

Pengembangan Sistem Analisis Sentimen Berbasis Java Pada Data Twitter Terhadap Omnibus Law Menggunakan Algoritma *Naïve Bayes* dan *K-Nearest Neighbor (K-NN)*

Elok Nur Hamdana¹, Muhammad Balya Iqbal Alfahmi²

^{1,2}Jurusan Teknik Informatika, Politeknik Negeri Malang

¹elok@polinema.ac.id, ²iqbalalfahmi@gmail.com

Abstrak

Undang-undang Cipta Kerja Omnibus law telah di sah kan, pengesahan tersebut mendapat penolakan dari berbagai elemen masyarakat. Hal itu disebabkan Omnibus Law UU Cipta Kerja, dinilai akan membawa dampak buruk bagi tenaga kerja atau buruh. Oleh karena itu peneliti melakukan analisis sentimen dalam bidang data mining terhadap UU Cipta Kerja Omnibus law pada media sosial *Twitter*. Penulis menggunakan metode *Naïve Bayes* dan *KNN* sebagai algoritma yang diterapkan dalam system berbasis *java* untuk memperbandingkan tingkat akurasi dari kedua metode tersebut. Peneliti menggunakan metode *web scrapping* untuk proses pengambilan data dari *twitter* secara *real time*. Hasil penelitian menunjukkan bahwa analisis sentimen terhadap data *Twitter* terhadap UU Cipta Kerja Omnibus law dengan menggunakan metode *Naïve Bayes* mencapai tingkat akurasi 75% dengan *class precision* untuk *pred. positive* adalah 57%, *pred negative* adalah 72%, dan *pred. neutral* adalah 54%. Lalu pada metode *KNN* tingkat akurasi mencapai 88%. Dimana *class precision* untuk *pred. positive* adalah 94%, *pred negative* adalah 61%, dan *pred. neutral* adalah 58%.

Kata kunci : *naive bayes, k-nearest neighbor, omnibus law*

1. Pendahuluan

Dewan Perwakilan Rakyat (DPR), Senin (5/10/2020), telah mengetok palu tanda disahkannya Omnibus Law RUU Cipta Kerja menjadi undang-undang. Pengesahan tersebut dilakukan dalam Rapat Paripurna ke-7 masa persidangan I 2020-2021 di Kompleks Parlemen, Senayan, Jakarta. Di sisi lain, pengesahan tersebut mendapat penolakan dari berbagai elemen masyarakat. Hal itu disebabkan Omnibus Law UU Cipta Kerja, dinilai akan membawa dampak buruk bagi tenaga kerja atau buruh (Rizal, 2020). Maka dari itu pada penelitian ini akan menganalisa untuk mengetahui opini dari masyarakat Indonesia mengenai UU Cipta Kerja Omnibus law yang baru saja di sahkan tersebut.

Biasanya masyarakat mengemukakan pendapatnya melalui sosial media. Salah satu aplikasi yang sering digunakan masyarakat Indonesia adalah *Twitter*. *Twitter* adalah sebuah *platform* untuk menyampaikan opini atau pendapat seseorang. Pertumbuhan pengguna *Twitter* di Indonesia sangat pesat dan menduduki peringkat 5 di dunia. (Utami, 2020). *Twitter* banyak digunakan orang untuk menyampaikan keluh kesahnya mulai dari keluh kesah mengenai kehidupan sehari-hari ataupun keluh kesah terhadap layanan yang diberikan baik dari pemerintah atau bidang lainnya. Oleh karena itu penelitian ini menggunakan *Twitter* sebagai media untuk mengambil data mengenai

opini masyarakat tentang UU cipta kerja omnibus law.

Berdasarkan studi literatur, penelitian mengenai Sentimen Analisis yang dilakukan oleh Irfan dkk tentang: “*Analisis Sentimen Kurikulum 2013 pada Twitter menggunakan Ensemble Feature dan Metode K-Nearest Neighbor*”. (Irfan, 2018). Dalam penelitian ini berdasarkan serangkaian pengujian yang sudah dilakukan, kombinasi fitur berdampak dalam meningkatkan akurasi metode *K-Nearest Neighbor (K-NN)* untuk menentukan opini positif atau negatif. Penggabungan fitur ini dapat melengkapi kelemahan masing masing fitur, sehingga hasil akhir akurasi yang didapatkan dengan menggabungkan kedua fitur tersebut mencapai 96%. Berbeda hal jika hanya menggunakan fitur secara independen saja, akurasi yang didapatkan hanya mencapai 80% pada fitur *Bag of Words (BoW)* dan 82% pada fitur *ensemble* tanpa *Bag of Words (BoW)*.

Penelitian lainnya yang dilakukan oleh Puspita Rani dan Widodo dengan judul: “*Perbandingan Metode KNN, Decision Tree dan Naïve Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS*” (Puspita, 2020). Hasil penelitian menunjukkan bahwa analisis sentimen terhadap data *Twitter* terhadap layanan BPJS dengan menggunakan metode *KNN* mencapai tingkat akurasi 95.58% dengan *class precision* untuk *pred. negative* adalah 45.00%, *pred. positive* adalah 0.00%, dan *pred. neutral* adalah 96.83%. Lalu pada metode *Decision Tree* tingkat akurasinya mencapai

96.13% dengan *class precision* untuk *pred. negative* adalah 55.00%, *pred. positive* adalah 0.00%, dan *pred. neutral* adalah 97.28%. Dan yang terakhir adalah metode *Naïve Bayes* yang mencapai akurasi 89.14% dengan *class precision* untuk *pred. negative* adalah 16.67%, *pred. positive* adalah 1.64%, dan *pred. neutral* adalah 98.40%.

Sentiment analisis adalah cabang dari *data mining*. *Data Mining* merupakan sebuah proses yang dapat mengekstrak informasi sehingga menghasilkan informasi yang sangat berharga. (Nurdin, 2017). Maka dapat dikatakan bahwa *data mining* merupakan proses untuk mencari informasi mengenai teknik tertentu. Pada peneliti ini menggunakan dua metode dalam analisis sentimen untuk memperbandingkan tingkat akurasi dari kedua metode tersebut yaitu metode *Naïve Bayes* dan KNN.

Naïve Bayes adalah metode *machine learning* untuk probabilitas. *Naïve Bayes* merupakan metode untuk klasifikasi *text* dengan kecepatan pemrosesan yang tinggi jika dalam data besar. (Fitriyyah, Safriadi, & Pratama, 2019). Ada juga yang berpendapat bahwa *Naïve Bayes* adalah metode yang digunakan untuk prediksi karena mengandung probabilistik sederhana yang diterapkan pada *teorema bayes* dengan ketergantungan yang kuat. (Sinaga, Sawaluddin, & Suwilo, 2020).

Selain menggunakan metode *Naïve bayes*, peneliti ini juga menggunakan metode KNN. Algoritma KNN adalah salah satu algoritma yang sudah *popular*. KNN termasuk ke dalam grup *instance-based learning*. Metode KNN merupakan teknik *lazy learning*. (Cahyanti, Rahmayani, & Ainy, 2020). Maksudnya adalah metode ini digunakan dalam klasifikasi data yang jaraknya dekat. Ada juga yang berpendapat bahwa algoritma KNN adalah algoritma pembelajaran yang banyak digunakan dalam sistem *cyber-fisiksosial* (CPSS) untuk menganalisis dan menambang data (*main data*). (Zhang, Chen, Liu, & Xi, 2020).

Berdasarkan latar belakang tersebut, maka dilakukannya penelitian ini untuk ke dua metode tersebut dengan mengintegrasikan *Twitter* sebagai *platform* untuk peneliti melakukan pengolahan data untuk mengetahui analisis sentimen terhadap UU cipta kerja omnibus law dengan *data mining*. Masalah dalam penelitian ini adalah bagaimana cara peneliti mengumpulkan data dari *Twitter* dan berapa tingkat akurasi dari metode *Naïve Bayes* dan KNN.

Dengan menggunakan pemrograman *java*, penulis memungkinkan untuk melakukan pengambilan data secara *online* dengan klasifikasi semi-*realtime* meskipun menggunakan metode KNN yang dikenal sebagai *supervised lazy classifier*. KNN yang akan digunakan menggunakan prinsip probabilitas yang juga sebagian digunakan dalam metode *Naïve Bayes* sebagai penentuan nilai titik x,

y, dan z untuk dijadikan acuan pencarian jarak terdekat.

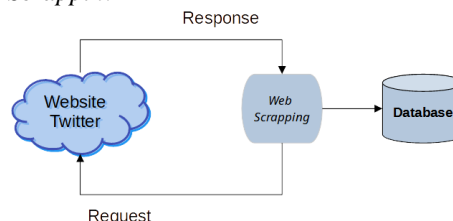
Penulis juga bertujuan untuk melakukan integrasi pengambilan keputusan melalui metode *Naïve Bayes* dan KNN ke dalam sistem perancangan dalam bentuk aplikasi *java* sehingga pengguna dapat memilih metode yang digunakan untuk klasifikasi data. Selain itu, tujuan penulis adalah untuk mengetahui cara mengumpulkan data *Twitter* secara *real time* dengan menggunakan metode *web scrapping* dan mengetahui tingkat akurasi dari kedua metode yang diteliti. Penelitian ini merupakan ide yang penulis buat sendiri sebagai bahan penelitian untuk menganalisis sentimen terhadap UU Cipta Kerja Omnibus Law dengan cara menarik data dari *Twitter* dan menganalisis data tersebut dengan menggunakan pemrograman *java* dan *Microsoft Excel*.

2. Metode/Perancangan

Dalam metode penelitian ini menjelaskan langkah-langkah yang dilakukan adalah sebagai berikut:

2.1 Pengumpulan Data

Pada penelitian ini pengumpulan data dimulai dengan penarikan data *tweet* dari *website twitter* yang kemudian disimpan ke database. Penarikan *data tweet* dilakukan dengan menggunakan metode *Web Scrapping*.



Gambar 1. Proses Pengambilan Data *Twitter*

Gambar 1 merupakan proses pengambilan data dari *Twitter* dengan menggunakan metode *Web Scrapping*. *Web Scrapping* merupakan suatu teknik untuk mengutip dan mengekstraksi data atau informasi dari suatu *website* dengan menggunakan *low-level HTML*. *Web Scrapping* ini mengambil data kotor secara *realtime* dari *website twitter*, yang selanjutnya akan dipilih menjadi data *tweet* bersih.

Data *tweet* ini akan disimpan di dalam database. Data yang diambil dari server *twitter* diperoleh dari *tweet* secara *real time* dengan kata kunci pencarian “Omnibus Law”. Setelah proses pengambilan data selesai dilakukan, data akan melalui proses *preprocessing* untuk mengambil hanya kata yang dibutuhkan sebagai bahan evaluasi klasifikasi. Data *tweet* akan digunakan dengan cara membagi seluruh data yang diambil menjadi 2 bagian, yaitu data latih sebanyak 80 % dan data uji sebanyak 20 %. Untuk proses pembuatan data latih, pengklasifikasian data dilakukan secara manual kedalam tiga kategori, yaitu positif, negatif, dan

netral. Untuk kategori positif, negatif, dan netral adalah klasifikasi manual terhadap data *tweet* latih yang telah diambil. Data *tweet* latih ini juga dilakukan proses *preprocessing* menggunakan fasilitas “Hitung Frekuensi Kata” pada menu “Data Latih” aplikasi hasil rancangan. Dengan fasilitas tersebut dapat diambil frekuensi kata yang terkait dengan kategori positif, negatif, maupun netral sebagai acuan klasifikasi *Naïve Bayes* maupun KNN.

2.2 Preprocessing

Preprocessing dilakukan untuk menghindari data yang kurang sempurna, gangguan pada data, dan data-data yang tidak konsisten. Tahapan pada *text preprocessing* yang dilakukan adalah:

1. Melakukan *filtering duplicate tweets* adalah tahapan dimana *tweet* yang mempunyai isi sama akan dihapus untuk menghindari duplikat isi dari *tweet*.
2. *Case folding* mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf “a” sampai dengan “z” yang diterima.
3. *Cleaning* tahap dimana karakter selain huruf dihilangkan dan dianggap *delimiter* dan menghapus juga *URL*, *mention* dan *hashtag*.
4. *Tokenizing/parsing* tahap pemotongan string *input* berdasarkan tiap kata yang menyusunnya.
5. *Filtering* adalah tahap mengambil kata-kata penting dari hasil token. Bisa menggunakan algoritma *stoplist* (membuang kata yang kurang penting) atau *wordlist* (menyimpan kata penting). Contoh *stopwords* adalah “yang”, “dan”, “di”, “dari”, “dengan” dan seterusnya (Triawati, 2009).

Output dari tahapan *Preprocessing* adalah berupa *tweet* bersih yang siap digunakan untuk proses selanjutnya

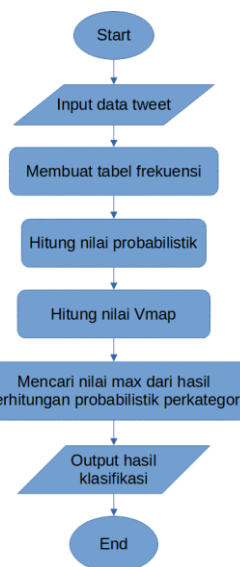
2.3 Klasifikasi Model

Perancangan *flowchart* bertujuan untuk memberi gambaran bagaimana proses klasifikasi *tweet* dengan perhitungan algoritma *Naïve Bayes Classifier* dan *K-Nearest Neighbor*,

– *Naïve Bayes Classifier*

Gambar 2 menunjukkan proses *Naïve Bayes Classifier* dimulai dengan menghitung frekuensi kata pada data *tweet* yang kemudian dilanjutkan dengan menghitung nilai probabilitas kategori dokumen dan nilai probabilitas kemunculan kata terhadap kategori, setelah itu nilai probabilitas tersebut dipergunakan untuk mendapatkan nilai *Vmap tweet* perkategori

Vmap merupakan nilai probabilitas total untuk masing-masing kategori positif, negatif, maupun netral dalam satu *tweet*. Masing-masing kategori (positif, negative, dan netral) memiliki nilai *Vmap* sehingga dapat ditentukan Nilai *Vmap* tertinggi sebagai acuan hasil dari klasifikasi *tweet* tersebut.

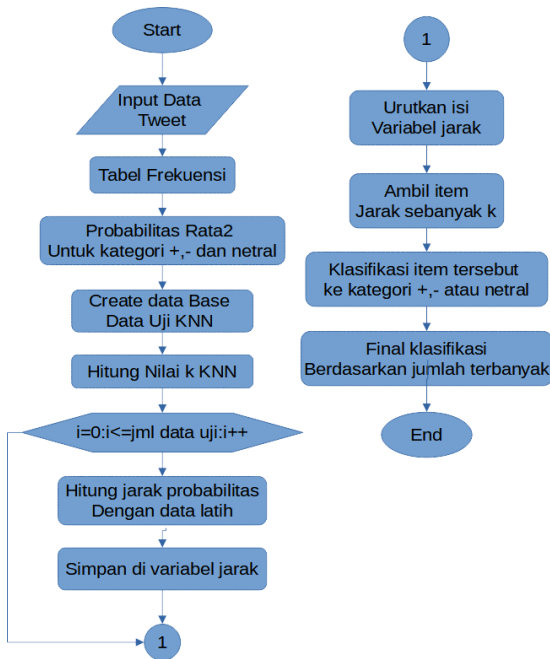


Gambar 2. Proses *Naïve Bayes Classifier*

– *K-Nearest Neighbor*

K-Nearest Neighbor (K-NN) merupakan metode pengelompokan yang mengelompokkan data baru berdasarkan jarak data baru itu ke beberapa data/tetangga (*neighbor*) terdekat. Untuk mengimplementasikan metode *K-Nearest Neighbor* (K-NN) dapat dengan melakukan langkah-langkah berikut ini:

1. Input data latih, dan konversi data tersebut menjadi data numerik 3 variabel berdasarkan klasifikasi positif, negatif, dan netral.
2. Klasifikasi nilai variabel diambil berdasarkan rata-rata probabilitas nilai kata untuk masing-masing kategori.
3. Input data uji melalui scrapping twitter dan lakukan konversi menjadi data numerik sebagaimana data latih.
4. Ambil nilai konstanta K berdasarkan jumlah data yang akan diuji nilai. Nilai K akan berubah secara otomatis bergantung dengan data latih yang digunakan.
5. Uji nilai data uji dengan melakukan pengulangan melalui elemen data latih.
6. Tentukan nilai jarak antara data uji dengan data latih hingga ke semua elemen data latih.
7. Urutkan elemen jarak dengan urutan dari nilai terkecil menuju terbesar.
8. Ambil elemen jarak tadi berdasarkan nilai K yang sudah ditentukan saat awal.
9. Ambil index elemen data yang diurutkan berdasarkan data uji.
10. Lakukan pengurutan kembali menjadi tiga kategori.
11. Elemen data jarak urut terbanyak menjadi keputusan prediksi positif, negatif, maupun netral yang akan diambil melalui metode KNN.
12. Lakukan pengujian data uji yang lain hingga seluruh data uji selesai klasifikasi.
13. Selesai.

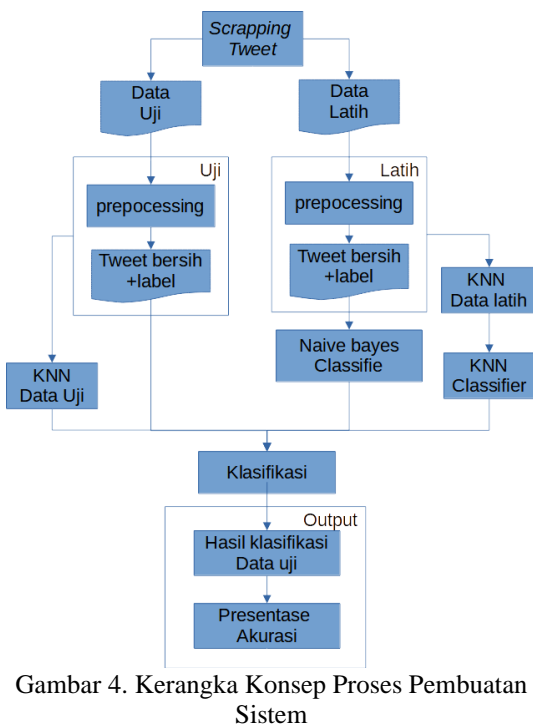


Gambar 3. Proses *K-Nearest Neighbor*

2.4 Kerangka Konsep Sistem

Kerangka konsep dibuat untuk menggambarkan tahapan proses yang dilakukan aplikasi sentimen analisis mulai dari pengambilan data, *preprocessing*, sampai perhitungan klasifikasi dan akurasi.

Gambar 4 menunjukkan kerangka konsep proses pembuatan sistem. Gambaran *output* sistem yang dibuat berupa jumlah kategori (positif, negative dan netral) hasil dari klasifikasi yang dilakukan oleh sistem terhadap data *tweet* yang diuji dan penyajian akurasi kebenaran klasifikasi kategori dalam bentuk *prosentase*.



Gambar 4. Kerangka Konsep Proses Pembuatan Sistem

3. Hasil dan Pembahasan

Hasil dan pembahasan mengenai sentimen analisis terhadap omnibus law dengan menggunakan metode *Naive Bayes* dan *KNN* yang peneliti lakukan dengan menggunakan *web scrapping* maka contoh *crawling* data yang sudah dilakukan *labeling* dan nilai akurasi dari setiap metode,

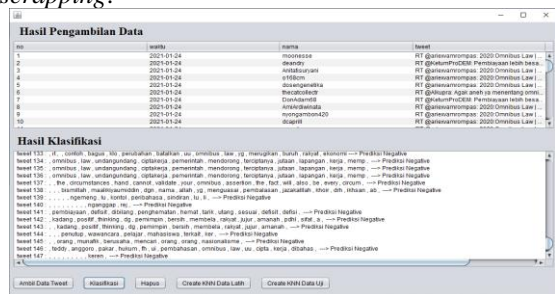
Pengujian sistem juga dilakukan dengan cara menjalankan aplikasi secara detail pada setiap menu yang ada, dengan tujuan untuk mengetahui menu atau fitur mana yang sudah berfungsi dengan baik maupun yang tidak berfungsi sesuai dengan sebagaimana mestinya.

Terdapat perlakuan berbeda dalam melakukan klasifikasi *Naive Bayes* dan *KNN*. Klasifikasi *KNN* membutuhkan konversi data latih menjadi data latih yang dapat diproses menggunakan metode *KNN* dengan menggunakan fasilitas “Create *KNN* Data Latih” pada menu “Klasifikasi Data Online”. Hal serupa juga perlu dilakukan untuk data yang akan diuji dengan menggunakan fasilitas “Create *KNN* Data Uji” pada menu yang sama. Fasilitas tersebut sebagaimana ditampilkan dalam Gambar 5 merujuk pada dua tombol di sebelah kanan bawah.

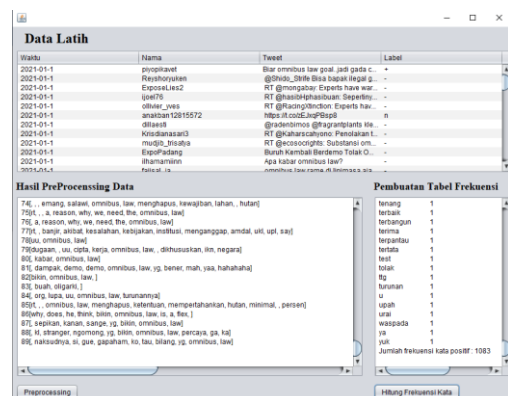
3.1 Hasil *Preprocessing*

Data mentah *twitter* yang diambil oleh aplikasi rancangan ditunjukkan dalam Gambar 5 yang selanjutnya akan dilakukan tahapan *preprocessing*.

Gambar 6 adalah hasil dari proses *preprocessing* data *tweet* yang diambil dari *website twitter* secara *real time* dengan menggunakan metode *web scrapping*.

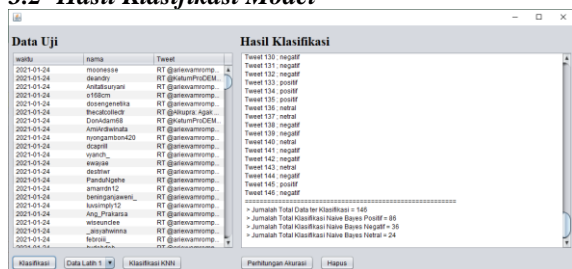


Gambar 5. Tampilan aplikasi dalam pengambilan data mentah *twitter*

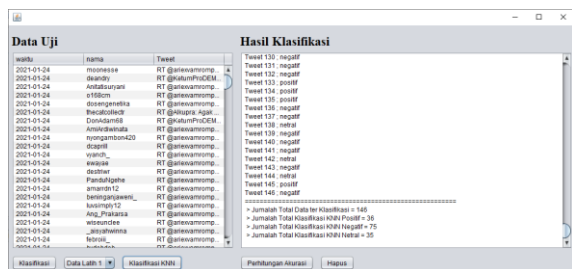


Gambar 6. Hasil *Preprocessing*

3.2 Hasil Klasifikasi Model



Gambar 7. Hasil Klasifikasi Naive Bayes



Gambar 8. Hasil Klasifikasi KNN

Gambar 7 merupakan hasil klasifikasi model perhitungan yang digunakan dengan menggunakan metode Naive Bayes sedangkan Gambar 8 merupakan hasil klasifikasi menggunakan metode KNN..

Setelah di dapatkan hasil pelabelan dari kedua metode tersebut, langkah selanjutnya adalah melakukan pengujian akurasi dari perhitungan output yang sudah di dihasilkan pada kedua metode tersebut dengan membandingkan dengan hasil yang dinilai secara manual. Hasil dari perhitungan akurasi tersebut adalah sebagai berikut.

3.3 Pengujian Akurasi

Pengujian akurasi dilakukan dengan membandingkan data hasil klasifikasi aplikasi rancangan dengan data klasifikasi manual yang telah dilakukan.

- Hasil Akurasi Naive Bayes

Tabel 1. Hasil Precision, Recall dan Akurasi Naive Bayes

K	Skenario Uji Naive Bayes		
	P	R	A
Positif	0.570	1.000	0.747
Negatif	0.722	0.473	0.733
Netral	0.542	0.310	0.394

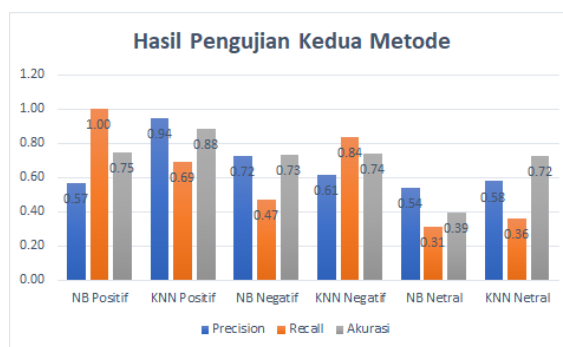
Dari Tabel 1 dapat dilihat hasil pada metode Naive Bayes menyatakan bahwa tingkat akurasi pada metode ini sebesar 74% untuk label positif, 73% untuk label negatif dan 39% untuk label netral.

- Hasil Akurasi K-Nearest Neighbor (K-NN)

Tabel 2. Hasil Precision, Recall dan Akurasi K-Nearest Neighbor (K-NN)

K	Skenario Uji K-Nearest Neighbor (K-NN)		
	P	R	A
Positif	0.944	0.694	0.884
Negatif	0.613	0.836	0.740
Netral	0.577	0.357	0.723

Dari Tabel 2 dapat dilihat hasil pada metode K-Nearest Neighbor (K-NN) menyatakan bahwa tingkat akurasi pada metode ini sebesar 88% untuk label positif, 74% untuk label negatif dan 72% untuk label netral.



Gambar 9. Grafik Hasil Pengujian Kedua Metode

Dari Gambar 9 dapat dilihat bahwa akurasi yang terbaik ada pada metode K-Nearest Neighbor (K-NN) untuk label positif.

4. Kesimpulan dan Saran

Tujuan dari penelitian yang dilakukan ini adalah untuk mengetahui tingkat akurasi dari dua metode berbeda yaitu Naive Bayes dan K-Nearest Neighbor (K-NN) dengan memanfaatkan aplikasi java hasil perancangan untuk melakukan klasifikasi.

Dalam penelitian ini dapat diketahui bahwa metode Naive Bayes dan KNN dapat digunakan dalam bidang data mining.

Analisis sentimen terhadap data twitter mengenai UU cipta kerja omnibus law mencapai tingkat akurasi 74%. Dimana class precision untuk pred. positive adalah 57%, pred negative adalah 72%, dan pred. neutral adalah 54% dalam metode Naive Bayes. Lalu tingkat akurasi mencapai 88%. Dimana class precision untuk pred. positive adalah 94%, pred negative adalah 61%, dan pred. neutral adalah 57% dalam metode KNN.

Dalam penelitian ini, dapat diketahui bahwa metode K-Nearest Neighbor (K-NN) adalah metode yang tingkat akurasinya lebih tinggi dibandingkan dengan metode Naive Baye dengan tingkat akurasi sebesar 88%. Metode KNN juga dapat dilakukan secara semi-realtime menggunakan aplikasi hasil perancangan. Semi-realtime yang dimaksud adalah bahwa data mentah dapat diambil secara langsung, sedangkan untuk pengambilan keputusan

membutuhkan dua langkah lagi yaitu melakukan “Create KNN Data Uji” lalu dilanjutkan dengan “Klasifikasi KNN”.

Sedangkan untuk saran pada penelitian selanjutnya sebaiknya jumlah data latih yang digunakan diharapkan lebih banyak dari sebelumnya untuk meningkatkan pemberian informasi yang berkaitan respon masyarakat terhadap UU Cipta Kerja Omnibus law dan agar dapat lebih meningkatkan keakuratan dari analisis yang di lakukan.

Saran berikutnya adalah langkah “Create KNN Data Uji” dapat dieliminasi dengan perubahan algoritma sehingga dapat menyederhanakan pengguna dalam menjalankan aplikasi. Dengan penyederhanaan tersebut dimungkinkan dapat diambil perbandingan yang adil dalam hal kecepatan eksekusi klasifikasi. Kecepatan eksekusi dapat ditampilkan dalam orde *mili-second* bersamaan dengan hasil klasifikasi.

Selain itu juga di sarankan untuk menggunakan metode lain maupun metode yang lebih mutakhir sehingga dapat mengetahui performa metode tersebut dibandingkan dengan *Naïve Bayes* dan KNN sebagaimana dalam penelitian ini.

Daftar Pustaka:

- Cahyanti, D., Rahmayani, A., & Ainy, S. (2020). *Analisis Performa Metode KNN Pada Dataset Pasien Pengidap Kanker Payudara*. 1(2), 39–43.
- Chandani Vinita, Wahono R S, Purwanto P, 2015, “Komparasi Algoritma Klasifikasi Machine Learning dan Feature Selection pada Analisis Sentimen Review Film”, *Journal of Intelligent System*, Vol 1, No. 1 February, pp 56-60
- Deviyanto Akhmad, Wahyudi M Didik R, “Penerapan Analisis Sentimen pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor”, *Jiska (Jurnal Informatika Sunan Kalijaga)*, Vol 3, No. 1, Mei 2018, pp 1-13
- Fitriyyah, S. N. J., Safriadi, N., & Pratama, E. E. (2019). *Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes*. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(3),279.
- Irfan M Rizzo, Fauzi M Ali, Tibyani, 2018, “Analisis Sentimen Kurikulum 2013 pada Twitter menggunakan Ensemble Feature dan Metode K-Nearest Neighbor”, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Vol 2, No.9, Sempember 2018, pp 30006-3014.
- Nurdin, A. D. (2017). *Penerapan Data Mining Untuk Menganalisis Penjualan Barang Dengan Menggunakan Metode Apriori Pada Supermarket Sejahtera Lhokseumawe*. Astika, D., Studi, P., Informatika, T., & Malikussaleh, U., Vol. 6 No.
- Pertiwi Melisa Winda, 2019, “Analisis Sentimen Opini Publik Mengenai Sarana dan Transportasi Mudik Tahun 2019 pada Twitter Menggunakan Algoritma Naïve Bayes, Neural Network, KNN dan SVM”, *Jurnal Inti Nusa Mandiri*, Vol 14, No. 1 Agustus 2019, pp 27-32
- Puspita Rani, Widodo Agus, 2020, “Perbandingan Metode KNN, Decision Tree, dan Naïve Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS”, *Jurnal Informatika Universitas Pamulang*, Vol.5, No. 4 Desember 2020, pp 646-654
- Rizal Jawahir G, 2020, “Ap itu Omnibus Law Cipta Kerja, isi dan dampaknya bagi Buruh?”, *Compas.com*, 06/10/2020, 10.45 wib, <https://www.kompas.com/tren/read/2020/10/06/104500965/apa-itu-omnibus-law-cipta-kerja-isi-dan-dampaknya-bagi-buruh?page=all#:~:text=KOMPAS.com%20%2D%20Dewan%20Perwakilan%20Rakyat,Kompleks%20Parlemen%2C%20Senayan%2C%20Jakarta>.
- Rozi Imam F, Hamdana Elok H, Alfahmi M B I, “Pengembangan Aplikasi Analisis Sentimen Twitter Menggunakan Metode Naïve Bayes Classifier (Studi Kasus Samsat Kota Malang)”, *Jurnal Informatika Polinema*, Jilid 4 Edisi 2, pp 149-149
- Sinaga, L. M., Sawaluddin, & Suwilo, S. (2020). *Analysis of classification and Naïve Bayes algorithm k-nearest neighbor in data mining*. *IOP Conference Series: Materials Science and Engineering*,725(1).
- Triawati, C. (2009). *Text Mining*. Bandung, Jawa Barat, Indonesia
- Utami, I. (2020). *Analisis Sistem Informasi Banjir Berbasis Media Twitter*. 9(1), 67–72.
- Undang-Undang (UU) Nomor 11 Tahun 2020, “Undang-Undang (UU) Tentang Cipta Kerja”, LN.2020/No.245, TLN No.6573, jdih.setneg.go.id:769hlm, <https://peraturan.bpk.go.id/Home/Details/14975/0/uu-no-11-tahun-2020>.
- Zhang, W., Chen, X., Liu, Y., & Xi, Q. (2020). *A Distributed Storage and Computation k-Nearest Neighbor Algorithm Based Cloud-Edge Computing for Cyber-Physical-Social Systems*. *IEEE Access*, 8, 50118–50130.