

Analisis Sentimen pada Steam Review Menggunakan Multinomial Naïve Bayes dan Seleksi Fitur Gini Index Text

Ragil Haditira, Danang Triantoro Murdiansyah*, Widi Astuti

Fakultas Informatika, Telkom University, Bandung
E-mail Korespondensi: danangtri@telkomuniversity.ac.id

History Artikel

Diterima : 09 Juli 2024 Disetujui : 30 September 2024 Dipublikasikan : 26 Oktober 2024

Abstract

Video game is one of the entertainment medias chosen by most people today, many of which are played through computer devices. On computer devices, many video games are obtained through one of the game distribution platforms, namely Steam. However, Steam has several shortcomings, including those related to Steam reviews. On Steam reviews, you can see the rating of the game, but the rating does not really show the actual quality or condition of the game. As one example, there are users who give a high rating to a game, but in the comments column the user actually mentions the shortcomings of the game. To reduce or anticipate unclear reviews for users who want to try or buy the game, sentiment analysis on reviews is used. In this research, the output produced is information on the results of sentiment classification in filtering reviews, using the Multinomial Naïve Bayes algorithm and combined with the Gini Index feature selection. Sentiment classification is divided into two classes, namely recommended and not recommended classes. In this study, to test the sentiment classification system, a dataset containing reviews in the form of review sentences from Steam is used. The test results using Multinomial Naïve Bayes and Gini Index, can achieve the best accuracy of 60.29%.

Keywords: *Steam Review, sentiment classification, Multinomial Naïve Bayes, Gini Index.*

Abstrak

Video *game* adalah salah satu media hiburan yang banyak dipilih oleh kebanyakan orang saat ini, banyak diantaranya dimainkan melalui perangkat komputer. Pada perangkat komputer, video *game* banyak didapatkan melalui salah satu platform distribusi *game*, yaitu *Steam*. Namun *Steam* memiliki beberapa kekurangan, diantaranya adalah terkait *Steam review*. Pada *Steam review* dapat dilihat rating dari game namun rating tersebut tidak benar-benar menunjukkan kualitas atau kondisi game sebenarnya. Sebagai salah satu contoh, ada pengguna yang memberikan rating tinggi pada suatu *game*, namun pada kolom komentar pengguna tersebut malah menyebutkan kekurangan dari *game*. Untuk mengurangi atau mengantisipasi ulasan (*review*) yang kurang jelas bagi para pengguna yang ingin mencoba atau membeli *game* tersebut, maka digunakan analisis sentimen pada ulasan. Pada penelitian ini keluaran yang dihasilkan adalah informasi hasil klasifikasi sentimen dalam memfilter *review*, dengan menggunakan algoritma Multinomial Naïve Bayes dan dikombinasikan dengan seleksi fitur Gini Index. Klasifikasi sentimen dibagi menjadi dua kelas, yaitu kelas direkomendasikan dan tidak direkomendasikan. Pada penelitian ini untuk menguji sistem klasifikasi sentimen yang dibangun, digunakan dataset yang berisi ulasan-ulasan berupa kalimat *review* dari Steam. Hasil pengujian dengan menggunakan Multinomial Naïve Bayes dan Gini Index, dapat menghasilkan akurasi paling baik sebesar 60.29%.

Kata Kunci: *Steam Review, klasifikasi sentimen, Multinomial Naïve Bayes, Gini Index.*

How to Cite: Haditira, Ragil, dkk (2024). Analisis Sentimen pada Steam Review Menggunakan Multinomial Naïve Bayes dan Seleksi Fitur Gini Index. *KOMPUTEK : Jurnal Teknik Universitas Muhammadiyah Ponorogo*, Vol 8 (2): 71-78

© 2024 Universitas Muhammadiyah Ponorogo. All rights reserved

ISSN 2614-0985 (Print)

ISSN 2614-0977 (Online)

PENDAHULUAN

Seiring perkembangan teknologi saat ini banyak platform penjualan dan pembelian *game* pada perangkat komputer, diantaranya adalah *Steam*, *Origin*, *Uplay*, *Epic Games Store*, dan lain sebagainya. Untuk mengenali apakah suatu *game* itu bagus atau buruk, penting mengetahui ulasan dari para pengguna atau pemain *game* tersebut. Sebagian pengguna dan berlaku sebagai pengulas (*reviewer*) *game* menyampaikan ulasan mereka pada kolom komentar terkait *game* yang diunduh, dan sudah mencoba apakah layak dibeli atau tidak. Sebagian ulasan terlihat jelas tergolong dalam ulasan yang direkomendasikan atau tidak direkomendasikan, namun terdapat ulasan lain yang tidak jelas kategorinya.

Salah satu teknik untuk mengklasifikasikan opini atau ulasan adalah dengan analisis sentimen (Liu, 2022). Dengan analisis sentimen, kita dapat mengenali apakah ulasan itu termasuk dalam golongan atau kelas tertentu, misalnya termasuk yang merekomendasikan atau tidak, termasuk positif atau negatif, dan lainnya. Salah satu pendekatan untuk melakukan analisis sentimen adalah dengan menggunakan *machine learning*. Berdasarkan hasil penelitian yang telah dilakukan (B. & M., 2016), dapat diketahui bahwa analisis sentimen dengan pendekatan *machine learning* memiliki berbagai kelebihan, diantaranya yaitu menghasilkan akurasi yang tergolong tinggi bila dibandingkan dengan metode lainnya.

Berdasarkan *paper* survey mengenai survey analisis sentimen yang berjudul “*A Survey of Sentiment Analysis Based on Machine Learning*” (Lin & Luo, 2020), pada penelitian ini digunakan metode MNB (*Multinomial Naïve Bayes*), yang akan diimplementasikan pada *dataset Steam review*. MNB memiliki karakteristik komputasi yang efisien dan dapat bekerja dengan baik dalam ruang fitur yang berdimensi tinggi, yang mana hal tersebut juga merupakan karakteristik dari

klasifikasi teks. MNB juga menangani secara langsung frekuensi dari *term* kata dan tetap dapat bekerja dengan baik pada *dataset* yang tidak besar.

Dalam penelitian (Chen, Huang, Tian, & Qu, 2009), dijelaskan bahwa metode *Naïve Bayes* sangat sensitif terhadap seleksi fitur sehingga jumlah fitur yang digunakan akan sangat berpengaruh pada metode ini. Pada penelitian lain mengenai fitur seleksi (Shang et al., 2007), metode *Gini Index* dilakukan perubahan menjadi *Gini Index Text* dengan KNN (*K Nearest Neighbor*) dan SVM (Support Vector Machines) sebagai *classifier*, hasil dari penelitian tersebut termasuk sangat baik, yaitu mencapai rata-rata akurasi 98.5% untuk KNN dan 93% pada SVM. Dengan pertimbangan tersebut, pada penelitian ini untuk meningkatkan akurasi dari metode MNB, digunakan *Gini Index Text* (Park, Kwon, & Kwon, 2010). Dengan menggunakan penggabungan seleksi fitur *Gini Index Text* dan MNB, diharapkan dapat menambah akurasi dibandingkan jika hanya menggunakan MNB saja.

KAJIAN PUSTAKA

Terdapat banyak penelitian mengenai analisis sentimen tentang topik-topik seperti komentar media sosial, produk, politik, dan banyak lagi. Terdapat banyak metode untuk mengklasifikasikan teks. Banyak peneliti yang mencoba untuk melakukan kombinasi metode agar mencapai performansi yang lebih baik. Seperti pada penelitian (Sahu & Ahuja, 2016), dimana peneliti membandingkan 6 teknik klasifikasi dengan menggunakan ekstraksi fitur *word*, *bigram*, dan *word* ditambah *adjective* dengan seleksi fitur *Information Gain*. Dengan menggunakan data *movie review* IMDB mendapatkan hasil terbaik dengan teknik *Random Forest* yaitu 88.95 %, sedangkan *Naïve Bayes* mendapatkan akurasi terendah yaitu 54.77%, peneliti mengusulkan untuk menerapkan konsep *Natural Language*

Processing (NLP) lebih dalam untuk prediksi polaritas dokumen yang lebih baik.

Pada penelitian (Lux, 2012) membandingkan beragam seleksi fitur untuk algoritma *Naïve Bayes*. Tujuan penelitiannya adalah untuk menentukan seleksi fitur yang akurat serta stabil dengan menggunakan data Newsgroups dan Reuters-21578. Pada penelitian ini *pre-processing* data tidak ditampilkan sehingga proses data yang dilakukan tidak dijelaskan sampai ke seleksi fitur. Hasil akurasi tertinggi didapatkan dengan seleksi fitur Chi Sebesar 81% untuk data Reuters-21578 dan 84% untuk Newsgroups. Untuk hasil yang stabil didapatkan dengan menggunakan metode *Term Frequency* (TF) dan *Domain Frequency* (DF).

Dalam penelitian (Jia & Sun, 2012), peneliti tersebut membandingkan seleksi fitur *Gini Index*, *Weight* formula, dan MIDF (*Word Frequency Mutual Information*) dengan teknik klasifikasi KNN dan *Naïve Bayes*. Tujuan dari penelitian tersebut adalah untuk membandingkan dan memvalidasi dengan fitur bobot baru berdasarkan *Gini Index* pada performa dari teknik klasifikasi. Hasil dari penelitian ini *Gini Index* mendapatkan performansi yang sangat baik yaitu 75% dengan KNN dan 95% dengan *Naïve Bayes*.

Pada penelitian (P. L. Varela, Martins, Aguiar, & Figueiredo, 2013), dibandingkan klasifikasi MNB (*Multinomial Naïve Bayes*) dan SVM dengan beragam seleksi fitur sebanyak 18 seleksi fitur. Untuk *pre-processing* data tidak dijelaskan dengan detail proses apa saja yang dilakukan. Hasil penelitian menunjukkan performa yang sama dengan menggunakan *Gini Index*, WLLR (*Weighted Log Likelihood Ratio*) dan CET (*Cross Entropy for Text*) pada MNNB.

Pada penelitian (Iqbal, 2018), peneliti membandingkan *Naïve Bayes*, *Maximum Entropy*, dan *Gradient Descent* dengan seleksi fitur *Gini Index*. Teori tidak dijelaskan dengan detail dan rumus yang ditampilkan hanya *Gini*

Index saja dan data yang digunakan tidak dijelaskan. Penelitian ini mendapatkan hasil akurasi untuk *Gradient Descent* 70.10% dan untuk *Naïve Bayes* 65.24%.

Pada penelitian (P. d Varela, 2012; P. L. Varela et al., 2013), peneliti membandingkan *Gini Index Text* dengan beragam seleksi fitur pada data IMDB *dataset*, *Spanish dataset*, dan *Portuguese dataset*. Data diekstraksi terlebih dahulu dengan *n-gram* dan dengan menggunakan teknik klasifikasi MNB, SVM, dan *Weighted SVM* (WSVM). Dari kesimpulan didapatkan bahwa seleksi fitur *Gini Index*, *domain frequency*, CET, dan CHI mendapatkan performansi terbaik yaitu 90% dengan SVM, dan 92% dengan MNB. Untuk klasifikasi SVM merupakan algoritma yang cepat dan dapat bekerja pada dimensi yang tinggi. WSVM berkerja lebih baik dibandingkan SVM jika tidak menggunakan seleksi fitur.

METODE PENELITIAN

2.1. *Gini Index* (*Gini Index Text*)

Gini Index umumnya digunakan untuk memisahkan atribut. Metode ini pertama kali digunakan pada *decision tree* dan berhasil meningkatkan presisi dari klasifikasi ini. Telah banyak penelitian untuk meningkatkan metode *Gini Index*, salah satunya adalah GIT (*Gini Index Text*) yang diperkenalkan oleh (Park et al., 2010). GIT dibuat untuk bekerja pada dokumen dengan fitur yang sangat banyak. Dengan metode GIT ini dapat mereduksi fitur dari *subset* fitur, serta dapat juga mempertahankan banyak fitur representatif. Terdapat beberapa formula GIT, salah satunya didefinisikan secara umum sebagai berikut.

$$GiniText(w) = \sum_{i=1}^m P(c_i | w)^2 \quad (1)$$

Dengan $P(c_i | w)$ adalah distribusi kategori atau kelas i pada munculnya kata, dan W adalah fitur (kata).

2.2. Naïve Bayes

Pada analisis sentimen dan klasifikasi teks, algoritma *Naïve Bayes* (Berrar, 2019) merupakan algoritma yang banyak digunakan. Terdapat beberapa jenis *Naïve Bayes*, yaitu *Multinomial Naïve Bayes*, *Naïve Bayes Multivariate Bernoulli*, *Naïve Bayes Gaussian*. Pada *Naïve Bayes* suatu fitur dianggap tidak bergantung pada fitur yang lainnya, yang dinamakan dengan asumsi naïve. *Naïve Bayes* didefinisikan secara umum sebagai berikut.

$$P = \frac{P(X|C)P(C)}{P(X)} \quad (2)$$

dengan X adalah himpunan dari fitur $X = \{x_1, \dots, x_j\}$ dan C menyatakan himpunan dari kelas atau kategori $C = \{c_1, \dots, c_k\}$.

2.2.1. Multinomial Naïve Bayes

Naïve Bayes classifier merupakan sebuah pengklasifikasi berdasarkan probabilitas sederhana yang mengaplikasikan teorema Bayes dengan asumsi ketidaktergantungan. Sebagaimana dituliskan sebelumnya, salah satu jenis *Naïve Bayes* adalah MNB (Chebil, Wedyan, Alazab, Alturki, & Elshaweesh, 2023). MNB memperhitungkan jumlah kata dalam dokumen dan mengasumsikan independensi kemunculan kata dalam dokumen. Dengan asumsi tersebut, kemungkinan tiap kejadian kata dalam dokumen adalah bebas tidak memperhitungkan urutan kata dan konteks kata dalam dokumen (Butar, Fauzi, & Indriati, 2019). Formula MNB adalah sebagai berikut.

$$P(w_i, c_k) = \frac{1 + \sum_{j=1}^N N_{ij} \delta_{jk}}{|V| + \sum_{i=1}^{|V|} \sum_{n=1}^N N_{in} \delta_{jk}}$$

dengan N adalah jumlah dari dokumen. bernilai 1 jika dokumen ke j milik kelas

dan bernilai 0 jika sebaliknya. adalah jumlah fitur yang terjadi pada dokumen ke $j + 1$ untuk menghindari nilai nol.

$|V|$ adalah jumlah kosakata dari sebuah dokumen ulasan.

2.3. Metrik Evaluasi

Setelah mendapatkan keluaran dari klasifikasi, keluaran tersebut akan diukur berdasarkan metrik evaluasi untuk mengetahui seberapa baik sistem yang dibangun. Salah satu metrik evaluasi yang digunakan adalah *confusion matrix*. Pada *confusion matrix* terdapat:

- TP (True Positive): jumlah ulasan positif yang diklasifikasikan positif.
- TN (True Negatif): jumlah ulasan negatif yang diklasifikasikan negatif.
- FP (False Positif): jumlah ulasan negatif yang diklasifikasikan positif.
- FN (False Negative): jumlah ulasan positif yang diklasifikasikan negatif.

Selain *confusion matrix*, digunakan juga metrik evaluasi lain, yaitu akurasi, *precision*, *recall*, dan *F1-Score* (Vujović & others, 2021).

Akurasi adalah perbandingan jumlah prediksi yang benar dengan total jumlah prediksi yang dibuat. Rumus akurasi adalah sebagai berikut.

$$\text{Akurasi} = \frac{TP + TN}{TP + FP + FN + TN} \quad (4)$$

Precision adalah seberapa banyak prediksi positif yang benar dari semua prediksi positif yang dibuat oleh model. Rumus *precision* adalah sebagai berikut.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

Recall adalah seberapa banyak prediksi positif yang berhasil ditemukan dari semua kasus positif yang sebenarnya ada. Rumus *recall* adalah sebagai berikut.

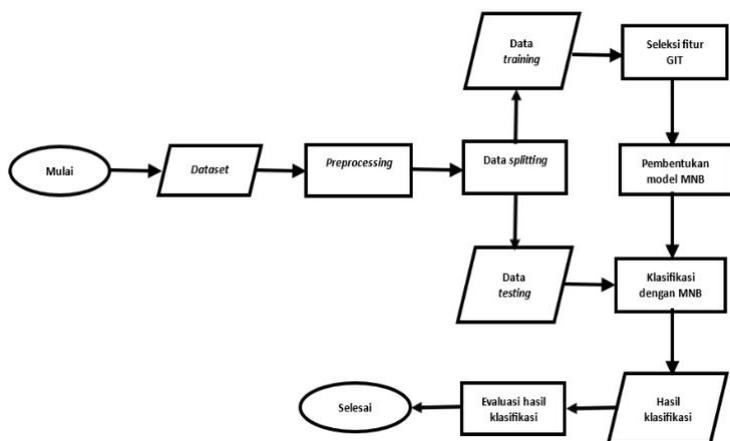
$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

F1 Score adalah *harmonic mean* dari *precision* dan *recall*, dan memberikan satu angka untuk mengevaluasi keseimbangan antara keduanya.

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

2.4. Sistem yang Dibangun

Rancangan sistem yang dibangun pada penelitian ini dapat dilihat pada diagram alir di Gambar 1. *Dataset* yang sudah didapatkan akan diproses melalui proses *pre-processing*, yaitu *tokenization*, *casefolding*, *stopword removal*, *stemming*, dan *lemmatization* (Chai, 2023). Kemudian *dataset* akan dibagi menjadi data *training* dan *data testing*. Untuk data *training*, fitur-fitur akan diseleksi dengan metode *Gini Index Text* (GIT). Setelah fitur-fitur diseleksi, model MNB (*Multinomial Naïve Bayes*) akan dibentuk atau dilatih dengan menggunakan data *training*. Setelah model MNB terbentuk, proses *testing* atau klasifikasi akan dilakukan dengan menggunakan data *testing*. Pada akhir proses, hasil klasifikasi akan dievaluasi dengan menggunakan metrik evaluasi akurasi, *precision*, *recall*, dan *F1-Score*.



Gambar 1. Diagram Alir Sistem Klasifikasi *Steam Review*.

HASIL DAN PEMBAHASAN

Dataset yang digunakan pada penelitian ini terdiri dari data kelas positif (ulasan merekomendasikan *game*) dan data kelas negatif (ulasan tidak merekomendasikan *game*) dari *Steam review*. Jumlah data yang digunakan adalah 4000 *records*, yang kemudian untuk data *training* dan *testing* akan dibagi berdasarkan rasio (data *training* berbanding data *testing*) 90:10, 80:20, dan

70:30.

Skenarion pengujian yang dilakukan terdiri dari 2 skenario utama, yang pertama yaitu pengujian dengan menggunakan MNB tanpa menggunakan seleksi fitur GIT, dan yang kedua pengujian dengan menggunakan MNB dengan seleksi fitur GIT. Masing-masing skenario dilakukan sebanyak 5 kali percobaan dan diambil rata-ratanya.

3.1. Hasil Pengujian Tanpa GIT (*Gini Index Text*)

Hasil pengujian menggunakan MNB dengan tanpa seleksi fitur GIT dapat dilihat pada Tabel 1.

Tabel 1. Hasil Pengujian Tanpa GIT.

Akurasi	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	Rasio
81.12%	81.95%	80.08%	81.00%	90:10
80.99%	83.02%	78.07%	80.46%	80:20
81.38%	83.30%	78.94%	81.04%	70:30

Berdasarkan Tabel 1, rata-rata hasil akurasi terbaik dengan menggunakan 5 kali percobaan menghasilkan rata-rata akurasi terbaik sebesar 81.38% yaitu dari rasio 70:30, rata – rata *precision* terbaik sebesar 83.30% yaitu dari rasio 70:30, rata-rata *recall* terbaik sebesar 80.08% yaitu dari rasio 90:10, dan rata-rata *F1-Score* sebesar 81.04% yaitu dari rasio 70:30.

3.2. Hasil Pengujian dengan GIT.

Hasil pengujian menggunakan MNB dengan seleksi fitur GIT dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pengujian dengan GIT.

Akurasi	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Rasio
60.29%	56.65%	90.44%	69.64%	90:10
59.69%	56.33%	91.46%	69.71%	80:20
58.33%	55.57%	89.15%	68.45%	70:30

Berdasarkan Tabel 2, rata-rata hasil akurasi

terbaik dengan menggunakan 5 kali percobaan menghasilkan rata-rata akurasi terbaik sebesar 60.29% yaitu dari rasio 90:10, rata – rata *precision* terbaik sebesar 56.65% yaitu dari rasio 90:10, rata-rata *recall* terbaik sebesar 91.46% yaitu dari rasio 80:20, dan rata-rata *F1-Score* sebesar 69.71% yaitu dari rasio 80:20.

Dari hasil pengujian pada penelitian ini, dapat dilihat bahwa performa sistem dengan menggunakan seleksi fitur GIT secara umum hasilnya lebih rendah daripada tanpa menggunakan seleksi fitur GIT. Dilihat dari metrik evaluasi akurasi, *precision*, dan *F1-Score*, performa MNB tanpa GIT lebih rendah daripada MNB dengan GIT. Hal tersebut kemungkinan disebabkan diantaranya karena *overfitting*. Selain itu, dapat juga disebabkan karena jika menggunakan seleksi fitur, secara tidak sengaja menghapus fitur yang saling melengkapi dalam klasifikasi. Hal ini dapat menyebabkan berkurangnya daya prediksi karena MNB bergantung pada kombinasi fitur tertentu untuk memprediksi kelas (Pudjihartono, Fadason, Kempa-Liehr, & O’Sullivan, 2022).

KESIMPULAN

Pada penelitian ini telah dibangun sistem klasifikasi ulasan untuk mengetahui sentimen dengan *dataset* dari *Steam review*. Penggunaan metode MNB (*Multinomial Naïve Bayes*) menghasilkan hasil terbaik pada scenario tanpa menggunakan fitur seleksi GIT (*Gini Index Text*), yaitu menghasilkan berturut-turut akurasi, *precision*, dan *F1-score* adalah 81.38%, 83.30%, dan 81.04%. Penggunaan metode MNB dengan seleksi fitur GIT secara umum tidak menghasilkan performa yang lebih baik daripada metode MNB tanpa seleksi fitur GIT. Hal tersebut kemungkinan disebabkan diantaranya karena *overfitting*, dan juga karena hilangnya fitur yang saling melengkapi dalam klasifikasi yang menyebabkan berkurangnya daya prediksi MNB.

DAFTAR PUSTAKA

- B., V., & M., B. (2016). Analysis of Various Sentiment Classification Techniques. *International Journal of Computer Applications*, 140(3), 22–27. <https://doi.org/10.5120/ijca2016909259>
- Berrar, D. (2019). *Bayes’ theorem and naive Bayes classifier*.
- Butar, T. M. E. Y. B., Fauzi, M. A., & Indriati, I. (2019). Penentuan Rating Review Film Menggunakan Metode Multinomial Naive Bayes Classifier dengan Feature Selection Berbasis Chi-Square dan Galavotti-Sebastiani-Simi Coefficient. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(1), 447–453.
- Chai, C. P. (2023). Comparison of Text Preprocessing Methods. *Natural Language Engineering*, 29(3), 509–553. <https://doi.org/10.1017/S1351324922000213>
- Chebil, W., Wedyan, M., Alazab, M., Alturki, R., & Elshaweesh, O. (2023). Improving semantic information retrieval using multinomial naive Bayes classifier and Bayesian networks. *Information*, 14(5), 272.
- Chen, J., Huang, H., Tian, S., & Qu, Y. (2009). Feature selection for text classification with Naïve Bayes. *Expert Systems with Applications*, 36(3, Part 1), 5432–5435. <https://doi.org/10.1016/j.eswa.2008.06.054>
- Iqbal, F. (2018). Sentiment Analysis Using Ensemble Learners and Gini Index. *International Journal of Engineering and Techniques*, 4(2), 586–591.
- Jia, X., & Sun, J. (2012). An Improved Text Classification Method Based on Gini Index. *Journal of Theoretical & Applied Information Technology*, 43(2).
- Lin, P., & Luo, X. (2020). A Survey of Sentiment Analysis Based on Machine Learning. In X. Zhu, M. Zhang, Y. Hong, & R. He (Eds.), *Natural Language Processing and Chinese*

Computing (pp. 372–387). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-60450-9_30

International Journal of Advanced Computer Science and Applications, 12(6), 599–606.

- Liu, B. (2022). *Sentiment Analysis and Opinion Mining*. Springer Nature.
- Lux, E. (2012). *Feature selection for text classification with Naive Bayes*. Retrieved from <https://dspace.cuni.cz/handle/20.500.11956/46007>
- Park, H., Kwon, S., & Kwon, H.-C. (2010). Complete Gini-Index Text (GIT) feature-selection algorithm for text classification. *The 2nd International Conference on Software Engineering and Data Mining*, 366–371.
- Pudjihartono, N., Fadason, T., Kempa-Liehr, A. W., & O’Sullivan, J. M. (2022). A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Frontiers in Bioinformatics*, 2, 927312. <https://doi.org/10.3389/fbinf.2022.927312>
- Sahu, T. P., & Ahuja, S. (2016). Sentiment analysis of movie reviews: A study on feature selection & classification algorithms. *2016 International Conference on Microelectronics, Computing and Communications (MicroCom)*, 1–6. <https://doi.org/10.1109/MicroCom.2016.7522583>
- Shang, W., Huang, H., Zhu, H., Lin, Y., Qu, Y., & Wang, Z. (2007). A novel feature selection algorithm for text categorization. *Expert Systems with Applications*, 33(1), 1–5. <https://doi.org/10.1016/j.eswa.2006.04.001>
- Varela, P. d. (2012). *Sentiment Analysis*. Lisboa: Instituto Superior Técnico.
- Varela, P. L., Martins, A., Aguiar, P., & Figueiredo, M. (2013). An empirical study of feature selection for sentiment analysis. *9th Conference on Telecommunications, Conftel, Castelo Branco*.
- Vujović, Ž. & others. (2021). Classification Model Evaluation Metrics.