.116	184.65	169.389	114.735

Dari perhitungan algoritma TF-IDF menyimpulkan bahwa kesamaan makna berita yang belum diketahui kebenarannya paling mendekati dengan berita merahputih.com dengan tingkat kepercayaan sebesar 23.11%.

#### 4 KESIMPULAN

Berdasarkan hasil perancangan sistem yang dilakukan oleh penulis tentang identifikasi berita hoax dengan menerapkan algoritma text mining, masih banyak kekurangan penulis dalam penyusunan skripsi ini. Berita hoax dapat diidentifikasi dengan cara membandingkan berita yang diidentifikasi dengan beberapa berita terpercaya. Penerapan algoritma text mining dalam mengidentifikasi berita hoax masih memiliki kelemahan dalam menilai kesamaan berita. Dengan adanya aplikasi identifikasi berita hoax diharapkan dapat membantu masyarakat dalam mengidentifikasi berita yang belum

diketahui kebenarannya. Dari perhitungan algoritma TF-IDF menyimpulkan bahwa kesamaan makna berita yang belum diketahui kebenarannya paling mendekati dengan berita merahputih.com dengan tingkat kepercayaan sebesar 23.11%.

## REFERENCES

- [1] A. W. Fathurrahman, M. Thoriqulhaq, and F. Arianto, "Penerapan Machine Learning untuk Pengklasifikasian Hoaks pada Platform Media Sosial," *Senada*, vol. 2022, no. Senada, pp. 66–68, 2022, [Online]. Available: <https://senada.upnjatim.ac.id/index.php/senada/article/view/48%0Ahttps://senada.upnjatim.ac.id/index.php/senada/article/download/48/26>.
- [2] N. I. Danu, "Identifikasi Berita Hoax Menggunakan Kombinasi Metode K-Nearest Neighbor (KNN) dan TF-IDF Berbasis Web Dengan Menggunakan Framework Codeigniter," 2021, [Online]. Available: <http://eprints.uwp.ac.id/id/eprint/3339/>.
- [3] Y. T. Handika, S. Defit, and G. W. Nurcahyo, "Text Mining Dalam Membandingkan Metode Naïve Bayes Dengan C.45 Dalam Mengidentifikasi Berita Hoax Pada Media Sosial," *Rang Tek. J.*, vol. 5, no. 1, pp. 116–123, 2022, doi: 10.31869/rtj.v5i1.2855.
- [4] E. Edward, "Identifikasi Berita HOAX Berbasis Web Menggunakan Algoritma C4.5," *J. Ilmu Komput. dan Sist. Inf.*, vol. 9, no. 1, p. 53, 2021, doi: 10.24912/jiksi.v9i1.11558.
- [5] H. Ashari, D. Arifianto, H. Azizah, and A. Faruq, "Perbandingan Kinerja Algoritma Multinomial Naive Bayes (MNB, Multivariate Bernoulli dan Rocchio Algorith) Dalam Klasifikasi Konten Berita Hoax Berbahasa Indonesia Pada Media Sosial," *Http://Repository.Unmuhjember.Ac.Id*, pp. 1–12, 2020.
- [6] R. K. Putri and M. Athoillah, "Identifikasi Berita Hoax Terkait Virus Corona Menggunakan Long Short-Term Memory," *Semin. Nas. Has. Ris. dan Pengabd.*, no. April, pp. 506–513, 2022, [Online]. Available: <https://snhrp.unipasby.ac.id/prosiding/index.php/snhrp/article/view/354/298>.
- [7] I. A. Ropikoh, R. Abdulhakim, U. Enri, and N. Sulistiyowati, "Penerapan Algoritma Support Vector Machine (SVM) untuk Klasifikasi Berita Hoax Covid-19," *J. Appl. Informatics Comput.*, vol. 5, no. 1, pp. 64–73, 2021, doi: 10.30871/jaic.v5i1.3167.
- [8] N. Agustina and M. Hermawati, "Implementasi Algoritma Naïve Bayes Classifier untuk Mendeteksi Berita Palsu pada Sosial Media," *Fakt. Exacta*, vol. 14, no. 4, pp. 1979–276, 2021, doi: 10.30998/faktorexacta.v14i4.11259.
- [9] N. T. L. Toruan, "Sistem Pendukung Keputusan Pemilihan Pembawa Acara Berita Terbaik Menerapkan Metode OCRA," *Bull. Comput. Sci. Res.*, vol. 1, no. 3, pp. 71–78, 2021.
- [10] Y. T. Handika, S. Defit, and G. W. Nurcahyo, "TEXT MINING DALAM MEMBANDINGKAN METODE NAÏVE BAYES DENGAN C. 45 DALAM MENGIDENTIFIKASI BERITA HOAX PADA MEDIA SOSIAL," *Rang Tek. J.*, vol. 5, no. 1, pp. 116–123, 2022.
- [11] A. G. Tammam, S. Sucipto, and R. Indriati, "Hoax Detection at Social Media With Text Mining Clarification SystemBased," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 3, no. 2, pp. 94–100, 2018.
- [12] M. Z. Hossain, M. N. Akhtar, R. B. Ahmad, and M. Rahman, "A dynamic K-means clustering for data mining," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 13, no. 2, pp. 521–526, 2019.
- [13] R. Wati, "Penerapan Algoritma Naive Bayes Dan Particle Swarm Optimization Untuk Klasifikasi Berita Hoax Pada Media Sosial," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 5, no. 2, pp. 159–164, 2020, doi: 10.33480/jitk.v5i2.1034.
- [14] M. Mashadi, *AKUNTABILITAS IDEAL*. 2018.
- [15] D. Maulina and R. Sagara, "Klasifikasi artikel hoax menggunakan support vector machine linear dengan pembobotan term frequency–Inverse document frequency," *J. Mantik Penusa*, vol. 2, no. 1, 2018.