

KOMPARASI ALGORITMA KLASIFIKASI UNTUK ANALISIS SENTIMEN *REVIEW* FILM BERBAHASA ASING

Satria Wira Yudha¹⁾, Mochamad Wahyudi²⁾

Program Studi Magister Ilmu Komputer
Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri
(STMIK Nusa Mandiri)

email: ¹⁾satriayudha678@gmail.com, ²⁾wahyudi@nusamandiri.ac.id

Jl. Damai No. 8, Warung Jati Barat (Margasatwa), Pasar Minggu, RT.1/RW.1,
Ragunan, Pasar Minggu, South Jakarta City, Jakarta 12540, Indonesia

Abstrak

Industri film secara global terus mengalami perkembangan, baik dari jumlah film yang dihasilkan, jumlah penonton, maupun jumlah perputaran uangnya. Data yang didapatkan dari *National Association of Theater Owners* (asosiasi pemilik bioskop di Amerika Serikat) menunjukkan bahwa ada perkembangan dari tahun 1987 dimana tiket bioskop terjual sebanyak 1,09 miliar tiket, menjadi 1,314 miliar tiket pada tahun 2016 untuk penjualan di wilayah Amerika Serikat dan Kanada. Saat ini, pendapat khalayak umum menjadi sumber yang penting dalam pengambilan keputusan seseorang akan suatu produk. Pendapat orang-orang dapat mengurangi ketidakpastian terhadap suatu produk tertentu dan membantu konsumen menyimpulkan kualitas suatu produk tertentu. Analisis sentimen adalah proses yang bertujuan untuk menentukan isi dari *dataset* yang berbentuk teks bersifat positif, negatif atau netral. Algoritma klasifikasi seperti Naïve Bayes (NB), Support Vector Machine (SVM), dan k-Nearest Neighbour (k-NN) diusulkan oleh banyak peneliti untuk digunakan pada analisis sentimen *review* film. Setelah melakukan beberapa kali pengujian hasil komparasi algoritma yang diujicobakan di beberapa *dataset* baik yang, berbahasa Inggris dan Jerman, SVM mendapatkan hasil yang terbaik dengan akurasi 91,92% dan AUC 0,981 pada *dataset* berbahasa Inggris, sedangkan pada dataset berbahasa Jerman menghasilkan akurasi 90,96% dan AUC 0,978.

Kata kunci: Teks Mining, Data Mining, Analisa Sentimen, Klasifikasi Teks

1. PENDAHULUAN

Industri film secara global terus mengalami perkembangan, baik dari jumlah film yang dihasilkan, jumlah penonton, maupun jumlah perputaran uangnya. Data yang didapatkan dari *National Association of Theater Owners* (asosiasi pemilik bioskop di Amerika Serikat) menunjukkan bahwa ada perkembangan dari tahun 1987 dimana tiket bioskop terjual sebanyak 1,09 miliar tiket, menjadi 1,314 miliar tiket pada tahun 2016 untuk penjualan di wilayah Amerika Serikat dan Kanada. Sedangkan untuk pendapatan *box office* di wilayah Amerika Serikat dan Kanada pada tahun 1987 sebesar \$4,25 miliar menjadi \$11,372 miliar pada tahun 2016, angka ini menunjukkan peningkatan yang sangat signifikan.

Agar dapat menonton film yang sesuai dengan keinginannya, banyak masyarakat yang membaca *review* dari suatu film sebelum membeli produk atau menonton suatu film tertentu. Pendapat orang-orang dapat mengurangi ketidakpastian terhadap suatu produk tertentu dan membantu konsumen menyimpulkan kualitas suatu produk tertentu. Saat ini, pendapat khalayak umum menjadi sumber yang penting dalam pengambilan keputusan seseorang akan suatu produk. Oleh karena itu proses analisa sentimen masyarakat menjadi sangat penting agar

dapat mengetahui seberapa akuratkah suatu *review* film tersebut.

Shweta Rana dan Archana Singh (Rana & Singh, 2016) melakukan penelitian tentang analisis perbandingan orientasi sentimen menggunakan teknik Support Vector Machine dan Naïve Bayes pada ulasan pengguna film. Hasil evaluasi algoritma Support Vector Machine dan Naïve Bayes di dataset ulasan pengguna film, ditunjukkan bahwa drama film memiliki tingkat akurasi yang tinggi (SVM 87,50%, NB 80%) diantara genre film yang lainnya (SVM : Action 75%, Adventure 72,50%, Drama 87,50%, Romantic 80%) (NB : Action 70%, Adventure 67,50%, Drama 80%, Romantic 77,50%). Hasil evaluasi menunjukkan bahwa teknik Support Vector Machine memiliki akurasi terbaik daripada Naïve Bayes.

Berdasarkan penelitian oleh Penelitian analisis sentimen juga dilakukan oleh Della Fitriani Budiono, Anto Satriyo Nugroho dan Afrizal Doewes (Budiono et al., 2017), mereka melakukan penelitian analisis sentimen terhadap Pemilu Gubernur DKI Jakarta 2017 diketahui bahwa hasil akurasi Mesin Multinomial Naïve Bayes dan Support Vector Machine tidak begitu berbeda satu sama lain yaitu 67,08% untuk Multinomial Naïve Bayes dan 68,87% untuk Support Vector Machine.

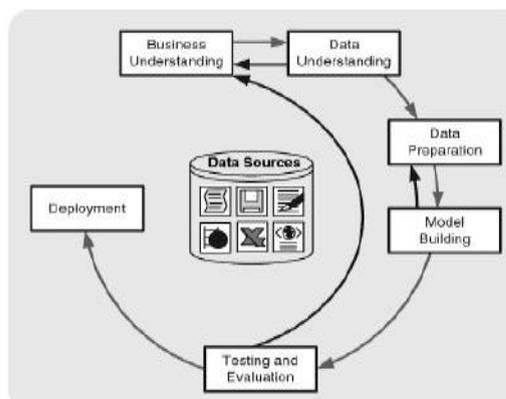
Tirana Noor Fatyanosa dan Fitra A. Bachtiar (Fatyanosa & Bachtiar, 2017) melakukan penelitian tentang perbandingan metode klasifikasi pada analisis sentimen media sosial Indonesia. Hasil percobaan menunjukkan bahwa Naive Bayes menghasilkan precision tertinggi (Netral 69,7%, Positif 23,6%, Negatif 6,7%), recall tertinggi (Netral 73,1%, Positif 69,8%, Negatif 9,9%), dan akurasi tertinggi (Netral 71,6%, Positif 56%, Negatif 21,6%) untuk sentimen netral dan positif. Tapi, Naïve Bayes tidak menghasilkan hasil yang baik untuk sentimen negatif.

Ghulam Asrofi Buntoro (Buntoro, 2017) melakukan analisis sentimen pada Pemilihan Gubernur DKI Jakarta 2017 Akurasi tertinggi didapat saat menggunakan metode klasifikasi Naïve Bayes Classifier (NBC), dengan nilai rata-rata akurasi mencapai 95%, nilai presisi 95%, nilai recall 95% nilai TP rate 96,8% dan nilai TN rate 84,6%.

Berdasarkan penelitian diatas, identifikasi masalah dalam penelitian ini adalah bagaimana menemukan metode algoritma yang memberikan model terbaik untuk mengklasifikasi dan menganalisa *review* film dan menganalisa sentimen *review* film apakah metode Naive Bayes, Support Vector Machine atau k-Nearest Neighbour yang memberikan hasil terbaik?

2. METODOLOGI PENELITIAN

Pemilihan metode data mining perlu dilakukan dengan tujuan untuk dapat memilih teknik data mining apa yang cocok untuk diterapkan, dan digunakan juga sebagai acuan untuk keseluruhan proses *text mining* dari *review* film. Saat ini terdapat beberapa metode yang bisa dipakai untuk melakukan *data mining*, diantaranya model CRISP, *generic model*, dan lain sebagainya. Proses *data mining* yang akan digunakan pada makalah yang akan dibuat adalah model referensi CRISP-DM, seperti terlihat pada gambar yang ada dibawah ini:



Gambar 1: CRISP-DM

Model CRISP-DM berisikan siklus kegiatan *data mining* yang terdiri atas enam fase. Gambar 1 menunjukkan seluruh tahap yang ada pada model CRISP-DM. Salah satu keuntungan

dari model ini adalah, tahapan fase dari model bukanlah sebuah tahapan yang kaku. Perpindahan maju dan mundur antara tiap fase yang berbeda bisa selalu dilakukan. Hal ini sesuai dengan sifat alami dari data mining itu sendiri, dimana proses *data mining* tidak selesai saat sebuah hasil di temukan, sebab proses *data mining* merupakan sebuah proses pembelajaran terus menerus. Pada penelitian kali menggunakan aplikasi yang bernama RapidMiner guna menguji beberapa algoritma komparasi seperti Naive Bayes (NB), Support Vector Machine (SVM), dan k Nearest Neighbour (k-NN)

3. HASIL DAN PEMBAHASAN

Data penelitian yang digunakan adalah sampel dari *review* berbahasa inggris beberapa film teratas Hollywood di IMDb, diantaranya Avengers: Infinity War, Incredibles 2 dan Mission: Impossible – Fallout. Selain merupakan film paling populer pada tahun 2018 (menurut website IMDb.com), pertimbangan memilih ketiga film tersebut adalah:

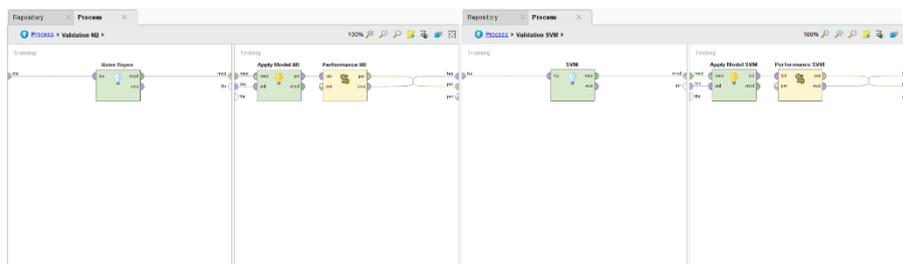
1. Avengers: Infinity War adalah film yang paling hype di tahun 2018 dengan penjualan \$2,003,241,872 dan jumlah *review* film 3.381 *user*
2. Incredibles 2 film kartun terpopuler di tahun 2018 dengan penjualan \$1,194,965,673 dan jumlah *review* film 674 *user*
3. Mission: Impossible - Fallout adalah salah satu film action terpopuler di tahun 2018 dengan penjualan \$777,937,480 dan jumlah *review* film 1.092 *user*

Adapun untuk pengambilan sampel dari populasi masing – masing *review* film menggunakan teknik Slovin dengan batas *error* sebesar 5 %, maka dari itu didapatkan:

Table 1: Tabel Detail *Review* Film

| No. | Nama Film | Jumlah <i>Review</i> | Total Sample | Sample <i>Review</i> Negatif | Sample <i>Review</i> Positif |
|--------------|-------------------------------|----------------------|--------------|------------------------------|------------------------------|
| 1. | Avengers: Infinity War | 3.381 | 358 | 179 | 179 |
| 2. | Incredibles 2 | 674 | 251 | 126 | 126 |
| 3. | Mission: Impossible - Fallout | 1.092 | 293 | 146 | 146 |
| TOTAL | | 8.606 | 902 | 451 | 451 |

Berikut ini merupakan contoh gambaran tahap desain model NB, SVM dan k-NN yang digunakan untuk proses awal pengolahan data:





Gambar 2: Desain Model Perbandingan Algoritma

Confusion matrix merupakan sebuah metode untuk evaluasi yang menggunakan tabel matrix. Berikut ini tabel matrix yang digunakan dalam data mining:

Table 2: Tabel Confusion Matrix

| <i>Correct</i> | <i>Classified as</i> | |
|----------------|-----------------------|-----------------------|
| | + | - |
| + | <i>True positive</i> | <i>False negative</i> |
| - | <i>False positive</i> | <i>True negative</i> |

Nilai akurasi dapat diperoleh dengan Persamaan 1. Nilai presisi menggambarkan jumlah data kategori positif yang diklasifikasikan secara benar dibagi dengan total data yang diklasifikasi positif. Presisi dapat diperoleh dengan Persamaan 2. Sementara itu, recall menunjukkan berapa persen data kategori positif yang terklasifikasikan dengan benar oleh sistem. Nilai recall diperoleh dengan Persamaan 3.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} * 100\% \quad (1)$$

$$Presisi = \frac{TP}{FP + TP} * 100\% \quad (2)$$

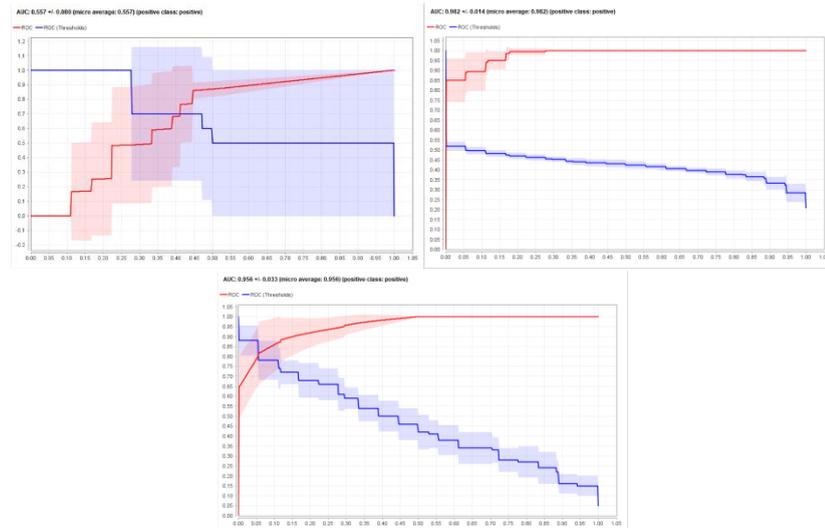
$$Recall = \frac{TP}{FN + TP} * 100\% \quad (3)$$

Model validasi biasa menggunakan 10 fold cross validation untuk learning dan pengujian data. Ini berarti membagi training data menjadi 10 bagian yang sama dan kemudian melakukan proses learning 10 kali. Peneliti juga harus mempertimbangkan terlebih nilai Area under curve (AUC) sebagai salah satu indikator untuk mengevaluasi kinerja dari classifier.

Table 3: Tabel Nilai AUC

| <i>AUC</i> | <i>Meaning</i> | <i>Symbol</i> |
|--------------------|---------------------------------|---------------|
| 0.90 – 1.00 | <i>excellent classification</i> | ↑ |
| 0.80 – 0.90 | <i>good classification</i> | ↗ |
| 0.70 – 0.80 | <i>fair classification</i> | → |
| 0.60 – 0.70 | <i>poor classification</i> | ↘ |
| < 0.60 | <i>failure</i> | ↓ |

ROC Curve adalah cara lain untuk menguji kinerja pengklasifikasi. Sebuah grafik ROC adalah plot dengan tingkat positif salah (FP) pada sumbu X dan tingkat positif benar (TP) pada sumbu Y. Titik (0,1) adalah klasifikasi sempurna yang mengklasifikasikan semua kasus positif dan kasus negatif dengan benar, karena tingkat positif salah (FP) adalah 0 (tidak ada), dan tingkat positif benar (TP) adalah 1. Titik (0,0) merupakan sebuah klasifikasi yang memprediksi setiap kasus menjadi negatif, sedangkan titik (1,1) sesuai dengan sebuah klasifikasi yang memprediksi setiap kasus menjadi positif. Titik (1,0) adalah klasifikasi yang tidak benar untuk semua klasifikasi.



Gambar 3: Contoh Hasil Kurva ROC
Setelah dilakukan percobaan menggunakan aplikasi RapidMiner, didapatkan hasil:

Tabel 10: Rekapitulasi Keseluruhan Hasil Percobaan

| No. | Dataset | NB | | | SVM | | | k-NN | | |
|-----|---|----------|-----------|---------|----------|-----------|---------|----------|-----------|---------|
| | | Accuracy | Precision | Recall | Accuracy | Precision | Recall | Accuracy | Precision | Recall |
| 1 | Review film | | | | | | | | | |
| | Avengers: Infinity War | 77.92% | 76.59% | 82.21 % | 91.63% | 95.37% | 87.83 % | 85.47% | 80.79% | 93.82 % |
| 2 | Review film | | | | | | | | | |
| | Incredibles 2 | 67.51% | 71.63% | 61.10 % | 80.54% | 87.72% | 73.07 % | 76.18% | 71.33% | 88.88 % |
| 3 | Review film | | | | | | | | | |
| | Mission: Impossible – Fallout | 80.77% | 82.62% | 78.18 % | 89.41% | 90.29% | 89.17 % | 83.23% | 77.84% | 95.14 % |
| 4 | Kompilasi review film Avengers: Infinity War, Incredibles 2 & Mission: Impossible - Fallout | | | | | | | | | |
| | Review film | 77.82% | 76.52% | 80.48 % | 88.14% | 90.49% | 85.38 % | 81.16% | 75.2% | 93.35 % |
| 5 | Review film | | | | | | | | | |
| | Avengers: Infinity War berbahasa Jerman | 83.21% | 77.44% | 93.76 % | 93.57% | 95.14% | 92.09 % | 77.07% | 70.09% | 96.66 % |

4. KESIMPULAN & SARAN

4.1 Kesimpulan

Kesimpulan yang didapatkan dari penelitian ini adalah bahwa, analisa review dari berbagai judul film baik yang berbahasa inggris maupun Jerman menunjukkan bahwa metode algoritma yang terbaik adalah metode SVM dibandingkan dengan metode NB dan k-NN, hal ini ditunjukkan dari beberapa hal diantaranya yaitu:

1. Nilai AUC

Pada perbandingan nilai AUC, metode SVM menduduki peringkat pertama dengan menunjukkan hasil tertinggi dengan nilai AUC 0,986 dengan kategori excellent classification.

2. Nilai Accuracy

Pada perbandingan nilai accuracy, metode SVM menduduki peringkat pertama dengan nilai accuracy yang didapatkan adalah 93,57% pada dataset review berbahasa Jerman film Avengers: Infinity War.

3. Nilai Precision

Pada perbandingan nilai precision, metode SVM menduduki peringkat pertama dengan nilai precision yang didapatkan adalah 95,37% pada dataset review film Avengers: Infinity War.

4. Nilai Recall

Pada perbandingan nilai recall, metode k-NN menduduki peringkat pertama dengan nilai recall yang didapatkan adalah 96.66% pada dataset review film berbahasa Jerman film Avengers: Infinity War.

Dari beberapa hal diatas, bisa didapatkan kesimpulan bahwa metode SVM adalah yang terbaik untuk digunakan pada analisa sentimen review film, hal ini tidak hanya pada dataset yang berbahasa Inggris, tetapi juga pada dataset berbahasa Jerman.

4.2 Saran-Saran

Adapun saran-saran yang dapat penulis sampaikan untuk penelitian-penelitian selanjutnya adalah:

1. Menggunakan metode pengklasifikasian lain sehingga bisa dilakukan penelitian yang berbeda dari yang sudah ada.
2. Menggunakan metode pemilihan fitur yang lain seperti Chi Square, Mutual Information, unigram, unigram + bigram, unigram + Part of Speech (POS), dan lain-lain agar hasilnya bisa lebih dibandingkan lagi.
3. Menggunakan dataset dari review film dengan bahasa yang lebih beragam lagi agar bisa menguatkan bahwa memang benar metode tersebut bisa digunakan di berbagai bahasa. Terutama nya adalah review film dalam bahasa yang kata-katanya diluar dari huruf alphabet, misalnya bahasa Arab, China, Jepang dan Korea.

Referensi

- Budiono, D.F., Nugroho, A.S., & Doewes. A. (2017). Twitter sentiment analysis of DKI Jakarta's gubernatorial election 2017 with predictive and descriptive approaches. 2017 International Conference on Computer, Control, Informatics and its Applications (IC3INA).
- Buntoro, G.A. (2018). Sentiment Analysis To Prediction Dki Jakarta Governor 2017 On Indonesian Twitter. International Journal of Science, Engineering and Information Technology, 2018.
- Chandani, V., Wahono, R.S., Purwanto. (2015). Komparasi Klasifikasi Machine Learning dan Feature Selection pada Analisis Sentimen Review Film. Journal of Intelligent System, Vol. 1, No.1. 56 – 60.
- Fatyanosa, T.N. & Bachtiar, F.A. (2017). Classification method comparison on Indonesian social media sentiment analysis. 2017 International Conference on Sustainable Information Engineering and Technology (SIET).
- Rana, S & Singh, A. (2016). *Comparative analysis of sentiment orientation using SVM and Naive Bayes techniques*. Dehradun: 2016 2nd International Conference on Next Generation Computing Technologies (NGCT) Date of Conference: 14-16 Oct. 2016.
- Wahyudi, M. & Putri, D. N. (2016). Algorithm Application Support Vector Machine With Genetic Algorithm Optimization Technique For Selection Features For The Analysis Of Sentiment On Twitter. Journal of Theoretical and Applied Information Technology 29th February 2016. Vol.84. No.3. ISSN: 1992-8645.