PEMETAAN OPINI PUBLIK TERHADAP PERUBAHAN KEBIJAKAN BPJS KESEHATAN DENGAN PENDEKATAN *SUPPORT VECTOR MACHINE(SVM)* DALAM ANALISIS SENTIMEN

Damayanti [1], Dendy Indriya Efendi [2], Dodi Solihudin [3], Cep Lukman Rohmat [4], Sandy Eka Permana [5]
[1,2,3] Teknik Informatika STMIK IKMI Cirebon,
[4] Rekayasa Perangkat Lunak STMIK IKMI Cirebon,
[5] Manajemen Informatika STMIK IKMI Cirebon
Jl. Perjuangan No.10B, Karyamulya, kec. Kesambi, Kota Cirebon, Jawa Barat 45135
damayantiimay947@gmail.com

ABSTRAK

Perubahan kebijakan iuran BPJS Kesehatan kerap menjadi topik hangat di dunia maya. Riset ini bertujuan untuk melakukan analisis sentimen terkait perubahan iuran BPJS Kesehatan dengan memanfaatkan algoritma *Support Vector Machine (SVM).* Fokus utama memahami respon masyarakat melalui komentar di media sosial, dimana *SVM* berperan dalam mengklasifikasikan sentimen teks, menyoroti pandangan masyarakat, dan mengidentifikasi potensi dampak kebijakan. Riset ini membincangkan aspek-aspek penting seperti analisis sentimen, implementasi *SVM*, dan kontribusi riset terhadap perkembangan metode analisis sentimen, pemahaman respons masyarakat terhadap perubahan kebijakan BPJS Kesehatan. *SVM* terbukti berhasil dengan tingkat akurasi mencapai 94.28%. Dalam mengevaluasi sentimen negatif, model *SVM* menunjukkan tingkat *presisi*, *recall*, dan *F1-score* mencapai 97% masing-masing. Sementara itu, untuk sentimen positif, *presisi* mencapai 35%, *recall* 42%, dan *F1-score* 39%. kesimpulan dari penelitian ini bahwa *SVM* memberikan kontribusi yang signifikan dalam menganalisis sentimen terkait kebijakan BPJS Kesehatan, khususnya dalam menghadapi sentimen negatif. Dampak praktisnya melibatkan peningkatan operasional BPJS Kesehatan berdasarkan evaluasi tanggapan masyarakat, dengan fokus peningkatan pelayanan dan komunikasi untuk mengurangi dampak negatif dan meningkatkan kepuasan pengguna. Penelitian ini menjadi dasar bagi pengambil kebijakan untuk keputusan yang responsif, sesuai harapan masyarakat. Latar belakang permasalahan menyoroti kompleksitas isu dan pentingnya pemahaman mendalam terhadap pandangan masyarakat terhadap perubahan kebijakan BPJS Kesehatan.

Kata kunci : Analisis sentiment, BPJS, *Support Vector Machine*

1. PENDAHULUAN

Perubahan kebijakan BPJS Kesehatan menjadi isu hangat dalam masyarakat sejak tahun 2020, terutama terkait kenaikan iuran untuk beberapa kelas peserta. Dampaknya cukup signifikan, khususnya bagi masyarakat yang harus membayar iuran yang lebih tinggi. Kebijakan ini diresmikan oleh Pemerintah dan menimbulkan beragam respons serta komentar dari masyarakat, terutama di media sosial, mencerminkan pentingnya pemahaman terhadap opini dan sentimen masyarakat mengenai kebijakan kesehatan. Oleh karena itu, penelitian ini bertujuan untuk menganalisis sentimen dan memprediksi komentar terhadap perubahan iuran BPJS Kesehatan menggunakan algoritma *Support Vector Machine (SVM)*[1]

Isu yang diidentifikasi dalam penelitian ini adalah bagaimana melakukan analisis sentimen terhadap komentar terkait perubahan kebijakan BPJS Kesehatan dengan menggunakan algoritma *Support Vector Machine* (*SVM*). Keterlibatan masyarakat yang tinggi, terlihat dari volume komentar yang mencapai angka yang signifikan, menunjukkan relevansinya dan nilai signifikan dalam menyelesaikan permasalahan kontroversial ini. Penerapan algoritma *SVM* menjadi sorotan karena keefektifannya dalam mengklasifikasikan teks. Oleh karena itu, penelitian ini berkontribusi pada pengembangan metode analisis sentimen dan mendukung pengambilan keputusan terkait perubahan kebijakan BPJS Kesehatan.

Penelitian sebelumnya oleh [2] telah menganalisis sentimen publik terhadap layanan BPJS Kesehatan menggunakan data *Twitter*. Hasilnya menunjukkan mayoritas sentimen netral, diikuti oleh sentimen negatif dan positif. Walau penelitian serupa telah dilakukan, penelitian ini tetap relevan karena difokuskan pada analisis sentimen terkait perubahan kebijakan BPJS Kesehatan dan menggunakan metode *SVM*.

BPJS Kesehatan, sebagai entitas yang mengatur program Jaminan Kesehatan Nasional di Indonesia, memiliki visi dan misi untuk memberikan cakupan medis berkualitas dan mendukung *Universal Health Coverage (UHC)* Analisis sentimen terhadap pelayanan BPJS Kesehatan, seperti yang dilakukan oleh [3] juga menggunakan metode *SVM* untuk mengklasifikasikan sentimen positif atau negatif dari masyarakat umum.

Penelitian ini memiliki tujuan utama dalam melakukan analisis sentimen terkait perubahan kebijakan kesehatan menggunakan metode *SVM*.

Selain itu, penelitian ini juga mencakup evaluasi pandangan masyarakat terhadap perubahan kebijakan tersebut, dengan harapan memberikan kontribusi pada perkembangan teknologi dan pengetahuan di bidang analisis sentimen dan ilmu komputer secara menyeluruh. Fokus pada respons masyarakat bertujuan untuk memberikan manfaat bagi pemerintah dan masyarakat dalam meningkatkan kualitas pelayanan kesehatan.

Pendekatan kualitatif-subjektif diterapkan dalam penelitian ini untuk mengevaluasi pandangan terhadap perubahan kebijakan BPJS Kesehatan dengan memanfaatkan algoritma *SVM*. Kombinasi pendekatan kualitatif dan *SVM* diharapkan dapat memberikan kesimpulan yang lebih akurat dalam mengevaluasi pendapat masyarakat terkait perubahan iuran BPJS Kesehatan. Hasil analisis sentimen diharapkan dapat memberikan informasi berharga bagi praktisi, peneliti, dan pemerintah dalam merumuskan kebijakan dan meningkatkan standar layanan kesehatan.

Penelitian ini juga berpotensi untuk pengembangan ilmu pengetahuan dan teknologi, terutama dalam mengembangkan algoritma analisis sentimen yang lebih akurat dan efisien. Temuan ini diharapkan dapat memberikan kontribusi pada perkembangan teknologi dan ilmu pengetahuan, terutama di bidang analisis sentimen dan ilmu komputer secara umum. Dengan mengikuti langkah-langkah analisis sentimen dan pemanfaatan *SVM*, penelitian ini diharapkan memberikan kontribusi signifikan pada pemahaman sentimen masyarakat terkait BPJS Kesehatan dan perubahan.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen atau *opinion mining*, yang mendapatkan popularitas sejak pertengahan 2000-an, merupakan bidang eksplorasi yang dinamis dalam pemrosesan bahasa. Melibatkan integrasi data mining dan *text mining*, tujuan dari analisis sentimen adalah mengembangkan perangkat terprogram untuk mengekstraksi dan memproses informasi subjektif dari masyarakat atau pakar melalui berbagai media. Fokus utamanya adalah menghasilkan informasi yang terstruktur untuk mendukung sistem pengambilan keputusan, dengan mengelompokan pesan dalam laporan, kalimat atau sudut pandang guna menentukan apakah pendapat tersebut bersifat positif atau negatif [4]

2.2. Text Mining

Text Mining adalah metode umum yang sering digunakan untuk mengeksplorasi data dan menganalisis informasi dari sumber teks yang belum mengidentifikasi pola yang relevan. Dalam prakteknya, text mining berfungsi sebagai alat dalam interaksi dinamis dengan memanfaatkan informasi teks. Proses pra-pemrosesan data melibatkan beberapa tahap, seperti *Cleaning* untuk menghapus informasi yang tidak diperlukan, *Case Folding* untuk menyamakan format huruf, *Tokenizing* untuk membagi kata-kata menjadi unit-token, *Stopword Removal* untuk menghilangkan kata-kata umum yang tidak signifikan, dan Lemmatization untuk mengubah kata-kata ke bentuk dasar. Langkah-langkah ini membantu menyiapkan data teks agar dapat diidentifikasi pola dan diekstraksi informasi dengan lebih efisien selama proses analisis *text mining* [5]

2.3. Support Vector Machine

Support Vector Machine (SVM) terkenal karena kemampuannya menghasilkan model klasifikasi yang efisien tanpa mengandalkan perhitungan pohon keputusan. Meskipun *SVM* tidak memanfaatkan pohon keputusan, tingkat keakuratannya cenderung meningkat seiring dengan kelengkapan informasi pada data pelatihan. Dalam konteks penilaian kelayakan kredit, *SVM* menjadi opsi untuk mengevaluasi apakah aplikasi kredit dapat dianggap layak atau tidak.

Implementasi *SVM* dalam penilaian kelayakan kredit melibatkan pengukuran kinerjanya dengan menggunakan metrik seperti *Area Under the Curve (AUC)* dan tingkat akurasi. *SVM* memiliki peran penting dalam memisahkan pelanggan yang berhak dan tidak berhak menerima kredit, diukur melalui *AUC* untuk menilai akurasi pemisahan kelas dan tingkat akurasi untuk prediksi yang tepat. Evaluasi kinerja algoritma *SVM* dalam penilaian kelayakan kredit dapat disederhanakan dengan meninjau hasil *AUC* dan tingkat akurasi dari tabel performa, menunjukkan seberapa efektif *SVM* dalam meramalkan keputusan kelayakan kredit berdasarkan data yang tersedia. [6] .

2.4. Klasifikasi

Klasifikasi ialah suatu proses dengan tujuan menemukan model atau fungsi yang mampu menggambarkan serta membedakan data ke dalam kelas-kelas tertentu. Tujuan klasifikasi adalah untuk mengenali pola yang memiliki nilai yang signifikan dari data yang berukuran cukup besar hingga sangat besar [7].

2.5. BPJS (Badan Penyelenggara Jaminan Sosial)

BPJS, yang merupakan singkatan dari Badan Penyelenggara Jaminan Sosial, didirikan untuk menjalankan Program Jaminan Sosial di Indonesia sesuai dengan ketentuan Undang-Undang No. 40 Tahun 2004 dan Undang-Undang No. 24 Tahun 2011 tentang Sistem Jaminan Sosial Nasional (SJSN). BPJS bertugas menggantikan beberapa lembaga jaminan sosial sebelumnya di Indonesia, termasuk Lembaga Asuransi Jaminan Kesehatan PT. Akses yang berubah status menjadi BPJS Kesehatan, dan Lembaga Jaminan Sosial Ketenagakerjaan PT. Jamsostek yang berubah menjadi BPJS Ketenagakerjaan. Proses transformasi dari PT. Akses dan PT. Jamsostek menjadi BPJS dilakukan secara

bertahap, di mana PT. Akses menjadi BPJS Kesehatan pada awal tahun 2014, sedangkan PT. Jamsostek berubah menjadi BPJS Ketenagakerjaan pada tahun 2015 [2]. Transformasi ini menjadi langkah strategis untuk menyelaraskan dan meningkatkan efisiensi pelaksanaan program jaminan sosial di Indonesia.

2.6. Knowledge Discopery in Database (KDD)

Knowledge Discovery in Databases (KDD) adalah suatu metode dalam *data mining* yang digunakan untuk melakukan proses pengolahan data yang telah ada dalam *database*. Tujuan utamanya adalah untuk mengolah data tersebut guna mendapatkan informasi baru yang lebih mudah dipahami [8].

2.7. Evaluasi

Proses evaluasi dengan menggunakan *cross-validation* dan perhitungan manual *confusion matrix* melibatkan pengukuran akurasi, *presisi*, *recall*, dan *F1-Score*. [9]

Akurasi mencerminkan sejauh mana model dapat mengklasifikasikan dengan tepat.

$$Accuracy = \frac{Tp+TN}{TP+TN+FP+FN} \quad (1)$$

Precision menggambarkan tingkat ketepatan antara data yang diminta dan hasil prediksi yang diberikan oleh model.

$$Precision = \frac{Tp}{TP+FP} \quad (2)$$

Recall menggambarkan keberhasilan model dalam menemukan kembali informasi yang relevan.

$$Recall = \frac{Tp}{TP+FN} \quad (3)$$

F1-Score menggambarkan perbandingan rata-rata precision dan recall yang dibobotkan. *Accuracy* cocok digunakan sebagai ukuran performa algoritma ketika *dataset* memiliki *false negatives* dan *false positives* yang mendekati. Namun, jika jumlahnya tidak dekat, lebih baik menggunakan *F1-Score* sebagai ukuran performa.

$$F1\text{-}Score = \frac{2.Precission.Recall}{Presisi+Recall} \quad (4)$$

2.8. Confusion Matrix

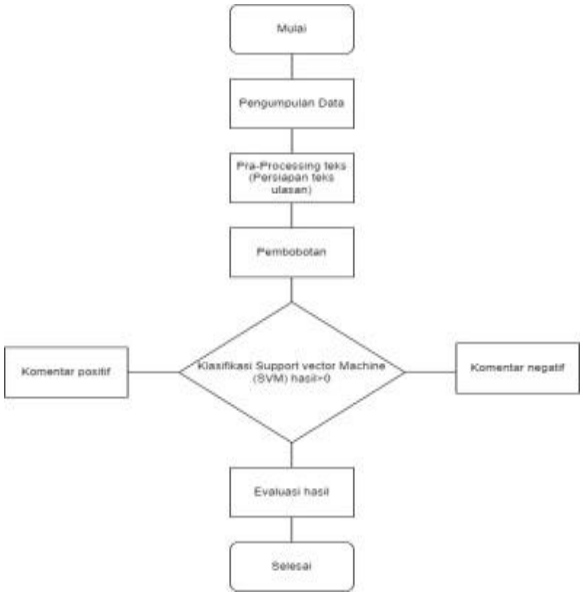
Confusion matrix adalah teknik dalam data mining yang seringkali dipakai untuk menilai performa suatu algoritma atau model klasifikasi. *Confusion matrix* memberikan informasi tentang keakuratan, presisi, dan tingkat kesalahan algoritma. Umumnya, metode ini diwujudkan dalam bentuk tabel yang menampilkan empat kemungkinan hasil klasifikasi, sebagaimana ditunjukkan berikut [10].

Tabel 1. *Confusion Matrix*

Actual Class	*Predicted Class*	
	+	-
+	*TP*	*TN*
-	*FP*	*FN*

3. METODE PENELITIAN

Metode untuk mendapatkan informasi dari *database* melibatkan eksplorasi data yang ada. Dalam *database*, terdapat tabel-tabel yang saling berkorelasi. Hasil informasi yang diperoleh dari proses ini dapat dijadikan sebagai dasar pengetahuan *(knowledge base)* untuk mendukung pengambilan keputusan [11].



Gambar 1. Alur Penelitian

3.1. Data Selection

Dalam tahap ini, yang merupakan bagian dari *Knowledge Discovery in Data (KDD)*, dilakukan pengumpulan data. Data tersebut dikumpulkan dari *Kaggle*, sebuah *platform* yang menyajikan berbagai dataset, kompetisi ilmu data, dan sumber daya lainnya untuk para praktisi ilmu data dan *machine learning.*

3.2. Pre-Processing

Pre-processing adalah tahap penting untuk membersihkan dan memperbaiki data yang tidak terstruktur dan kaya akan karakter. Tujuan utama dari tahap pra-pemrosesan adalah mengatasi kekacauan tersebut. Dalam langkah ini, beberapa tindakan harus diambil, termasuk pembersihan, konversi huruf ke huruf kecil, tokenisasi, dan *filtering*.

3.3. Transformation

Dalam tahap transformasi, data mengalami perubahan bentuk agar dapat diolah selama proses data mining. Pada langkah ini, data diproses dan dimodifikasi dengan tujuan memastikan penggunaannya yang efisien dalam menghasilkan pemahaman atau pola yang relevan melalui proses *data mining*.

3.4. Data Mining

Dalam tahap data mining, proses analisis sentimen dengan menggunakan algoritma *Support Vector Machine (SVM)* melibatkan serangkaian

langkah-langkah, dimulai dari pengumpulan dan seleksi data ulasan hingga evaluasi kinerja model. Data ulasan diperoleh dari berbagai sumber dengan memastikan keberagaman pendapat dalam dataset. Proses selanjutnya melibatkan pembersihan dan normalisasi data untuk mengatasi duplikasi dan menjaga konsistensi teks. Melalui tokenisasi dan pembentukan fitur seperti vektor kata atau matriks *TF-IDF*, ulasan diwakili secara numerik. Setelah pembagian dataset menjadi data pelatihan dan pengujian, model *SVM* dilatih dengan menentukan parameter optimal dan membentuk *hyperplane*. Validasi dan evaluasi model, dengan menggunakan metrik seperti akurasi dan *presisi*, memberikan wawasan mendalam tentang kemampuan model dalam mengklasifikasikan sentimen ulasan. Hasil dan interpretasi membantu memahami preferensi pengguna, dan jika diperlukan, model dapat diperbaiki melalui iterasi untuk meningkatkan kinerja. Secara keseluruhan, proses ini tidak hanya menciptakan model yang handal, tetapi juga memberikan pemahaman mendalam tentang nuansa dan preferensi yang mungkin terlewatkan tanpa analisis yang terstruktur.

3.5. Evaluation

Pada tahap evaluasi ini, *confusion matrix* akan digunakan untuk menilai kinerja akurasi dari algoritma *Support Vector Machine (SVM)*, dengan evaluasi yang mencakup akurasi, *presisi*, dan *recall*. Confusion matrix merupakan salah satu metode untuk mengukur performa model klasifikasi.

4. HASIL DAN PEMBAHASAN

Pada tahap ini, ditampilkan data yang telah diperoleh pada tahap pengambilan data. Data tersebut mencakup data sekunder, yaitu data perubahan kebijakan BPJS dengan entitas *Comment* dan Label. Jumlah data yang dikumpulkan dalam penelitian ini sebanyak 3060 data. Dalam proses pemilihan data untuk analisis, peneliti menggunakan dataset yang diambil dari *Kaggle*, sebuah situs yang menyediakan kumpulan data yang beragam. Fokus utama penelitian adalah pada ulasan perubahan kebijakan BPJS kesehatan, yang dapat diakses melalui tautan https://www.kaggle.com/datasets/aeworld/sentimen-id-bpjs.

Tabel 2. Tabel *Dataset* Komentar BPJS

No	Comment	Label
1.	Bertambah lagi beban pengeluaran jadi punya hutang bulanan sekeluarga..	Negatif
2.	Intinya setiap warga negara saat ini wajib setor kpd pemerintah, untuk membantu membayar hutang negara	Negatif
3.	berkedok bpjs? kalau mau jadi preman tidak usah pake embel2 program	Negatif
….	….	…..
3058	Gimana kok jadi ruwet gini...	Negatif
3059	Setuju sebagai dasar Pelayanan Rakyat Kecil, Sebagai dasar menentukan BLT, Bebas Pajak Jual beli Tanah, Bebas/pengurang Bayar Iuran BPJS JHT dan Kesehatan	Positif
3060	Kebijakan yang membuat masyarakat semakin terpuruk. Seandai nya juga mau berobat, blm tentu BPJS ITU BS DI GUNAKAN. HAL UTAMA DALAM BEROBAT ADALAH KEPASTIAN, JIKA MENGGUNAKAN BPJS. nyawa kita berujung maut. Contoh nya di daerah saya dan itu terjadi dengan saya sendiri. SUNGGUH KONYOL KEBIJAKAN PEMERINTAHAN JOKOWI SAAT INI. SEMOGA MATA KALIAN TERBUKA LEBAR dengan kebijkan kalian ini	Negatif

4.1. Case Folding

Dalam langkah transformasi huruf, langkah ini melibatkan konversi teks dari huruf besar ke huruf kecil dalam dokumen yang sebelumnya menggunakan huruf kapital. Proses ini bertujuan untuk menghapus perbedaan antara huruf besar dan huruf kecil dalam teks.

Tabel 3. Case Folding

Comment	***Case Folding***
Bertambah lagi beban pengeluaran jadi punya hutang bulanan sekeluarga..	bertambah lagi pengeluaran jadi punya hutang bulanan sekeluarga..
Setuju sebagai dasar Pelayanan Rakyat Kecil, Sebagai dasar menentukan BLT, Bebas Pajak Jual beli Tanah, Bebas/pengurang Bayar Iuran BPJS JHT dan Kesehatan	setuju sebagai dasar pelayanan rakyat kecil, sebagai dasar menentukan blt, bebas pajak jual beli tanah, bebas/pengurang bayar iuran bpjs jht dan kesehatan
Intinya setiap warga negara saat ini wajib setor kpd pemerintah, untuk membantu membayar hutang negara	intinya setiap warga negara saat ini wajib setor kpd pemerintah, untuk membantu membayar hutang negara

4.2. Cleaning Text

Langkah selanjutnya adalah membersihkan data, yang melibatkan penghapusan karakter seperti simbol, angka, tanda baca, dan emotikon.

Tabel 4. *Cleaning Text*

Comment	***Cleaning Text***
bertambah lagi pengeluaran jadi punya hutang bulanan sekeluarga..	bertambah lagi pengeluaran jadi punya hutang bulanan sekeluarga
setuju sebagai dasar pelayanan rakyat kecil, sebagai dasar menentukan blt, bebas pajak jual beli tanah,	setuju sebagai dasar pelayanan rakyat kecil sebagai dasar menentukan blt bebas pajak jual beli

Comment	*Cleaning Text*
bebas/pengurang bayar iuran bpjs jht dan kesehatan	tanah bebas pengurang bayar iuran bpjs jht dan kesehatan
intinya setiap warga negara saat ini wajib setor kpd pemerintah, untuk membantu membayar hutang negara	intinya setiap warga negara saat ini wajib setor kpd pemerintah untuk membantu membayar hutang negara

4.3. Tokenize

Tokenisasi dilakukan dengan menggunakan spasi sebagai karakter pemisah setiap kata. Tiap *tweet* terdiri dari serangkaian kata yang saling terhubung dan dipisahkan oleh spasi.

Tabel 5. *Tokenize*

Comment	*Tokenize*
bertambah lagi pengeluaran jadi punya hutang bulanan sekeluarga	[bertambah, lagi, pengeluaran, jadi, punya, hutang, bulanan, sekeluarga]
setuju sebagai dasar pelayanan rakyat kecil sebagai dasar menentukan blt bebas pajak jual beli tanah bebas pengurang bayar iuran bpjs jht dan kesehatan	[setuju, sebagai, dasar, pelayanan, rakyat, kecil, sebagai, dasar, menentukan, blt, bebas, pajak, jual, beli, tanah, bebas, pengurang, bayar, iuran, bpjs, jht, dan, kesehatan]
intinya setiap warga negara saat ini wajib setor kpd pemerintah untuk membantu membayar hutang negara	[intinya, setiap, warga, negara, saat, ini, wajib, setor, kpd, pemerintah, untuk, membantu, membayar, hutang, negara]

4.4. Penghapusan Stopword

Setiap kata dalam *tweet* akan diperiksa. Jika *tweet* tersebut mengandung kata sambung, kata depan, kata ganti, atau kata yang tidak relevan dengan analisis sentimen, maka kata-kata tersebut akan dihilangkan

Tabel 6. Stopword

Comment	*Stopword*
bertambah lagi pengeluaran jadi punya hutang bulanan sekeluarga	bertambah pengeluaran jadi punya hutang bulanan sekeluarga
setuju sebagai dasar pelayanan rakyat kecil sebagai dasar menentukan blt bebas pajak jual beli tanah bebas pengurang bayar iuran bpjs jht dan kesehatan	setuju landasan pelayanan masyarakat kecil menentukan bantuan langsung tunai (blt) pajak jual beli tanah pengurangan pembayaran iuran bpjs jaminan hari tua (jht) kesehatan
intinya setiap warga negara saat ini wajib setor kpd pemerintah untuk membantu membayar hutang negara	Intinya warga negara saat ini wajib setor kepada pemerintah membantu membayar hutang negara

4.5. Lemmatized

Langkah *lemmatization* dilakukan untuk menghilangkan imbuhan sehingga kata dapat diubah menjadi bentuk dasarnya.

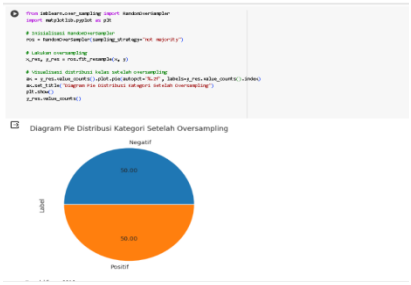
Tabel 7. Lemmatized

Comment	*Lemmatized*
bertambah lagi pengeluaran jadi punya hutang bulanan sekeluarga	tambah beban keluar hutang bulan keluarga
setuju sebagai dasar pelayanan rakyat kecil sebagai dasar menentukan blt bebas pajak jual beli tanah bebas pengurang bayar iuran bpjs jht dan kesehatan	setuju landas layan masyarakat kecil tentu bantuan langsung tunai (blt) pajak jual beli tanah kurang bayar iuran bpjs jamin hari tua (jht) kesehatan.
Intinya warga negara saat ini wajib setor kepada pemerintah membantu membayar hutang negara	Inti warga negara saat ini wajib setor kepada pemerintah membantu bayar hutang negara

4.6. Undersampling dan Oversampling

Penanganan ketidakseimbangan kelas dalam dataset adalah langkah kritis untuk memastikan performa optimal pada model machine learning, khususnya dalam tugas klasifikasi. Dua pendekatan utama yang sering digunakan melibatkan oversampling, yaitu peningkatan sampel pada kelas minoritas, dan undersampling, yaitu pengurangan sampel pada kelas mayoritas. Pendekatan ini efektif dalam mengatasi bias model terhadap kelas mayoritas, sekaligus meningkatkan kemampuan model dalam mengklasifikasikan kelas minoritas. Sebagai contoh, jika kelas negatif memiliki 2912 sampel dan kelas positif hanya 148 sampel, teknik ini dapat diterapkan untuk mencapai keseimbangan yang diinginkan.

Gambar 2. *Undersampling*



Gambar 3. *Oversampling*

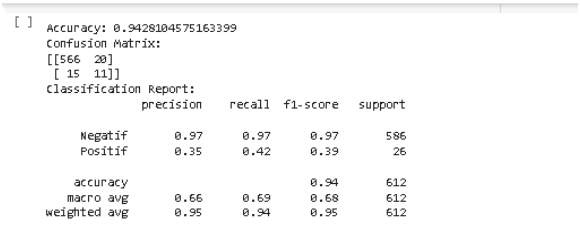
4.7. Pembobotan TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) adalah tahapan metode pembobotan yang digunakan dalam pemrosesan teks untuk memberikan bobot pada setiap kata.

Gambar 4. *TF-IDF*

4.8. Evaluasi

Model yang dievaluasi berhasil mencapai tingkat akurasi sebesar 94.28%. Dalam penanganan kelas Positif, akurasi prediksi sekitar 35%, dan model berhasil mengidentifikasi sekitar 42% dari total kasus Positif. Meskipun begitu, hasil evaluasi diperkuat oleh *confusion matrix*, yang memberikan rincian lebih mendalam mengenai prediksi yang tepat dan prediksi yang salah. Sementara itu, kelas Negatif menunjukkan kinerja yang sangat baik dengan tingkat *precision, recall*, dan *F1-score* mencapai 97%.

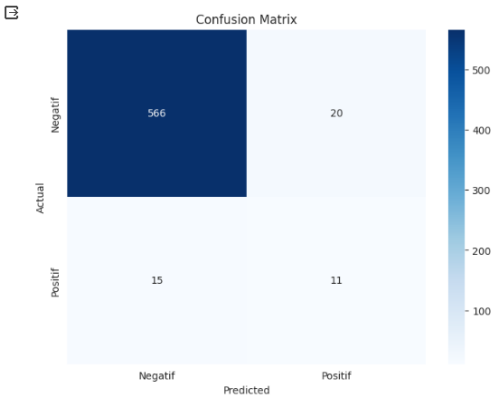


```
Accuracy: 0.9428104575163399
Confusion Matrix:
[[566  20]
 [ 15  11]]
Classification Report:
              precision    recall  f1-score   support

     Negatif       0.97      0.97      0.97       586
     Positif       0.35      0.42      0.39        26

    accuracy                           0.94       612
   macro avg       0.66      0.69      0.68       612
weighted avg       0.95      0.94      0.95       612
```

Gambar 5. Hasil Evaluasi dan Hasil



Gambar 6. Confusion Matrix

Confusion matrix memberikan gambaran prediksi model untuk setiap kelas. Dari evaluasi tersebut, terdapat 566 True Negative (TN), menunjukkan bahwa model secara tepat memprediksi kasus-kasus yang sebenarnya termasuk dalam kelas Negatif. Sebanyak 20 kasus False Positive (FP) mengindikasikan kesalahan model dalam memprediksi kasus-kasus yang seharusnya kelas Negatif sebagai Positif. Selanjutnya, terdapat 15 kasus False Negative (FN), menandakan bahwa model keliru memprediksi kasus-kasus yang sebenarnya kelas Positif sebagai kelas Negatif. Di sisi lain, terdapat 11 True Positive (TP), menunjukkan bahwa model berhasil memprediksi dengan benar kasus-kasus yang seharusnya kelas Positif. Informasi ini memberikan gambaran lengkap mengenai kinerja model dalam mengklasifikasikan masing-masing kelas.

5. KESIMPULAN DAN SARAN

Penelitian ini menyimpulkan bahwa algoritma *Support Vector Machine (SVM)* berhasil menganalisis sentimen terhadap perubahan kebijakan BPJS Kesehatan dengan tingkat akurasi mencapai 94.28%. Meskipun model ini telah mencapai tingkat akurasi yang memadai, terdapat peluang untuk meningkatkan performa, terutama dalam mengenali sentimen positif. Peningkatan kemampuan *SVM* dapat diarahkan dengan menggabungkan analisis sentimen dan data ulasan pengguna. Evaluasi model menunjukkan bahwa dalam menangani kelas positif, akurasi prediksi sekitar 35%, dan model berhasil mengidentifikasi sekitar 42% dari total kasus positif. Confusion matrix memperkuat temuan evaluasi dengan menunjukkan performa yang sangat baik pada kelas negatif, mencapai tingkat *precision, recall*, dan *F1-score* sebesar 97%. Dalam menangani kelas positif, terdapat 20 kasus *False Positive* dan 15 kasus *False Negative*, mencerminkan tantangan dalam memprediksi kasus positif. Secara keseluruhan, hasil evaluasi ini menjadi landasan penting dalam merumuskan kesimpulan dan memberikan dasar untuk saran pengembangan model di masa mendatang.

DAFTAR PUSTAKA

[1] D. D. Nada, S. Soehardjoepri, dan R. M. Atok, "Perbandingan Analisis Sentimen Mengenai BPJS pada Media Sosial Twitter Menggunakan Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM)," *J. Sains Dan Seni ITS*, 2023, [Daring]. Tersedia pada: https://ejurnal.its.ac.id/index.php/sains_seni/article/view/96330

[2] A. R. Kardian dan D. Gustiana, "Analisis Sentimen Berdasarkanan Opini Pengguna pada Media Twitter Terhadap BPJS Menggunakan Metode Lexicon Based dan Naïve Bayes Classi er Twitter Text Mining," *J. Ilm. KOMPUTASI*, vol. 20, hal. 39–52, 2021.

[3] F. Wulandari, P. A. Jusia, dan J. Jasmir, "Klasifikasi Data Mining Untuk Mendiagnosa Penyakit ISPA Menggunakan Metode Naïve Bayes Pada Puskesmas Jambi Selatan," *Jurnal Ilmiah Mahasiswa Sistem Informasi*. 2020.

[4] L. A. Andika, P. A. N. Azizah, dan R. Respatiwulan, "Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier,"

Indones. J. Appl. Stat., vol. 2, no. 1, hal. 34, 2019, doi: 10.13057/ijas.v2i1.29998.

[5] H. C. Husada dan A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, hal. 18–26, 2021, doi: 10.34148/teknika.v10i1.311.

[6] Syafi'i, O. Nurdiawan, dan G. Dwilestari, "Penerapan Machine Learning Untuk Menentukan Kelayakan Kredit Menggunakan Metode Support Vektor Machine," *J. Sist. Inf. dan Manaj.*, vol. 10, no. 2, hal. 1–6, 2022.

[7] H. D. Wijaya dan S. Dwiasnati, "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat," *J. Inform.*, vol. 7, no. 1, hal. 1–7, 2020, doi: 10.31311/ji.v7i1.6203.

[8] D. P. Utomo dan M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, 2020, [Daring]. Tersedia pada: http://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/2080

[9] W. Hidayat, M. Ardiansyah, dan A. Setyanto, "Pengaruh Algoritma ADASYN dan SMOTE terhadap Performa Support Vector Machine pada Ketidakseimbangan Dataset Airbnb," *Edumatic J. Pendidik. Inform.*, vol. 5, no. 1, hal. 11–20, 2021, doi: 10.29408/edumatic.v5i1.3125.

[10] M. A. Ramadhan dan M. I. Wahyudin, "Analisis Sentimen Mengenai Keberhasilan Indonesia di Ajang Thomas Cup 2020 (Studi Kasus Media Sosial Twitter) Menggunakan Metode Naïve Bayes dan Decision Tree," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 6, no. 4, hal. 505–511, 2022, doi: 10.35870/jtik.v6i4.560.

[11] T. Raudya, B. Irawan, dan A. Faqih, "Perbandingan implementasi 2 algoritma svm dan k-nn dalam pengklasifikasian kepuasan pengguna smart e-learning," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 3, hal. 2057–2062, 2023.