

Komparasi Algoritma Klasifikasi pada Analisis Sentimen Klaim Manfaat Jaminan Hari Tua

Comparison of Classification Algorithms toward Sentiment Analysis on Claim of Old Age Guaranteed

¹Ahmad Al Kaafi*, ²Hilda Rachmi, ³Suparni

^{1,3}Teknologi Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika

²Sistem Informasi Kampus Kota Bogor, PSDKU Kota Bogor, Universitas Bina Sarana Informatika
Jalan Kramat Raya No.98, Kwitang, Kec. Senen, Kota Jakarta Pusat, DKI Jakarta, Indonesia

*e-mail: ahmad.akf@bsi.ac.id

(received: 25 Juli 2022, revised: 23 Oktober 2022, accepted: 9 November 2022)

Abstrak

Pro dan kontra terhadap aturan klaim Jaminan Hari Tua (JHT) yang tertuang pada Permenaker No 2 Tahun 2022 memang tak dapat dihindari di Indonesia, terutama pada media sosial twitter. Tentunya hal ini menjadi hal yang patut untuk dianalisis sebagai pengambil kebijakan pemerintah agar tidak salah langkah, salah satunya menggunakan teknik sentiment analisis. Tahapan penelitian yang dilakukan diantaranya pengumpulan data, pre-processing, model yang digunakan, eksperimen dan penilaian model, serta validasi dan evaluasi hasil. Peneliti melakukan crawling 1000 data tweet berbahasa Indonesia terhadap opini pro dan kontra terkait klaim Jaminan Hari Tua menggunakan bahasa pemrograman R dan dilakukan Text Cleaning Process. Pada tahap selanjutnya Pre-processing dilakukan menggunakan Case Folding, Stemming, Tokenizing dan Stopword Removal. Penelitian ini menghasilkan sentimen masyarakat Indonesia sangat bahagia akan pernyataan presiden yang meminta aturan Permenaker No 2 Tahun 2022 direvisi, kemudian diuji menggunakan 3 algoritma klasifikasi dan penerapan seleksi fitur Particle Swarm Optimization. Hasil pengujian sentimen analisis klaim manfaat JHT didapat nilai tertinggi pada algoritma Support Vector Machine dengan penambahan Particle Swarm Optimization yakni nilai akurasi sebesar 82.92%, precision sebesar 84.53%, recall sebesar 80.90% dan nilai AUC sebesar 0.892.

Kata kunci: Support Vector Machine, Analisis Sentimen, Jaminan Hari Tua.

Abstract

The pros and cons of the Old Age Security (JHT) claim rules contained in the Minister of Manpower Regulation No. 2 of 2022 are unavoidable in Indonesia, especially on Twitter social media. Of course, this is something that deserves to be analyzed as government policy makers so as not to make a wrong move, one of which is using sentiment analysis techniques. The stages of the research carried out include data collection, pre-processing, the model used, experimentation and model assessment, as well as validation and evaluation of results. The researcher crawled 1000 Indonesian-language tweets on the pros and cons of claims related to Old Age Security using the R programming language and carried out the Text Cleaning Process. In the next stage, pre-processing is carried out using Case Folding, Stemming, Tokenizing and Stopword Removal. This research resulted in the sentiments of the Indonesian people being very happy with the president's statement asking for the regulation of the Minister of Manpower No. 2 of 2022 to be revised, then tested using 3 classification algorithms and the application of Particle Swarm Optimization feature selection. The results of the sentiment testing of the JHT benefit claim analysis obtained the highest value on the Support Vector Machine algorithm with the addition of Particle Swarm Optimization, namely the accuracy value of 82.92%, precision of 84.53%, recall of 80.90% and AUC value of 0.892..

Keywords: Support Vector Machine, Sentiment Analysis, Old Age Guaranteed

1 Pendahuluan

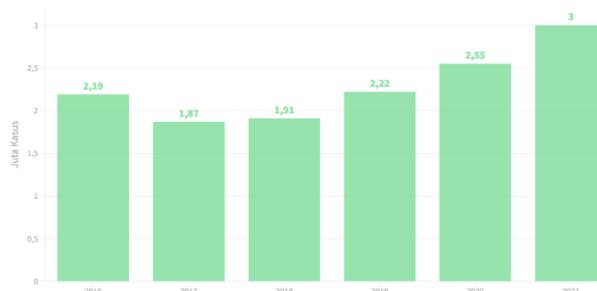
Indonesia merupakan salah satu negara dengan jumlah penduduk terbanyak di dunia, yakni menempati posisi keempat terbanyak. Dalam mensejahterakan rakyat pemerintahan Indonesia selalu berupaya melaksanakan berbagai cara yang berupaya meningkatkan kesejahteraan rakyat. Kesejahteraan yang dimaksud tersebut harus dapat dinikmati secara berkelanjutan, adil, dan merata menjangkau seluruh rakyat, salah satunya dengan dibuatnya suatu jaminan sosial untuk meningkatkan perlindungan sosial bagi mereka yang bekerja dengan skema funded social security [1].

Berdasarkan hasil sensus 2020 (BPS), jumlah penduduk di Indonesia yang berusia 15-64 tahun adalah sebanyak 70,72%, jumlah penduduk yang berada pada usia 0-14 tahun sebanyak 1,2% dan usia 61 ke atas sebanyak 9,78% [2]. Jumlah penduduk usia produktif mendominasi struktur kependudukan di Indonesia, artinya Indonesia sebenarnya sedang berada dalam periode bonus demografi, di mana jumlah penduduk usia produktif (15-64 tahun) mencapai 70,72 persen penduduk Indonesia[3][4].

Seiring berkembangnya teknologi dan menurunnya fungsi kognitif dalam tubuh sehingga para tenaga kerja akan mengalami penuaan. Penuaan penduduk ditandai dengan semakin meningkatnya proporsi penduduk usia lanjut (lansia) dan biasanya digambarkan dengan perubahan struktur piramida penduduk dari piramida penduduk muda menjadi piramida penduduk tua. Penuaan secara alami berhubungan dengan penurunan kemampuan fisik seseorang yang umumnya berhubungan dengan penurunan tingkat kesehatan yang akan berdampak pada aspek sosial maupun ekonomi.

Tenaga kerja mempunyai arti dan peranan yang penting dalam pelaksanaan pembangunan nasional pada umumnya, dan dalam peningkatan produksi dan produktivitas pada khususnya [5]. Oleh karena itu pemerintah membuat sejumlah aturan guna memberi perlindungan jaminan sosial terhadap tenaga kerja. Pemerintah menyelenggarakan program jaminan sosial dalam rangka melindungi seluruh masyarakat Indonesia sehingga semua terjamin pemenuhan kebutuhan dasar dan dapat hidup dengan layak [6]. Jaminan sosial bertujuan memberikan perlindungan dasar guna memenuhi kebutuhan bagi tenaga kerja dan keluarganya dengan memberikan kepastian arus penerimaan penghasilan keluarga [7][8].

Di dalam suasana pandemi yang belum usai, banyak para tenaga kerja yang mengalami pemutusan kerja. Hal tersebut membuat para tenaga kerja untuk memutuskan mencairkan Jaminan Hari Tua (JHT) guna memenuhi kebutuhan sehari-hari.



Gambar 1. Jumlah Klaim JHT BPJS Ketenagakerjaan (2016-2021) [9]

Pada tahun 2021 di Indonesia mengalami kenaikan jumlah kasus klaim program Jaminan Hari Tua (JHT) menjadi 3 juta atau 17,34% dari tahun sebelumnya yang sebesar 2,55 juta kasus [9]. Hal tersebut dipicu juga karena adanya perubahan sistem layanan klaim JHT menjadi daring, sehingga mempermudah masyarakat dalam mengajukan klaim. Data pada bulan Desember 2021 menunjukkan bahwa tingkat kesuksesan klaim JHT dengan sistem daring maupun luring telah mencapai 95% [9].

Berdasarkan laporan yang dikeluarkan BPJS Ketenagakerjaan (25/2), klaim JHT sepanjang 2021 mencapai Rp 36,42 triliun yang pada tahun sebelumnya, klaim JHT baru mencapai Rp 33,1 triliun [10].

Februari 2022, Kementerian Ketenagakerjaan menjadi sorotan setelah mengubah mekanisme pencairan JHT yang tertuang dalam Peraturan Menaker Nomor 2 Tahun 2022. Permenaker Nomor 2 Tahun 2022 mencabut Permenaker Nomor 19 Tahun 2019, berdasarkan aturan lama bahwa manfaat JHT dapat diberikan kepada peserta yang mengundurkan diri dan dibayarkan secara tunai setelah melewati masa tunggu 1 bulan [11] [12]. JHT hanya bisa dicairkan jika pekerja memasuki usia pensiun 56 tahun, cacat total, atau meninggal dunia [9]. Pejabat Pengganti Sementara Deputy Direktur

Bidang Hubungan Masyarakat dan Antar Lembaga BPJS Ketenagakerjaan Dian Agung Senoaji menegaskan peserta masih bisa mencairkan sebagian dana jaminan hari tua (JHT) meski belum berusia 56 tahun [13]. Tentu saja ini hal ini menimbulkan pro dan kontra dikalangan masyarakat terkait pencairan JHT. Tidak sedikit masyarakat melakukan protes akan peraturan baru atas klaim JHT yang dinilainya mempersulit dalam hal pencairan. Disisi lain banyak masyarakat yang mendukung akan hal mekanisme baru pencairan JHT yang tertuang dalam permenaker no 2 tahun 2022 itu, karena sesuai dengan dasar dalam penamaan kesejahteraan yang digunakan yaitu Jaminan Hari Tua berarti jaminan tersebut digunakan kelak saat di hari tua atau usia pensiun.

Opini dan pendapat masyarakat terkait pro dan kontra terkait klaim JHT pun bermunculan di media sosial salah satunya twitter. Perubahan regulasi tata cara pencairan dana JHT yang tertuang pada Permenaker Nomor 2 Tahun 2022, salah satunya pencairan dana JHT hanya dapat dilakukan apabila pemegang kuasa telah mencapai usia 56 tahun menjadi perbincangan hangat di media sosial Twitter dikarenakan perubahan tersebut tidak memihak masyarakat. Penelitian ini dilakukan untuk mengetahui bagaimana memetakan dan menganalisa jaringan komunikasi yang terbentuk pada media sosial twitter agar pemangku kepentingan dalam merumuskan kebijakan dapat sesuai dengan harapan stakeholder. Seperti data survey [14] pada tahun 2021 sebanyak 200.000 orang melakukan tweeting setiap menitnya. Membaca seluruh opini yang ada di twitter tidaklah efektif karena membutuhkan waktu yang lama, namun jika sebagian yang dibaca akan menjadi tidak tepat sasaran. Untuk itu dibutuhkan teknik yang tepat dalam menambang opini tersebut sehingga opini dapat diklasifikasikan secara otomatis dalam class positif dan negatif. Teknik tersebut adalah sentiment analisis. Sentimen analisis atau dikenal juga dengan istilah opinion mining merupakan proses untuk memahami, mengekstrak, dan mengolah data berbasis teks untuk mendapatkan informasi yang terkandung dalam opini tersebut [15]

Tujuan Penelitian ini adalah mengetahui besaran akurasi dari sentimen analisis yang dihasilkan menggunakan method improvement melalui penerapan seleksi fitur Particle Swarm Optimization pada metode K-Nearest Neighbor, Naïve Bayes, dan Support Vector Machine, agar dapat memberikan informasi sentimen masyarakat terkait klaim JHT pada Permenaker Nomor 2 Tahun 2022 di media sosial twitter. Selain untuk mengetahui tingkat akurasi, analisis sentimen pro dan kontra terkait klaim JHT pada Permenaker Nomor 2 Tahun 2022 juga bertujuan untuk dapat membantu pemerintah dalam mengambil kebijakan dengan mengetahui tingkat sentimen masyarakat dilihat dari data sampel komentar twitter dengan perbandingan antara sentimen positif maupun negatif.

2 Tinjauan Literatur

Penelitian terdahulu telah dilakukan terkait dengan analisis sentimen terhadap opini masyarakat dalam penggunaan Mobile-JKN [16]. Penelitian bertujuan untuk melihat apakah respons masyarakat positif, negatif, atau netral terhadap aplikasi mobile-JK. Hasilnya dari 528 tweet, 222 tweets (42%) positif, 153 tweets (29%) adalah netral dan 152 tweets (29%) negative.

Penelitian serupa juga telah dilakukan terkait analisis sentimen masyarakat terhadap transportasi online dan membandingkan nilai akurasi SVM dan SVM-PSO [17]. Data penelitian diperoleh dari scraping tweet dari 1/1/2019 hingga 15/10/2019 sebanyak 1.852 data tweet yang dibagi menjadi data testing 1.130 tweet dan training 722 tweet. Proses analisis menggunakan tools RapidMiner. Hasil analisis sentimen positif menggunakan SVM adalah sebesar 62% dan sentimen negatif sebesar 38%, sedangkan pada SVM-PSO, opini positif sebesar 53% dan negatif 47%. Tingkat akurasi dari penelitian ini pada SVM sebesar 95,46% dan AUC 0,979 (excellent classification), sedangkan pada SVM-PSO sebesar 96,04% dan AUC 0,993 (excellent classification).

Pada penelitian berikutnya terkait Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19 [18]. Penelitian tersebut menggali respon dari masyarakat terhadap tindakan vaksinasi di media sosial twitter sebanyak 845 tweet, dengan menggunakan dua kata kunci, “vaksinmerahputih” dan “vaksinsinovac”. Setelah diuji data tersebut diklasifikasikan menggunakan metode SVM dan Naïve Bayes. Hasil klasifikasi dari metode Naïve Bayes mendapatkan rata-rata akurasi 85,59%, sedangkan SMV sebesar 84,41%. Hasil sentimen pada metode Naïve Bayes dengan kata kunci “vaksinsinovac” mendapatkan sentimen positif 66% dan negatif 34%, sedangkan “vaksinmerahputih” memperoleh sentimen positif 89% dan negatif 11%. Metode SVM

dengan kata kunci “vaksinsinovac” mendapatkan sentimen positif 96% dan negatif 4%, sedangkan “vaksinmerahputih” mendapatkan sentimen positif 98% dan negatif 2%.

Penelitian lain[15] membandingkan algoritma klasifikasi algoritma K-Nearest Neighbor, Naive Bayes, Support Vector Machine dengan penerapan seleksi fitur Generate Algorithm dan Particle Swarm Optimization untuk meningkatkan akurasi. Dari hasil penelitian tersebut, nilai akurasi yang ditunjukkan 3 algoritma dengan penerapan seleksi fitur Generate Algorithm adalah 65,62%, 62,14%, dan 68,51%. Sedangkan dengan penerapan seleksi fitur Particle Swarm Optimization sebesar 66,21%, 69,83%, dan 73,07% sehingga didapatkan rata-rata peningkatan akurasi pada 3 algoritma yang digunakan sebesar 4,28%. Akurasi terbesar didapatkan pada algoritma Support Vector Machine dengan penambahan Particle Swarm Optimization pada nilai 73,07%.

Berdasarkan permasalahan dan penelitian terkait berhasil diperoleh opini masyarakat untuk mengetahui kecenderungan sentimen yang terjadi. Pada penelitian ini digunakan beberapa algoritma klasifikasi guna mendapatkan hasil paling baik dalam memberikan akurasi terhadap analisis sentimen [15]. Dari penelitian sebelumnya yang telah dijelaskan diatas, maka peneliti menggunakan algoritma klasifikasi K-Nearest Neighbor, Naive Bayes, Support Vector Machine untuk analisis sentimen. Selain itu peneliti juga menerapkan Particle Swarm Optimization sebagai seleksi fitur yang dapat meningkatkan akurasi algoritma pada penelitian sentimen masyarakat terkait klaim JHT ini.

3 Metode Penelitian

Analisis sentimen atau opinion mining merupakan cabang ilmu dari data mining yang biasanya digunakan untuk menganalisis data tekstual berupa sebuah opini yang mengandung polaritas sehingga nantinya menghasilkan sebuah informasi yang memiliki nilai positif, negatif, atau netral [17] [18]. Tujuan dari opinion mining adalah menentukan polaritas sekumpulan teks dalam sebuah tulisan yang ada di dokumen, kalimat, paragraf, dan lain-lain cenderung mengarah ke arah negatif, netral, atau positif [15].

Proses pada text mining meliputi tahap selection, preprocessing, implementation, dan evaluation. Data yang digunakan pada text mining akan diolah dengan Knowledge Discovery Database (KDD) yang meliputi pengumpulan dan pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Untuk penelitian biasanya data yang digunakan tidak sedikit. Oleh karena itu bila analisis data dilakukan secara manual akan memakan waktu yang cukup lama [15].

Dalam penelitian ini objek yang diteliti adalah opini masyarakat mengenai klaim JHT di Indonesia melalui media sosial twitter. Data yang digunakan berupa tweet (mention, reply, like, retweet) berbahasa Indonesia. Pada penelitian ini peneliti melakukan beberapa langkah metode untuk mendapatkan dan menganalisa data. Berikut ini langkah penelitian yang dilakukan, diantaranya:

a. Pengumpulan Data

Pengumpulan data dilakukan dengan teknik *crawling*. *Crawling* data dilakukan dengan menggunakan tools RStudio. Peneliti mengekstraksi data *tweet* sejak awal Februari sampai dengan 23 Februari 2022 sebanyak 1000 *tweet* dengan kata kunci JHT, Klaim JHT yang berbahasa Indonesia pada media sosial twitter. Semua *tweet* yang diperoleh akan disimpan pada file .csv untuk dianalisis. Dengan bahasa pemrograman R dilakukan juga Text Cleaning Process, seperti Text Subbing, Penggantian Teks (HTML dan URL), Replace Emoticons, Emojis, Mentions Dan Hashtags, Replace Slang Words, Text Stripping, Remove Duplicate Text dan Word Cloud.

b. Pre-processing

Tahap pre-processing adalah tahapan dimana dilakukan seleksi data agar data yang akan digunakan menjadi lebih terstruktur. Peneliti melakukan tahap pre-processing menggunakan tools RStudio. Tahap pre-processing pada penelitian ini, diantaranya:

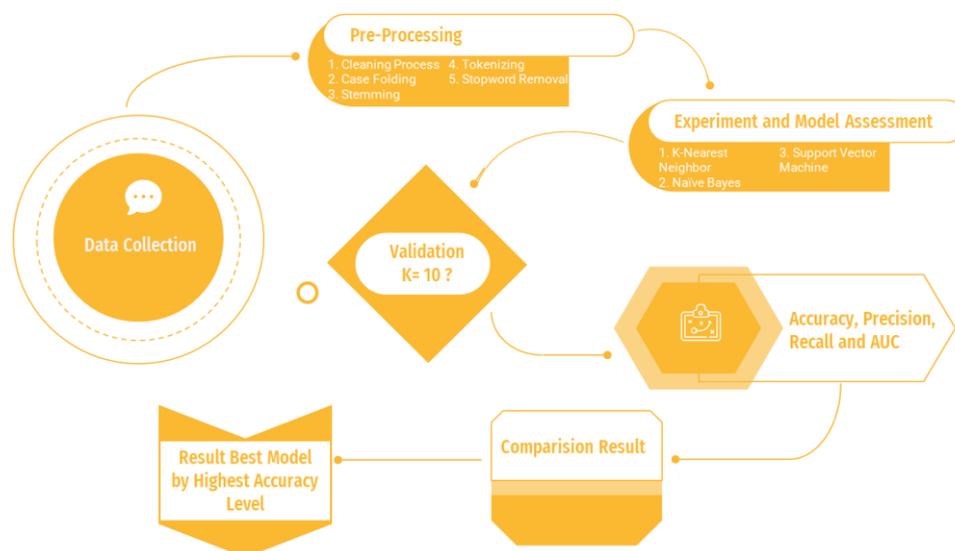
1. *Cleaning Process*

Proses pembersihan teks dalam tweet seperti menghapus substring, menghapus HTML & URL, menghapus emoji, menghapus mentions & hashtags, dan menghapus kata slang dengan mengimport leksikon Bahasa Indonesia.

2. *Case folding*

Proses mengkonversi keseluruhan teks dalam tweet menjadi bentuk standar (huruf kecil atau *lowercase*).

3. *Stemming*
Proses menemukan kata dasar dari sebuah kata. Menghilangkan imbuhan dari sebuah kata.
 4. *Tokenizing*
Proses mengubah tweet menjadi token. Token adalah kata-kata yang dipisahkan oleh spasi dalam teks.
 5. *Stopword Removal*
Proses mengambil kata-kata penting dan membuang kata-kata yang tidak penting pada *tweet* (*stopwords*) misalnya kata penghubung seperti "dan", "atau", "kemudian" dan seterusnya atau kata-kata yang tidak berpengaruh pada proses klasifikasi.
- c. Eksperimen dan Penilaian Model atau Metode
RapidMiner Studio versi 9.9 digunakan untuk memproses eksperimen data dengan menerapkan beberapa algoritma klasifikasi seperti Naïve Bayes, K-Nearest Neighbor, dan Support Vector Machine (SVM) serta penggunaan seleksi fitur pada setiap algoritmanya.
- d. Validasi dan Evaluasi Hasil
Validasi akan dibagi menjadi 10 bagian dan proses training serta testing akan dilakukan sebanyak 10 kali. Evaluasi hasil dan penilaian akurasi dilakukan dengan *confusion matrix* serta kurva ROC untuk menilai hasil AUC. Setelah mendapatkan hasil dari setiap algoritma kemudian akan dilihat algoritma mana yang memberikan hasil akurasi tertinggi. Disamping itu juga dilihat penggunaan dari seleksi fitur *Particle Swarm Optimization* (PSO) apakah dapat meningkatkan akurasi dari setiap algoritma. Seleksi fitur merupakan metode untuk penganalisaan data dengan tujuan memilih karakter/fitur yang optimal dan mengabaikan karakter/fitur tidak berpengaruh[19].
Langkah penelitian yang telah dideskripsikan dapat dilihat pada Gambar 2 dengan urutan mulai dari pengumpulan data, pre-processing data dalam beberapa tahap, pengujian model sampai didapatkan model terbaik dengan tingkat akurasi tertinggi.



Gambar 2. Kerangka Penelitian

4 Hasil dan Pembahasan

Dari total dataset yang diambil dengan crawling data tweet menggunakan tools RStudio sejak Februari sampai 23 Februari 2022 yakni sebanyak 1000 sample tweet dengan kata kunci JHT atau Klaim JHT yang berbahasa Indonesia pada media sosial twitter.

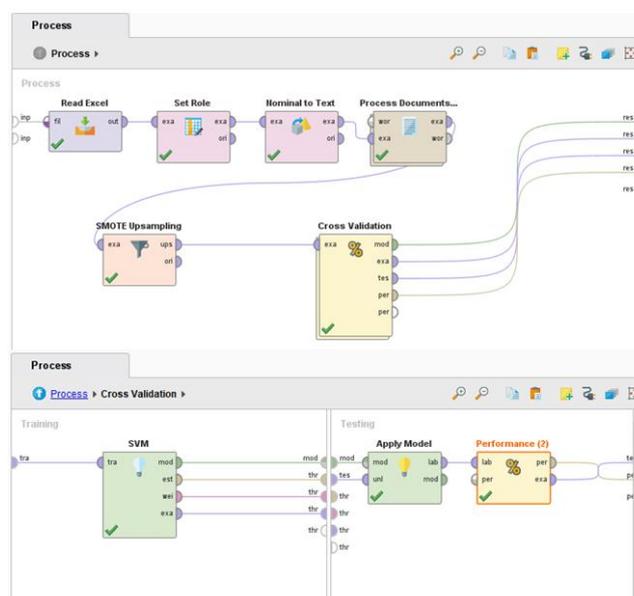
Sebelum dilakukan proses pengklasifikasian teks dilakukan beberapa tahap proses preprocessing yang meliputi proses *cleaning*, *case folding*, *stemming*, *tokenizing* dan *stopword removal*.

Tahap preprocessing dilakukan guna pengolahan data mentah menjadi koleksi data yang siap[20]. Tahap ini bertujuan membersihkan data, karena sebagian besar tweet yang sudah diekstrak belum terstruktur dikarenakan data tweet merupakan komunikasi informal sehingga banyak ditemukan kesalahan ketik, penggunaan kata-kata gaul atau adanya konten yang tidak diinginkan seperti URLS, stopwords, emoji dll.

Tabel 1. Tahapan Preprocessing

Proses	Output
<i>Teks asli</i>	Asosiasi Serikat Pekerja Indonesia (ASPEK Indonesia) menduga kendala dana jadi alasan pemerintah dan BPJS Ketenagakerjaan menunda pencairan dana Jaminan Hari Tua (JHT) pekerja hingga usia 56 tahun. #jht https://t.co/z6iOOFLeX5.....
<i>Cleaning Process</i>	Asosiasi Serikat Pekerja Indonesia (ASPEK Indonesia) menduga kendala dana jadi alasan pemerintah dan BPJS Ketenagakerjaan menunda pencairan dana Jaminan Hari Tua (JHT) pekerja hingga usia 56 tahun.
<i>Case Folding</i>	asosiasi serikat pekerja indonesia aspek indonesia menduga kendala dana jadi alasan pemerintah dan bpjs ketenagakerjaan menunda pencairan dana jaminan hari tua jht pekerja hingga usia tahun
<i>Stemming</i>	asosiasi serikat kerja indonesia aspek indonesia duga kendala dana jadi alas pemerintah dan bpjs ketenagakerjaan tunda cair dana jamin hari tua jht kerja hingga usia tahun
<i>Tokenizing</i>	"asosiasi" "serikat" "kerja" "indonesia" "aspek" "indonesia" "duga" "kendala" "dana" "jadi" "alasan" "pemerintah" "dan" "bpjs" "ketenagakerjaan" "tunda" "cair" "dana" "jamin" "hari" "tua" "jht" "kerja" "hingga" "usia" "tahun"
<i>Stopword Removal</i>	"asosiasi" "serikat" "kerja" "indonesia" "aspek" "indonesia" "duga" "kendala" "dana" "alasan" "pemerintah" "bpjs" "ketenagakerjaan" "tunda" "cair" "dana" "jamin" "tua" "jht" "kerja" "usia"

Pada tahap berikutnya menggunakan tools RapidMiner untuk menerapkan beberapa algoritma klasifikasi seperti Naïve Bayes, K-Nearest Neighbor, dan Support Vector Machine serta penggunaan seleksi fitur pada setiap algoritmanya. Setiap data akan diuji dengan menggunakan cross validation dengan k = 10 kali dimana data akan dibagi menjadi dua bagian yaitu data training dan testing. Kemudian mencari akurasi yang tertinggi pada setiap algoritmanya dan menggunakan Particle Swarm Optimization untuk seleksi fitur.



Gambar 3. Proses Pengolahan Data Pada Rapid Miner

Percobaan pertama setelah melakukan pre-processing dan menerapkan cross validation kemudian menggunakan algoritma klasifikasi tanpa adanya penambahan seleksi fitur yang digunakan.

Tabel 2. Hasil Akurasi Analisis Sentimen Dengan Algoritma Klasifikasi

No	Proses	Output			
		Akurasi	Precision	Recall	AUC
1	<i>K-Nearest Neighbor</i>	53.31%	51.76%	98.49%	0.791
2	<i>Naïve Bayes</i>	67.39%	64.20%	79.01%	0.881
3	<i>Support Vector Machine</i>	78.64%	78.73%	79.02%	0.873

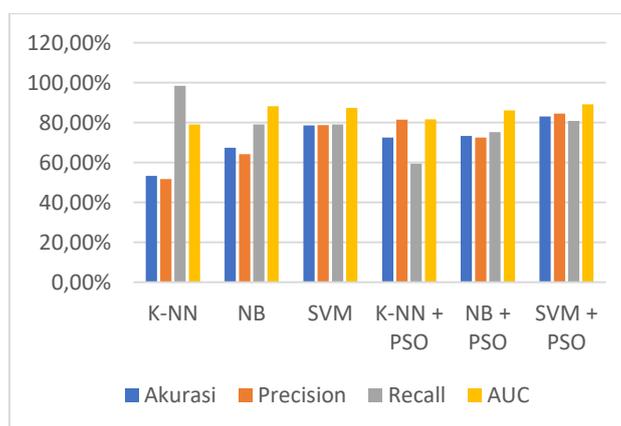
Hasil percobaan pertama yang terangkum pada Tabel 2 menghasilkan tingkat akurasi terbesar pada penggunaan algoritma Support Vector Machine dengan nilai 78,64%, dan precision 78,73%. Untuk nilai recall terbesar didapat pada algoritma K-Nearest Neighbor dengan hasil 98,49%, sedangkan AUC terbesar didapat pada algoritma Naïve Bayes dengan hasil 0,881.

Percobaan kedua dilakukan dengan menggunakan data dan tahap pre-processing yang sama untuk setiap algoritma serta penerapan cross validation pada percobaan pertama, namun ditambahkan seleksi fitur.

Tabel 3. Hasil Akurasi Analisis Sentimen Dengan Algoritma Klasifikasi Dan Particle Swarm Optimization

No	Proses	Output			
		Akurasi	Precision	Recall	AUC
1	<i>K-Nearest Neighbor</i>	72.49%	81.42%	59.33%	0.817
2	<i>Naïve Bayes</i>	73.25%	72.57%	75.24%	0.861
3	<i>Support Vector Machine</i>	82.99%	84.53%	80.90%	0.892

Pada Tabel 3 terlihat nilai akurasi yang ditunjukkan 3 algoritma adalah 72.49%, 73.25%, dan 82.99%. Percobaan ini menunjukkan hasil akurasi, *precision*, *recall* dan AUC terbesar pada penggunaan algoritma *Support Vector Machine*.



Gambar 4. Grafik Perbandingan

Berdasarkan Gambar 4 di atas, bahwa setelah menambahkan fitur seleksi dapat meningkatkan nilai akurasi pada setiap algoritma yang digunakan. Namun ada beberapa hasil yang mengalami penurunan setelah menerapkan seleksi fitur, seperti hasil recall pada algoritma K-Nearest Neighbor dan Naïve Bayes serta AUC pada algoritma Naïve Bayes.

Disamping mencari akurasi pada setiap algoritma, akan disajikan pula tebaran kata-kata yang digunakan dalam sebuah teks pada tweet terkait JHT. Adapun kata-kata yang sering muncul diantaranya: JHT, kerja, cair, usia, dana, uang, atur, jamin, buruh, tua, pemerintah dll seperti pada Gambar 5.

Ucapan Terima Kasih

Terimakasih kepada semua tim yang terlibat pada penyusunan penelitian ini.

Referensi

- [1] Skola, "BPJS Ketenagakerjaan: Sejarah Singkat, Jenis, Visi dan Misi Halaman all - Kompas.com." [Online]. Available: <https://www.kompas.com/skola/read/2020/01/15/160000469/bpjs-ketenagakerjaan--sejarah-singkat-jenis-visi-dan-misi?page=all>. [Accessed: 20-Feb-2022].
- [2] Badan Pusat Statistik, "Hasil Sensus Penduduk 2020," 2021.
- [3] A. Kristianus, "70,7% Penduduk Indonesia Usia Produktif," 2021. [Online]. Available: <https://investor.id/business/234726/707-penduduk-indonesia-usia-produktif>. [Accessed: 21-Feb-2022].
- [4] S. Indrayani, "Kementerian Komunikasi dan Informatika," 2021. [Online]. Available: <https://www.kominfo.go.id/content/detail/33004/angkatan-kerja-produktif-melimpah/0/artikel>. [Accessed: 21-Feb-2022].
- [5] "Undang-Undang Republik Indonesia Nomor 13 Tahun 2003 Tentang Ketenagakerjaan." [Online]. Available: <https://jdih.kemenkeu.go.id/fulltext/2003/13TAHUN2003UU.htm>. [Accessed: 21-Oct-2021].
- [6] I. Lestari and H. Hirawati, "Analisis Sistem Klaim Jaminan Hari Tua (Jht) Pada Badan Penyelenggara Jaminan Sosial (BPJS) Ketenagakerjaan Cabang Cilacap," vol. 4, no. 1, pp. 74–79, 2015.
- [7] *Undang-Undang Republik Indonesia Nomor 24 Tahun 2011 Tentang Badan Penyelenggara Jaminan Sosial*, no. November. Indonesia: Kementerian Sekretariat Negara Republik Indonesia, 2011, pp. 1–68.
- [8] J. Abdullah, "Bentuk-Bentuk Jaminan Sosial Dan Manfaatnya Bagi Tenaga Kerja Dalam Hukum Ketenagakerjaan Indonesia," *YUDISIA J. Pemikir. Huk. dan Huk. Islam*, vol. 9, no. 1, pp. 121–135, 2018.
- [9] A. Karnadi, "Jumlah Klaim JHT BPJS Ketenagakerjaan Naik 17,34% pada 2021," 2022. [Online]. Available: <https://dataindonesia.id/bursa-keuangan/detail/jumlah-klaim-jht-bpjs-ketenagakerjaan-naik-1734-pada-2021>. [Accessed: 20-Feb-2022].
- [10] BPJS Ketenagakerjaan, "Laporan Keuangan dan Laporan Pengelolaan Program Tahun 2020," 2021.
- [11] Menteri Ketenagakerjaan Republik Indonesia, *Peraturan Menteri Ketenagakerjaan Republik Indonesia Nomor 2 Tahun 2022 Tentang Tata Cara Dan Persyaratan Pembayaran Manfaat Jaminan Hari Tua*. Indonesia: Kementerian Hukum Dan Hak Asasi Manusia Republik Indonesia, 2022, pp. 1–9.
- [12] N. Putri and M. Abiyu, "Masalah Penetapan Manfaat Jaminan Hari Tua (Jht) Setelah Terbitnya Peraturan Menteri Ketenagakerjaan Nomor 2 Tahun 2022," *Rewang Rencang J. Huk. Lex Gen.*, vol. 3, no. 5, pp. 395–411, 2022.
- [13] "BPJS Ketenagakerjaan." [Online]. Available: <https://www.bpjsketenagakerjaan.go.id/berita/27943/Polemik-Aturan-Baru-JHT,-Begini-Penjelasan-BPJS-Ketenagakerjaan>. [Accessed: 18-Feb-2022].
- [14] H. Patel, "What Happen in an Internet Minute - Bond High Plus," 2021. [Online]. Available: <https://www.bondhighplus.com/2021/04/14/what-happen-in-an-internet-minute>. [Accessed: 20-Feb-2022].
- [15] H. Rachmi, S. Suparni, and A. Al Kaafi, "Analisis Sentimen Sistem Ganjil Genap Kota Bogor," *J. ELTIKOM*, vol. 5, no. 2, pp. 92–99, 2021.
- [16] M. Nurmalasari, N. Anggita Temesvari, S. Ni, and matul Maula, "Analisis Sentimen terhadap Opini Masyarakat dalam Penggunaan Mobile-JKN untuk Pelayanan BPJS Kesehatan Tahun 2019," *Heal. Inf. Manag. J. ISSN*, vol. 8, no. 1, pp. 2655–9129, 2020.
- [17] V. K. S. Que, A. Iriani, and H. D. Purnomo, "Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 9, no. 2, pp. 162–170, 2020.
- [18] B. Laurensz and Eko Sedyono, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi <http://sistemasi.ftik.unisi.ac.id>

- dalam Upaya Mengatasi Pandemi Covid-19,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 2, pp. 118–123, 2021.
- [19] P. Arsi, R. Wahyudi, and R. Waluyo, “Optimasi SVM Berbasis PSO pada Analisis Sentimen Wacana Pindah Ibu Kota Indonesia,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 231–237, 2021.
- [20] P. Arsi and R. Waluyo, “Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM),” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021.