

Analisis Sentimen Terhadap Pembobolan Data pada Twitter dengan Algoritma Naive Bayes

Ahmad Turmudi Zy¹⁾, Agung Nugroho²⁾, Ahmad Rivaldi³⁾, Irfan Afriantoro^{*)4)}

¹⁾²⁾³⁾⁴⁾Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa

Correspondence author : irfanafriantoro@pelitabangsa.ac.id, Bekasi, Indonesia

DOI: <https://doi.org/10.37012/jtik.v8i2.1240>

Abstrak

Perkembangan teknologi informasi kini sangat cepat dan jauh berbeda dengan masa awal kehadirannya. Era globalisasi telah menempatkan peranan teknologi informasi ke dalam suatu posisi yang sangat strategis karena dapat menghadirkan suatu dunia tanpa batas, jarak, ruang, dan waktu serta dapat meningkatkan produktivitas serta efisiensi. Twitter merupakan media sosial yang mudah digunakan untuk penyebaran informasi secara cepat dan luas. Sejak ramainya kasus Bjorka hal itu memicu banyak masyarakat yang mengkritik di berbagai media sosial salah satu diantaranya media sosial Twitter sehingga kritik atau opini tersebut dapat dimanfaatkan untuk melakukan analisis sentimen. Berdasarkan hal tersebut, diperlukan sebuah metode yang dapat secara otomatis melakukan klasifikasi opini ke dalam kategori positif dan negatif melalui proses analisis sentimen. Proses analisis sentimen dilakukan dengan proses data preprocessing, pembobotan kata menggunakan metode TF-IDF, penerapan algoritma, dan pembahasan atas hasil klasifikasi. Metode klasifikasi data yang digunakan dalam penelitian ini adalah *Naive Bayes Classifier* (NBC). Data tersebut akan diproses menggunakan *text mining* dan klasifikasi menggunakan algoritma Naive Bayes. Metode tersebut menghasilkan tingkat dan hasil yang cukup baik. Klasifikasi dapat memberikan kemudahan bagi pengguna untuk melihat opini positif dan negatif. Berdasarkan pengujian yang telah dilakukan, hasil klasifikasi terbaik diperoleh dengan nilai accuracy, precision, dan recall tertinggi yang mendapatkan hasil dengan nilai accuracy 98.33%, precision 100.00%, dan recall sebanyak 97.13%.

Kata Kunci : Analisis Sentimen, Naive Bayes, Twitter, Opini, Klasifikasi.

Abstract

The development of information technology is now very fast and much different from the early days of its presence. The era of globalization has placed the role of information technology in a very strategic position because it can present a world without boundaries, distance, space, and time and can increase productivity and efficiency. Twitter is a social media that is easy to use for the dissemination of information quickly and widely. Since the hectic Bjorka case, it has triggered many people to criticize on various social media, one of which is Twitter social media so that these criticisms or opinions can be used to conduct sentiment analysis. Based on this, we need a method that can automatically classify opinions into positive and negative categories through a sentiment analysis process. The sentiment analysis process is carried out by preprocessing data, word weighting using the TF-IDF method, applying the algorithm, and discussing the classification results. The data classification method used in this study is the Naive Bayes Classifier (NBC). The data will be processed using text mining and classification using the Naive Bayes algorithm. This method produces a fairly good rate and results. Classification can make it easy for users to see positive and negative opinions. Based on the tests that have been carried out, the best classification results are obtained with the highest accuracy, precision, and recall values which get results with 98.33% accuracy values, 100.00% precision, and 97.13% recall.

Keywords: Sentiment Analysis, Naive Bayes, Twitter, Opinion, Classification.

PENDAHULUAN

Perkembangan teknologi informasi kini sangat cepat dan jauh berbeda dengan masa awal kehadirannya. Era globalisasi telah menempatkan peranan teknologi informasi ke dalam suatu posisi yang sangat strategis karena dapat menghadirkan suatu dunia tanpa batas, jarak, ruang, dan waktu serta dapat meningkatkan produktivitas serta efisiensi. Teknologi informasi telah merubah pola hidup masyarakat secara global dan menyebabkan perubahan sosial budaya, ekonomi, dan kerangka hukum yang berlangsung secara cepat dengan signifikan.

Teknologi informasi saat ini menjadi pedang bermata dua karena selain memberikan kontribusi bagi peningkatan kesejahteraan, kemajuan dan peradaban manusia, sekaligus menjadi sarana efektif perbuatan melawan hukum. Kemajuan teknologi informasi terutama pada bidang komputer dan internet terbukti telah memberikan dampak positif bagi kemajuan kehidupan manusia. Perlu digaris bawahi, dibalik kelebihan dan kemudahan yang ditawarkan oleh komputer dan internet, ternyata memiliki sisi gelap yang dapat menghancurkan kehidupan dan budaya manusia itu sendiri. Informasi sudah dianggap sebagai “power” yang diartikan sebagai kekuatan dan kekuasaan yang sangat menentukan nasib manusia itu sendiri. Saat ini ketergantungan masyarakat akan teknologi informasi semakin tinggi sehingga semakin tinggi pula resiko yang dihadapi. Seiring perkembangan teknologi internet, mengakibatkan munculnya kejahatan baru yang disebut dengan *new cybercrime* melalui jaringan internet.

Berawal pada tahun 2003 banyak kejahatan-kejahatan (*cybercrime*) yang bermunculan dengan memanfaatkan kemajuan dari teknologi informasi, seperti kejahatan *carding (credit card fraud)*, *ATM/EDC skimming*, *hacking*, *cracking*, *phising (internet banking fraud)*, *malware (virus/worm/trojan/bots)*, *cybersquatting*, pornografi, perjudian online, *transnasional crime* (perdagangan narkoba, mafia, terorisme, *money laundering*, *human trafficking*, *underground economy*). Kesemua tindak pidana tersebut bisa dengan mudah dan efektif dilakukan dengan memanfaatkan kemajuan teknologi informasi itu sendiri. Tidak hanya itu tindak pidana (*cybercrime*) yang berpotensi dilakukan dengan mudah dan efektif dengan memanfaatkan perkembangan teknologi dan informasi juga pada sektor pengelolaan data dan informasi khususnya pada pengelolaan data pribadi yang membutuhkan perlindungan data. Sebab dengan kemajuan teknologi informasi dan

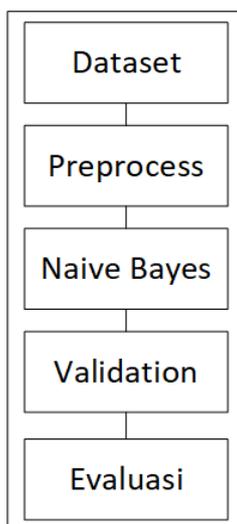
komunikasi tersebut membuat batas privasi makin tipis sehingga berbagai data-data pribadi semakin mudah untuk tersebar.

Salah satu kasus yang sekarang terjadi di Indonesia yaitu pembobolan data yang dilakukan oleh seorang Hacker yang dijuluki “Bjorka” dengan username pada twitter @bjorkanesian. Akhir-akhir ini masyarakat Indonesia dihebohkan dengan kemunculan seorang hacker yang bernama Bjorka. Pasalnya hacker bernama Bjorka ini berhasil membongkar data negara termasuk identitas para pejabat tinggi[4]. Kebocoran data ini bukanlah hal yang baru untuk Indonesia hingga muncul istilah Indonesia adalah negara opensource dikarenakan kebocoran data adalah hal yang tidak pertama kali terjadi. Bjorka merupakan nama akun seorang yang diduga hacker dari sebuah forum khusus. Hacker Bjorka berhasil menyita perhatian publik dalam kurun waktu beberapa waktu terakhir. Popularitasnya semakin meningkat setelah Bjorka mengunggah data-data rahasia milik pemerintah yang berhasil ia dapatkan[6]. Tak berhenti sampai disitu, hacker Bjorka bahkan mengancam Presiden Joko Widodo atau Jokowi. Ia juga bahkan melakukan *doxing* (menyebarkan data tanpa izin) beberapa orang penting di Indonesia, misalnya Kemenkominfo Johnny G Plate, Menteri BUMN Erick Thohir, dan yang terbaru Mendagri M Tito Karnavian serta Gubernur DKI Jakarta Anies Baswedan. Pembahasan tentang Bjorka yang semakin memanas dan semakin meresahkan menjadikan pihak Istana turun tangan untuk mengatasi serangan dari hacker Bjorka tersebut. Salah satu langkah yang diambil oleh Presiden Republik Indonesia, Presiden Joko Widodo untuk memburu hacker Bjorka adalah dengan membentuk tim khusus atau timsus lintas lembaga negara.

Sejak ramai nya kasus tersebut hal itu memicu banyak masyarakat yang mengkritik diberbagai media sosial salah satu diantaranya media sosial Twitter. Twitter merupakan media sosial yang mudah digunakan untuk penyebaran informasi secara cepat dan luas. Pada data-data Twitter tersebut penulis melakukan analisis sentimen tentang pendapat atau komentar warganet terhadap peristiwa yang terjadi. Analisis Sentimen ini mengelompokkan pendapat atau komentar Positif dan Negatif untuk kemudian dibandingkan banyaknya masing-masing kelompok dengan perhitungan algoritma klasifikasi Naive Bayes. Berdasarkan uraian diatas, perlu dilakukan penelitian berjudul “Analisis Sentimen Terhadap Pembobolan Data Pada Twitter Dengan Algoritma Naive Bayes”.

METODE

Kerangka pemikiran yakni alur pikiran yang dijadikan sebagai skema atau dasar pemikiran untuk memperkuat indikator yang melatar belakangi penelitian ini. Dalam kerangka pemikiran ini dijelaskan alur-alur dalam menyelesaikan penelitian. Penjelasan yang disusun akan menggabungkan antara teori dengan masalah yang diangkat dalam penelitian ini. Kerangka pemikiran yang diusulkan dapat dilihat pada gambar dibawah ini:



Gambar 1. Kerangka Pemikiran

Pada penelitian ini *dataset* yang digunakan berasal dari *tweet* atau komentar di twitter dan data yang akan diambil dengan *query* Bjorka. Data tweet diambil dan kemudian dilakukan pemrosesan menggunakan perangkat lunak rapidminer, ditahap awal akan dilakukan *preprocess* yang didalam proses tersebut ada tahap yang harus dilalui satu per satu. Selanjutnya proses *validation* dengan membagi porsi data untuk *training* dan *testing*. Kemudian diuji dengan algoritma Naïve Bayes.

Objektivitas pada penelitian ini adalah untuk mengetahui hasil klasifikasi berdasarkan model yang diusulkan dengan Teknik *performance classification*. Kemudian untuk tahap terakhir akan dilakukan evaluasi menggunakan *confussion matrix* untuk melihat hasil akurasi algoritma.

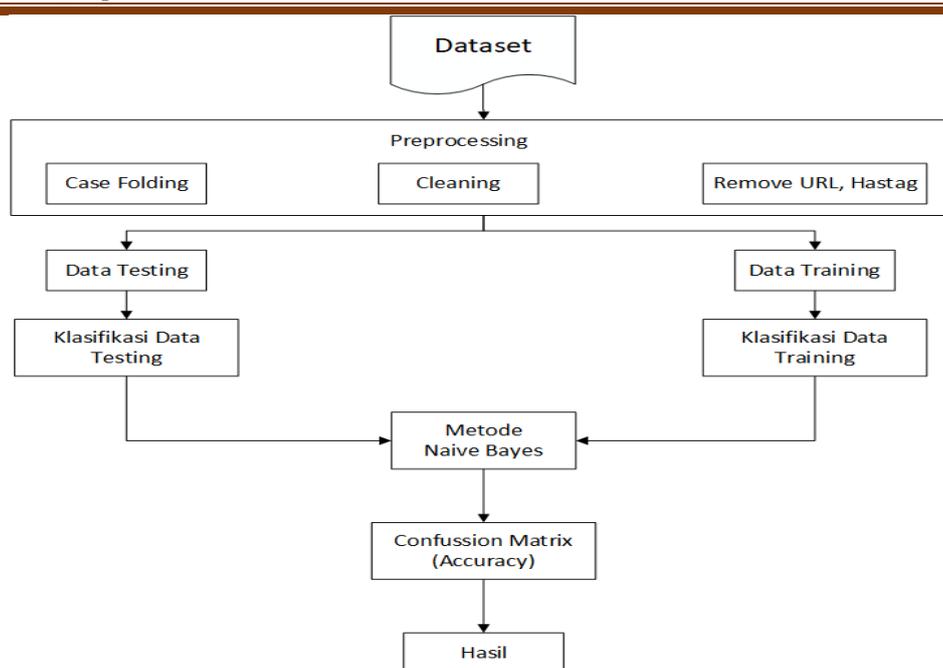
Objek Penelitian adalah permasalahan yang diteliti dan dibahas dalam penelitian. Pada penelitian diangkat permasalahan tentang Pembobolan Data. Akhir-akhir ini kasus pembobolan data yang sedang terjadi ini mendapat sorotan dari masyarakat luas tentang pro kontra maupun respon terhadap pembobolan data yang sering terjadi ini. Namun banyaknya

pendapat masyarakat mengenai kasus ini yang berupa respon positif ataupun negatif. Oleh sebab itu, objek penelitian yang diambil adalah respon dari masyarakat terhadap Pembobolan Data tersebut.

Subjek penelitian adalah individu atau kelompok yang terlibat dalam permasalahan. Adapun subjek penelitian ini adalah seorang *Hacker* bernama Bjorka yang melakukan penyerangan pembobolan data tersebut. Data yang diambil untuk penelitian ini merupakan data dari akun media sosial twitter dengan keyword pencarian “bjorka” dan data yang diambil berupa komentar maupun *tweet* pada twitter.

Jenis data yang digunakan dalam penelitian ini adalah data kualitatif yang merupakan data yang berbentuk kata maupun kalimat. Kata-kata tersebut nantinya akan diproses untuk mengetahui sentimen yang terdapat didalamnya. Pada penelitian ini, data berasal dari komentar dan tweet yang berhubungan dengan objek penelitian. Data yang digunakan pada penelitian ini merupakan data yang diperoleh dari *twitter*, baik berupa tweet maupun komentar yang diambil melalui proses penarikan atau *crawling* menggunakan API Twitter dengan memanfaatkan aplikasi RapidMiner. Data yang sudah terkumpul nantinya akan dibagi menjadi dua bagian yaitu data *training* dan data *testing*.

Pada penelitian ini metode yang diusulkan dalam analisis sentimen untuk mengetahui sentimen tentang Pembobolan Data, dengan menggunakan metode klasifikasi Naive Bayes, pada metode ini terdapat dua proses antara lain proses *training* dan proses *testing*. Pada langkah awal akan dilakukan preprocessing data kemudian memberikan sentimen (*labeling*) terhadap *review text dataset*, setelah itu data yang telah diberi sentimen dikumpulkan antara sentimen positif dan sentiment negatif menjadi data *training* dan data *testing*, selanjutnya akan dilakukan model klasifikasi Naive Bayes dan mendapatkan hasil *Confusion Matrix* setelah itu akan dihitung hasil *Accuracy* dari Naive Bayes tersebut.



Gambar 2. Metode Yang Diusulkan

HASIL DAN PEMBAHASAN

Setelah dilakukan analisa, dilanjutkan pembahasan mengenai hasil dari pengujian yang telah dilakukan. Untuk mengetahui hasil nilai *accuracy* serta prediksi yang dapat digunakan untuk mengetahui tentang sentimen atau opini masyarakat terhadap kasus pembobolan data dengan algoritma Naive Bayes.

Pada penelitian ini jumlah *dataset* yang digunakan sebanyak 550 data *tweet* yang terdiri dari *tweet* positif dan *tweet* negatif. Proses *crawling* data dengan kata kunci “bjorka” sebagai objek utama penelitiannya. Proses pengumpulan data dilakukan dengan proses *crawling* data yang menggunakan perangkat lunak RapidMiner. Jumlah data yang diperoleh dari proses tersebut sebanyak 3000 data. Setelah itu dilakukan proses *preprocessing* atau seleksi data, dataset yang diperoleh sebanyak 550 data. Untuk lebih jelasnya dapat dilihat pada tabel dataset berikut ini:

Tabel 1. *Dataset*

Kata Kunci Twitter	Data Awal	Data Akhir	Positif	Negatif
Bjorka	3000	550	285	265

Crawling data menggunakan RapidMiner untuk mendapatkan dataset yang akan digunakan untuk penelitian. *Crawling* data dengan keyword “bjorka”, berikut adalah hasil dari *crawling* data:

Tabel 2. Hasil *Crawling* Data

Row No.	Id	Cre...	From-User	From-User-Id	To-User	To-User-Id	Language	Source	Text
1	1574168216...	Sep ...	detikcom	69183155	?	-1	in	<a href="http...	Seseorang
2	1571524008...	Sep ...	Pro Gamers	533359500	?	-1	in	<a href="http:/...	Itu yg memt
3	1571155602...	Sep ...	Susi Pudjast...	2606229566	?	-1	in	<a href="http:/...	🤔🤔🤔🤔🤔
4	1574320591...	Sep ...	qin	190833859	?	-1	in	<a href="http:/...	RT @remot
5	1574320542...	Sep ...	bakomstrasu...	1422739125...	?	-1	in	<a href="http:/...	Susah seka
6	1574320528...	Sep ...	Virsha Alma'i	472680319	?	-1	in	<a href="http:/...	RT @remot
7	1574320294...	Sep ...	ris	1248579173...	?	-1	in	<a href="http:/...	RT @remot
8	1574320040...	Sep ...	DVL	145123423	?	-1	in	<a href="http...	RT @remot
9	1574320027...	Sep ...	Desnat	8289741995...	detikcom	69183155	in	<a href="http...	@detikcom
10	1574319923...	Sep ...	Pam	92878321	Braderhood16	1205991650...	in	<a href="http:/...	@Braderho
11	1574319842...	Sep ...	Bakomstra Pati	32057838	?	-1	in	<a href="http:/...	Susah seka
12	1574319808...	Sep ...	Erriza Syahpu...	232763422	?	-1	in	<a href="http:/...	RT @Presic
13	1574319741...	Sep ...	abdu somad	3268667966	?	-1	in	<a href="http:/...	https://t.co/Y
14	1574319709...	Sep ...	Demokrat Jat...	32058332	?	-1	in	<a href="http:/...	Susah seka
15	1574319680...	Sep ...	Aka	1555850322...	?	-1	in	<a href="http:/...	RT @remot
16	1574319552...	Sep ...	30	628662378	?	-1	in	<a href="http:/...	RT @remot

Pada tahap berikutnya dilakukan penggabungan antara atribut data yang akan di uji dengan data training yang sebelumnya sudah kita tentukan sentimen nya secara manual. Adapun tahapannya sebagai berikut:

1. *Retrive Data Training*

Fungsi dari *Retrive* ini agar nantinya atribut yang ada pada *Data Training* ini dapat digabungkan di *Data Testing*.

2. *Union*

Ketika dataset antara data *training* dan data *testing* tidak lengkap maka *Union* akan menggabungkan atribut keduanya dan mengisinya sebagai *Missing Values* atau kolom yang tidak ada nilainya.

3. *Replace Missing Values*

Operator ini digunakan untuk mengganti nilai yang tidak ada pada data ini menjadi nilai nol. Setelah selesai di filter maka data ini siap untuk diproses. Langkah selanjutnya adalah melakukan *sentiment analisis* pada data *testing*.

Dalam melakukan pengujian dataset ini, data dibagi menjadi dua yaitu data *training* dan data *testing*. Pada pengujian ini akan dihitung tingkat *accuracy*, *precision*, dan *recall*

terhadap data *testing* menggunakan Rapidminer untuk mengetahui seberapa akurat hasil pengujian data tersebut. Pada penelitian ini dilakukan satu kali pengujian untuk menghitung tingkat *accuracy*, *precision*, dan *recall* dengan rasio data training 55% dan data testing 45%, yaitu sebagai berikut:

Akurasi merupakan tingkat kedekatan antara hasil prediksi dengan hasil fakta. Prediksi adalah tingkat ketetapan antara informasi yang diminta oleh user dengan jawaban yang diberikan oleh sistem. Setelah data sudah terbagi kemudian ditentukan nilai prediksinya. Berikut ini nilai prediksi pada pengujian data testing 45%.

Tabel 3. *Accuracy Data Testing*

accuracy: 98.33%

	true Negatif	true Positif	class precision
pred. Negatif	126	5	96.18%
pred. Positif	0	169	100.00%
class recall	100.00%	97.13%	

Prediksi dari data sentimen negatif yang terdeteksi negatif adalah 126 dan yang salah adalah 0, sedangkan data sentimen positif yang terdeteksi positif adalah 169 dan yang salah adalah 5. Hasil dari *accuracy* Naïve Bayes dengan Rapidminer adalah 98.33%.

Adapun hasil *accuracy* diperoleh dengan cara menjumlah data *true positive* (TP) ditambah dengan *true negative* (TN) dibagi dengan total jumlah data *testing*. Berikut cara perhitungan manual dari *accuracy*:

$$\begin{aligned} \text{Akurasi} &= \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \\ &= \frac{169+126}{169+126+0+5} \times 100\% \\ &= \frac{295}{300} \times 100\% \\ &= 98.33\% \end{aligned}$$

Presisi merupakan pengukuran seberapa ketepatan informasi yang diberikan oleh sistem. Pada pengujian ini akan dihitung *precision* terhadap data testing menggunakan Rapidminer. Berikut ini hasil *precision* pengujian Algoritma Naive Bayes dengan data testing.

Tabel 4. Precision Data Testing

precision: 100.00% (positive class: Positif)

	true Negatif	true Positif	class precision
pred. Negatif	126	5	96.18%
pred. Positif	0	169	100.00%
class recall	100.00%	97.13%	

Prediksi dari data sentimen negatif yang terdeteksi negatif adalah 126 dan yang salah adalah 0, sedangkan data sentimen positif yang terdeteksi positif adalah 169 dan yang salah adalah 5. Hasil dari precision Naïve Bayes dengan Rapidminer adalah 100.00%.

Adapun hasil *precision* diperoleh dengan cara data *true positive* (TP) ditambah dengan *true negative* (TN) dibagi dengan total jumlah data *true positive* ditambah dengan *false positive* (FP). Berikut cara perhitungan manual dari *precision*:

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP+FP} \times 100\% \\
 &= \frac{169}{169+0} \times 100\% \\
 &= \frac{169}{169} \times 100\% \\
 &= 100.00\%
 \end{aligned}$$

Recall merupakan pengukuran tingkat keberhasilan pengambilan data yang relevan. Pada pengujian ini akan dihitung *recall* terhadap data *testing* menggunakan Rapidminer. Berikut ini hasil *recall* pengujian Algoritma Naive Bayes dengan data *testing*:

Tabel 5. Recall Data Testing

recall: 97.13% (positive class: Positif)

	true Negatif	true Positif	class precision
pred. Negatif	126	5	96.18%
pred. Positif	0	169	100.00%
class recall	100.00%	97.13%	

Prediksi dari data sentimen negatif yang terdeteksi negatif adalah 126 dan yang salah adalah 0, sedangkan data sentimen positif yang terdeteksi positif adalah 169 dan yang salah adalah 5. Hasil dari *recall* Nave Bayes dengan Rapidminer adalah 97.13%.

Adapun hasil *recall* diperoleh dengan cara data *true positive* (TP) dibagi dengan total jumlah data *true positive* (TP) ditambah dengan *false negative* (FN). Berikut cara perhitungan manual dari *recall*:

KESIMPULAN DAN REKOMENDASI

Berdasarkan penelitian yang telah dilakukan maka dapat disimpulkan bahwa penggunaan algoritma Naive Bayes dengan menggunakan *operator performances* menghasilkan tingkat dan hasil yang cukup baik. Nilai *accuracy*, *precision*, dan *recall* tertinggi dengan hasil nilai *accuracy* sebanyak 98.33%, *precision* sebanyak 100%, dan *recall* sebanyak 97.13% dan menghasilkan hasil opini berbentuk positif terhadap objek penelitian yaitu pembobolan data pada twitter sebesar 45%.

Dari penelitian yang dilakukan, dibuat rekomendasi untuk kedepannya yaitu agar bisa menambahkan fitur lainnya untuk mendapatkan hasil akurasi yang lebih tinggi, dan juga perlu untuk melakukan pengujian menggunakan metode atau algoritma lain agar dapat memperoleh perbandingan tingkat akurasi dalam melakukan analisis sentimen.

REFERENSI

- R. Aswandi, P. Rofifah, N. Muchsin, and M. Sultan, "Perlindungan Data Dan Informasi Pribadi Melalui Indonesian Data Protection System (IDPS)." [Online]. Available: <https://www.hukumonline.com/berita/baca/lt5d1c3962e01a4/perlindungandata-pribadi-tersebar>
- M. H. Rumlus and H. Hartadi, (2020). "Kebijakan Penanggulangan Pencurian Data Pribadi dalam Media Elektronik," Jurnal HAM, vol. 11, no. 2, p. 285, Aug. 2020, doi: 10.30641/ham.2020.11.285-299.
- Siti Nieke Noviyanti, (2022). "Viral! Hacker Bjorka Bikin Geger Indonesia Sebut Berhasil Bongkar Data Negara, Siapa Dia Sebenarnya?," journalsoreang.com, Sep. 20, 2022.
- Satria Dwi Kurniawan, (2022). "Bjorka vs Pemerintah," kompasiana.com, Sep. 15 2022.
- Agatha Vidya Nariswari, (2022). "Beragam Respons Pemerintah Terkait Ulah Hacker Bjorka, Bakal Bentuk Timsus," suara.com, Sep. 12, 2022.
- Devira Prastiwi, (2022). "6 Tanggapan Berbagai Pihak soal Kemunculan Hacker Bjorka," liputan6.com, Sep. 12, 2022.
- Chyntia Sami Bhayangkara, (2022). "Grusa-grusu Pemerintah Buru Bjorka, Pengalihan Kasus Sambo?," suara.com, Sep. 17, 2022.
- A. Akhtar, S. Javaid, S. Ejaz, and R. S. Satti, (2019). "Data Analysis of Educational Websites Using RapidMiner," . Available: <http://preston.edu.pk/courses.php>

- G. Gupta and G. S. Bhathal, (2018). "Sentiment Analysis Of English Tweets Using Data Mining". Sentiment Analysis. BookRix.
- W. A. Prabowo and C. Wiguna, (2021). "Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM," JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 5, no. 1, p. 149, Jan. 2021, doi: 10.30865/mib.v5i1.2604.
- Tsalis Annisa, (2021). "Mengenal peran sentiment analysis beserta cara kerjanya," *ekrut.com*, Nov. 10, 2021.
- N. Ruhjana, "Analisis Sentimen Terhadap Penerapan Sistem Plat Nomor Ganjil/Genap Pada Twitter Dengan Metode Klasifikasi Naive Bayes." [Online]. Available: www.situs.com
- A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," Jurnal Teknoinfo, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.
- A. Yadav, C. K. Jha, A. Sharan, and V. Vaish, "Sentiment analysis of financial news using unsupervised approach," in *Procedia Computer Science*, 2020, vol. 167, pp. 589–598. doi: 10.1016/j.procs.2020.03.325.
- R. Nofitri and N. Irawati, "Integrasi Metode Neive Bayes Dan Software Rapidminer Dalam Analisis Hasil Usaha Perusahaan Dagang," JURTEKSI (Jurnal Teknologi dan Sistem Informasi), vol. 6, no. 1, pp. 35–42, Dec. 2019, doi: 10.33330/jurteksi.v6i1.393.
- D. B. Srisulistiowati¹, M. Khaerudin², S. Rejeki³, and U. Bhayangkara Jakarta, "Sistem Informasi Prediksi Penjualan Alat Tulis Kantor Dengan Metode Fp-Growth (Studi Kasus Toko Koperasi Sekolah Bina Mulia)."
- O. Somantri, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Analisis Sentimen Penilaian Tempat Tujuan Wisata Kota Tegal Berbasis Text Mining", [Online]. Available: www.google.com/maps
- M. Allahyari et al., "A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques," Jul. 2017, [Online]. Available: <http://arxiv.org/abs/1707.02919>
- E. Sabna, "Penerapan Text Mining Untuk Pengelompokan Penelitian Dosen," Jurnal Ilmu Komputer, vol. 9, no. 2, pp. 161–164, Oct. 2020, doi: 10.33060/jik/2020/vol9.iss2.183.
- N. Ruhjana, "Analisis Sentimen Terhadap Penerapan Sistem Plat Nomor Ganjil/Genap Pada Twitter Dengan Metode Klasifikasi Naive Bayes." [Online]. Available: www.situs.com