Analisis Sentimen Stakeholder atas Layanan haiDJPb pada Media Sosial Twitter Dengan Menggunakan Metode Support Vector Machine dan Naïve Bayes

[1]Muhammad Luthfiy Kurniawan Harsono, [2]Yuris Alkhalifi, [3]Nurajijah, [4]Windu Gata
[1,2,3,4]Magister Ilmu Komputer/ STMIK Nusa Mandiri Jakarta
[1,2,3,4]Jl. Kramat Raya Nomor 18, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450 Indonesia
email : 14002324@nusamandiri.ac.id, [2]14002360@nusamandiri.ac.id, [3]nurajijah.nja@nusamandiri.ac.id, [4]windugata@nusamandiri.ac.id

ABSTRACT
Creation of a special account on social media Instagram and Twitter that aims to accommodate the delivery of questions, input, criticism and suggestions from stakeholders around the ongoing business processes and the use of applications in the framework of budget planning, commitment making, disbursement of the state budget, accounting receipts and financial reporting based on the fact that social media cannot be separated from community activities because the presence of digital devices (smartphones) and affordable internet access makes various groups of people able to obtain information quickly and easily. Organizations have an interest in getting benchmarks for services that have been provided in order to improve the quality of services going forward based on tweets data obtained from Twitter social media. This study discusses the process of collecting and processing tweet data on the @haiDJPb account in order to perform stakeholder sentiment analysis of haiDJPb services on twitter social media using Support Vector Machine and Naïve Bayes algorithms and the accuracy of the Support Vector Machine algorithm is 74.55 % and 77.18% for the Naïve Bayes algorithm.

Keywords **–** Naïve Bayes Algoritm , Support Vector Machine Algoritm, haiDJPb, Twitter

ABSTRAK
Pembuatan akun khusus pada media sosial instagram dan twitter yang bertujuan untuk menampung penyampaian pertanyaan, masukan, kritik dan saran dari stakeholder seputar proses bisnis berjalan serta penggunaan aplikasi dalam rangka perencanaan anggaran, pembuatan komitmen, pencairan APBN, pembukuan penerimaan serta pelaporan keuangan didasarkan pada fakta bahwa media sosial tidak dapat dipisahkan dari aktivitas masyarakat dikarenakan keberadaan perangkat digital (smartphone) dan akses internet yang terjangkau membuat berbagai kalangan masyarakat dapat memperoleh informasi dengan cepat dan mudah. Organisasi mempunyai kepentingan untuk mendapatkan tolak ukur atas layanan yang telah diberikan demi peningkatan kualitas layanan kedepannya berdasarkan data tweets yang didapatkan dari media sosial twitter. Penelitian kali ini membahas tentang proses pengumpulan dan pengolahan data tweet pada akun @haiDJPb dalam rangka melakukan analisis sentimen stakeholder atas layanan haiDJPb pada media soasial twitter menggunakan algoritma Support Vector Machine dan Naïve Bayes dan didapatkan hasil nilai akurasi untuk algoritma Support Vector Machine adalah 74,55% dan 77,18% untuk algoritma Naïve Bayes.

Kata Kunci **–** Algoritma Naïve Bayes , Algoritma Support Vector Machine, haiDJPb, Twitter

1. Introduction

Dengan semakin banyaknya proses bisnis yang menjadi inisiatif dan tanggung jawab pada institusi Direktorat Jenderal Perbendaharaan (DJPb), sebagai salah satu eselon 1 pada Kementerian Keuangan, serta semakin banyaknya aplikasi yang digunakan oleh stakeholder dalam rangka perencanaan anggaran, pembuatan komitmen, pencairan APBN, pembukuan penerimaan serta

pelaporan keuangan, maka sebagai konsekuensinya semakin banyak pula pertanyaan, masukan, kritik dan saran atas proses bisnis berjalan serta penggunaan aplikasi yang disediakan.

Pimpinan memandang penting untuk membuat portal khusus sebagai sarana penyaluran aspirasi dari stakeholder dalam rangka menampung pertanyaan, masukan, kritik dan saran atas pelaksanaan pekerjaan yang berhubungan dengan perencanaan anggaran, pembuatan komitmen, pencairan APBN, pembukuan penerimaan serta pelaporan keuangan, baik yang menyangkut proses bisnis berjalan maupun penggunaan aplikasi yang disediakan. Maka lahirlah portal "hai.djpb.go.id" yang pada kemudian hari bertransformasi menjadi "hai.kemenkeu.go.id" untuk mengakomodir pertanyaan seputar proses bisnis berjalan serta penggunaan aplikasi terkait eselon 1 selain DJPb pada Kementerian Keuangan. Penyampaian pertanyaan, masukan, kritik dan saran dapat dilakukan secara langsung melalui akses ke portal dengan memilih menu "buat tiket" maupun menyampaikannya melalui email resmi hai.djpbn@kemenkeu.go.id.

Pada pelaksanaannya, akses ke portal maupun berkirim email dipandang belum memadai. Untuk itu dibuatlah akun khusus pada media sosial Instagram "hai.djpb" dan twitter "@haiDJPb". Pembuatan akun pada media sosial ini tidak terlepas dari kebiasaan masyarakat kita yang tidak dapat dipisahkan dari aktivitas di media sosial dikarenakan keberadaan perangkat digital (smartphone) dan akses internet yang terjangkau membuat berbagai kalangan masyarakat dapat memperoleh informasi dengan cepat dan mudah.[1]

Media sosial twitter dipilih dikarenakan pada tahun 2010, pengguna twitter di seluruh dunia mencapai 106 juta pengguna dan meningkat sebanyak 300 ribu pengguna setiap harinya. Twitter mendapatkan lebih dari 3 juta request setiap hari. Indonesia menduduki peringkat ke-8 dalam mengakses twitter atau sebesar 2% dari lalu lintas situs di seluruh dunia. Pengguna twitter mengirimkan tweet sebanyak 55 juta pesan setiap harinya.[2]

Organisasi berkepentingan untuk mendapatkan informasi/ feedback dari stakeholder atas layanan yang diberikan melalui penggunaan media sosial sebagai sarana menampung pertanyaan, masukan, kritik dan saran dari stakeholder,. Pemberian label "positif" dan "negatif" yang dilakukan oleh pihak yang memiliki reputasi di bidangnya dapat membantu institusi untuk mendapatkan tolak ukur atas layanan yang telah diberikan serta membantu dalam peningkatan kualitas layanan kedepannya. Dalam pelaksanaannya, pengelompokan feedback untuk pemberian label sulit dilakukan jika hanya mengandalkan quey biasa. Query adalah standard query language untuk mendefinisikan dan memanipulasi database yang didukung oleh database server.[3]

Pemilihan query yang kurang spesifik akan menghasilkan pencarian yang tidak relevan. Hasil pelabelan pada peringkat awal belum tentu relevan, sehingga dapat dinyatakan pencarian dengan query biasa tidaklah efektif. Jadi dibutuhkan pengelompokan untuk mengatasi kendala tersebut. Permasalahan yang muncul adalah bagaimana sistem dapat melakukan pengelompokan yang relevan untuk memenuhi kebutuhan pengguna.[3]

Oleh karena itu, akan dibuat klasifikasi dengan menggunakan metode Support Vector Machine (SVM) dan Naïve Bayes (NB). Diharapkan dengan dukungan dari dua metode tersebut, dapat membantu institusi dalam melakukan penentuan feedback berdasarkan tingkat kesamaan.

Berikut adalah beberapa batasan masalah dari sistem yang dibuat. Feedback yang dikumpulkan bersumber dari cuitan di media sosial twitter dengan kata kunci "haiDJPb" dan diproses menggunakan tools RapidMiner. Proses transformasi teks menggunakan metode Transform Cases, Remove http, Remove @, Tokenize, Filter Tokens (by length) dan Filter Stopwords (Dictionary). Pembobotan token menggunakan algoritma TF-IDF. Bentuk masukan sistem berupa file dengan format excel yang bersumber dari proses web crawling ke media sosial twitter. Bentuk keluaran sistem adalah nilai accuracy, precission dan recall dari Performace Vector (Performance-SVM) dan Performance Vector (Performance-NB).

Penelitian serupa sudah pernah dilakukan sebelumnya, diantaranya yang dilakukan oleh Agnes Rossi Trisna Lestari, Rizal Setya Perdana, M. Ali Fauzi, mahasiswa Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya. Mereka melakukan pengklasifikasian opini berupa teks sentimen pada dokumen twitter berbahasa Indonesia untuk memprediksi kekuatan cagub dan cawagub pada Pilkada DKI tahun 2017.[4] Penelitian lainnya tentang analisis kepribadian seseorang berdasarkan tweets atau posting yang ada pada twitter yang dilakukan oleh Mohammad Zoqi Sarwani dan Wayan Firdaus Mahmudi. Proses analisis dilakukan dengan mengklasifikasikan

tweets kedalam 4 kelas, yaitu Guardian, Artisan, Idealist dan Rasional yang kemudian data yang didapat diolah dengan menggunakan algoritma Naïve Bayes Classifer (NBC).[5]

Penelitian yang sama juga dilakukan oleh Yessivha Imanuela Claudy, Rizal Setya Perdana dan M. Ali Fauzi. Perbedaan dari penelitian sebelumnya yaitu terletak pada tujuan pelaksanaan tes adalah untuk mengetahui karakter calon karyawan dan algoritma yang digunakan adalah K-Nearest Neighbor (KNN).[6] Sedangkan penelitian yang dilakukan oleh Windu Gata dan Purnomo memanfaatkan text mining untuk melakukan pengecekan terhadap penyalahgunaan aplikasi atau pengiriman SMS pada LKBN ANTARA menggunakan algoritma K-Nearest Neighbor.[7] Sementara penelitian yang dilakukan oleh Rizal Setya Perdana dan M. Ali Fauzi adalah melakukan analisis sentimen tingkat kepuasan pengguna penyedia layanan telekomunikasi seluler Indonesia pada twitter dengan metode Support Vector Machine dan Lexicon Based Features dimana hasilnya adalah metode tersebut baik digunakan untuk klasifikasi analisis sentimen.[8]

Penelitian lainnya dilakukan oleh Windu Gata, Hasan Basri, Rais Hidayat, Yuyun Elizabeth Patras, Baharuddin Baharuddin, Rhini Fatmasari, Siswanto Tohari dan Nia Kusuma Wardhani tentang penerapan algoritma Naïve Bayes, Random Forest, C4.5 pada game online untuk memprediksi prestasi belajar dimana hasilnya adalah algoritma Naïve Bayes menghasilkan akurasi yang lebih tinggi dibandingkan dengan algoritma Random Forest dan C4.5 dalam memprediksi hubungan antara pelajar bermain game online dengan prestasi belajarnya.[9]

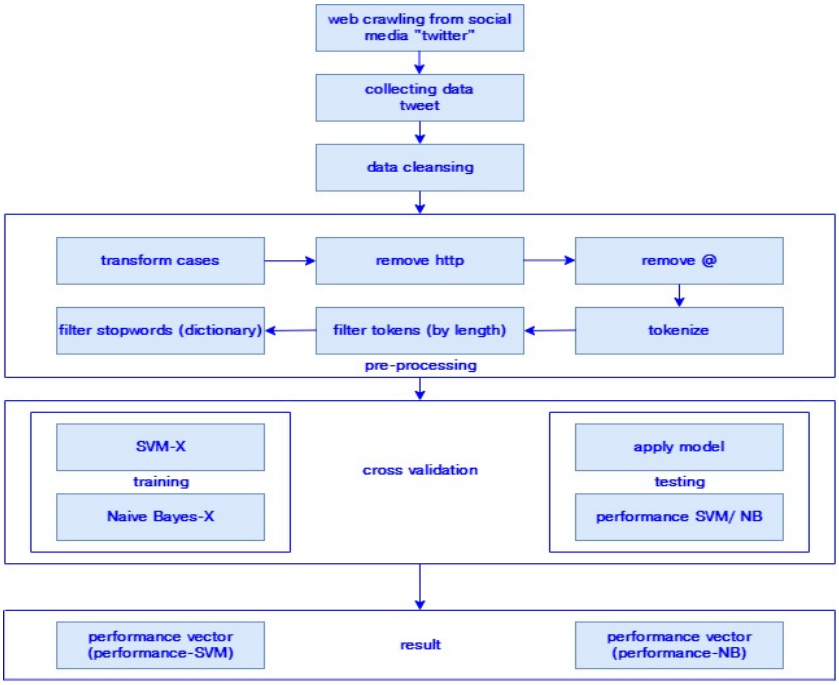
Sementara penelitian yang dilakukan oleh Deden Rustiana dan Nina Rahayu memanfaatkan tweet twitter untuk menganalisis sentimen pasar otomotif mobil dan membandingkannya dengan data pencarian merk mobil terlaris dari GAIKINDO dan mengolahnya menggunakan algoritma Naïve Bayes dan menghasilkan analisis dengan akurasi yang tinggi.[10] Penelitian lainnya tentang implementasi K-Nearest Neighbor dalam pengklasifikasian follower twitter yang menggunakan Bahasa Indonesia yang dilakukan oleh Muhammad Rivki dan Mukhairil Bachtiar dimana mereka mengklasifikasikan follower dari pengguna twitter berdasarkan ketertarikannya dengan menggunakan text mining dan mengolah data yang didapatkan dengan menggunakan algoritma K-Nearest Neighbor untuk melakukan promosi di twitter.[11]

Penelitian serupa dilakukan oleh Winda Estu Nurjanah, Rizal Setya Perdana dan Mochammad Ali Fauzi tentang analisis sentimen terhadap tayangan televisi berdasarkan opini masyarakat pada media sosial twitter menggunakan metode K-Nearest Neighbor dan pembobotan jumlah retweet dimana hasilnya adalah tingkat akurasi menggunakan pembobotan tekstual adalah 82,50% sedangkan pembobotan non tekstual adalah 60% dan penggabungan keduanya mencapai 83,33%.[12] Penelitian lainnya yang menjadi referensi penulis adalah analisis kecenderungan informasi dengan menggunakan metode text mining (studi kasus: akun twitter @detikcom) yang dilakukan oleh Syaifudin Karyadi, Hasbi Yasin dan Moch. Abdul Mukid dimana mereka mengelompokkan 500 tweet dari akun twitter @detikcom menggunakan metode K-Means Clustering untuk mengetahui kecenderungan topik pemberitaan dan mengetahui topik yang paling sering muncul.[13]

Tujuan dari penelitian ini adalah untuk mendapatkan tolak ukur atas layanan yang telah diberikan serta membantu dalam peningkatan kualitas layanan kedepannya bagi organisasi berdasarkan data tweets yang didapatkan dari media sosial twitter serta membandingkan penggunaan algoritma Support Vector Machine dan Naïve Bayes dalam melakukan proses pengolahan data.

2. Research Method

Metode penelitian yang akan digunakan pada penelitian ini sebagaimana ditunjukan pada gambar 1 dibawah ini :



Gambar 1. Metodologi Penelitian Sentimen Stakeholder atas layanan HaiDJPb Pada Media Sosial Twitter dengan Menggunakan Metode Support Vector Machine dan Naïve bayes.

Proses diawali dari pengumpulan cuitan di media sosial twitter dengan menggunakan metode web crawling dengan kata kunci "haiDJPb". Sebelum data diproses menggunakan tools RapidMiner, dilakukan proses cleansing terlebih dahulu dengan menghilangkan duplikasi dan retweet (RT). Proses dilanjutkan dengan pre-processing menggunakan metode Transform Cases, Remove http, Remove @, Tokenize, Filter Tokens (by length) dan Filter Stopwords (Dictionary). Pembobotan token menggunakan algoritma TF-IDF. Proses selanjutnya adalah cross validation menggunakan algortima Support Vector Machine dan Naïve Bayes. Bentuk keluaran sistem adalah nilai accuracy, precission dan recall dari Performace Vector (Performance-SVM) dan Performance Vector (Performance-NB).

3. Result and Analysis

3.1 Pengumpulan Data

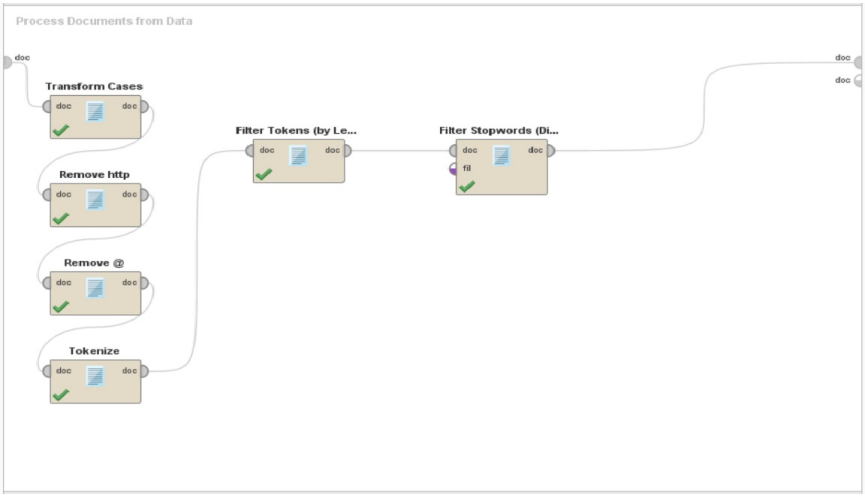
Proses dimulai dengan *web crawling* menggunakan *tools* RapidMiner 9.6 untuk mendapatkan data *tweet* dari media sosial twitter (*data collection*) dengan parameter *query* "haidjpb" dan didapatkan 146 tweet.

Gambar 2. Proses Pengumpulan Data dari Twitter

3.2 *Data Cleansing*

Setelah mendapatkan data *tweet*, langkah pertama adalah dilakukan *data cleansing* terlebih dahulu dengan tujuan untuk menghilangkan duplikasi yang terjadi. Dari 146 *tweet,* 12 *tweet* termasuk duplikasi dan dilakukan penghapusan sehingga menyisakan *134 tweet*. Proses dilanjutkan dengan menghilangkan *retweet* (RT) pada data *tweet*.

3.3 *Preprocessing*



Gambar 3. Tahap *Preprocessing*

3.3.1 *Case Folding*

Proses dilanjutkan dengan tahap *pre-processing* menggunakan aplikasi RapidMiner 9.6 dengan parameter *vector creation* "TF-IDF" agar data siap untuk diproses. Pada tahap *pre-processing*, dilakukan proses *transform cases* (*to Lower case*) untuk mengubah teks menjadi huruf kecil semua.

3.3.2 *Remove Url*

Proses dilanjutkan dengan *remove* http untuk menghilangkan "http" pada teks. Proses ini dilakukan karena sering munculnya *URL* pada data twitter yang membuat data tidak efektif dan tidak memiliki arti. Penyebab munculnya alamat web atau URL ini adalah karena banyaknya pengguna mempromosikan situs tertentu agar pengguna lain langsung bisa masuk pada halaman *web* yang dimaksud.

3.3.3 *Remove Username*

Bertujuan menghapus nama pengguna twitter yang diawali dengan simbol "@" karena dianggap tidak penting dan perlu dihilangkan.

3.3.4 *Remove Punctuation* Dan *Clean Number*

Proses dilanjutkan dengan menghapus semua karakter non alphabet, misalnya simbol, spasi, angka dan lain-lain menggunakan fungsi *tokenize* pada RapidMiner. Meskipun dalam penulisan komentar selalu menyertakan sebuah angka di setiap awal atau akhir kalimat untuk menunjukkan bahwa kalimat tersebut diulang-ulang maka dalam Bahasa Indonesia yang baik itu merupakan hal yang salah. Begitu juga pada sebuah penelitian, apabila menemukan sebuah kata yang menggunakan tambahan angka maka perlu dihapus.[14]

3.3.5 *Clean One Character Or More*

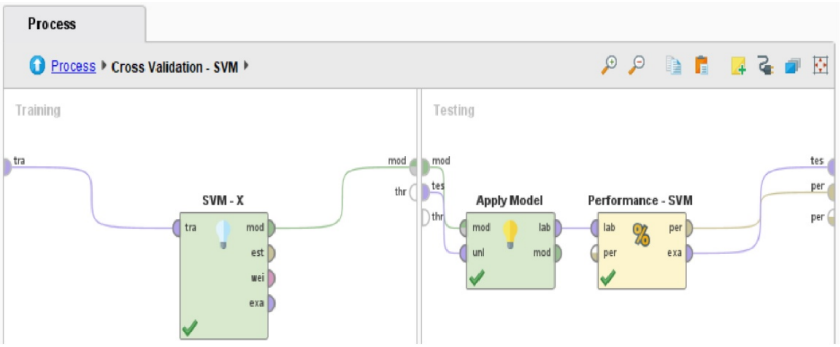
Proses selanjutnya adalah menghapus kata dengan jumlah huruf tertentu melalui fungsi *filter tokens* (*by length*) dengan parameter *min chars* 4 dan *max chars* 25 untuk membatasi jumlah huruf pada kata minimal 4 dan maksimal 25 pada teks.

3.3.6 *Remove Stopword*

Stopword diproses pada sebuah kalimat jika mengandung kata-kata yang sering keluar dan dianggap tidak penting seperti waktu, penghubung dan lain sebagainya. Untuk itu perlu dilakukan penghapusan. Untuk melakukan proses penghapusan kata ini diperlukan sebuah data atau daftar kata yang diinginkan untuk dihapus.[14]

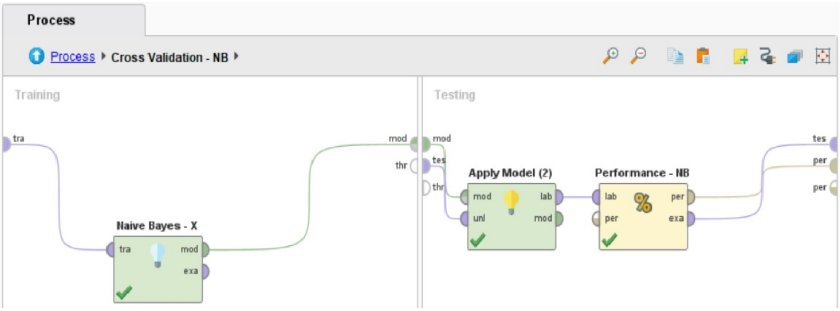
3.4 *Cross Validation*

Langkah berikutnya dilakuakan proses *cross validation* menggunakan algoritma SVM dan NB yang meliputi proses *training* dan *testing* (*apply model* dan *performance*-SVM/ *performance*-NB). Setelah itu dilakukan proses pemodelan dengan menggunakan Algoritma SVM dan NB untuk mendapatkan nilai *accuracy, precision* dan *recall* berdasarkan *performance vector* (*performance*-SVM) dan *performance vector* (*performance*-NB).



Gambar 4. *Cross Validation Process*-SVM

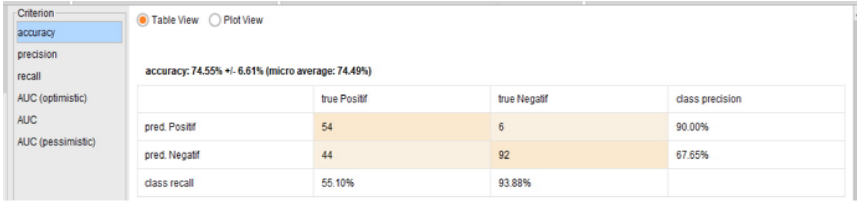
Terdiri dari 2 tahap, yaitu tahap *training* dan *testing*. Pada tahap *training*, data dilatih menggunakan algoritma SVM. Kemudian pada tahap *testing*, data diproses melalui *Apply Model* dan *Performance Test*.



Gambar 5. *Cross Validation Process*-NB

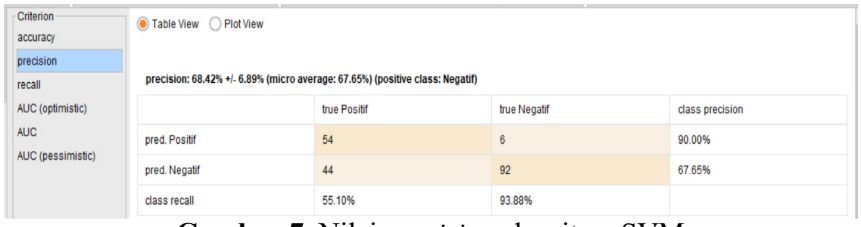
Terdiri dari 2 tahap, yaitu tahap *training* dan *testing*. Pada tahap *training*, data dilatih menggunakan algoritma NB. Kemudian pada tahap *testing*, data diproses melalui *Apply Model* dan *Performance Test*.

Dari proses *cross validation* yaitu *training* dan *testing*, algoritma SVM menghasilkan nilai *accuracy* 74,55%, *precision* 68,42% dan *recall* 93,89%.



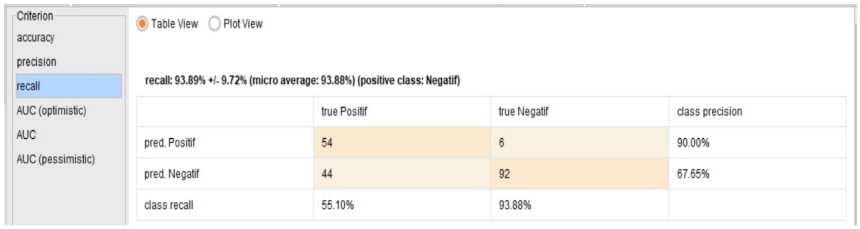
Gambar 6. Nilai *accuracy* algoritma SVM

Nilai *accuracy* yang didapatkan algoritma SVM adalah 74,55% dengan margin +/- 6,61% dengan nilai rata-rata mikro sebesar 74,49%. Dengan demikian nilai *accuracy* maksimal adalah 81,16% dan nilai *accuracy* minimal adalah 67,94%.



Gambar 7. Nilai *precision* algoritma SVM

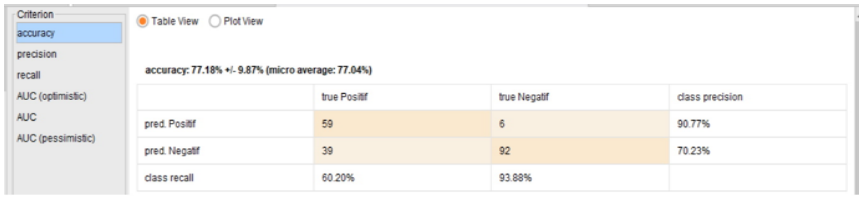
Nilai *precision* yang didapatkan algoritma SVM adalah 68,42% dengan margin +/- 6,89% dengan nilai rata-rata mikro sebesar 67,65% (kelas positif: negatif). Dengan demikian nilai *precision* maksimal adalah 75,31% dan nilai *precision* minimal adalah 61,53%.



Gambar 8. Nilai *recall* algoritma SVM

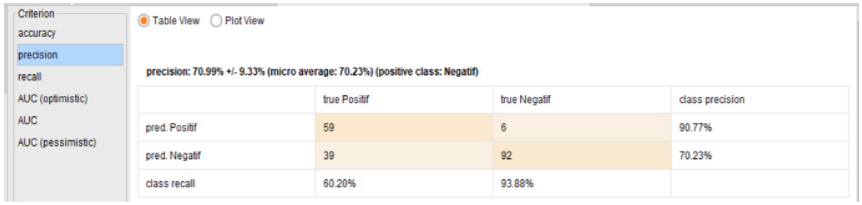
Nilai *recall* yang didapatkan algoritma SVM adalah 93,89% dengan margin +/- 9,72% dengan nilai rata-rata mikro sebesar 93,88% (kelas positif: negatif). Dengan demikian nilai *recall* maksimal adalah 103,61% dan nilai *recall* minimal adalah 84,17%.

Sedangkan proses *cross validation* pada algoritma NB menghasilkan nilai *accuracy* 77,18%, *precision* 70,99% dan *recall* 93,89%.



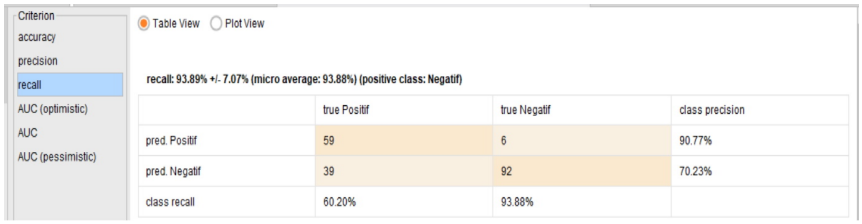
Gambar 9. Nilai *accuracy* algoritma NB

Nilai *accuracy* yang didapatkan algoritma NB adalah 77,18% dengan margin +/- 9,87% dengan nilai rata-rata mikro sebesar 77,04%. Dengan demikian nilai *accuracy* maksimal adalah 87,05% dan nilai *accuracy* minimal adalah 67,31%.



Gambar 10. Nilai *precision* algoritma NB

Nilai *precision* yang didapatkan algoritma NB adalah 70,99% dengan margin +/- 9,33% dengan nilai rata-rata mikro sebesar 70,23% (kelas positif: negatif). Dengan demikian nilai *precision* maksimal adalah 80,32% dan nilai *precision* minimal adalah 61,66%.



Gambar 11. Nilai *recall* algoritma NB

Nilai *recall* yang didapatkan algoritma NB adalah 93,89% dengan margin +/- 7,07% dengan nilai rata-rata mikro sebesar 93,88% (kelas positif: negatif). Dengan demikian nilai *recall* maksimal adalah 100,96% dan nilai *recall* minimal adalah 86,82%.

Berdasarkan analisis dan pengujian yang dilakukan, maka didapatkan hasil sebagai berikut:

1. Dari 134 tweet yang dikumpulkan, 98 tweet memiliki label positif dan 36 tweet memiliki label negatif.
2. Implementasi algoritma SVM pada penelitian ini menghasilkan nilai *accuracy* 74,55%, *precision* 68,42% dan *recall* 93,89%.
3. Implementasi algoritma NB pada penelitian ini menghasilkan nilai nilai *accuracy* 77,18%, *precision* 70,99% dan *recall* 93,89%.
4. Nilai *accuracy* dan *precision* pada algoritma NB lebih tinggi dibandingkan dengan algoritma SVM. Sementara nilai *recall* pada algoritma SVM sama dengan algoritma NB, akan tetapi margin pada algoritma SVM lebih besar daripada algoritma NB yaitu sebesar +/- 9,72% dibandingkan dengan +/- 7,07% sehingga nilai *recall* maksimal untuk algoritma SVM adalah 103,61% sedangkan pada algoritma NB adalah sebesar 100,96%.

Hasil penelitian akan semakin akurat apabila data latih yang digunakan dalam pembelajaran berjumlah lebih banyak, akan tetapi dapat juga mengurangi keakuratan apabila kata-kata yang terdapat pada data *tweet* yang dikumpulkan bias atau bermakna ganda.

4. Conclusion

Berdasarkan analisis dan pengujian yang dilakukan, maka kesimpulan yang dapat diambil adalah feedback atas layanan haiDJPb yang didapatkan dari tweet stakeholder telah menunjukkan hasil yang memuaskan. Masih terdapatnya tweet dengan label negatif menunjukkan bahwa kinerja organisasi masih harus ditingkatkan lagi. Sementara penggunaan algoritma SVM dan NB dalam proses pengolahan data menunjukkan hasil yang berbeda dimana nilai accuracy dan precision algortima NB lebih tinggi dibandingkan dengan algoritma SVM, sementara nilai recall kedua algoritma sama.

References

[1] "Fenomena Media Sosial Dalam Penyebaran Informasi." [Online]. Available: https://kulonprogokab.go.id/v3/portal/web/view_berita/4586/Fenomena-Media-Sosial-Dalam-Penyebaran-Informasi. [Accessed: 24-Apr-2020].

[2] "Infographic: Twitter Statistics, Facts & Figures | Digital Buzz Blog." [Online]. Available: http://www.digitalbuzzblog.com/infographic-twitter-statistics-facts-figures/. [Accessed: 28-Apr-2020].

[3] A. Indranandita, B. Susanto, and A. Rahmat, "Sistem Klasifikasi Dan Pencarian Jurnal Dengan Menggunakan Metode Naive Bayes Dan Vector Space Model," J. Inform., vol. 4, no. 2, 2011, doi: 10.21460/inf.2008.42.48.

[4] A. Rossi, T. Lestari, R. Setya Perdana, and M. A. Fauzi, "Analisis Sentimen Tentang Opini Pilkada DKI 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Näive Bayes dan Pembobotan Emoji," J. Pengemb. Teknol. Inf. dan Ilmu Komput., vol. 1, no. 12, pp. 1718–1724, 2017.

[5] M. Z. Sarwani and W. F. Mahmudy, "Analisa Twitter Untuk Mengetahui Karakter," Semin. Nas. Sist. Inf. Indones., no. November, pp. 2–3, 2015.

[6] Y. I. Claudy, R. S. Perdana, and M. A. Fauzi, "Klasifikasi Dokumen Twitter Untuk Mengetahui Karakter Calon Karyawan Menggunakan Algoritme K-Nearest Neighbor (KNN)," J. Pengemb. Teknol. Inf. dan Ilmu Komput., vol. 2, no. 8, pp. 2761–2765, 2018.

[7] W. Gata, "Akurasi Text Mining Menggunakan Algoritma K-Nearest Neighbour pada Data Content Berita SMS," vol. 6, pp. 1–13, 2017.

[8] R. S. Perdana and M. A. Fauzi, "Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter dengan Metode Support Vector Machine dan Lexicon Based Features Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler ," no. October, 2017.
[9] W. Gata et al., "Algorithm Implementations Naïve Bayes, Random Forest. C4.5 on Online Gaming for Learning Achievement Predictions," vol. 258, no. Icream 2018, 2019, doi: 10.2991/icream-18.2019.1.
[10] N. Rahayu, "ANALISIS SENTIMEN PASAR OTOMOTIF MOBIL : TWEET TWITTER MENGGUNAKAN NAÏVE BAYES," vol. 8, no. 1, pp. 113–120, 2017.
[11] M. Rivki, M. Bachtiar, T. Informatika, F. Teknik, U. K. Indonesia, and J. Dipati, "Jurnal Sistem Informasi (Journal of Information Systems). 1 / 13 (2017), 31-37 DOI : http://dx.doi.org/10.21609/jsi.v13i1.500," vol. 13, pp. 31–37, 2017.
[12] W. E. Nurjanah, R. S. Perdana, and M. A. Fauzi, "Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan M," no. October, 2017.
[13] S. Karyadi and H. Yasin, "1 , 2 , 3," vol. 5, pp. 763–770, 2016.
[14] S. Mujilahwati, "PRE-PROCESSING TEXT MINING PADA DATA TWITTER," 2016.