

ANALISIS SENTIMEN DATA PADA BPJS KESEHATAN DENGAN METODE BACKPROPAGATION NEURAL NETWORK

SENTIMENT ANALYSIS OF BPJS KESEHATAN DATA WITH BACKPROPAGATION NEURAL NETWORK METHOD

Ricky Bima Pratama Ardikar, Budhi Irawan, S.Si., M.T.2, Casi Setianingsih, S.T., M.T.3

1,2,3Prodi S1 Teknik Komputer, Fakultas Teknik Elektro, Universitas Telkom

rickybima@student.telkomuniversity.ac.id, budhiirawan@telkomuniversity.co.id,

setiacasie@telkomuniversity.co.id

Abstrak

Pada zaman sekarang kemajuan teknologi semakin pesat memungkinkan segala sesuatu menjadi lebih mudah untuk diakses menggunakan internet. Semakin banyaknya pengguna internet di Indonesia menyebabkan bermunculan media sosial yang bisa digunakan untuk menyampaikan opini, bertukar informasi, mengunggah video dan foto, salah satunya yaitu media sosial Twitter. Pada media sosial Twitter pertukaran informasi yang cepat menjadi kelebihan dari media sosial tersebut hal itu sering digunakan dalam menyampaikan berita maupun opini berupa kritik dan saran seperti kepada lembaga pemerintah, contohnya setiap ada isu kenaikan iuran ataupun terkait pelayanan pada BPJS Kesehatan selalu menjadi perdebatan para penggunanya. BPJS Kesehatan merupakan lembaga pemerintah yang menjamin kesehatan masyarakat Indonesia, dalam hal ini para pekerja negeri sipil (PNS) maupun pekerja swasta diwajibkan untuk mendaftar asuransi tersebut, selain pekerja juga adanya asuransi untuk masyarakat yang tidak mampu. Belum adanya sistem yang dapat menganalisis opini pengguna bersifat positif maupun negatif yang disampaikan oleh pengguna BPJS Kesehatan tersebut, maka akan dibuat sebuah sistem analisis sentimen dengan metode klasifikasi Backpropagation Neural Network. Dengan adanya sistem tersebut dapat membantu menganalisis opini pengguna BPJS Kesehatan pada media sosial Twitter. Dari hasil penelitian pada tugas akhir ini dalam proses analisis sentimen pada tweet pengguna Twitter mendapatkan accuracy sebesar 87.14%, precision sebesar 86.97%, recall sebesar 87.14%, dan F1 score sebesar 87.05%.

Kata kunci : Backpropagation Neural Network, BPJS Kesehatan, Sentiment Analysis, Twitter

Abstract

In this day and age technological advances are rapidly making everything possible easier to access using the internet. The increasing number of internet users in Indonesia has led to the emergence of social media that can be used to express opinions, exchange information, upload videos and photos, one of which is Twitter social media. On social media Twitter the rapid exchange of information is an advantage of social media. It is often used in delivering news and opinions in the form of criticism and suggestions such as to government agencies, for example, every time there is an issue of rising fees or related services to BPJS Kesehatan, it is always being debated by its users. BPJS Kesehatan is a government agency that guarantees the health of Indonesian people, in this case civil servants (PNS) and private workers are required to register for insurance, in addition to workers there is also insurance for people who can not afford. The absence of a system that can analyze both positive and negative user opinions submitted by users of the BPJS Kesehatan, a sentiment analysis system will be made using the Backpropagation Neural Network classification method. With this system, it can help analyze BPJS Kesehatan user opinions on social media Twitter. From the results of this final project in the process of sentiment analysis on Twitter users' tweets get an accuracy of 87.14%, a precision of 86.97%, a recall of 87.14%, and an F1 score of 87.05%.

Keywords: Backpropagation Neural Network, BPJS Kesehatan, Sentiment Analysis, Twitter

1. Pendahuluan

BPJS Kesehatan sudah terbentuk sejak tahun 1968 dengan nama yang berbeda. Jika dulu banyak orang yang mengenalnya sebagai nama PT Askes [1]. Namun sekarang ini berganti nama menjadi BPJS sejak tahun 2014 [1]. Melalui Program Jaminan Kesehatan Nasional-Kartu Indonesia Sehat (JKN-KIS) yang diselenggarakan oleh BPJS Kesehatan, negara hadir di tengah kita untuk memastikan seluruh penduduk Indonesia terlindungi oleh jaminan kesehatan yang komprehensif, adil, dan merata [2].

Adapun dari kebijakan – kebijakan BPJS Kesehatan sekarang banyak yang mengkritisi bahkan menolak, hal itu salah satunya dikarenakan adanya kenaikan iuran BPJS Kesehatan. Presiden Konfederasi Serikat Pekerja Indonesia (KSPI) Said Iqbal menyatakan bahwa kalangan buruh akan tetap menggelar aksi unjuk rasa untuk menyampaikan aspirasi pada hari ini, Rabu, 2 Oktober 2019 [3]. "Besok kaum buruh akan tetap melakukan aksi besar-besaran di 10 provinsi. Khusus di Jabodetabek, aksi akan di DPR RI," kata Said dalam siaran pers konfederasi yang diterima di Jakarta, Selasa, 1 Oktober 2019 [3]. Dalam unjuk rasa tersebut, serikat buruh akan menyampaikan tiga tuntutan utama, yaitu menolak revisi Undang-Undang Nomor 13 Tahun 2003 tentang Ketenagakerjaan, menolak kenaikan iuran BPJS Kesehatan, dan menuntut revisi Peraturan Pemerintah Nomor 78 tahun 2015 tentang Pengupahan [3]. Maupun pengakuan dari pihak BPJS Kesehatan sendiri, hal ini disampaikan Kepala BPJS Kesehatan Cabang Ambon, Afliana Latumakulita, dalam rapat koordinasi bersama Pemerintah Provinsi Maluku di Kantor Gubernur Maluku, Kota Ambon, Rabu (18/9/2019). "Dalam rapat dengan pemprov Maluku tadi, saya sampaikan dari sisi kualitas pelayanan belum merata baik, dari sisi sumber daya manusia maupun sarana-prasarana," kata Alfiana kepada wartawan di kantor Gubernur Maluku, Kota Ambon. Menurutnya, dari 205 unit puskesmas, hanya 160-unit yang ditempatkan dokter gigi. Bukan saja itu, terdapat 30 puskesmas yang belum memiliki dokter umum [4]. Dalam praktiknya BPJS Kesehatan masih mengalami kerugian. Untuk mengurangi defisit BPJS Kesehatan, Ikatan Dokter Indonesia (IDI) sempat mengusulkan skema sistem terpisah mirip di Thailand. Masyarakat mampu menggunakan asuransi swasta dan BPJS Kesehatan untuk masyarakat miskin. Peningkatan iuran untuk BPJS Kesehatan yang baru baru ini santer terdengar menuai pro-kontra. Terlepas dari itu, pemerintah perlu meninjau kemampuan bayar kelompok masyarakat miskin, kualitas pelayanan kesehatan yang diberikan selama ini mengingat selalu defisit serta efektivitas pelayanan yang ada seperti fasilitas dan plafon yang diberikan.

Untuk mempermudah dalam menganalisa kualitas layanan maupun program yang ada pada BPJS Kesehatan, maka akan dibuat sebuah sistem berbasis web menggunakan sentiment analysis dengan metode BPNN (Backpropagation Neural Network) untuk klasifikasi teks data yang ada di komentar media sosial Twitter BPJS Kesehatan yang mana data itu akan dikelompokkan berdasarkan sentimen positif, sentiment netral, maupun sentimen negatif. Diharapkan dengan adanya sistem ini dapat membantu BPJS Kesehatan untuk lebih meningkatkan lagi kualitas layanan maupun program – program nya untuk masyarakat, selain itu juga adanya pertimbangan dalam membuat sebuah kebijakan.

2. Dasar Teori

2.1 Sentiment Analysis

Analisis sentimen adalah penambangan teks kontekstual yang mengidentifikasi dan mengekstrak informasi subjektif dalam materi sumber, dan membantu bisnis untuk memahami sentimen sosial dari merek, produk atau layanan mereka sambil memantau percakapan online. Namun, analisis aliran media sosial biasanya dibatasi hanya untuk analisis sentimen dasar dan metrik berbasis hitungan[9].

Guna menghasilkan opini yang dibutuhkan, sentiment analysis tidak hanya harus bisa mengenali opini dari teks. Proses yang juga disebut sebagai opini mining ini juga perlu bekerja dengan mengenali tiga aspek berikut:

1. Subjek: topik apa yang sedang dibicarakan.
2. Polaritas: apakah opini yang diberikan bersifat positif, netral, negatif.
3. Pemegang opini: seseorang yang mengeluarkan opini tersebut.

Sentiment analysis kemudian akan membedakan teks menjadi dua kategori, yakni fakta dan opini. Fakta merupakan ekspresi objektif mengenai sesuatu. Sementara opini adalah ekspresi subjektif yang menggambarkan sentimen, perasaan, maupun penghargaan terhadap suatu hal.

2.2 Term Frequency – Inverse Document Frequency (TF-IDF)

TF-IDF merupakan metode untuk menghitung bobot setiap kata yang paling umum digunakan pada information retrieval. Metode ini juga terkenal efisien, mudah dan memiliki hasil yang akurat [11]. Metode TF-IDF ini akan menghitung nilai Term Frequency (TF) dan Inverse Document Frequency (IDF) pada setiap kata (token) di setiap dokumen dalam korpus atau word cloud. Term frequency (tf) adalah frekuensi kemunculan sebuah term dalam dokumen yang bersangkutan, sedangkan Inverse document frequency (idf) adalah suatu statistik yang mengkarakteristikan sebuah term dalam keseluruhan koleksi dokumen.

2.3 Backpropagation Neural Network

Algoritma backpropagation termasuk metode pelatihan supervised learning yang telah terdapat pasangan input target serta di desain untuk operasi multi-layer perceptron. Sesuai dengan namanya, algoritma ini melakukan 2 tahap perhitungan, yaitu perhitungan maju untuk menghitung error antara keluaran aktual dan target, dan perhitungan mundur yang mempropagasikan balik error tersebut untuk memperbaiki bobot-bobot pada semua neuron yang ada.

3. Pembahasan

3.1 Gambaran Umum Sistem

Proses dari sistem akan berjalan, dimulai dari Komentar pada Twitter akan diambil dengan menggunakan library python GetOldTweets3 format data yang diambil adalah format JSON (Java Script Object Notation) dan kemudian diubah menjadi bentuk CSV. Setelah proses pengambilan data kemudian masuk ke tahap Pre-Processing agar data bisa di klasifikasikan. Selanjutnya dilakukan proses pembobotan dengan TF-IDF dimana setiap kata dijadikan ke bentuk vector. Setelah setiap kata diberikan bobotnya, maka masuk ke proses klasifikasi dengan metode Back Propagation Neural Network. Hasil dari klasifikasi akan berbentuk prediksi tentang komentar pengguna twitter apakah bernilai sentiment positif ataupun negative, dan juga akan menggambarkan berapa persentasi positif dan negatif.

3.2 Perancangan Sistem

Sistem pendeteksi ujaran yang dibuat dalam penelitian tugas akhir ini adalah sistem analisis sentiment twitter berbasis web. Dan data tersebut sudah di verifikasi oleh **Balai Bahasa Provinsi Jawa Barat** total data sebanyak 1400 data. Perancangan sistem Aplikasi Deteksi Sentimen Positif dan Negatif dalam bentuk dokumentasi *unified modeling language* yang terdiri dari *use case diagram*, *sequence diagram*, dan *class diagram*.

4. Implementasi dan Pengujian

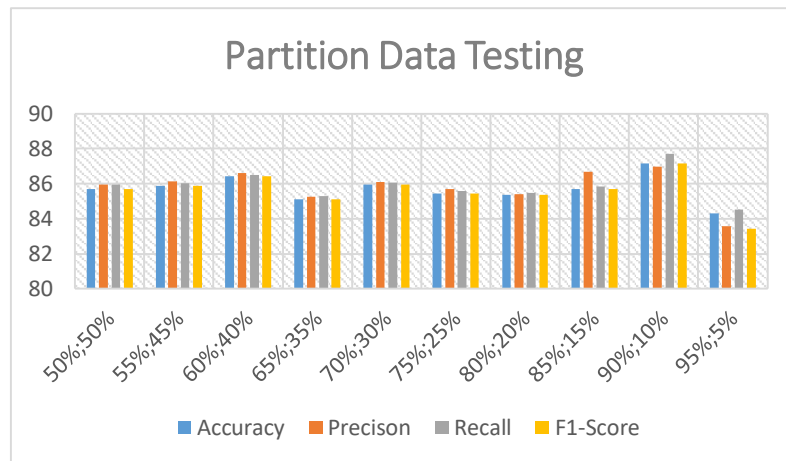
4.1 Pengujian Partisi Data

Pengujian ini bertujuan untuk mengetahui partisi data mana yang terbaik. Partisi dilakukan sebanyak 10 kali dengan membagi data uji dan data latih dengan porsi yang berbeda.

Tabel 4.1 Hasil Pengujian Partisi Data

Pengujian ke-	Data Latih	Data Uji	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	50%	50%	85.71	85.94	85.94	85.71
2	55%	45%	85.87	86.12	86.04	85.87
3	60%	40%	86.43	86.61	86.51	86.42
4	65%	35%	85.10	85.27	85.31	85,10
5	70%	30%	85.95	86.11	86.07	85.95
6	75%	25%	85.43	85.59	85.58	85.43
7	80%	20%	85.36	85.42	85.47	85.36
8	85%	15%	85.71	86.69	85.84	85.69
9	90%	10%	87.14	86.97	87.71	87.14
10	95%	5%	84.29	83.58	84.52	83.42

Pada gambar diatas dapat dilihat bahwa pengujian partisi pada urutan kesembilan dengan partisi 90% data latih dan 10% data uji mendapatkan tingkat accuracy 87,14%. Berikut ini merupakan grafik dari hasil partisi data secara keseluruhan:



Gambar 4. 1 Grafik Pengujian Partisi Data

Pengujian partisi pada urutan kesembilan dengan partisi 90% data latih dan 10% data uji mendapatkan tingkat accuracy 87,14%, precision 86,97%, recall 87,71% dan F1 score 87,14% tertinggi dibandingkan porsi partisi data yang lain.

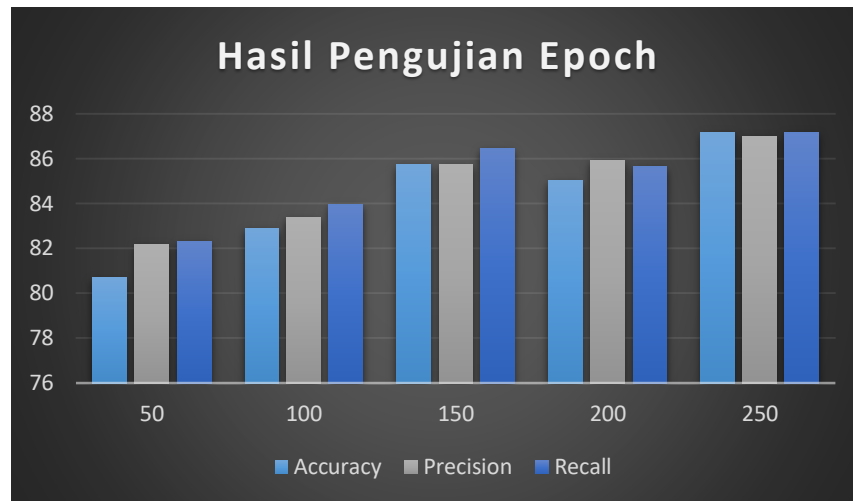
4.2 Pengujian Epoch

Pada pengujian Epoch akan menggunakan nilai Accuracy, Precision, Recall dan F-1 Score tertinggi. Pada pengujian ini juga menggunakan parameter learning rate sebesar 0.01.

Tabel 4.2 Hasil Pengujian Epoch

Pengujian ke-	Learning Rate	Epoch	Accuracy (%)	Precision (%)	Recall (%)
1	0,01	50	80,71	82,16	82,29
2		100	82,86	83,37	83,96
3		150	85,71	85,71	86,46
4		200	85	85,91	85,63
5		250	87,14	86,97	87,14

Berdasarkan hasil pengujian epoch selama 5 kali pengujian didapatkan hasil terbaik yaitu pada epoch = 250. Berdasarkan tabel pengujian epoch diatas, dapat dilihat grafik hasil pengujian epoch dibawah ini:

Gambar 4. 2 Grafik Hasil Pengujian *Epoch*

Pada gambar 4.3 diatas dapat dilihat pada grafik hasil pengujian epoch. Dari hasil pengujian dapat disimpulkan jika semakin besar epoch maka akan semakin bagus. Oleh karena itu maka pada metode Backpropagation Neural Network nilai epoch yang terbaik adalah 250 karena mendapatkan accuracy = 87.14%, precision = 86.97% dan recall = 87.14%.

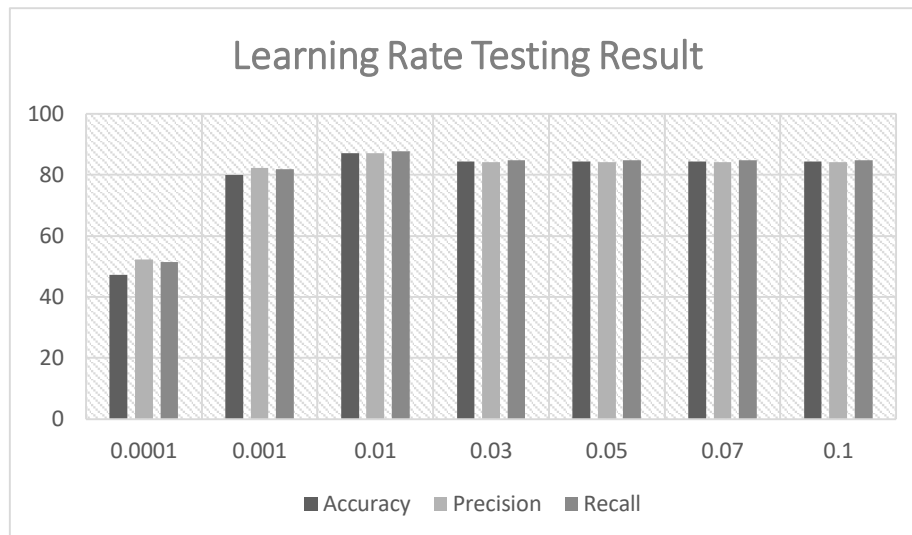
4.3 Pengujian *Learning Rate*

Pada pengujian *Learning Rate* akan menggunakan nilai Accuracy, Precision, Recall dan F-1 Score terbaik pada saat pengujian partisi data. Pada pengujian ini juga menggunakan epoch sebanyak 250.

Tabel 4.3 Hasil Pengujian *learning rate*

Pengujian ke-	Epoch	<i>Learning rate</i>	Accuracy (%)	Precision (%)	Recall (%)
1	250	0.0001	47,14	52,23	51,46
2		0.001	80	82,17	81,88
3		0.01	87,14	86,97	87,71
4		0.03	84,29	84,11	84,79
5		0.05	84,29	84,11	84,79
6		0.07	84,29	84,11	84,79
7		0.1	84,40	84,15	84,69

Grafik hasil pengujian *learning rate* bisa dilihat dibawah ini:



Gambar 4. 3 Grafik Pengujian Learning Rate

Pada gambar 4.3 dapat dilihat grafik hasil dari pengujian menggunakan parameter *learning rate*. Jika kita amati dari grafik diatas bisa disimpulkan jika semakin kecil *learning rate* maka *accuracy*, *precision*, *recall* dan *F1 score* akan memperoleh nilai yang tertinggi. Penggunaan *learning rate* untuk menentukan seberapa cepat sistem belajar dari data latih yang diperoleh maka dapat dilihat jika *learning rate* yang kecil bisa membuat sistem pada performa terbaik.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisa yang telah dilakukan dari Tugas Akhir ini, maka dapat ditarik kesimpulan sebagai berikut:

1. Sistem analisis sentimen dalam Bahasa Indonesia pada tweet pengguna BPJS Kesehatan menggunakan metode klasifikasi Back Propagation Neural Network berbasis web berhasil mengklasifikasikan komentar Twitter pengguna BPJS Kesehatan dengan hasil 49.10% sentiment negatif dan 50,90% positif.
2. Performansi kinerja sistem BPJS Kesehatan dengan metode klasifikasi Back Propagation Neural Network berdasarkan pengujian alpha didapatkan rata-rata precision sebesar 86.97%, recall sebesar 87.14%, dan Accuracy sebesar 87,14%.
3. Berdasarkan Pengujian Beta, web analisis sentiment twitter dengan Metode Backpropagation Neural Network dinyatakan valid dan reliabilitas, sehingga pengguna dapat mengetahui hasil analisis sentiment.

5.2 Saran

1. Hasil dari tugas akhir ini dirasa kurang maksimal, adapun beberapa pengembangan yang harus terus dilakukan dalam penelitian ini diantaranya:
2. Menambahkan dataset agar kalimat yang bisa dikombinasikan semakin bervariasi dan sudah diverifikasi oleh Balai Bahasa.
3. Menambahkan clustering agar mampu mengelompokkan pembahasan.

DAFTAR PUSTAKA

- [1] T. Davidson, D. Warmlesley, M. Macy, and I. Weber, "Automated Hate Speech Detection and the Problem of Offensive Language," no. Icwsml, pp. 512–515, 2017.
- [2] Hukum Online, "Sanksi Pidana Penghinaan di Sosial Media." [Online]. Available: <https://www.hukumonline.com/klinik/detail/ulasan/lt56d7218a32d8f/sanksi-bagi-pem-bully-di-media-sosial/>.
- [3] Aliandu, 2012, Analisis Sentimen Tweet Berbahasa Indonesia di Twitter, Tesis, Fakultas MIPA, Pasca Sarjana Ilmu Komputer, Universitas Gadjah Mada, Yogyakarta.
- [4] Maulana Aziz A., and Saniati, "Analisis Sentimen Twitter Menggunakan Backpropagation Neural Network," Jurnal TEKNOINFO, pp. 23 - 27, 2016.
- [5] Zhang, X.F, Huang, H.Y, Zhang K.L. 2009. KNN Text Categorization Algorithm Based on Semantic Centre. 2009 International Conference on Information Technology and Computer Science
- [6] D. Suyanto, Data Mining Untuk Klasifikasi dan Klasterisasi Data. Penerbit Informatika Bandung.
- [7] DataScienceIndonesia, " Perbedaan Supervised and Unsupervised Learning" [Online]. Available: <https://datascience.or.id/article/Perbedaan-Supervised-and-Unsupervised-Learning-5a8fa6e6>.
- [8] W. Koehrsen, "Beyond Accuracy: Precision and Recall Choosing the right metrics for classification tasks." [Online]. Available: <https://towardsdatascience.com/beyond-accuracy-precision-and-recall-3da06bea9f6c>.
- [9] Toward DS, "Sentiment Analysis: Concept, Analysis and Applications." [Online]. Available: <https://towardsdatascience.com/sentiment-analysis-concept-analysis-and-applications-6c94d6f58c17>.
- [10] C. Yi, "An English POS Tagging Approach Based on Maximum Entropy," 2015 Int. Conf. Intell. Transp. Big Data Smart City, pp. 81–84, 2015.
- [11] A. F. Wicaksono and P. Ayu, "HMM Based Part-of-Speech Tagger for Bahasa Indonesia," Proceedings 4th Int. MALINDO (Malay Indones. Lang. Work., no. January 2010, pp. 1–7, 2010.
- [12] Media Ekrut, "Mengenal sentiment analysis dan perannya dalam data media sosial." [Online]. Available: <https://www.ekrut.com/media/sentiment-analysis-adalah/>.
- [13] Han Jiawei, Kamber Micheline, Pei Jian, "DATA MINING Concepts and Techniques", pp. 8, 2016.