KLASIFIKASI SMS SPAM DENGAN KOMPARASI METODE SVM DAN NAÏVE BAYES

Muhamad Hajat Syafii[1]
[1] Universitas Nusa Mandiri
Jl. Margonda Raya No. 545, Depok, Jawa Barat
[1]14002612@nusamandiri.ac.id

ABSTRACT

Technological developments force most of the population throughout the world, including in Indonesia, to be able to take advantage of these advances. One of the technologies in question is the internet and gadgets. The rapid development of smartphones has not changed the function of one of the service providers, namely the short message service (SMS). SMS is currently still used to send messages to users who already know each other and to people who don't know each other, with many purposes including to offer products or services. It is a problem to find a way to classify incoming SMS as spam SMS or not. SMS classification in this study uses the Naïve Bayes algorithm and Support Vector Machine (SVM). The purpose of this research is to compare and find out which algorithm is better in classifying spam SMS and also to find out which SMS is received as spam or not spam. The results showed that the Naïve Bayes Algorithm has higher accuracy than the SVM algorithm, with a value of 0.94. Precision, recall, f1-score and accuracy values for the Naïve Bayes algorithm also have the best performance compared to the SVM algorithm for SMS Spam classification.

Keywords – SMS Spam Classification, Naive Bayes Algorithm, SVM Algorithm, Spam SMS

I. PENDAHULUAN

Perkembangan teknologi saat ini sangat berperan dalam perkembangan kehidupan manusia. Kondisi ini memaksa sebagian besar penduduk diseluruh dunia termasuk di Indonesia sendiri untuk dapat menggunakan kemajuan tersebut. Salah satu teknologi yang dimaksud adalah internet dan Gadget. Pengguna smartphone di Indonesia berada di urutan keempat dalam daftar dunia setelah Cina, India dan Amerika Serikat, dengan jumlah sebesar 192,15 juta pengguna smartphone pada tahun 2022 [1].

Perkembangan smartphone yang pesat tidak merubah fungsi dari salah satu layanan provider yaitu layanan pesan singkat atau *short message service* (SMS). SMS saat ini masih dipergunakan untuk mengirimkan pesan kepada pengguna yang sudah saling kenal maupun kepada orang yang belum dikenal, dengan banyak tujuan diantaranya untuk menawarkan produk atau jasa. SMS juga saat ini banyak digunakan oleh orang yang tidak bertanggungjawab dalam melakukan tindak kejahatan, kondisi ini berarti tidak di inginkan oleh pengguna smartphone. Dalam kondisi ini maka sms yang masuk merupakan SMS spam dan tidak di inginkan oleh pengguna.

Berdasarkan permasalahan tersebut maka diperlukan cara agar dapat diklasifikasikan SMS yang masuk termasuk SMS spam atau bukan. Dengan demikian diperlukan cara atau teknik untuk mengklasifikasi SMS yang masuk.

Pengklasifikasian SMS dalam penelitian ini menggunakan algoritma Naïve Bayes dan *Support Vector Machine* (SVM). Perbandingan hasil yang di lihat adalah dari hasil precision, recall, dan f-score dengan bahasa pemrograman python dan menggunakan tool Jupyter Notebook.

Tujuan dalam penelitian ini adalah untuk membandingkan dan mengetahui algoritma yang lebih baik dalam pengklasifikasian SMS spam dan juga untuk mengetahui SMS yang diterima spam atau bukan spam.

II. TINJAUAN PUSTAKA

A. *Short Message Service* (SMS) dan Spam

Short Message Service merupakan fasilitas yang digunakan untuk mengirim dan menerima pesan teks ke dan dari telepon selular. SMS adalah protokol komunikasi yang memungkinkan pertukaran pesan teks singkat sesama telepon selular [2].

S*pam* adalah aktivitas mengirimkan pesan secara elektronik kepada orang lain secara terus-menerus, dan dengan jumlah yang banyak tanpa dikehendaki oleh si penerimanya [3].

B. Algoritma Naive Bayes

Naive bayes merupakan bagian dari *Machine Learning* yang memanfaatkan konsep probabilitas dan konsep statistik dalam prediksi atau menemukan variabel baru pada suatu permasalahan. Thomas Bayes adalah orang yang pertama kali mengemukakan Algoritma Naïve Bayes.

Implementasi rumus dari *Teorema Bayes* adalah sebagai berikut [4] :

$$P(A|B) = \frac{P(A)}{P(B)} P(B|A)$$

Keterangan :
A,B : Kejadian
$P(A|B)$: Peluang kejadian A bila B terjadi
$P(B|A)$: probabilitas B yang ditentukan A adalah benar
$P(A), P(B)$: probabilitas bebas A dan B

Teorema Naïve Bayes adalah hasil kombinasi *Teorema Naïve* dan *Teorema Bayes*, dengan demikian diperlukan penambahan pada kelas dari sampel data yang dianalisis, sehingga implementasi rumus dasar *Naive Bayes* menjadi seperti berikut [4]:

$$P(C|F_1, \dots, F_n) = \frac{P(C)P(F_1, \dots, F_n|C)}{P(F_1, \dots, F_n)}$$

Keterangan : Terdapat variabel C yang mewakili suatu kelas, sementara variabel F1....Fn mewakili berbagai karakteristik atau feature yang dibutuhkan dalam proses klasifikasi suatu objek. Kesimpulan dari implementasi rumus dasar *Teorema Naive Bayes* adalah bahwa peluang masuknya suatu sampel data dengan memiliki berbagai karakteristik yang ada dalam suatu kelas C (*Posterior*) merupakan peluang munculnya kelas C (Sebelum masuknya sampe tersebut atau *Prior*), dikalikan dengan peluang kemunculan karakteristik sampel data pada kelas C atau *Likehood*, dibagi dengan peluang kemunculan dari berbagai karakteristik dari suatu data secara global atau *Evidence*.

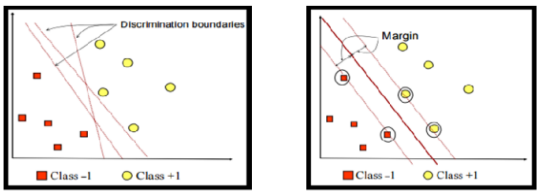
Bentuk sederhana implementasi rumus dasar *Teorema Naive Bayes* [4] :

$$Pasterior = \frac{prior \text{ x } likelihood}{evidence}$$

Nilai *Evidence* selalu tetap pada setiap kelas dalam suatu data. Nilai - nilai yang mewakili *Posterior* yang nantinya akan dilakukan proses perbandingan dengan nilai *Posterior* kelas yang lain untuk menentukan hasil klasifikasi dari suatu sampel data.

C. *Support Vector Machine (SVM)*

Support Vector Machine adalah metode yang digunakan untuk menjalankan proses *Pattern Recognition*, metode ini terbilang bukan metode terbaru lagi dalam *machine learning* namun metode ini sangat penting dan sering digunakan dalam berbagai riset teknologi [5]. SVM merupakan salah satu metode *Machine Learning* yang memiliki prinsip kerja *Structural Risk Minimization* yang bertujuan untuk mencari sebuah *Hyperplane* atau pemisah antar 2 buah kelas pada suatu input space. Berikut contoh implementasi pencarian *Hyperplane* terbaik antar 2 kelas yang berbeda pada suatu *input space* :



Gambar 1. *Hyperplane* pada *Support Vector Machine*

Melalui representasi gambar tersebut dijelaskan bahwa pada gambar kiri terdapat beberapa *pattern* yang terbentuk yang merupakan keanggotaan dari 2 buah kelas yang berbeda yaitu +1 dan -1. Beberapa pattern yang terbentuk yang keanggotaannya oleh kelas -1 ditandai oleh warna merah (kotak), sementara pattern milik kelas +1 ditandai oleh warna kuning(Lingkaran). Suatu proses klasifikasi objek yang menggunakan metode *Support Vector Machine* berusaha untuk mencari sebuah *Hyperplane* atau pemisah berupa garis yang memisahkan antar kelas, berbagai *Hyperplane* yang terbentuk dalam representasi gambar tersebut dinamakan *Discrimination Boundaries*.

Hyperplane atau pemisah yang terbaik dimana memisahkan antar kedua kelompok kelas dapat diukur dengan mengukur margin atau jarak antara *Hyperplane* yang terbentuk dengan pattern terdekat dari kedua kelas yang ada. *Pattern* yang terdekat dengan garis pemisah atau *Hyperplane* disebut dengan *Support Vector*. Pada representasi gambar kanan diperlihatkan bahwa adanya garis solid yang menunjukkan *Hyperplane* terbaik yang terletak pada tengah - tengah antara kedua kelas yang ada. Usaha untuk mencari garis pemisah atau *Hyperplane* antar kedua kelas yang ada ini merupakan inti dari konsep metode *Support Vector Machine*.

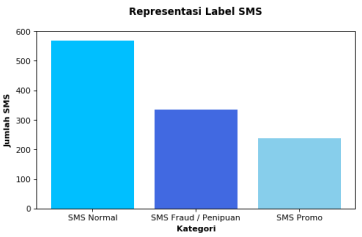
C. Penelitian Terkait

Studi literatur yang relevan dibutuhkan untuk bahan acuan dan pertimbangan dalam penelitian. Beberpaa penelitian yang sudah dilakukan yaitu :

1. Pengujian sistem menggunakan metode Multinomial Naïve Bayes (MNB) pada identifikasi SMS berbahasa Indoensia yang mengandung spam dan tidak mengandung spam mampu melakukan klasifikasi SMS berbahasa Indonesia yang mengandung spam dan tidak mengandung spam [6].
2. Naive Bayes Classifier algorithm dapat dipergunakan sebagai sarana untuk memfiltrasi SMS yang masuk ke dalam inbox pengguna dan dapat memberikan hasil filtrasi yang cukup akurat untuk menyaring SMS yang masuk [7].

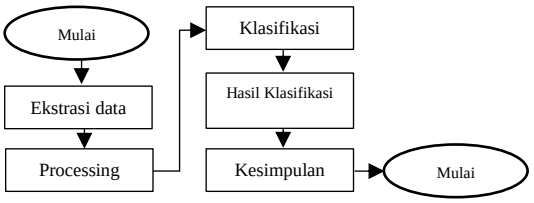
III. METODE PENELITIAN

Data yang digunakan pada penelitian ini merupakan dataset SMS berbahasa Indonesia dengan jumlah 1143 record, yang terdiri dari SMS normal berjumlah 569, SMS penipuan berjumlah 335 dan SMS promosi berjumlah 239.



Gambar 2. Sebaran Tipe Data SMS

Klasifikasi SMS spam dilakukan dengan menggunakan Algoritma Naïve Bayes dan SVM (*Support Vector Machine*). Proses data klasifikasi menggunakan bahasa pemrograman python, dengan bantuan tool jupyter notebook. Alur penelitian digambarkan sebagai berikut :



Gambar 3. Alur Penelitian

Sebelum klasifikasi dilakukan, tahapan menyiapkan dataset dalam fomat CSV file atau *Comma Separated Values File*. Tahap selanjutnya yaitu ektraksi data yang berguna untuk meningkatkan kualitas dataset. Berikutnya adalah tahap processing dataset sebelum tahap klasifikasi data. Klasifikasi data dilakukan dengan implementasikan algoritma untuk mendapatkan nilai dari setip algoritma. Berikutnya adalah membandingkan hasil algoritma terbaik dengan melihat *precision, recall, f1-score* dan *accuracy*.

IV. HASIL DAN PEMBAHASAN

A. Tahap Persiapan dan Klasifikasi

Tahap awal dalam pengolahan data yaitu menyiapkan dataset, dan memastikan bahwa dataset dalam bentik file CSV. Tujuan dilakukannya tahap ini adalah agar data yang digunakan lebih presisi dan memudahkan dalam menjalankan klasifikasi.

Tahap selanjutnya yaitu proses klasifikasi dengan menggunakan tool jupyter notebook. Data dieksplor dengan menerapkan teknik aritmatika dan teknik grafis dalam meringkas dan mempermudah pemahaman data.

	Teks	label
0	[PROMO] Beli paket Flash mulai 1GB di MY TELKO...	2
1	2.5 GB/30 hari hanya Rp 35 Ribu Spesial buat A...	2
2	2016-07-08 11:47:11.Plg Yth, sisa kuota Flash ...	2
3	2016-08-07 11:29:47.Plg Yth, sisa kuota Flash ...	2
4	4.5GB/30 hari hanya Rp 55 Ribu Spesial buat an...	2

Gambar 4. Isi SMS

B. Hasil Perbandingan Algoritma

Pada penelitian ini proses pengolahan data SMS yang dilakukan menggunakan pengkodean dengan tool jupyter notebook. Pengujian yang dilakukan adalah untuk melihat hasil perbandingan dari kedua algoritma yaitu Algoritma Naïve Bayes dan *Support Vector Machine*. Pengujian dilakukan terhadap nilai *precision, recall, f1-score* dan nilai *accuracy*.

	precision	recall	f1-score	support
SMS Normal	1.00	0.95	0.97	140
SMS Fraud / Penipuan	0.88	0.93	0.90	81
SMS Promo	0.88	0.92	0.90	65
accuracy			0.94	286
macro avg	0.92	0.93	0.93	286
weighted avg	0.94	0.94	0.94	286

Gambar 5. Hasil Algoritma Naïve Bayes

	precision	recall	f1-score	support
SMS Normal	0.95	0.99	0.97	140
SMS Fraud / Penipuan	0.90	0.91	0.91	81
SMS Promo	0.93	0.85	0.89	65
accuracy			0.93	286
macro avg	0.93	0.92	0.93	286
weighted avg	0.93	0.93	0.93	286

Gambar 6. Hasil Algoritma SVM

Nilai *precision* merupakan nilai kecocokan antara dokumen sebenarnya dengan hasil prediksi klasifikasi. Berdasarkan output terlihat bahwa nilai *precision* untuk sms normal dengan nilai tertinggi diperoleh dengan Algoritma Naïve Bayes yaitu sebesar 1.00, sedangkan untuk sms fraud dan sms promo dengan metode SVM yaitu masing masing sebesar 0.90 dan 0.93.

Nilai *recall* merupakan nilai proporsi dokumen yang dapat diprediksi dengan tepat dari total jumlah dokumen secara keseluruhan dalam suatu kelas. Berdasarkan output terlihat bahwa nilai *recall* untuk sms normal dengan nilai tertinggi diperoleh dengan algoritma SVM yaitu sebesar 0.99, sedangkan untuk sms fraud dan sms promo dengan Algoritma Naïve Bayes yaitu masing masing sebesar 0.91 dan 0.92.

Nilai *f1-score* merupakan keseimbangan antara *precision* dan *recall.* Kelas SMS normal nilai f1-score pada algoritma *Naïve Bayes* dan SVM memiliki nilai yang sama yaitu sebesar 0.97, sedangkan untuk sms fraud nilai 0.91 pada algoritma SVM dan sms promo pada Algoritma Naïve Bayes dengan nilai 0.90.

Accuracy adalah rasio prediksi benar (negative dan positif) pada keseluruhan data. Nilai akurasi memberikan jawaban pada pertanyaan yang benar dan tidak benar pada hasil prediksi. Nilai akurasi tertinggi diperoleh pada Algoritma Naïve Bayes dengan nilai akurasi sebesar 0.94 dan pada Algoritma SVM sebesar 0.93. Dengan demikian klasifikasi data sms menggunakan Algoritma Naïve Bayes memiliki hasil prediksi lebih tinggi daripada algoritma *Support Vector Machine*.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Algoritma *Naïve Bayes* mempunyai *accuracy* lebih tinggi dibandingkan dengan algoritma SVM, dengan nilai sebesar 0.94. Nilai *precision, recall, f1-score* dan *accuracy* untuk algoritma *Naïve Bayes* juga memiliki kemampuan terbaik dibandingkan algoritma SVM untuk klasifikasi SMS Spam.

B. Saran

Untuk mendapatkan hasil yang sempurna maka disarankan pada penelitian selanjutnya melakukan perbandingan beberapa algoritma yang lain secara bersamaan (Algoritma Naïve Bayes, *Decision Tree*, SVM, dan *logistic regression*).

VI. REFERENSI

[1] dataindonesia, “Pengguna Smartphone Indonesia Terbesar Keempat Dunia pada 2022.” https://dataindonesia.id/digital/detail/pengguna-smartphone-indonesia-terbesar-keempat-dunia-pada-2022.

[2] techopedia, “Short Message Service (SMS),” 2019. https://www.techopedia.com/definition/24275/short-message-service--sms.

[3] M. Huda, *Keamanan Informasi*. nulisbuku.com, 2020.

[4] M. Ambika, G. Raghuraman, dan L. SaiRamesh, “Enhanced decision support system to predict and prevent hypertension using computational intelligence techniques,” *Soft Comput.*, vol. 24, no. 17, hal. 13293–13304, 2020, doi: 10.1007/s00500-

020-04743-9.

[5] M. Nour dan K. Polat, "Automatic Classification of Hypertension Types Based on Personal Features by Machine Learning Algorithms," *Math. Probl. Eng.*, vol. 2020, hal. 1–14, 2020, doi: 10.1155/2020/2742781.

[6] H. Herwanto, N. L. Chusna, dan M. S. Arif, "Klasifikasi SMS Spam Berbahasa Indonesia Menggunakan Algoritma Multinomial Naïve Bayes," *J. Media Inform. Budidarma*, vol. 5, no. 4, hal. 1316, 2021, doi: 10.30865/mib.v5i4.3119.

[7] B. Indiarto, "Klasifikasi Sms Spam Dengan Metode Naive Bayes Classifier Untuk Menyaring Pesan Melalui Selular," *J. Telemat. MKOM*, vol. 8, no. 2, hal. 167–172, 2016.