

ANALISIS SENTIMEN PLATFORM YOUTUBE MENGENAI PRO KONTRA TERHADAP NATURALISASI PEMAIN SEPAKBOLA TIMNAS INDONESIA MENGUNAKAN METODE *NAÏVE BAYES*

Vera Puteri Prasetyo¹, Dahlan²,

^{1,2}Politeknik TEDC Bandung; Jl. Politeknik Jl. Pesantren No.2, Cibabat, Kec. Cimahi Utara, Kota Cimahi, Jawa Barat 40513; telp (022) 6645951

Keywords:

Naive Bayes;
Media Sosial;
Analisis
Sentimen.

Correspondent Email:

veraputriprasetyov@gmail.com



JITET is licensed under
a Creative Commons
Attribution-NonCommercial
4.0 International License

Abstrak. Naturalisasi pemain sepak bola asing ke Timnas Indonesia menimbulkan pro dan kontra di masyarakat, yang tercermin dalam komentar di YouTube. Penelitian ini bertujuan menganalisis sentimen warganet menggunakan algoritma Naïve Bayes. Data diambil dari 509 komentar dan dianalisis dengan RapidMiner. Hasil terbaik diperoleh dari satu kali pengujian dengan akurasi 72,55%, *precision* 83,58%, dan *recall* 76,71%. Sebagian besar komentar tergolong positif terhadap naturalisasi, dengan 268 data berlabel positif dan 80 berlabel negatif. terhadap naturalisasi. Temuan ini menunjukkan efektivitas *Naïve Bayes* dalam klasifikasi sentimen media sosial dan mengindikasikan dukungan publik terhadap kebijakan naturalisasi pemain.

Abstract. The naturalization of foreign football players into the Indonesian national team has sparked both support and criticism among the public, as reflected in YouTube comments. This study aims to analyze netizen sentiment using the Naïve Bayes algorithm. A total of 509 comments were collected and analyzed using RapidMiner. The best results were obtained from a single test with an accuracy of 72.55%, a precision of 83.58%, and a recall of 76.71%. Most comments were classified as positive toward naturalization, with 268 labeled as positive and 80 as negative. These findings demonstrate the effectiveness of the Naïve Bayes algorithm in classifying social media sentiment and indicate public support for the naturalization policy of football players.

1. PENDAHULUAN

Salah satu cara untuk mendapatkan kewarganegaraan adalah melalui proses naturalisasi. Di tempat lain, naturalisasi berarti mendapatkan kewarganegaraan bagi penduduk asing setelah memenuhi syarat yang ditetapkan oleh undang-undang. Secara umum, naturalisasi adalah proses mengubah status kewarganegaraan seseorang karena dengan mengajukan permohonan kepada pemerintahan negara yang terkait dan memenuhi persyaratan yang telah diatur dalam Undang-Undang Nomor

12 Tahun 2006 tentang Kewarganegaraan Republik Indonesia [1].

Hari-hari ini, sering terjadi hujatan tentang fakta bahwa pemain tim nasional Indonesia telah dinaturalisasi dan Menurut bola “Pemain yang awalnya ingin membela Timnas Indonesia menjadi takut karena netizen menyukai serangan seperti itu ada kemungkinan bahwa naturalisasi adalah salah satu topik yang paling sensitif dan kontroversial bagi masyarakat Indonesia. Pendapat, sikap, dan emosi seseorang adalah pusat analisis sentimen. Kemajuan dalam pemrosesan bahasa

alami dan kemajuan dalam penambangan *teks*, data, *web*, dan pembelajaran mesin telah mendorong pertumbuhan pesat di bidang ini. Sangat umum untuk menggunakan analisis sentimen dalam analisis produk, konsumen, jasa, kesehatan, industri, dan keuangan [2].

Media sosial sekarang menjadi alat penting untuk berkomunikasi dengan orang lain dengan cepat dan mudah berkat kemajuan teknologi informasi digital. Menurut Instruksi Presiden Nomor 3 Tahun 2019 tentang Percepatan Pembangunan Sepakbola Nasional, program naturalisasi adalah bagian dari upaya pemerintah untuk meningkatkan sepak bola nasional. Dengan peningkatan peringkat FIFA dari 172 ke 134, naturalisasi diharapkan dapat meningkatkan kualitas dan performa Timnas Indonesia di tingkat internasional [3].

Youtube adalah salah satu situs media sosial yang paling populer di Indonesia. Salah satu keunggulan YouTube adalah tingkat interaktivitas yang tinggi. Tidak jarang ada pendapat yang berbeda dalam kolom komentar YouTube yang rumit. Akibatnya, kolom komentar YouTube menjadi sumber data penting untuk analisis sentimen.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen adalah mengumpulkan, memahami, dan menganalisis data teks untuk menentukan sentimen dalam sebuah opini, apakah itu positif, negatif. Metode ini digunakan untuk mengekstraksi opini, memahami, dan menganalisis data teks secara objektif untuk menginterpretasikan informasi yang disajikan dalam sebuah opini [4].

Dalam data mining, *sentiment analysis* menemukan dan mengekstrak informasi dari teks untuk memahami perasaan sosial dalam teks [5].

2.2 Platform Youtube

Di YouTube, fitur komentar memungkinkan interaksi yang aktif di mana pengguna dapat memberikan pendapat atau bertanya tentang konten video, memicu diskusi. Berbagai perspektif yang berbeda dan perasaan sering kali muncul dalam diskusi ini. Tidak jarang ada pendapat yang berbeda dalam kolom komentar YouTube yang rumit.

2.3 Data Mining

Tujuan utama data mining adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. Istilah "data mining", yang juga dikenal sebagai "penemuan pengetahuan", digunakan untuk mendapatkan pengetahuan yang tersembunyi dari kumpulan data yang sangat besar [6].

Singkatnya, data mining adalah suatu proses pencarian data yang dapat secara otomatis mendapatkan sebuah model dari database yang sangat besar [7].

2.4 Algoritma Naïve Bayes

Naïve Bayes menghasilkan klasifikasi kelas dengan membandingkan nilai posterior dari kelas yang ada. Metode ini bekerja secara fitur independen, yang berarti bahwa satu fitur dalam data tidak berhubungan dengan ada atau tidaknya fitur lain dalam data yang sama [8].

2.5 Klasifikasi

Metode ini menemukan pola untuk membedakan kelas data satu sama lain dan menentukan objek mana yang termasuk dalam kategori tertentu dengan mempertimbangkan sifat dan tingkah laku dari kelompok yang telah didefinisikan sebelumnya [9].

2.6 Clustering

Clustering didefinisikan sebagai pengelompokan sejumlah objek atau data ke dalam kelompok atau kelompok, sehingga setiap kelompok berisi data yang seminimal mungkin dan objek yang berbeda dari kelompok lainnya [10].

2.7 Sepakbola

Sepakbola adalah permainan di mana pemain bersaing untuk menyepak bola dan memasukkan bola ke gawang lawan dan mencegah gol [11].

Selama bertahun-tahun, sepak bola telah mengalami banyak transformasi dan perkembangan dari bentuknya yang sederhana dan primitif sampai menjadi permainan olahraga modern yang sangat digemari dan disenangi banyak orang [12].

2.8 Naturalisasi

Naturalisasi adalah proses memperoleh kewarganegaraan bagi orang non-pribumi (asing) untuk di berikan kewarganegaraan [13].

2.9 RapidMiner

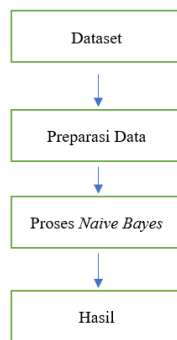
RapidMiner adalah *platform* perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. Ini menyediakan beragam alat untuk implementasi, evaluasi, dan pemodelan data.

Bahkan orang tanpa pengalaman pemrograman dapat dengan mudah membangun dan menguji berbagai model dengan *RapidMiner* [14].

Dengan menggunakan operator-operator, *Rapid Miner* memudahkan pengguna melakukan perhitungan data yang sangat besar [15].

3. METODE PENELITIAN

Metode penelitian ini terdiri dari 4 tahapan yaitu *Dataset*, *Preparasi data*, *Proses Naive Bayes* dan Hasil pada Gambar 3.1



Gambar 3.1 Diagram Alur Penelitian

3.1 Dataset Komentar Pro kontra Naturalisasi Timnas Indonesia

Dataset yang digunakan dalam penelitian ini adalah data komentar Pro Kontra yang diperoleh langsung dari komentar di youtube. Data tersebut mencakup 510 data komentar yang berasal dari komentar. Data yang diperoleh terdiri dari 4 *field* yaitu *Users Ids*, *Timestamps*, *Comments*, dan Keterangan. Pada Tabel 1 terdapat *dataset* yang digunakan.

Tabel 1 Sampel Data Komentar Youtube

No	Users IDs	Timestamps	Comments	Keterangan
1.	UCBgcsVvew5LB4XQwyk3bw	2024-12-10T06:07:06Z	Naturalisasi menghambat perkembangan pemain lokal.	Negatif
2.	UCODVf5jigFw6z750AiyugTQ	2024-12-09T04:19:34Z	Seharusnya kita lebih fokus membina pemain muda sejak dini.	Negatif
3.	UCTkfcyOpz2hMC5ThCJF-Q	2024-12-05T13:28:06Z	Pemain naturalisasi tidak selalu loyal kepada timnas.	Negatif
4.	UC7QU0yocq4sBwmE7hEa2dA	2024-12-05T03:10:35Z	Banyak pemain naturalisasi yang hanya datang untuk uang.	Negatif
5.	UCNKs4LAtuzzg2vbxah7yMOw	2024-12-04T01:55:00Z	Timnas seharusnya berisi pemain yang benar-benar berasal dari Indonesia.	Negatif
6.	UCkPbSdaKvw7zeSNiGvxcNA	2024-10-06T12:04:52Z	Pemain lokal harus mendapatkan prioritas karena mereka yang akan membangun sepak bola Indonesia.	Negatif
7.	UCkPbSdaKvw7zeSNiGvxcNA	2024-10-06T08:58:18Z	Jika negara lain melakukannya, tidak ada salahnya Indonesia juga mengadopsi sistem ini.	Negatif
8.	UCmOS6QoAyl_cprUsx7K8LA	2024-10-06T02:57:48Z	Naturalisasi bukan berarti menutup kesempatan pemain lokal, justru membuat mereka lebih termotivasi.	Positif
9.	UCuf-iOyBH3xMnCF1qeKgYA	2024-10-05T14:29:36Z	Fisik dan stamina pemain naturalisasi bisa membantu timnas bermain lebih konsisten.	Positif
10.	UCFAAaoQulxuyymwzSR7Bpg	2024-10-04T23:19:11Z	Banyak pemain naturalisasi yang tetap menjaga budaya Indonesia.	Positif
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-
4505.	UCR17Dap4qm_r58CgZr_rfw	2024-04-24T15:57:51Z	Bikin petisi, gebukin towel di gbk	Negatif

3.2 Preparasi Data

Tujuannya adalah untuk mendapatkan data yang bersih dan siap untuk di analisis lebih lanjut. Setelah mendapatkan data, pada tahap ini langkah selanjutnya adalah melakukan preparasi atau pembersihan data. Data akan diproses kembali untuk menghilangkan informasi yang dianggap kurang penting dalam proses pengolahan data selanjutnya. Jumlah *dataset* dalam penelitian ini terdiri dari 510 dengan 4 *field*. Pada Tabel 2 terdapat preparasi data yang digunakan.

Tabel 2 Preparasi Data

No	Users IDs	Timestamps	Comments	Keterangan
1	UCBgcsVvew5LB4XQwyk3bw	2024-12-10T06:07:06Z	Naturalisasi menghambat perkembangan pemain lokal.	Negatif
2	UCODVf5jigFw6z750AiyugTQ	2024-12-09T04:19:34Z	Seharusnya kita lebih fokus membina pemain muda sejak dini.	Negatif
3	UCTkfcyOpz2hMC5ThCJF-Q	2024-12-05T13:28:06Z	Pemain naturalisasi tidak selalu loyal kepada timnas.	Negatif
4	UC7QU0yocq4sBwmE7hEa2dA	2024-12-05T03:10:35Z	Banyak pemain naturalisasi yang hanya datang untuk uang.	Negatif
5	UCNKs4LAtuzzg2vbxah7yMOw	2024-12-04T01:55:00Z	Timnas seharusnya berisi pemain yang benar-benar berasal dari Indonesia.	Negatif
6	UCkPbSdaKvw7zeSNiGvxcNA	2024-10-06T12:04:52Z	Pemain lokal harus mendapatkan prioritas karena mereka yang akan membangun sepak bola Indonesia.	Negatif
7	UCkPbSdaKvw7zeSNiGvxcNA	2024-10-06T08:58:18Z	Jika negara lain melakukannya, tidak ada salahnya Indonesia juga mengadopsi sistem ini.	Negatif
8	UCmOS6QoAyl_cprUsx7K8LA	2024-10-06T02:57:48Z	Naturalisasi bukan berarti menutup kesempatan pemain lokal, justru membuat mereka lebih termotivasi.	Positif
9	UCuf-iOyBH3xMnCF1qeKgYA	2024-10-05T14:29:36Z	Fisik dan stamina pemain naturalisasi bisa membantu timnas bermain lebih konsisten.	Positif
10	UCFAAaoQulxuyymwzSR7Bpg	2024-10-04T23:19:11Z	Banyak pemain naturalisasi yang tetap menjaga budaya Indonesia.	Positif
-	-	-	-	-
-	-	-	-	-
510.	UCIAMe9gZ69PgnIUOVZ1b6OA	2024-04-25T02:57:00Z	Naturalisasi membuka peluang bagi pemain keturunan Indonesia di luar negeri untuk membela Garuda.	positif

3.3 Hasil dan Pembahasan

Dengan menggunakan algoritma *Naive Bayes*, penulis akan melakukan analisis sentimen data Komentar Youtube pro kontra Naturalisasi Timnas Indonesia dan menggunakan hasil dari proses *RapidMiner* untuk menghasilkan nilai akurasi dari variabel yang ditargetkan dalam proses penggalian data.

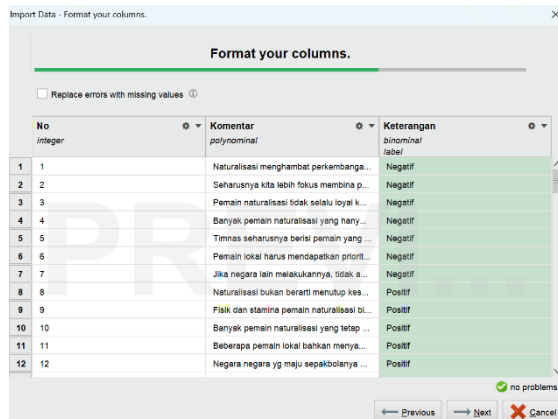
4. HASIL DAN PEMBAHASAN

4.1 Proses Pengujian *Naive Bayes* Pada *Rapid Miner*

Tahapan ini merupakan proses dimulainya pengujian pada data dengan menggunakan *Rapidminer*, berikut langkah awal untuk melakukan pengujian *naive bayes* pada *Rapidminer*:

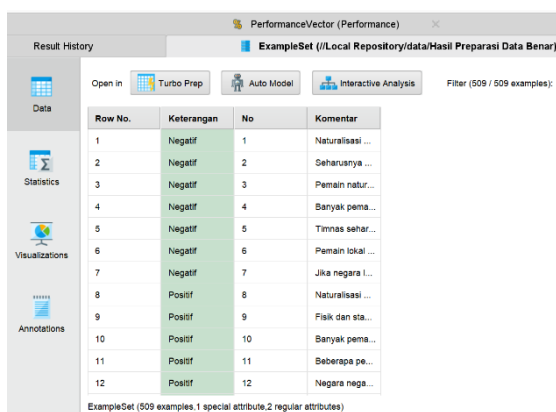
1. Import Data

Dilakukan *import* data yang akan diuji dan mengubah *type* menjadi binominal apabila kolom hanya berisi 2 (dua) karakter, integer untuk kolom yang berisi nomor saja mengubah *role* pada kolom kondisi menjadi label, atribut ini berperan sebagai atribut target untuk operator pembelajaran atau sering disebut dengan “*variable target*”.



Gambar 4.1 Proses *Import* Data

Pada Gambar 4.2 disajikan hasil *import* data yang telah mendeteksi sebanyak 509 data dan selanjutnya akan dilakukan proses *split* data.

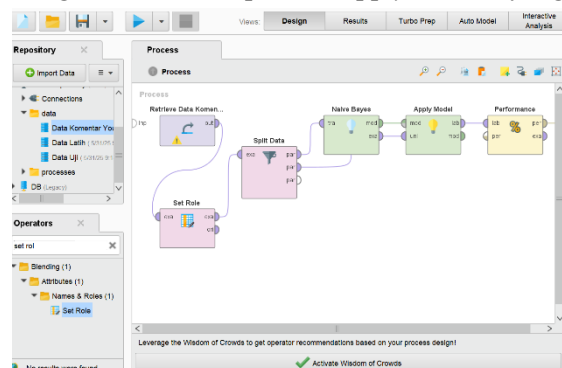


Gambar 4.2 Hasil *Import* Data

2. Split Data

Ditambahkan dengan cara *drag & drop* atau dengan cara *copy paste*, setelah itu data dibagi menjadi 2 (dua) dengan operator *split* data. Untuk data *training* sebesar 80% dan untuk data *testing* sebesar 20%.

Lalu operator *Naive Bayes* untuk dihubungkan dengan data dan operator *Apply Model* yang



berfungsi untuk mengecek data dari model *Naive Bayes* dan tambahkan operator *performance* yang berfungsi untuk mengetahui sebaran data serta akurasi dari data yang telah diujikan.

Gambar 4.3 Proses *Split* Data

Pada pengolahan data komentar youtube tentang pro kontra naturalisasi pemain Timnas Indonesia menggunakan algoritma Naive Bayes dilakukan pengujian sebanyak 1 (satu) kali, penulis mendapatkan hasil dari 5 atribut dengan nilai pred.Negatif sebanyak 18 data dan nilai pred.Positif sebanyak 56 data dengan *accuracy* pengujian sebesar 72.55% Hasil *accuracy* dapat dilihat pada Gambar 4.4

	true Negatif	true Positif	class precision
pred. Negatif	18	17	51.43%
pred. Positif	11	56	83.58%
class recall	62.07%	76.71%	

Gambar 4.4 Hasil *Accuracy* Algoritma *Naive Bayes*

Selain hasil akurasi yang didapat, adapun hasil dari *Precision* pada pengujian ini sebesar 83.58%. *Precision* adalah proporsi prediktif positif artinya dari semua prediksi yang dikategorikan sebagai “Positif” dan “Negatif”

	true Negatif	true Positif	class precision
pred. Negatif	18	17	51.43%
pred. Positif	11	56	83.58%
class recall	62.07%	76.71%	

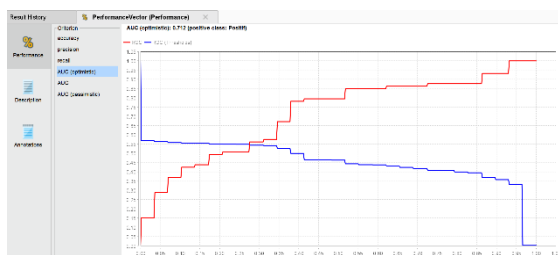
dikatakan benar. Hasil *precision* dapat dilihat pada Gambar 4.5

Gambar 4.5 Hasil *Precision* Algoritma *Naive Bayes*

Adapun hasil dari *Recall* pada pengujian ini sebesar 76.71%. *Recall* adalah metrik evaluasi yang mengukur kemampuan model untuk menemukan semua contoh yang relevan dalam data.

Gambar 4.6 Hasil *Recall* Algoritma *Naive Bayes*

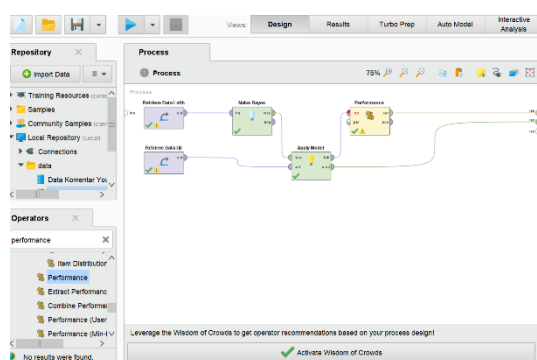
Pada Gambar 4.7 menunjukkan grafik AUC (*Area Under Curve*) dan kurva ROC (*Receiver Operating Characteristic*) untuk sebuah model klasifikasi. AUC yang dihasilkan sebesar 0.712, ini menunjukkan bahwa model memiliki kemampuan prediksi yang kurang sempurna atau dapat dikatakan performa menurun. Garis merah merupakan performa model diberbagai *threshold* klasifikasi, semakin mendekati sudut kiri atas (1.00) artinya semakin performa modelnya. Garis biru merupakan titik-titik *threshold* pada kurva ROC.



Gambar 4.7 Grafik AUC (*Area Under Curve*) Algoritma *Naive Bayes*

4.2 Simple Distribution Naive Bayes

Dalam mengklasifikasi komentar youtube mengenai Pro Kontra tentang Naturalisasi Timnas Indonesia, proses pengujian ini dilakukan untuk mengetahui hasil *simple distribution* yang ada pada proses *data mining*. Ditambahkan operator *Naive Bayes* untuk dihubungkan dengan data latih, setelah itu penulis menambahkan operator *Apply Model*



yang berfungsi untuk mengecek data latih dari model *Naive Bayes* dan dihubungkan dengan data *testing* yang dimana data *testing* ini berfungsi sebagai perbandingan data latih yang ada pada *Apply Model*, setelah itu ditambahkan operator *performance* yang berfungsi untuk mengetahui sebaran data serta akurasi dari data latih yang telah diujikan menggunakan data *testing*.

Gambar 4.8 Proses Pengujian Menggunakan Algoritma *Naive Bayes*

Hasil model distribusi algoritma *naive bayes* dari proses *data mining* digunakan untuk melihat nilai probabilitas. Dapat dilihat nilai yang dihasilkan dari algoritma *naive bayes* pada atribut keterangan nilai *class* "Positif" sebanyak 0,736 sedangkan *class* "Negatif" sebanyak 0,264 ditunjukkan pada gambar 4.9

Simple Distribution

Distribution model for label attribute Keterangan

Class Positif (0.736)
2 distributions

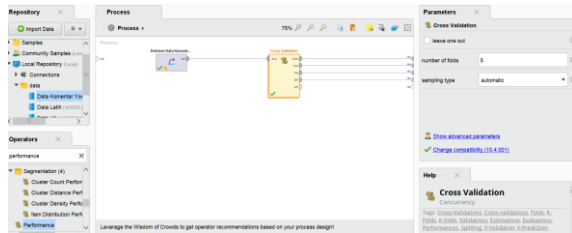
Class Negatif (0.264)
2 distributions

Gambar 4.9 Hasil *Simple Distribution* Algoritma *Naive Bayes*

4.3 Proses Validasi Menggunakan Cross Validation

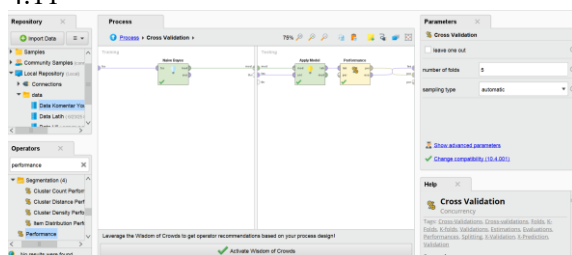
Dalam Proses ini menggunakan *Cross Validation* dengan algoritma *Naive Bayes*. *Cross Validation* adalah teknik yang digunakan untuk mengevaluasi model *machine learning* agar dapat mengukur performa model secara akurat. Pengujian dilakukan dengan menggunakan jumlah fold yang berbeda yaitu:

1. Pengujian Menggunakan 5 *Cross Validation*. Teknik ini akan dilakukan pengujian sebanyak 5 kali agar menghasilkan nilai yang lebih optimal.



Gambar 4.10 Proses Cross Validation 5 fold

Pada pengujian ini menggunakan data yang diseleksi fitur *cross validation naïve bayes*, *apply model*, dan *performance*. Untuk melihat persentase *accuracy*, *precision*, *recall*, dan *number of fold* di isi dengan angka 5. Proses 5 fold cross validation dapat dilihat pada Gambar 4.11



Gambar 4.11 Proses Algoritma Di Dalam Cross Validation 5 fold

Berikut adalah hasil pengujian menggunakan Algoritma Naïve Bayes dan Cross Validation. Tab PerformanceVector menampilkan berbagai metrik untuk

mengevaluasi kinerja model klasifikasi. Penulis mendapatkan hasil dari 5 atribut dengan nilai pred. Positif sebanyak 268 dan nilai pred Negatif sebanyak 80 dengan *accuracy* pengujian sebesar 67,78%.

accuracy: 67.78% +/- 5.81% (micro average: 67.78%)				
	true Negatif	true Positif	class precision	
pred. Negatif	77	98	44.00%	
pred. Positif	66	268	80.24%	
class recall	53.85%	73.22%		

Gambar 4.12 Hasil Accuracy 5 Fold Cross Validation

Adapun hasil dari *Precision* pada pengujian ini sebesar 80,12%.

precision: 80.12% +/- 1.97% (micro average: 80.24%) (positive class: Positif)				
	true Negatif	true Positif	class precision	
pred. Negatif	77	98	44.00%	
pred. Positif	66	268	80.24%	
class recall	53.85%	73.22%		

Gambar 4.13 Hasil Precision 5 Fold Cross Validation

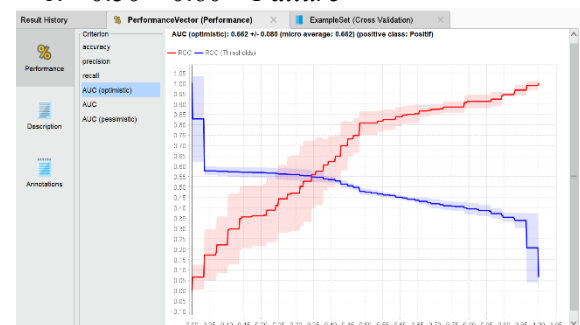
Hasil dari *Recall* pada pengujian ini sebesar 73,23%.

recall: 73.23% +/- 8.48% (micro average: 73.22%) (positive class: Positif)				
	true Negatif	true Positif	class precision	
pred. Negatif	77	98	44.00%	
pred. Positif	66	268	80.24%	
class recall	53.85%	73.22%		

Gambar 4.14 Hasil Recall 5 Fold Cross Validation

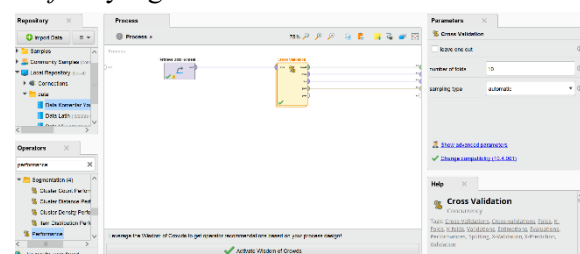
Dari analisis data diatas menggunakan RapidMiner dengan pengukuran Naive Bayes, hasil AUC-1.00 masuk dalam kategori sangat baik (klasifikasi baik). Tingkat akurasi AUC dibagi menjadi lima bagian, yaitu:

- 0.90 – 1.00 = *Excellent Classification*
- 0.80 – 0.90 = *Good Classification*
- 0.70 – 0.80 = *Fair Classification*
- 0.60 – 0.70 = *Poor Classification*
- 0.50 – 0.60 = *Failure*



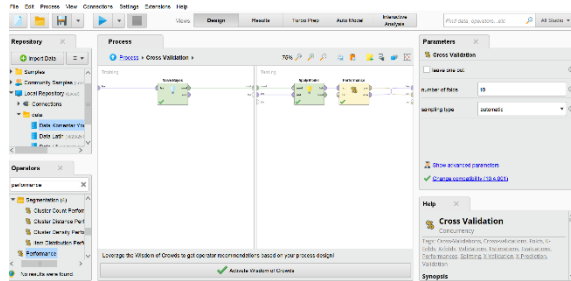
Gambar 4.14 Grafik AUC (Area Under Curve) 5 Fold Cross Validation

2. Pengujian Menggunakan 10 Cross Validation. Pada pengujian ini cara perhitungannya sama dengan yang dilakukan sebelumnya yaitu 5 fold cross validation. Namun yang membedakan hanya *number of fold* yang bernilai 10.



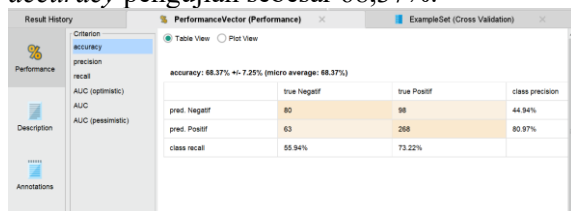
Gambar 4.15 Proses Cross Validation 10 fold

Pengujian ini menggunakan data yang diseleksi fitur, *cross validation*, *naive bayes*, *apply model* dan *performance*. Untuk melihat persentase *accuracy*, *precision*, *recall* dan *number of fold* diisi dengan angka 10.



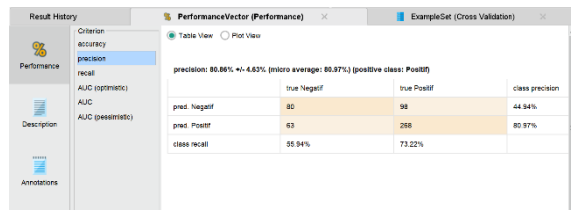
Gambar 4.16 Proses Algoritma Di Dalam *Cross Validation* 10 fold

Berikut adalah hasil pengujian menggunakan Algoritma *Naive Bayes* dan *Cross Validation*. Penulis mendapatkan hasil dari 5 atribut dengan nilai pred.Positif sebanyak 268 dan nilai pred.Negatif sebanyak 80 dengan *accuracy* pengujian sebesar 68,37%.



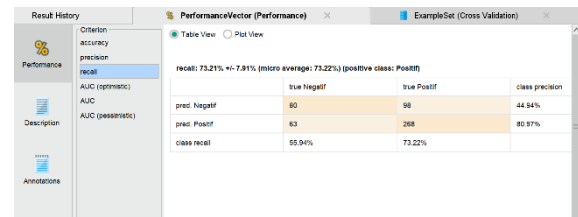
Gambar 4.17 Hasil *Accuracy* 10 Fold *Cross Validation*

Adapun hasil dari *Precision* pada pengujian ini sebesar 80,86%. Hasil *precision* dapat dilihat pada Gambar 4.18



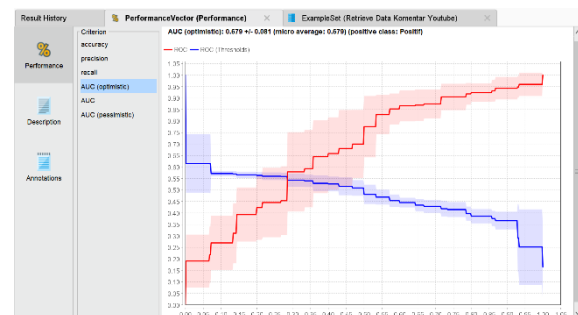
Gambar 4.18 Hasil *Precision* 10 Fold *Cross Validation*

Hasil dari *Recall* pada pengujian ini sebesar 73,21%.



Gambar 4.19 Hasil *Recall* 10 Fold *Cross Validation*

Dari analisis data diatas menggunakan *Rapidminer* dengan pengukuran *Naive Bayes*, hasil AUC 1.0.



Gambar 4.20 Grafik AUC (Area Under Curve) 10 Fold *Cross Validation*

4.4 Hasil Evaluasi Pengujian *Naive Bayes*

Dari hasil preparasi data komentar youtube. Setelah dilakukan analisis klasifikasi *naive bayes* pada tools *RapidMiner*. Berikut perbandingan hasil dari klasifikasi *naive bayes* menggunakan 3 *options test* yang disajikan pada Tabel di bawah ini:

Tabel 4.1 Model Evaluasi

Model Evaluasi	Accuracy	Precision	Recall
1 (satu) kali pengujian	72,55%	83,58%	76,71%
5 Fold <i>Cross Validation</i>	67,78%	80,12%	73,23%
10 Fold <i>Cross Validation</i>	68,37%	80,86%	73,21%

Dari hasil evaluasi didapatkan *Accuracy* sebesar 72,55%, *Precision* 83,58%, dan *Recall* 76,71%. Pada pengujian di atas dengan menggunakan *RapidMiner* terdapat 268 data yang diprediksi “Positif” dan 80 data yang diprediksi “Negatif” pada komentar Youtube tentang Naturalisasi pemain sepak bola Timnas Indonesia.

5. KESIMPULAN

- a. Hasil analisis sentimen positif lebih dominan dibandingkan dengan sentimen negatif hal ini menunjukkan bahwa mayoritas masyarakat memiliki pandangan yang mendukung terhadap naturalisasi pemain asing untuk tim sepak bola nasional Indonesia.
- b. Dengan menggunakan algoritma *naive bayes*, penelitian ini dapat mengidentifikasi dan mengklasifikasikan sentimen dengan akurasi yang cukup baik dengan hasil evaluasi model pada nilai *accuracy* sebesar 72.55%, *precision* sebesar 83.58% dan *recall* sebesar 76.71%.

DAFTAR PUSTAKA

- [1] M. A. Syahrin, "Naturalisasi Dalam Hukum Kewarganegaraan :," *J. Thengkyang*, vol. 2, no.1, pp. 36–53, 2019, [Online]. Available: https://www.researchgate.net/profile/M-Syahrin/publication/335976295_Naturalisasi_dalam_Hukum_Kewarganegaraan_Memahami_Konsep_Sejarah_dan_Isu_Hukumnya/links/5d885a67a6fdcc8fd6107118/Naturalisasi-dalam-Hukum-Kewarganegaraan-Memahami-Konsep-Sejarah-dan-Isu-
- [2] B. Franko, N. Wilyanto, H. Irsyad, U. Multi, and D. Palembang, "Analisis Sentimen Terhadap Naturalisasi Pemain pada Youtube Menggunakan Decision Tree dan Naive Bayes Sentiment Analysis of Player Naturalization on Youtube Using Decision Trees and Naive Bayes," vol. 03, no. September, pp. 8–16, 2024, doi: 10.57203/session.v3i1.2024.8-16.
- [3] A. S. Adhityas *et al.*, "Analisis Sentimen Terhadap Program Naturalisasi Timnas Indonesia Pada X Menggunakan Algoritma," vol. 12, no. 2, pp. 2642–2650, 2025.
- [4] M. G. Andriawan and T. Ernawati, "Penggunaan Algoritma Naïve Bayes Dan Support Vector Machine Untuk Analisis Sentimen Konflik Palestina Dan Israel Pada Platform X," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, pp. 3222–3230, 2024, doi: 10.23960/jitet.v12i3.4943.
- [5] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 10, no. 1, pp. 34–40, 2022, doi: 10.23960/jitet.v10i1.2262.
- [6] W. Purba, W. Siawin, and . H., "Implementasi Data Mining Untuk Pengelompokkan Dan Prediksi Karyawan Yang Berpotensi Phk Dengan Algoritma K-Means Clustering," *J. Sist. Inf. dan Ilmu Komput. Prima (JUSIKOM PRIMA)*, vol. 2, no. 2, pp. 85–90, 2019, doi: 10.34012/jusikom.v2i2.429.
- [7] S. P. Dewi, N. Nurwati, and E. Rahayu, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor," *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 639–648, 2022, doi: 10.47065/bits.v3i4.1408.
- [8] H. M. Siregar, "Bulletin of Information Technology (BIT) Implementasi Metode Naive Bayes Pada Perancangan Aplikasi Sistem Pakar Diagnosa Bronkiektasis," *Bull. Inf. Technol.*, vol. 1, no. 3, pp. 112–121, 2020.
- [9] I. Romli and A. T. Zy, "Penentuan Jadwal Overtime Dengan Klasifikasi Data Karyawan Menggunakan Algoritma C4.5," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 4, no. 2, p. 696, 2020.
- [10] T. Alfina and B. Santosa, "Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Membentuk Cluster Data (Studi Kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS)," *Anal. PerbandinganMetode Hierarchical Clust. K-means dan Gabungan Keduanya dalam Clust. Data*, vol. 1, no. 1, pp. 1–5, 2012.
- [11] S. Irianto, "Standardisasi Kecakapan Bermain Sepakbola Untuk Siswa Sekolah Sepakbola (Ssb) Ku 14-15 Tahun Se-Daerah Istimewa Yogyakarta," *J. Olahraga Prestasi*, vol. 7, no. 7, pp. 44–50, 2011.
- [12] A. Atradinal and R. Sepriani, "Pemulihan Kekuatan Otot Pada Atlet Sepakbola," *J. MensSana*, vol. 2, no. 2, p. 99, 2017, doi: 10.24036/jm.v2i2.86.
- [13] G. K. Annas and N. M. Hazzar, "Analisis Persamaan Hak Kewarganegaraan Bagi Pemain Naturalisasi Sepakbola Di Indonesia," *Wicarana*, vol. 2, no. 2, pp. 127–143, 2023, doi: 10.57123/wicarana.v2i2.37.
- [14] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, "Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell," *JATI (Jurnal)Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 101–104, 2023, doi: 10.36040/jati.v7i1.6094.
- [15] D. Novianti, "Implementasi Algoritma Naïve Bayes Pada Data Set Hepatitis Menggunakan Rapid Miner," *Paradig. - J. Komput. dan Inform.*, vol. 21, no. 1, pp. 49–54, 2019, doi: 10.31294/p.v21i1.4979.