

# Analisis Sentiment pada Ulasan Film Dengan Optimasi Ensemble Learning

Andreyestha<sup>1</sup>, Agus Subekti<sup>2</sup>

<sup>1</sup>Universitas Bina Sarana Informatika  
e-mail: andreyes2505@bsi.ac.id

<sup>2</sup>Pusat Penelitian Elektronik dan Telekomunikasi  
Lembaga Ilmu Pengetahuan Indonesia (LIPI)  
e-mail: agus.subekti@lipi.go.id

## Abstract

Dalam dunia hiburan khususnya film, kini situs web ulasan film menjadi media bagi orang-orang untuk memberikan penilaian mengenai seberapa bagus film tersebut. Mereka tidak harus menjadi pakar dalam dunia perfilman untuk menilai kualitas dari film yang mereka saksikan, semua orang dapat memberikan penilaian. Sentimen yang ditemukan dalam komentar, umpan balik atau kritik memberikan indikator yang berguna untuk berbagai tujuan dan dapat dikategorikan berdasarkan polaritas, polaritas tersebut cenderung akan dicari tahu apakah secara keseluruhan positif atau negatif. Algoritma Naïve Bayes dan Random Forest merupakan algoritma yang dapat memberikan hasil analisa klasifikasi sesuai yang diharapkan pada penelitian ini, analisa akan dilakukan dengan membandingkan beberapa kombinasi algoritma untuk diuji pada Polarity Dataset 2.0 dari Cornell University, diantaranya yaitu Algoritma tersebut akan dikombinasikan dengan seleksi fitur Chi Square, Adaboost, dan Voting. Dari hasil pengujian yang didapat algoritma AdaBosst dan Voting mampu meningkatkan akurasi dari metode Naïve Bayes (NB) and Random Forest (RF). Model yang diusulkan dengan Chi Square + Voting 2 (RF + SVM) memiliki nilai akurasi 84,6%, dan model ini memiliki nilai akurasi yang lebih tinggi. Ketika diuji pada ulasan film berbahasa Indonesia algoritma Voting 2 (RF + SVM) mampu menghasilkan akurasi hingga 92%.

**Kata Kunci:** ulasan film, ensemble learning, adaboost, voting

## Abstract

*In the world of entertainment, especially movies, now the movie review website has become a media for people to give an assessment of how good the film is. They don't have to be experts in the movie industry to judge the quality of the movies they watch, everyone can give an assessment. Sentiments found in comments, feedback or criticism provide useful indicators for various purposes and can be categorized based on polarity, polarity is likely to be sought out whether overall positive or negative. Naïve Bayes and Random Forest algorithms are algorithms that can provide the results of classification analysis as expected in this study, the analysis will be done by comparing several combinations of algorithms to be tested on Polarity Dataset 2.0 from Cornell University, including the algorithm will be combined with Chi Square feature selection, Adaboost, and Voting. From the test results obtained by AdaBosst and Voting algorithms can improve the accuracy of the Naïve Bayes (NB) and Random Forest (RF) methods. The model proposed with Chi Square + Voting 2 (RF + SVM) has an accounting value of 84.6%, and this model has a higher accuracy value. When tested on movie reviews in Indonesian language the Voting 2 algorithm (RF + SVM) is able to produce accuracy up to 92%.*

**Keywords:** movie review, ensemble learning, adaboost, voting

## Pendahuluan

Media online memberikan dampak yang signifikan bagi kehidupan, pemanfaatannya banyak tersebar diberbagai bidang. Mulai dari media sosial,

forum, E-Commerce, media hiburan, transportasi, dan masih banyak lagi. Dalam dunia hiburan khususnya film, kini situs web ulasan film menjadi media bagi orang-orang untuk memberikan penilaian mengenai

seberapa bagus film tersebut. Mereka tidak harus menjadi pakar dalam dunia perfilman untuk menilai kualitas dari film yang mereka saksikan, semua orang dapat memberikan penilaian.

Informasi yang ditulis oleh pengguna pada situs web ulasan film bisa berupa pujian, kritik, atau saran dari film tersebut. Situs web ulasan populer seperti Amazon, TripAdvisor dan IMDb memberikan media untuk komunitas online dengan berfungsi sebagai outlet gratis untuk pengalaman pasca-pembelian publik. Secara khusus, IMDb memungkinkan penggunaannya untuk menulis ulasan film yang terdiri dari peringkat numerik pada skala satu hingga sepuluh ditambah dengan komentar tekstual (Banerjee & Chua, 2018). Sentimen yang ditemukan dalam komentar, umpan balik atau kritik memberikan indikator yang berguna untuk berbagai tujuan dan dapat dikategorikan berdasarkan polaritas, polaritas tersebut cenderung akan dicari tahu apakah secara keseluruhan positif atau negatif (Dey, Chakraborty, Biswas, Bose, & Tiwari, 2016).

Analisis sentimen dapat didefinisikan sebagai proses yang dapat mengotomatiskan penambangan sikap, pendapat, pandangan dan emosi dari teks, ucapan, kicauan, dan sumber data melalui Natural Language Processing (NLP). Analisis sentimen melibatkan pengelompokan pendapat dalam teks ke dalam kategori seperti "positif" atau "negatif" atau "netral" (A. Kharde & Sonawane, 2016). Tugas dasar dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau fitur/tingkat aspek apakah pendapat yang dikemukakan dalam dokumen, kalimat atau fitur entitas/aspek. Sentiment mengacu pada fokus topik tertentu, pernyataan pada suatu topik dapat berbeda makna dengan pernyataan yang sama pada subjek yang berbeda. Contohnya adalah hal yang baik untuk mengatakan alur film tidak terprediksi, tapi adalah hal yang tidak baik jika 'tidak terprediksi' dinyatakan pada kemudi dari kendaraan. Bahkan pada produk tertentu, kata-kata yang sama dapat menggambarkan makna kebalikan, contoh adalah hal yang buruk untuk waktu start-up pada kamera digital jika dinyatakan "lama", namun jika "lama" dinyatakan pada usia baterai maka akan menjadi hal positif (Syahfitri Kartika Lidya, Opim Salim Sitompul, 2015).

Algoritma Naïve Bayes dan Random Forest merupakan algoritma yang dapat memberikan hasil analisa klasifikasi sesuai yang diharapkan pada penelitian ini, analisa akan dilakukan dengan membandingkan beberapa kombinasi algoritma untuk diuji pada Polarity Dataset 2.0 dari Cornell University, diantaranya yaitu Algoritma tersebut akan dikombinasikan dengan seleksi fitur Chi Square, Adaboost, dan Voting. Dari kombinasi algoritma tersebut hasil akurasi akan didapat sehingga akan diketahui kombinasi algoritma yang lebih baik dan dapat meningkatkan hasil akurasi dari penelitian sebelumnya (Baid, Gupta, & Chaplot, 2017) (Palkar, Gala, Shah, & Shah, 2016).

### Metode Penelitian

Data yang digunakan merupakan ulasan film dari situs IMDb dalam bentuk teks yang dikumpulkan oleh Cornell University versi 2.0 dengan jumlah 2000 ulasan film yang terdiri dari 1000 ulasan positif dan 1000 ulasan negatif. Data ini terdiri dari file html yang tidak diproses, tidak berlabel dari arsip IMDb di newsgroup <http://reviews.imdb.com/Reviews>. File html asli tidak memiliki format yang konsisten - ulasan mungkin tidak memiliki peringkat penulis, dan ketika itu, peringkat dapat muncul di tempat yang berbeda dalam file dalam bentuk yang berbeda. Klasifikasi file ditentukan berdasarkan peringkat pertama yang dapat diidentifikasi (Cornell, 2019).

Melihat hasil penelitian sebelumnya, akurasi yang didapat dengan menggunakan metode Naive Bayes dan Random Forest akan dicoba untuk dioptimalkan. Dataset diuji menggunakan Chi Square dengan Adaboost dan Voting terhadap metode Naive Bayes Classifier dan Random Forest yang ditambahkan K-Fold Cross Validation dengan menggunakan Polarity Dataset berisi ulasan film dari situs IMDb, data yang di uji berjumlah 2000 ulasan dengan rincian 1000 ulasan positif dan 1000 ulasan negatif, kemudian mengevaluasi hasil akurasi, Precision, Recall dan ROC Area dari algoritma tersebut untuk mengetahui kombinasi pengujian yang tepat pada data yang di uji menggunakan algoritma Naive Bayes dan Random Forest terhadap analisa sentimen pada ulasan film yang diharapkan dapat menghasilkan akurasi lebih baik dari penelitian sebelumnya (Baid et al., 2017) (Palkar et al., 2016).

## Hasil dan Pembahasan

Penelitian dilakukan dengan menggunakan laptop dengan processor AMD A9-9420 Radeon R5, 5 Compute Cores 2C+3G 3.00GHz dengan RAM 4GB, dan system operasi Windows 10 Professional 64bit. Sedangkan dalam melakukan eksperimen didukung dengan tools WEKA 3.8.1.

Penelitian ini menggunakan polarity dataset 2.0 dari cornell university yang berisi ulasan film dari situs IMDb dengan jumlah 2000 ulasan yang terdiri dari 1000 ulasan positif dan 1000 ulasan negatif.

### 1. Tahapan Proses Text Mining

Dalam text mining terdapat tahapan yang dilakukan, dan tahapan tahapan tersebut saling berkaitan antara satu dengan yang lainnya, berikut untuk penjabarannya:

1. Pengumpulan data  
Pada tahapan ini, data yang didapat merupakan ulasan film dari situs IMDb yang dikumpulkan oleh Cornell University (Cornell, 2019).
2. Pre-processing data  
Pre-processing data melibatkan operasi pendahuluan yang membantu mengubah data agar mudah digunakan sebelum tugas analisis sentimen yang sebenarnya dapat dilakukan. Beberapa hal yang dilakukan pada tahap data pre-processing adalah transformasi ke dalam bentuk yang lebih informatif. Dengan tahap yang dilakukan berdasarkan atas proses berikut (Palkar et al., 2016).
  - a. Remove numbers  
Angka tidak memainkan peran vital dalam menentukan orientasi teks dan karenanya, mereka dihapus dari data..
  - b. Remove punctuation  
Dalam percobaan, tanda baca dihilangkan dari teks yang dipertimbangkan. Meskipun emotikon yang dibuat menggunakan tanda baca dapat membantu dalam memprediksi sentimen, mereka bahkan dapat digunakan dengan cara sarkastik. Penggunaan emotikon tidak dipertimbangkan dan mendasarkan percobaan hanya pada analisis kata-kata yang diperoleh dari kalimat.
  - c. Remove stop words

Kata-kata seperti kata ganti, preposisi, kata sambung, dll. Yang sering muncul dalam data tetapi tidak menyampaikan konten yang berarti atau informasi penting mengenai nilai sentimen kalimat disebut kata berhenti. Stop word removal dapat membantu mengurangi kebutuhan memori saat mengklasifikasikan ulasan.

- d. Strip whitespaces  
Whitespaces tidak memiliki tujuan yang berarti dalam tugas dan dengan demikian dilepaskan dari teks aslinya.
  - e. Convert to lower case  
Untuk menjaga konsistensi dan memetakan kata-kata terlepas dari kasusnya, kalimat tersebut dikonversi menjadi huruf kecil.
  - f. Stemming  
Banyak kata-kata yang berasal dari kata dasar, berguna untuk menghapus awalan dan akhiran kata-kata menjadi kata dasarnya saja. Stemming dapat secara signifikan mengurangi beban memori selama pelatihan dan klasifikasi.
3. Feature selection  
Feature selection atau seleksi fitur adalah proses menghapus fitur yang berlebihan, sambil mempertahankan fitur-fitur yang memiliki kemampuan disambiguasi tinggi. Penggunaan fitur dimensi yang lebih tinggi seperti bigrams dan trigram menghadirkan masalah, yaitu dari jumlah fitur yang meningkat dari 300.000 menjadi sekitar 11.000.000. Sebagian besar fitur ini berlebihan dan berisik. Termasuk akan mempengaruhi efisiensi dan akurasi (Narayanan, Arora, & Bhatia, 2013). Penggunaan Feature selection dalam penelitian yang diusulkan karena digunakan secara integrasi agar akurasi algoritma naïve bayes dan random forest dapat meningkat. Penelitian ini akan melihat penggunaan satu algoritma Feature selection yaitu:
    - a. Chi Square merupakan metode yang digunakan untuk menghitung tingkat dependensi fitur. Dalam fitur algoritma Chi Square disaring dengan mengukur derajat asosiasi antara dirinya dan kategori yang

pasti (Zheng, Wang, & Gao, 2018). Dalam pemrosesan teks, biasanya dua kelas digunakan untuk mengukur tingkat ketergantungan antara sebuah kata  $t$  dan kelas tertentu  $c$  antara kedua tabel tersebut. Fungsi atas chi square dapat dilihat pada persamaan (1) (Pratama, Sari, & Adikara, 2018):

$$x^2(t, c) = \frac{N(AD-CB)^2}{|F| + \sum_{w \in D} TF(w, C)} \quad (1)$$

Dengan keterangan

$t$  = Term

$c$  = Kelas / Katagori

$N$  = Jumlah data Training

$A$  = Jumlah dokumen pada kelas  $c$  yang memuat  $t$

$B$  = Jumlah dokumen bukan kelas  $c$  yang memuat  $t$

$C$  = Jumlah dokumen pada kelas  $c$  yang tidak memuat  $t$

$D$  = Jumlah dokumen bukan kelas  $c$  yang tidak memuat  $t$

Untuk mendapatkan nilai Chi Square dari suatu kata pada sebuah kelas, dibutuhkan nilai Chi Square tunggal dari kata tersebut. Agar mendapatkan nilai Chi Square tunggal dari suatu kata dapat menjumlahkan nilai Chi Square dari masing-masing kelas. Nilai Chi Square tunggal tiap kata di dapatkan melalui algoritme pada persamaan (2).

$$x^2(t) = \sum_c^k = 1^{x^2(t,c)} \quad (2)$$

Setelah diketahui, dilakukan pengurutan kata berdasarkan nilai Chi Square tertinggi hingga terendah. Hal ini menandakan bahwa semakin besar nilai Chi Square, semakin dependen suatu fitur, dan semakin penting fitur tersebut untuk digunakan dalam proses klasifikasi.

## 2. Ensemble Learning

Pada pendekatan algoritma dan ensemble memiliki tujuan yang sama, yaitu memperbaiki algoritma pengklasifikasi tanpa mengubah data, sehingga dapat dianggap ada 2 pendekatan saja, yaitu pendekatan level data dan pendekatan level algoritma. Dengan membagi menjadi 2 pendekatan dapat mempermudah fokus objek perbaikan, pendekatan level data difokuskan pada pengolahan awal data, sedangkan pendekatan level algoritma difokuskan pada perbaikan algoritma atau menggabungkan (ensemble) (Saifudin & Wahono, 2015). Penelitian ini akan melihat penggunaan dua algoritma Ensemble Learning yang diantaranya:

### 1. Adaptive Boosting (AdaBoost)

AdaBoost adalah kependekan dari Adaptive Boosting merupakan algoritma machine learning yang dirumuskan oleh Yoav Freund and Robert Schapire. AdaBoost secara teoritis dapat secara signifikan digunakan untuk mengurangi kesalahan dari beberapa algoritma pembelajaran yang secara konsisten menghasilkan kinerja pengklasifikasi yang lebih baik (Saifudin & Wahono, 2015). Algoritma adaboost dirancang khusus untuk menyelesaikan masalah klasifikasi yang berfungsi untuk meningkatkan akurasi di setiap algoritma yang lemah.

Algoritma adaboost ditunjukkan sebagai berikut :

Given:  $(x_1, y_1), \dots, (x_n, y_n)$  where  $y_i \in \{+1, -1\}$

Initialize weights:  $D_1(i) = 1/n$  ( $i = 1, \dots, n$ )

For  $t = 1, \dots, T$ :

1. Choose a weak classifier  $h_t$  which minimizes the weighted error:

$$h_t = \arg \min_{h_j \in \mathcal{H}} \epsilon_j = \sum_{i=1}^n D_t(i) I[y_i \neq h_j(x_i)]$$

2. If  $\epsilon_t = \min_j \epsilon_j > 1/2$ , set  $T = t - 1$  and stop loop.
3. Choose an  $\alpha_t$ .
4. Update the weights:

$$D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

where  $Z_t$  is a normalization factor assuring  $D_{t+1}$  is a distribution.

The strong classifier is:

$$H(x) = \text{sign}(\sum_{t=1}^T \alpha_t h_t(x))$$

Sumber : (Hu & Hu, 2005)

**Gambar 3.1** Algoritma Adaboost

## 2. Voting

Penggabungan keputusan atau voting adalah sekumpulan proses integrasi dari beberapa metode untuk menghasilkan kesimpulan yang memiliki standar sama. Metode penggabungan keputusan ada dua, yaitu metode suara terbanyak dan meta-learning (Raharjo & Quafafou, 2015). Jika sebuah ensemble terdiri dari pengklasifikasi yang salah mengklasifikasikan instance yang berbeda, saling melengkapi, maka anggotanya dikatakan "beragam," dan kombinasi prediksi mereka dapat mencapai kinerja di atas salah satu dari mereka secara individual (Gomes, Barddal, Enembreck, & Bifet, 2017).

Voting dengan aturan Average Probability dapat dilihat pada persamaan (3).

$$H(X) = \text{arg}i=1, \dots, n \max(LCi(X)) \quad (3)$$

Dimana :

$n$  = Jumlah class

$LCi(X)$  = Probabilitas terkomputasi terbesar dalam Persamaan 1 untuk menentukan hasil akhir Dari ensemble classifier

Parameter  $LCi(X)$  dijelaskan pada persamaan (4).

$$LCi(X) = \sum_{j=1}^m \frac{1}{m} P_j(w_i | x) \quad (4)$$

$m$  = Jumlah base learner dalam ensemble classifier

## 3. Klasifikasi

Klasifikasi merupakan proses pengelompokan data diantaranya data yang digunakan tersebut mempunyai kelas label atau target. Sehingga algoritma- algoritma untuk menyelesaikan masalah klasifikasi dikategorisasikan ke dalam supervised learning atau pembelajaran yang terawasi, dan tertuang dalam bentuk model Model - model tersebut disebut Classifier. Jadi, Classifier inilah yang akan digunakan untuk menyusun kelas-kelas yang terkandung di dalam data. Ada banyak jenis algoritma klasifikasi.

## 4. Naïve Bayes

Naïve Bayes (NB) adalah metode yang digunakan dalam statistika untuk menghitung peluang dari suatu hipotesis, Naïve Bayes menghitung peluang suatu kelas berdasarkan pada atribut yang dimiliki dan menentukan kelas yang memiliki probabilitas paling tinggi (Socrates, Akbar, Akbar, Arifin, & Herumurti, 2017). Hukum Bayes diberi nama sesuai dengan matematikawan Inggris Thomas Bayes. Klasifikasi Naïve Bayes menganggap bahwa ada atau tidak adanya fitur tertentu dari suatu kelas tidak terkait dengan ada atau tidak adanya fitur lain.

Metode Naive Bayes berdasar pada algoritme dengan teknik klasifikasi yang telah melalui uji efisiensi dan efektifitas dalam suatu database dengan jumlah data yang besar.

Dasar dari teorema Bayes dinyatakan dalam Persamaan (5) (Socrates et al., 2017).

$$cNB = \text{argmax} P(c) \prod_{x \in X} P(x|c) \quad (5)$$

Keterangan:

$cNB$  = Kelas hasil penghitungan Naïve Bayes.

$P(c)$  = Probabilitas kelas pada data latih.

$P(x|c)$  = Peluang fitur  $x$  muncul pada dokumen  $c$ .

$x \in X$  = jumlah fitur unik pada dokumen.

Probabilitas parameter  $P(c_j)$  dijelaskan pada Persamaan (6).

$$P(c_j) = \frac{N_j}{N} \quad (6)$$

Keterangan:

$P(c_j)$  = Probabilitas masing-masing kelas pada data latih.

$N_j$  = Jumlah dokumen pada suatu kelas pada data latih.

$N$  = Jumlah dokumen pada data latih.

## 5. Random Forest

Random Fores (RF) merupakan metode pembelajaran untuk klasifikasi dan regresi. Metode tersebut membangun sejumlah pohon keputusan pada waktu

pelatihan. Untuk mengklasifikasikan case baru, ia mengirimkan case baru ke masing-masing pohon. Setiap pohon melakukan klasifikasi dan menghasilkan kelas. Kelas output dipilih berdasarkan voting mayoritas yaitu jumlah maksimum kelas serupa yang dihasilkan oleh berbagai pohon dianggap sebagai output dari Random Forest (Baid et al., 2017).

Metode standar seperti pohon regresi melibatkan pemisahan node menggunakan split terbaik dalam semua variabel sementara di Random Forest setiap node dibagi dengan mempertimbangkan subset prediktor yang dipilih secara acak terbaik. Strategi ini kuat jika dilihat sebagai properti yang terlalu pas. Di atas itu, user friendly karena hanya terdiri dari dua variabel yang terlibat dalam algoritma (Palkar et al., 2016).

Tahapan penyusunan dan pendugaan menggunakan RF adalah (Palkar et al., 2016):

1. Tahap I
  - a. Tahapan bootstrap : tarik contoh acak dengan pemulihan berukuran n dari gugus data training
  - b. Tahapan random sub-setting : susun tree berdasarkan data tersebut, namun pada setiap proses pemisahan pilih secara acak m < d peubah penjelas, dan lakukan pemisahan terbaik.
  - c. Ulangi langkah a-b sebanyak k kali sehingga diperoleh k buah tree acak
2. Lakukan pendugaan gabungan berdasarkan k buah tree tersebut (misal menggunakan majority Voting untuk kasus klasifikasi, atau rata-rata untuk kasus regresi).Evaluasi Hasil Kinerja Model Usulan

### 5. Evaluasi Hasil Kinerja Model Usulan

Dalam pengevaluasian model digunakan k-fold validation dan confusion matrix dengan membandingkan hasil klasifikasi yang dilakukan oleh system dengan hasil klasifikasi yang sebenarnya. Pengukuran akurasi dengan confusion matrix dapat dilihat pada tabel 1.

Tabel 1. Confusion Matrix

Class Aktual	Class Prediksi		
		Yes	no
	Yes	TP	FN
No	FP	TN	

Dengan formula penghitungan sebagai berikut :

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN}$$

$$Sensitivity = Recall = TPRate = \frac{TP}{TP + FN}$$

$$Specificity = TNRate = \frac{TN}{FP + TN}$$

$$FPRare = \frac{FP}{FP + TN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$G - mean = \sqrt{Sensitivity \times Specificity}$$

$$F - Measure = \frac{2 \times Recall \times Precision}{(Recall + Precision)}$$

$$AUC = \frac{1 + TPRate + FPRate}{2}$$

Pada segmen pertama, pengujian dilakuakn terhadap model NB, dan RF dengan preprocess Chi Square. Hasil pengujian segmen pertama diperoleh hasil akurasi tertinggi *Random Forest* sebesar 82,25% (table 2).

Tabel 2. Hasil Pengujian dengan Chi Square

NO	KRITERIA	NB	RF
1	Akurasi	79,50%	82,25%
2	SENSITIVITY = RECALL = TPrate	0,758	0,82
3	FPrate	0,156	0,18
4	Precision	0,865	0,82
5	F-Measure	0,808	0,82
6	ROC	0,88	0,89
7	G-Mean	0,799	0,82
8	AUC	0,957	1,001

Pada segmen Kedua, pengujian dilakuakn terhadap model NB, dan RF dengan preprocess *Chi Square* dan algoritma *AdaBoost*. Hasil pengujian segmen kedua diperoleh hasil akurasi tertinggi *Random Forest* sebesar 83,25% (table 3).

Tabel 3. Hasil Pengujian dengan Chi Square + AdaBoost

NO	KRITERIA	NB	RF
1	Akurasi	81,65%	83,25%
2	SENSITIVITY = RECALL = TPrate	0,805	0,83
3	FPrate	0,17	0,162
4	Precision	0,835	0,84
5	F-Measure	0,819	0,83
6	ROC	0,89	0,89
7	G-Mean	0,816	0,832
8	AUC	0,988	0,99

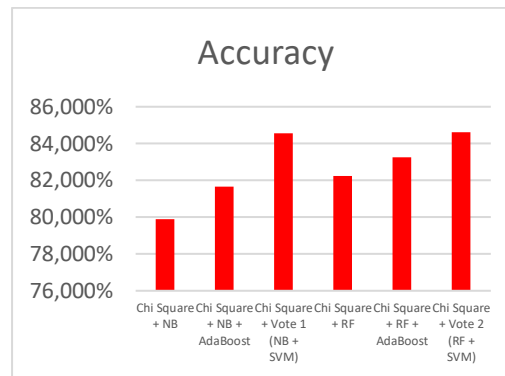
Pada segmen Ketiga, pengujian dilakukan terhadap model NB, dan RF dengan preprocess *Chi Square* dan algoritma *Voting*. Dalam algoritma *Voting* (1), NB akan dikombinasikan dengan *Support Vectore Machine*. Sedangkan algoritma *Voting* (2), RF akan dikombinasikan dengan *Support Vectore Machine* Hasil pengujian segmen kedua diperoleh hasil akurasi tertinggi *Voting 2 (RF + SVM)* sebesar 84,6%

Tabel 4. Hasil Pengujian dengan Voting (+ SVM RBF Kernel)

NO	KRITERIA	Voting 1 (NB+SVM)	Voting 2 (RF+SVM)
1	Akurasi	84,6%	84,6%
2	SENSITIVITY = RECALL = TPrate	0,843	0,843
3	FPrate	0,152	0,152
4	Precision	0,849	0,849
5	F-Measure	0,846	0,846
6	ROC	0,905	0,911
7	G-Mean	0,846	0,846
8	AUC	0,99	0,99

Dari hasil klasifikasi dengan mengintegrasikan seleksi fitur *Chi Square* dan ensemble *AdaBoost* dan *Voting* dari sample polarity dataset 2.0 ulasan film, diketahui ensemble learning khususnya *AdaBoost* dan *Voting* dapat meningkatkan hasil akurasi dari algoritma *Naïve Bayes* dan *Random Forest*. Algoritma *Voting 2 (RF+SVM)* memiliki nilai akurasi serupa dengan *Voting 1 (NB+SVM)* yaitu sebesar 84,6% (gambar 2), namun *ROC Area* pada

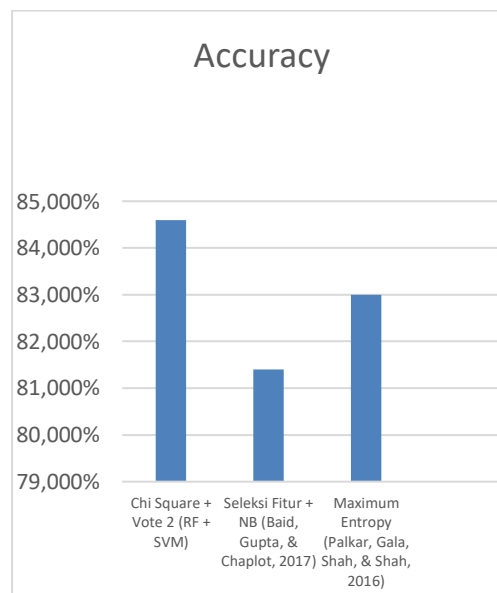
*Voting 2* lebih baik yaitu sebesar 0,911 (tabel 4).



Gambar 2. Perbandingan akurasi

## 6. Perbandingan Penelitian Terkait

Pada penelitian sebelumnya sudah dilakukan pengujian dengan algoritma *Naïve Bayes*, *Support Vectore Machine*, *Random Forest* dan *Meximum Entropy*. Didapatkan hasil bahwa metode yang diusulkan dengan seleksi fitur *Chi Square*, ensemble *Voting* dengan algoritma *Random Forest* memiliki nilai akurasi lebih tinggi (gambar 3).



Gambar 3. Perbandingan Dengan Penelitian Sebelumnya

## Kesimpulan

Penelitian ini mencoba mengoptimalkan metode *Naive Bayes* dan *Random Forest* yang dikombinasikan

dengan seleksi fitur Chi Square, algoritma AdaBoost, dan Voting untuk memberikan hasil akurasi yang lebih baik dari penelitian sebelumnya, polarity dataset 2.0 diuji untuk mendapatkan hasil akurasi tertinggi. Dalam hasil pengujian yang dilakukan, metode Naive Bayes yang ditambahkan seleksi fitur Chi Square mendapat akurasi 79,5% tidak dapat melampaui akurasi tertinggi sebelumnya (Baid, Gupta, & Chaplot, 2017) mendapat akurasi sebesar 81,4%.

Dari hasil pengujian yang didapat algoritma AdaBoost dan Voting mampu meningkatkan akurasi dari metode Naive Bayes dan Random Forest. Pengujian algoritma AdaBoost yang dikombinasikan dengan Random Forest memiliki hasil akurasi lebih baik yaitu 83,25% dibanding kombinasi dengan Naive Bayes yaitu 81,65%. Begitu juga dengan hasil kombinasi algoritma Voting, Hasil akurasi algoritma Voting 1 dan 2 memiliki hasil serupa yaitu 84,6%, Voting 2 (RF + SVM) memiliki hasil lebih baik dibanding Voting 1 (NB+SVM) pada sisi ROC Area.

Model yang diusulkan dengan Chi Square + Voting 2 (RF + SVM) memiliki nilai akurasi 84,6%, dan model ini memiliki nilai akurasi yang lebih tinggi dari penelitian sebelumnya sebesar 81,4% (Baid, Gupta, & Chaplot, 2017) dengan model seleksi fitur + Naive Bayes dan 83% dengan algoritma Maximum Entropy.

## Referensi

- A. Kharde, V., & Sonawane, S. S. (2016). Sentiment Analysis of Twitter Data: A Survey of Techniques. *International Journal of Computer Applications*, 139(11), 5–15. <https://doi.org/10.5120/ijca2016908625>
- Baid, P., Gupta, A., & Chaplot, N. (2017). Sentiment Analysis of Movie Reviews using Machine Learning Techniques, 179(December 2017), 45–49. <https://doi.org/10.5120/ijca2017916005>
- Banerjee, S., & Chua, A. Y. K. (2018). Tracing the growth of IMDb reviewers in terms of rating, readability and usefulness. *2018 4th International Conference on Information Management, ICIM 2018*, 57–61. <https://doi.org/10.1109/INFOMAN.2018.8392809>
- Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment Analysis of Review Datasets Using Naive Bayes' and K-NN Classifier. *International Journal of Information Engineering and Electronic Business*, 8(4), 54–62. <https://doi.org/10.5815/ijieeb.2016.04.07>
- Gomes, H. M., Barddal, J. P., Enembreck, F., & Bifet, A. (2017). A Survey on Ensemble Learning for Data Stream Classification. *ACM Computing Surveys*, 50(2), 1–36. <https://doi.org/10.1145/3054925>
- Hu, W., & Hu, W. (2005). Network-based intrusion detection using adaboost algorithm. *Proceedings - 2005 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2005, 2005*, 712–717. <https://doi.org/10.1109/WI.2005.107>
- IMDb, 2019. What Is IMDb. [https://help.imdb.com/article/imdb/general-information/what-is-imdb/G836CY29Z4SGNMK5?ref\\_=hel\\_psect\\_cons\\_1\\_1#](https://help.imdb.com/article/imdb/general-information/what-is-imdb/G836CY29Z4SGNMK5?ref_=hel_psect_cons_1_1#).
- Narayanan, V., Arora, I., & Bhatia, A. (2013). Fast and accurate sentiment classification using an enhanced Naive Bayes model. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8206 LNCS, 194–201. [https://doi.org/10.1007/978-3-642-41278-3\\_24](https://doi.org/10.1007/978-3-642-41278-3_24)
- Palkar, R. K., Gala, K. D., Shah, M. M., & Shah, J. N. (2016). Comparative Evaluation of Supervised Learning Algorithms for Sentiment Analysis of Movie Reviews. *International Journal of Computer Applications*, 142(1), 975–8887. Retrieved from <http://www.ijcaonline.org/archives/volume142/number1/palkar-2016-ijca-909660.pdf>
- Pratama, N. D., Sari, Y. A., & Adikara, P. P. (2018). Analisis Sentimen Pada Review Konsumen Menggunakan Metode Naive Bayes Dengan Seleksi Fitur Chi Square Untuk Rekomendasi Lokasi Makanan Tradisional. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 2(9), 2982–2988.
- Raharjo, A. B., & Quafafou, M. (2015).



- Penggabungan Keputusan Pada Klasifikasi Multi-Label. *JUTI: Jurnal Ilmiah Teknologi Informasi*, 13(1), 12. <https://doi.org/10.12962/j24068535.v13i1.a384>
- Saifudin, A., & Wahono, R. S. (2015). Penerapan Teknik Ensemble untuk Menangani Ketidakseimbangan Kelas pada Prediksi Cacat Software. *Journal of Software Engineering*, 1(1), 28–37.
- Socrates, I. G. A., Akbar, A. L., Akbar, M. S., Arifin, A. Z., & Herumurti, D. (2017). Optimasi Naive Bayes Dengan Pemilihan Fitur Dan Pembobotan Gain Ratio. *Lontar Komputer: Jurnal Ilmiah Teknologi Informasi*, 7(1), 22. <https://doi.org/10.24843/lkjiti.2016.v07.i01.p03>
- Syahfitri Kartika Lidya, Opim Salim Sitompul, S. E. (2015). Sentiment Analysis Pada Teks Bahasa Indonesia Menggunakan Support Vector Machine ( Svm ) dan K - Nearest Neighbour (K-NN). *Seminar Nasional Aplikasi Teknologi Informasi 2015 (SNATI 2015)*, 1–8.