Rafiqa Cahyani, Indri Sudanawati Rozas, Nita Yalina

**Analisis Sentimen Pada Media Sosial Twitter Terhadap Tokoh Publik Peserta Pilpres 2019**

**[[1]](#footnote-1) *Abstract*—** **In Indonesia, Twitter has become a microblogging site that is very popular with the public. People are used to expressing their opinions or opinions on social media including on Twitter, this is used in response to an event or events that occur in their environment. In this study various processes were carried out, namely crawling data from Twitter, pre-processing text data, Naïve Bayes Classifier for text probability to sentiment, and Simple Wighted Product Model for ranking sentiments. From the results of these experiments can be obtained the best accuracy results using the third proportion with random data collection by getting an average accuracy value of 99.75%. As for ranking sentiments for each public figure, it can be seen that the first rank is KH Ma'ruf Amin, the second is Sandiaga Salahudin Uno, the third is Joko Widodo, and the last is Prabowo Subianto.**

**MATICS** : Jurnal Ilmu Komputer dan Teknologi Informasi ISSN : 1978-161X(p); 2477-2550(e)

Volume 9, No. 1(2017), pp 33-40 DOI : 10.18860/mat.v9i1.XXXX

**Received** : January25th 2016; **Accepted** : March 9th 2016 ; **Avalaible Online** : April 5th 2016

*Index Terms*— Twitter, Public Figure, Naïve Bayes Clasifier, Simple Wighted Product Model.

***Abstract---Di Indonesia twitter menjadi salah satu situs microblogging yang sangat digemari oleh masyarakat. Masyarakat sudah terbiasa dalam menyampaikan pendapat atau opininya dalam media sosial termasuk dalam twitter, hal ini digunakan dalam menanggapi suatu kejadian atau peristiwa yang terjadi di lingkungan mereka. Pada penelitian ini dilakakan berbagai proses yaitu crawling data dari twitter, pre prosesing data teks, Naïve Bayes Classifier untuk probabilitas teks terhadap sentimen, dan Simple Wighted Product Model untuk perangkingan sentimen. Dari hasil percobaan tersebut dapat diperoleh hasil akurasi terbaik yaitu menggunakan proporsi ketiga dengan pengambilan data secara acak dengan mendapatkan nilai rata rata akurasi sebesar 99,75%. Sedangkan untuk perangkingan sentimen terhadap masing masing tokoh publik dapat di ketahui yang memperoleh peringkat pertama yaitu KH Ma’ruf Amin , yang kedua yaitu Sandiaga Salahudin Uno, yang ketiga yaitu Joko Widodo, dan yang terakhir yaitu Prabowo Subianto***

*Kata Kunci*— Twitter, Tokoh Publik, Naïve Bayes Clasifier, Simple Wighted Product Model

T

# Introduction

Tahun 2019 merupakan tahun politik, dimana akan dilaksanakannya pemilihan presiden dan wakil presiden Indonesia. Dalam pelaksanaan pemilihan presiden dan wakil presiden terdapat berbagai opini masyarakat dengan sentimen positif dan negatif pada para tokoh publik yang mencalonkan dirinya untuk menjadi Presiden dan Wakil Presiden Republik Indonesia. Adapun tokoh publik tersebut adalah Bapak Joko Widodo, KH Ma’ruf Amin, Prabowo Subianto, dan Sandiaga Salahudin Uno. Pada penelitian kali ini peneliti akan meneliti sejauh mana citra positif yang dimiliki oleh keempat tokoh publik tersebut berdasarkan sentimen dari masarakat. Kelas sentimen terbagi menjadi dua polaritas yaitu sentimen negatif dan sentimen positif.

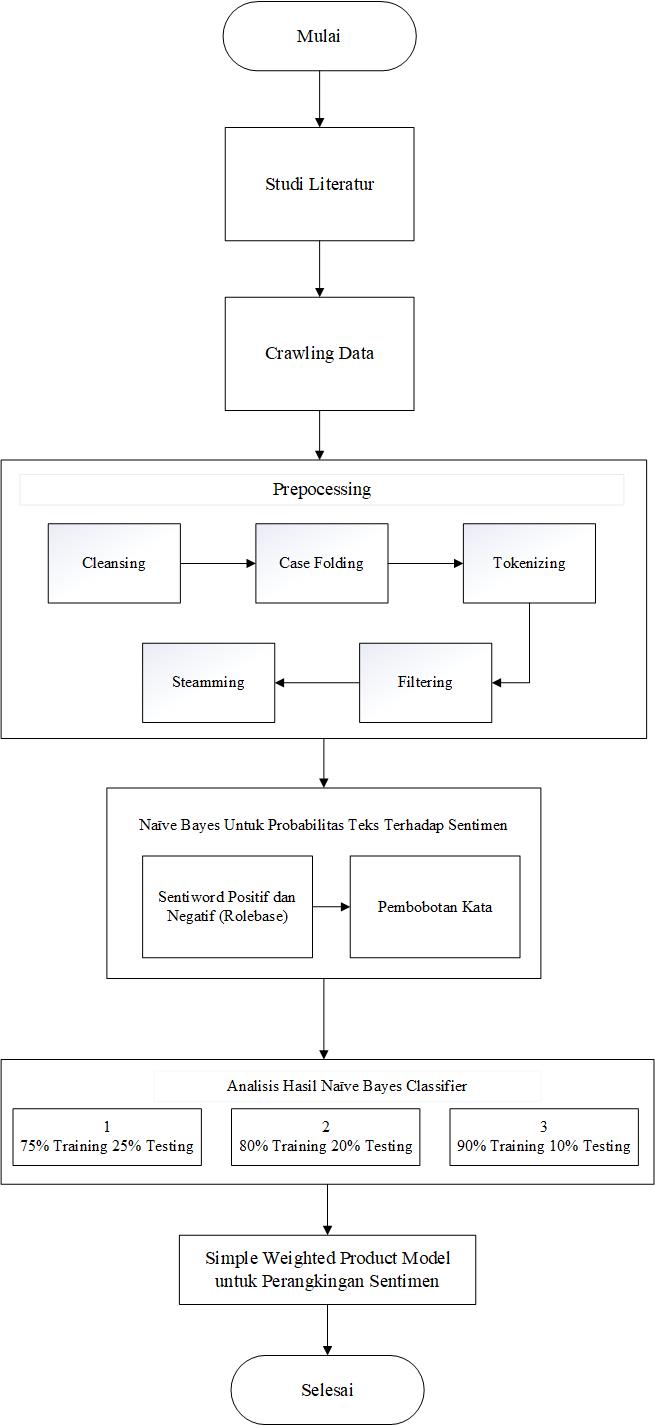
Penelitian yang akan dilakukan adalah pengklasifikasian topik sentimen dan sentimen dari informasi teks twitter yang membahas mengenai tokoh publik menggunakan metode *Naïve Bayes Classifier dan Simple Weighted Product Model*. *Metode Naïve Bayes Clasifier* juga digunakan dalam menentukan sentimen dari *tweet* sehingga dapat diketahui sentimen positif, negatif yang menjadi kelas sentimen pada teks twitter tersebut. Sedangkan *Simple Weighted Product Model* digunakan dalam menentukan rangking atau urutan dari para tokoh publik tang bersangkutan.

Untuk pengambilan data twitter menggunakan API (*Application Programming Interface*) dari twitter. Pemberian label kelas merupakan masalah yang cukup sulit, pada data tanpa label dengan jumlah yang sangat besar. Oleh karena itu digunakan metode pelabelan otomatis yaitu dengan memberikan label secara otomatis berdasarkan penanda tertentu yaitu dengan menggunakan sentiword posititif dan sentiword negative Bahasa Indonesia yang didapatkan dari github. Metode klasifikasi yang digunakan adalah metode klasifikasi *Naive Bayes Classifier*. Metode *Naive Bayes Classifier* merupakan salah satu metode klasifikasi yang memiliki performa yang cukup baik untuk mengklasifikasikan data teks. Dan digunakan *Simple Wighted Product Model* sebagai suatu medote untuk memperoleh nilai rangking terbaik dari keempat tokoh tersebut.

# mETODE PENELITIAN

Data pada penelitian ini berupa opini berbahasa Indonesia yang di dapatkan dari akun twitter para tokoh publik yang bersangkutan. Dari data yang tersedia dan setelah melalui proses penghapusan data duplikat dan data netral didapatkan data dari Joko Widodo sebanyak 5.755, KH Ma’ruf Amin sebanyak 6.177, Prabowo Subianto sebanyak 7130, dan Sandiaga Salahudin Uno sebanyak 5521. Data ini adalah data yang digunakan sebagai data klasifikasi.

Adapunlangkah langkah untuk perancangan analisis sentiment yang dibahas adalah sebagai berikut:



Gambar 2.1 Metodelogi Penelitian

A. Studi Literatur

Langkah yang pertama yaitu studi literatur yang digunakan untuk mencari data dan informasi yang dibutuhkan dari beberapa sumber .

B. Crawling Data

Langkah pertama yaitu dengan melakukan crawling data twitter. Pada crawling data twitter ini harus mempunyai akuntwiiter terlebih dahulu langkah selanjutnya yaitu registrasi di https://dev.twitter.com/apps/new untuk mendapatkan kode akses Api twitter. Setelah mendapatkan kode akses Api twitter, selanjutnya digunakan software Rstudio dan V Tag 6.0 untuk melakukan crawling data twitter.

C. Pre-Prossecing Data

Langkah yang ketiga yaitu pre-proses data yang berguna untuk merubah dataset yang belum diolah menjadi dataset yang siap untuk diklasifikasikan. Adapun beberapa tahapannya adalah proses *cleansing, case folding, tokeninzing, filtering dan stemming*.

D. Naïve Bayes Untuk Probabilitas Teks Terhadap Sentimen

Probabilitas sentimen merupakan proses penentuan teks twitter yang dihitung untuk didapatkan nilai tertinggi dari beberapa kategori dalam sentimen. Seperti halnya untuk menghitung tweet terhadap topik, perhitungan probabilitas tweet terhadap sentimen juga menggunakan rumus dimana untuk menentukan probabilitas sentimen memerlukan nlai p(ci) dan p(wkj.ci) untuk probabilitas awal sentimen didapatkan dari jumlah tweet yang tergabung dalam topik ci dibagi dengan semua topik.

p(wkj.ci) =



Dan untuk menghitung probabilitas teks terhadap sentiment juga dibuatkan wordcloud positif dan negative dari masing masing tokoh publik

E. Analisis Hasil Naïve Bayes Classifier

Pada sistem klasifikasi menggunakan Naïve Bayes Classifier, tahapan awal yang dilakukan adalah dengan membagi data menjadi dua bagian yaitu data training dan data testing. Pembagian data ini akan dilakukan secara acak dan urut, hal ini dilakukan karena untuk mengetahui tingkat akurasi yang dihasilkan dari kedua cara tersebut mana yang menghasilkan akurasi yang lebih tinggi dari keempat calon publik yang bersangkutan.Penelitian ini membagi data menjadi 3 kali proporsi.

1. 75% data training dan 25% data testing

2. 80% data training dan 20% data testing

3. 90% data training dan 10% data testing

Setelah proses sistem klasifikasi berakhir dari data training yang tersedia. Proses selanjutnya yang harus dilakukan adalah pengujian terhadap sistem klasifikasi yang dibuat. Evaluasi merupakan tahapan yang paling menentukan, karena bisa dilihat dari proses evaluasi tersebut apakah sistem klasifikasi yang sudah dibuat mempunyai peforma yang baik atau sebaliknya. Confusion matrix adalah salah satu cara untuk melakukan pengujian pada sistem klasifikasi. Confusion Matrix juga bisa dipakai untuk menghitunng tingkat akurasi, presisisi, dan recall.

F. Simple Weighted Product Model Untuk Perangkingan Sentimen

Pada penelitian ini pengunaan metode Simple Weighted Product Model digunakan untuk menghitung nilai sentimen yang ada pada topik setiap tokoh publik sehingga didapatkan urutan-urutan nama tokoh publik yang mendapatkan nilai sentimen terbaik sampai dengan sentimen terburuk. Langkah awal yaitu dengan memilih alternatif, kemudian memasukkan nilai bobot kepentingan dari setiap kriteria untuk mendapatkan hasil rangking.



Nilai akhir dari perhitungan akan digunakan untuk perangkingan. Semakin besar nilai akhir suatu alternatif (tokoh publik) maka akan mendapatkan citra positif terbaik dari masing masing tokoh publik.

# hasil dan pembahasan

Dari penggunaan data dari Joko Widodo sebanyak 14.526, KH Ma’ruf Amin sebanyak 13.207, Prabowo Subianto sebanyak 17.003, dan Sandiaga Salahudin Uno sebanyak 12.533. Data ini adalah data yang digunakan sebagai data klasifikasi. Dan setelah melalui proses pre-processing data maka langkah selanjutnya adalah pembobotan kata menggunakan sentiword positif dan

negatif yang didapatkan dari Github untuk masing masing tokoh publik. Perhitungan term atau kata menggunakan kata hasil preprocessing dan menghasilkan bobot masing-masing kata yang dibawah ini:

1. Pembobotan Kata

Pembobotan kata untuk Para tokoh publik peserta pilpres 2019, pembobotan kata ditentukan dari banyaknya kemunculan kata yang ada pada tweet yang bersangkutan. Hal ini juga menjadi bahan untuk pembuatan wordcloud sentimen positif dan negatif dari tokoh publik yang bersangkutan. Wordcloud ini dibuat untuk mengetahui kata sentimen positif dan negatif mana yang paling sering digunakan masayarakat untuk menilai tokoh publik yang mereka perbincangkan.

A. Pembobotan Kata Joko Widodo



Would Cloud Sentimen Positif untuk Joko Widodo



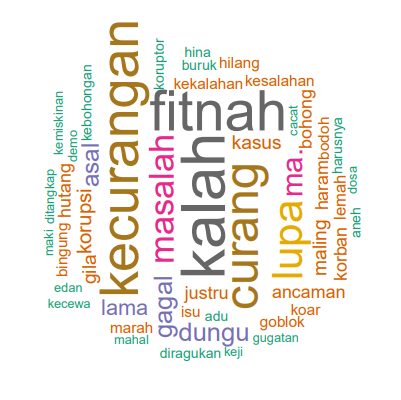
Would Cloud Sentimen untuk Joko Widodo

Dapat dilihat dari gambar diatas bahwa bobot probabilitas kata sentimen positif dan negatif yang paling sering digunakan masyarakat dalam memerikan tweet atau opini positif dan negatif sesuai dengan kamus sentimen positif dan negatif yang didapat dari Github untuk Joko Widodo. Terdapat 1.884 baris kata positif yang sering digunakan masyarakat yang pertama yaitu kata “banyak” yang digunkan sebanyak 320 kali. Setelah itu kata “kerja” yang digunakan sebanyak 308 kali, dan seterusnya. Selanjutnya, Terdapat 2056 baris kata yang sering digunakan masyarakat yang pertama yaitu kata “kalah” yang digunkan sebanyak 132 kali. Setelah itu kata “kecurangan” yang digunakan sebanyak 94 kali.

B. Pembobotan Kata KH. Ma’ruf Amin



Would Cloud Sentimen Positif untuk KH.Ma’ruf Amin



Would Cloud Sentimen Negatif untuk KH.Ma’ruf Amin

Dapat dilihat dari gambar diatas bahwa bobot probabilitas kata positif dan negatif yang paling sering digunakan masyarakat dalam memerikan tweet atau opini positif dan negatif sesuai dengan kamus positif dan negative yang didapat dari Github untuk KH.Ma’ruf Amin. Terdapat 780 baris kata yang sering digunakan masyarakat yang pertama yaitu kata “maju” yang digunkan sebanyak 430 kali. Setelah itu kata “semoga” yang digunakan sebanyak 422 kali, dan seterusnya. Selanjutnya, terdapat 657 baris kata yang sering digunakan masyarakat yang pertama yaitu kata “kalah” yang digunkan sebanyak 120 kali. Setelah itu kata “fitnah” yang digunakan sebanyak 108 kali.

C. Pembobotan Kata Untuk Prabowo Subianto



WouldCloud Sentimen Positif untuk Prabowo Subianto



WouldCloud Sentimen Negatif untuk Prabowo Subianto

Dapat dilihat dari gambar diatas bahwa bobot probabilitas kata sentimen positif dan negatif yang paling sering digunakan masyarakat dalam memerikan tweet atau opini positif dan negatif sesuai dengan kamus sentimen positif dan negatif yang didapat dari Github untuk Prabowo Subianto. Terdapat 1664 baris kata yang sering digunakan masyarakat yang pertama yaitu kata “semoga” yang digunkan sebanyak 408 kali. Setelah itu kata “menang” yang digunakan sebanyak 375 kali, dan seterusnya. Selanjutnya, terdapat 1707 baris kata yang sering digunakan masyarakat yang pertama yaitu kata “kalah” yang digunkan sebanyak 270 kali. Setelah itu kata “fitnah” yang digunakan sebanyak 146 kali.

D. Pembobotan Kata Untuk Sandiaga Uno



WouldCloud Sentimen Positif untuk Sandiaga Uno



WouldCloud Sentimen Negatif untuk Sandiaga Uno

Dapat dilihat dari gambar diatas bahwa bobot probabilitas kata positif dan negatif yang paling sering digunakan masyarakat dalam memerikan tweet atau opini positif dan negatif sesuai dengan kamus positif dan negative yang didapat dari Github untuk Sandiaga Salahudin Uno. Terdapat 1257 baris kata yang sering digunakan masyarakat yang pertama yaitu kata “semoga” yang digunkan sebanyak 387 kali. Setelah itu kata “menang” yang digunakan sebanyak 336 kali, dan seterusnya. Selanjutnya, terdapat 1277 baris kata yang sering digunakan masyarakat yang pertama yaitu kata “kalah” yang digunkan sebanyak 163 kali. Setelah itu kata “kecurangan” yang digunakan sebanyak 112 kali.

2. Naïve Bayes Untuk Probabilitas Teks Terhadap Sentimen

Setelah semua kata yang ada pada data set digunakan sebagai data training sudah diberikan bobot untuk masing masing tokoh publik, maka selanjutnya adalah pembuatan sistem klasifikasi dengan menggunakan metode Naïve Bayes Classifier.

Pada sistem klasifikasi, tahapan yang pertama yang dilakukan adalah dengan membagi data menjadi dua yaitu data training dan data testing. Pembagian data ini akan dilakukan secara acak dan urut, hal ini dilakuakn karena untuk mengetahui tingkat akurasi yang dihasilkan dari kedua cara tersebut mana yang menghasilkan akurasi yang lebih tinggi dari keempat calon publik yang bersangkutan.

Pada penelitian kali ini peneliti akan membagi data menjadi 3 kali proporsi.

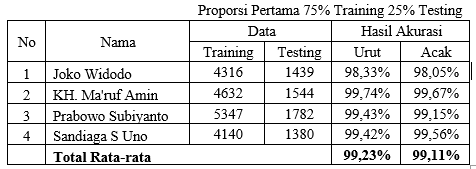
A. 75% data training dan 25% data testing

B. 80% data training dan 20% data testing

C. 90% data training dan 10% data testing

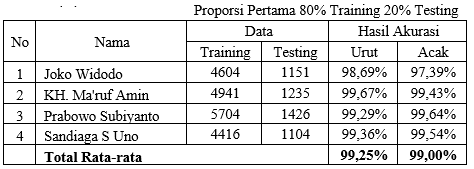
Setelah sistem klasifikasi sudah selesai dibuat dari data training yang ada. Langkah selanjutnya yang harus dilakukan adalah pengujian terhadap sistem klasifikasi yang dibuat. Evaluasi terhadap sistem ini merupakan tahapan yang menentukan apakah dari sistem klasifikasi yang sudah dibuat mempunyai peforma yang baik atau tidak. Untuk melakukan pengujian pada sistem klasifikasi, biasanya digunakan confusion matrix. Confusion Matrix ini bisa digunakan untuk menghitung akurasi, presisisi, dan recall. Untuk itu akan akan dibahas mengenai presentasi rata rata nilai akurasi dari proporsi pertama sampai dengan yang ketiga baik pengambilan data secara urut maupun acak yang dimiliki oleh setiap tokoh publik yang bersangkutan.

A. Presentasi Nilai Akurasi Pada Proporsi Pertama



Dalam tabel diatas dapat diketahui bahwa hasil rata rata akurasi pada proporsi pertama oleh keempat tokoh tersebut yaitu pada pengambilan data urut diperoleh rata rata akurasi sebesar 99,23% sedangkan dari pengambilan data secara acak pada proporsi pertama diperoleh rata rata akurasi sebesar 99,11%.

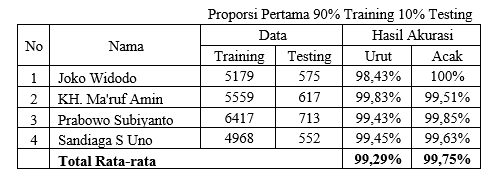
B. Presentasi Nilai Akurasi Pada Proporsi Kedua



Dalam tabel diatas dapat diketahui bahwa hasil rata rata akurasi pada proporsi kedua oleh keempat tokoh tersebut yaitu pada pengambilan data urut diperoleh rata rata akurasi sebesar 99,25% sedangkan dari pengambilan

data secara acak pada proporsi pertama diperoleh rata rata akurasi sebesar 99,00%.

C. Presentasi Nilai Akurasi Pada Proporsi Ketiga

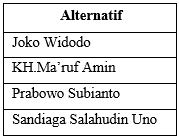


Dalam tabel diatas dapat diketahui bahwa hasil rata rata akurasi pada proporsi ketiga oleh keempat tokoh tersebut yaitu pada pengambilan data urut diperoleh rata rata akurasi sebesar 99,29% sedangkan dari pengambilan data secara acak pada proporsi pertama diperoleh rata rata akurasi sebesar 99,75%.

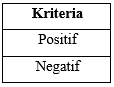
Dari ketiga percobaan yang dilakukan maka dapat diambil kesimpulan jika metode Naïve Bayes Classifier ini cocok diterapkan pada proporsi ketiga dengan pengambilan data secara acak. Hasil rata rata yang diperoleh adalah 99,75%, hasil rata rata ini merpakan hasil yang paling tinggi diperoleh dibandingkan dengan percobaan yang lain.

3. Perhitungan Menggunakan Metode Simple Weighted Product Model

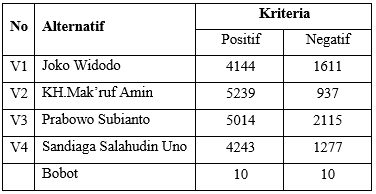
Pada proses ini akan menghasilkan nilai rangking tiap-tiap tokoh publik dengan nilai sentimen yang didapatkan dengan menggunakan metode Metode Simple Weighted Product Model. Informasi sentimen tokoh politik diperoleh dari hasil klasifikasi tweet tokoh publik. Penggunaan metode Metode Simple Weighted Product Model dalam penentuan peringkat tokoh publik. Langkah pertama yaitu menentukan nilai alternatif dari setiap tokoh tersebut.



Yang kedua yaitu menentukan kriteria perhitungan SWP dari masing masing tokoh



Yang ketiga yaitu menentukan nilai kriteria setiap alternatif perhitungan SWP



Selanjutnya yaitu melakukan proses perhitungan menggunakan metode Simple Weightes Product (SWP) akan dilakukan normalisasi bobot.



W1 = 10 / (10+10) = 10 / 20 = 0,5

W2 = 10 / (10+10) = 10 / 20 = 0,5

Selanjutnya yaitu menghitung vector S



S1 = (4144 0,5) (1611 -0,5) = 1,603844

S2 = (5239 0,5) (937 -0,5) = 2,364582

S3 = (5014 0,5) (2115 -0,5) = 2,539703

S4 = (4243 0,5) (1277 -0,5) = 1,822809

Tahap selanjutnya yaitu menghitung vector perangkingan



V1 = 1,603844 / (1,603844 + 2,364582 + 2,539703 + 1,822809) = 0,218777

V2 = 2,364582 / (1,603844 + 2,364582 + 2,539703 + 1,822809) = 0,322548

V3 = 2,539703 / (1,603844 + 2,364582 + 2,539703 + 1,822809) = 0,210028

V4 = 1,822809 / (1,603844 + 2,364582 + 2,539703 + 1,822809) = 0,248648

Dari proses perhitungan nilai sentiment dengan bobot topik diatas, didapatkan nilai akhir sentimen pada masing – masing tokoh publik dengan nilai tertinggi yakni 0,322548 (KH. Maruf Amin), 0,248648 (Sandiaga Salahudin Uno), 0,218777 (Joko Widodo) dan nilai terendah yakni 0,210028 (Prabowo Subianto). Maka kesimpulan didapatkan bahwa KH. Maruf Amin memiliki nilai citra positif tertinggi, disusul oleh Sandiaga Salahudin Uno kemudian Joko Widodo, dan yang terakhir yakni Prabowo Subianto.

# KESIMPULAN

Berdasarkan hasil pembuatan sistem klasifikasi dengan menggunakan Naïve Bayes Classifier dan menggunakan metode perhitungan Simple Weighted Product Model terhadap semua data yang diperoleh dari twitter, maka dapat disimpulkan bahwa:

1. Hasil yang di dapatkan menunjukkan bahwa metode Naïve Bayes Classifier cocok diterapkan di ketiga proporsi yang ada, karena sama sama mendapatkan nilai akurasi yang cukup tinggi. Tetapi setelah dilakukan rata rata maka percobaan yang paling sesuai yaitu menggunakan proporsi ketiga dengan pengambilan data secara acak dengan memperoleh hasil rata rata sebesar 99,75%.
2. Dengan menggunakan metode perhitungan Simple Weighted Product Model maka hasil yang didapatkan yaitu tokoh publik yang mendapatkan urutan (rangking) dalam sentimen masyarakat terkait citra positif yang dimiliki tokoh publik yang pertama adalah KH. Ma’ruf Amin dengan mendapatkan nilai 0,322548, setelah itu adalah Sandiaga Salahudin Uno yang mendapatkan nilai 0,248648, lalu Joko Widodo yang mendapatkan nilai 0,218777 dan nilai terendah yakni Prabowo Subianto dengan nilai 0,210028.

# SARAN

Berdasarkan hasil penelitian dari analisis sentimen teks pada media sosial twitter menggunakan Naïve Bayes Classifier dan dengan menggunakan metode perhitungan Simple Weighted Product Model yang telah dilakukan, diperlukan penelitian yang lebih lanjut terkait dengan analisis sentimen tokoh publik ini. Berikut saran yang dapat disampaikan dari hasil kesimpulan diatas:

1. Pada Proses klasifikasi menggunakan Naive Bayes sangat bergantung dari data latih yang digunakan. Penggunaan data latih sebaiknya menggunakan proporsi data yang sama dari setiap kelasnya serta memastikan kelas dari data latih tersebut benar.
2. Pada proses crawling data twitter, data yang didapatkan dari masing masing tokoh publik banyak terdapat data outlier (data yang tidak sesuai). Oleh sebab itu crawling data sebaiknya dilakukan menggunakan cara lain atau dengan mengganti kata kunci yang lebih sesuai sehingga tidak didapatkan banyak data outlier
3. Untuk mendapatkan hasil yang signifikan, perlu dilakukan eksperimen penggunaan leksikon polaritas dan multi kata, menggunakan strategi yang sangat mendasar berdasarkan pencarian lemmas polaritas dalam teks / tweet. Jika tweet tidak mengandung setidaknya satu lemma juga ditemukan dalam leksikon polaritas eksternal, maka tweet tersebut tidak memiliki polaritas apa pun
4. Hasil akurasi yang didapatkan dalam penelitian kali ini belum tentu sama tingginya dengan hasil akurasi jika menggunakan metode yang berbeda. Hal ini bisa menjadi rekomendasi untuk penelitian selanjutnya untuk melakukan penelitian dengan Teknik yang sama tetapi menggunakan metode yang berbeda.

# References

[1] Asosiasi Penyelenggara Jasa Internet Indonesia – APJII, “Penetrasi & Perilaku Pengguna Internet Indonesia – survey 2019,” p. 34, 2019

.

[2] Abdillah, S. (2011). *Penyakit Stroke Dengan Klasifikasi Data Mining Pada RS Santa Maria Pemalang*.

[3] Ariadi, D., & Fithriasari, K. (2015). Klasifikasi Berita Indonesia Menggunakan Metode Naive Bayesian Classification dan Support Vector Machine dengan Confix Stripping Stemmer. *JURNAL SAINS DAN SENI ITS Vol. 4, No.2*.

[4] Balagatabi, Z. N., & Balagatabi, H. N. (2012). *Comparison of Decision Tree and SVM Methods in Classification of Researcher ’ s Cognitive Styles in Academic Environment*. *1*(January), 31–43.

[5] Borcard, D., Gillet, F., Legendre, P., Borcard, D., Gillet, F., & Legendre, P. (2011). Unconstrained Ordination. In *Numerical Ecology with R*. https://doi.org/10.1007/978-1-4419-7976-6\_5

[6] Chen, L. F., Su, C. T., & Chen, K. H. (2011). An improved particle swarm optimization for feature selection. *Intelligent Data Analysis*, *16*(2), 167–182. https://doi.org/10.3233/IDA-2012-0517

[7] Copestake, A. (2007). *Natural Language Processing Outline of today ’ s lecture*. (October).

[8] Crawley, M. J. (2007). The R Book. Chichester: John Wiley & Sons. In *The R Book*. https://doi.org/10.1002/9780470515075

[9] Goodfellow, I. J., Bengio, Y., & Courville, A. (2016). Applied math and machine learning basics. *Deep Learning*, 25–26.

[10] Hearst, M. A., & Hall, S. (n.d.). *P99-1001*.

[11] Juniawan, I. (2009). *Klasifikasi Dokumen Teks Berbahasa Indonesia Menggunakan Minor Component Analysis*.

[12] Lestari, A. R. T., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Tentang Opini Pilkada Dki 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Näive Bayes dan Pembobotan Emoji. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *1*(12), 1718–1724.

[13] Manning, C. D., Raghavan, P., & Schutze, H. (2008). *Introduction to Information Retrieval*. https://doi.org/10.1017/CBO9780511809071

[14] Mitchell, T. M. (2009). Machine learning in ecosystem informatics and sustainability. *IJCAI International Joint Conference on Artificial Intelligence*. https://doi.org/10.1007/978-3-540-75488-6\_2

[15] Prasetyo, E. (2014). Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab. In *Penerbit Andi*. https://doi.org/10.1017/CBO9781107415324.004

[16] Rini, D. C., Farida, Y., & Puspitasari, D. (2016). Klasifikasi Menggunakan Metode Hybrid Bayessian-Neural Network (Studi Kasus: Identifikasi Virus Komputer). *Jurnal Matematika “MANTIK,”* *1*(2), 38. https://doi.org/10.15642/mantik.2016.1.2.38-43

[17] Saraswati, N. W. S. (2013). Naïve Bayes Classifier Dan Support Vector Machines Untuk Sentiment Analysis. *Seminar Nasional Sistem Informasi Indonesia*, 586–591. Retrieved from http://is.its.ac.id/pubs/oajis/index.php/file/download\_file/512

1. Manuscript received March 22, 2007. (Write the date on which you submitted your paper for review.) This work was supported in part by Informatics Engineering Department of Maulana Malik Ibrahim Islamic State University. Paper titles should be written in uppercase and lowercase letters, not all uppercase. Avoid writing long formulas with subscripts in the title; short formulas that identify the elements are fine (e.g., "Nd–Fe–B"). Do not write "(Invited)" in the title. Full names of authors are preferred in the author field, but are not required. Put a space between authors' initials.

   F. A. Author is with the Informatic Engineering Departement of Maulana Malik Ibrahim Islamic State University , Malang, Indonesia (corresponding author provide phone 0341-000000; email [author@uin-malang.ac..id](mailto:author@uin-malang.ac..id))

   S. B. Author, Jr., was with Institut Technology of Bandung, Bandung, Indonesia. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar. colostate.edu).

   T. C. Author is is the Informatic Engineering Departement of Maulana Malik Ibrahim Islamic State University, Malang, Indonesia (corresponding author provide phone 0341-000000; email [author@uin-malang.ac..id](mailto:author@uin-malang.ac..id) [↑](#footnote-ref-1)