

**ANALISIS SENTIMEN OPINI MASYARAKAT TENTANG INSTITUSI
POLRI BERDASARKAN TWITTER MENGGUNAKAN METODE
SUPPORT VECTOR MACHINE (SVM)**

SKRIPSI



Oleh:

SAID AHMAD SIROJUDIN

NIM : 701190051

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTHAN THAHA
SAIFUDDIN JAMBI**

2023

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

ANALISIS SENTIMEN OPINI MASYARAKAT TENTANG INSTITUSI POLRI BERDASARKAN TWITTER MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)

SKRIPSI



Oleh:

SAID AHMAD SIROJUDIN

NIM : 701190051

**Skripsi ini ditulis untuk memenuhi
sebagian persyaratan untuk mendapatkan gelar Sarjana**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTHAN THAHA
SAIFUDDIN JAMBI
2023**

HALAMAN PERSETUJUAN



KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI SULTHAN THAHA SAIFUDDIN JAMBI
FAKULTAS SAINS DAN TEKNOLOGI
Jalan Lintas Jambi-Muaro Bulian Km. 16 Simpang Sungai Duren Kab. Muaro Jambi
36363 Email: fst.uinjambi.ac.id

NOTA DINAS

Jambi, 05 Juni 2023

Kepada,
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sulthan Thaha Saifuddin di Jambi

Assalamualaikum.wr.wb

Dengan ini diberitahukan bahwa saya telah melakukan bimbingan, arahan dan koreksi naskah skripsi dengan :

Judul : Analisis Sentimen Opini Masyarakat Tentang Institusi Polri Berdasarkan
Twitter Menggunakan Metode Support Vector Machine (SVM)
Nama : Said Ahmad Sirojudin
NIM : 701190051
Jurusan : Sistem Informasi

Saya memandang bahwa naskah skripsi tersebut sudah dapat diajukan kepada Fakultas Sains dan Teknologi UIN Sulthan Thaha Saifuddin untuk diujikan dalam Sidang Munaqasyah.

Wassalamualaikum.wr.wb

Pembimbing I,

Dr. Try Susanti, S.Si., M.Si
NIP/NIDN: 197603032005012005

Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber aslinya;
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah,
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi



KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI SULTHAN THAHA SAIFUDDIN JAMBI
FAKULTAS SAINS DAN TEKNOLOGI
Jalan Lintas Jambi-Muaro Bulian Km. 16 Simpang Sungai Duren Kab. Muaro Jambi
36363 Email: fst.uinjambi.ac.id

NOTA DINAS

Jambi, 05 Juni 2023

Kepada,
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sulthan Thaha Saifuddin di Jambi

Assalamualaikum.wr.wb

Dengan ini diberitahukan bahwa saya telah melakukan bimbingan, arahan dan koreksi naskah skripsi dengan :

Judul : Analisis Sentimen Opini Masyarakat Tentang Institusi Polri Berdasarkan
Twitter Menggunakan Metode Support Vector Machine (SVM)
Nama : Said Ahmad Sirojudin
NIM : 701190051
Jurusan : Sistem Informasi

Saya memandang bahwa naskah skripsi tersebut sudah dapat diajukan kepada Fakultas Sains dan Teknologi UIN Sulthan Thaha Saifuddin untuk diujikan dalam Sidang Munaqasyah.

Wassalamualaikum.wr.wb

Pembimbing II,

Mhd. Theo Ari Bangsa, M.Cs
NIP/NIDN: 2007089205

Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber aslinya;
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah,
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

HALAMAN PENGESAHAN



KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI
SULTHAN THAHA SAIFUDDIN JAMBI
FAKULTAS SAINS DAN TEKNOLOGI

Jl. Lintas Jambi Ma. Bulian Km. 16 Sei. Duren Kec. Jaluko Kab. Muaro Jambi 36363
Telp/Fax : (0741) 584118 Web: <https://uinjambi.ac.id>/email : mail@uinjambi.ac.id

PENGESAHAN SKRIPSI / TUGAS AKHIR

Nomor : B- 651 /D. VII/PP.009/08/2023

Skrripsi/ Tugas Akhir dengan Judul : "Analisis Sentiment Opini Masyarakat Terhadap Tentang
Institusi Polri Berdasarkan Twitter Menggunakan Metode
Support Vector Machine (SVM)"

Yang dipersiapkan dan disusun oleh

Nama : Said Ahmad Sirojudin

NIM : 701190051

Telah dimunaqasahkan pada : 27 Juni 2023

Nilai Munaqasah : 88,91

Dan dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sulthan Thaha Saifuddin
Jambi.

TIM MUNAQASAH :

Ketua Sidang

↓

Dr. Try Susanti, S.Si., M.Si

NIP. 197603032005012005

Penguji I

Andreo Yudertha, ST., M. Eng

NIP. 198907262020121006

Penguji II

Yeri Ramadhani, M. Kom

NIDN. 2014068504

Pembimbing I

↓

Dr. Try Susanti, S.Si., M.Si

NIP. 197603032005012005

Pembimbing II

Mhd. Theo Ari Bangsa, M. Cs

NIDN. 2007089205

Sekretaris Sidang

Mhd. Theo Ari Bangsa, M. Cs

NIDN. 2007089205

Jambi, 07-08- 2023

Dekan Fakultas Sains dan Teknologi



Iskandar, S. Ag., M.Pd., M.S.I., M.H., Ph. D

NIP. 197512242009121001

Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah,
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

PERSEMBAHAN

Skripsi ini kupersembahkan kepada kedua orang tuaku yang tercinta Ayahanda dan Ibunda yang telah mengasuhku mulai dari lahir hingga dewasa sekarang ini, semoga kedua orang tuaku selalu mendapat rahmat dari Allah Subhanahuwa Ta'ala, Amin.

Terimakasih juga kepada saudara/i kandung yang tercinta dan terimakasih atas dukungan dan do'anya sehingga saya dapat menyelesaikan studi pendidikan di perguruan tinggi ini. Sahabat-sahabat seperjuangan khususnya Program Studi Sistem Informasi, umumnya Fakultas Sains dan Teknologi UIN Sulthan Thaha Saifuddin Jambi serta orang-orang yang mencintai ilmu pengetahuan.

Terimakasih untuk semua yang telah membantuku dalam penyelesaian skripsi ini, semoga Allah SubhanahuwaTa'ala selalu memberikan taufiq dan hidayah kepada kitasemua. Amin yaa Robbal 'Alamin.

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

MOTTO

”Wahai orang-orang yang beriman! Jadilah kamu sebagai penegak keadilan karena Allah (ketika) menjadi saksi dengan adil. Dan janganlah kebencianmu terhadap suatu kaum mendorong kamu untuk berlaku tidak adil. Berlaku adillah. Karena (adil) itu lebih dekat kepada takwa. Dan bertakwalah kepada Allah, sungguh, Allah Mahateliti terhadap apa yang kamu kerjakan.”
(Q.S Al-Maaida:8)

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suntha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suntha Jambi

PERNYATAAN KEASLIAN NASKAH

PERNYATAAN KEASLIAN NASKAH

Yang bertandatangan dibawah ini:

Nama : Said Ahmad Sirojudin
NIM : 701190051
Jurusan : Sistem Informasi

Menyatakan bahwa skripsi yang berjudul:

ANALISIS SENTIMEN OPINI MASYARAKAT TENTANG INSTITUSI POLRI BERDASARKAN TWITTER MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)

Merupakan Skripsi yang saya susun sebagai syarat untuk memperoleh gelar Sarjana dari Fakultas Sains dan Teknologi Universitas Islam Negeri Sulthan Thaha Saifuddin Jambi adalah hasil karya sendiri.

Adapun bagian-bagian tertentu dalam penulisan skripsi yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Apabila dikemudian hari ditemukan seluruh atau sebagian skripsi bukan hasil karya saya sendiri atau terindikasi adanya unsur plagiat dalam bagian-bagian tertentu, saya bersedia menerima sanksi sesuai dengan peraturan dan perundang- undangan yang berlaku.

Jambi, Agustus 2023



Said Ahmad Sirojudin

Nim. 701190051

Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah,
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suttha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suttha Jambi

ABSTRACT

Twitter is one of the social media sites that is currently trending in Indonesia. People were busy talking about various topics, including the police. The purpose of this research is to analyze public opinion sentiment about the National Police Institution based on Twitter using the support vector machine method. The research phase began with the crawling of as much as 6,925 tweets from 3 keywords, namely National Police Institutions, National Police Public Relations, and Regional Police. Then proceed with the preprocessing stage consisting of cleaning, case folding, tokenizing, and filtering. Next is the feature extraction stage using TF-IDF, and finally the classification and evaluation stages. From the results of manual data testing (73:27), it was found that accuracy was 70.66%, precision was 70.68%, and recall was 99.76%. The second data test (82:18) obtained an accuracy of 86%, precision of 86.21%, and recall of 99.71%. From the results of manual data testing (82:18), it was found that accuracy was 70.66%, precision was 70.68%, and recall was 99.76%. The second data test (82:18) obtained an accuracy of 86%, precision of 86.21%, and recall of 99.71%. From the results of the data system test (80:20), accuracy is 87.55%, positive precision is 87.53%, negative precision is 88.24%, positive recall is 99.48%, and negative recall is 21.43%. From the results of data system testing (60:40), accuracy is 86.89%, positive precision is 86.84%, negative precision is 88.46%, positive recall is 99.61%, and negative recall is 16.43%. Testing the split data validation system (80:20) with an accuracy of 87.55% and testing the overall data cross validation system using k-fold 5 with an accuracy of 86.673%.

Kata kunci : *Police Institution, Data Mining, Support Vector Machine*

ABSTRAK

Twitter merupakan salah satu media sosial yang saat ini selalu menjadi pusat trending di Indonesia. Masyarakat ramai berbincang berbagai macam topik termasuk salah satunya tentang polisi. Tujuan dari penelitian ini adalah melakukan Analisis sentiment opini masyarakat tentang institusi polri berdasarkan twitter menggunakan metode support vector machine. Tahapan penelitian dimulai dengan tahapan crawling data Twitter sebanyak 6.925 data dari 3 kata kunci yaitu *Institusi polri*, *Humas polri* dan *Polda*. Kemudian dilanjutkan dengan tahap Preprocessing terdiri dari (cleansing, case folding, tokenizing dan filtering). Selanjutnya tahap Ekstrasi fitur menggunakan tf-idf dan terakhir tahap klasifikasi dan evaluasi. Dari hasil pengujian manual data (73:27) didapatkan Accuracy sebanyak 70.66%, Precision 70.68% dan Recall 99.76%. Pengujian data kedua (82:18) didapatkan Accuracy sebanyak 86%, Precision 86.21% dan Recall 99.71%. Dari hasil pengujian manual data (82:18) didapatkan Accuracy sebanyak 70.66%, Precision 70.68% dan Recall 99.76%. Pengujian data kedua (82:18) didapatkan Accuracy sebanyak 86%, Precision 86.21% dan Recall 99.71%. Dari hasil pengujian sistem data (80:20) Accuracy sebanyak 87.55%, Precision Positif 87.53%, Precision Negatif 88.24%, Recall Positif 99.48%, Recall Negati 21.43%. Dari hasil pengujian sistem data (60:40) Accuracy sebanyak 86.89%, Precision Positif 86.84%, Precision Negatif 88.46%, Recall Positif 99.61%, Recall Negati 16.43%. Pengujian sistem split validation data (80:20) akurasi 87.55% dan Pengujian sistem cross validation data keseluruhan menggunakan k-fold 5 akurasi sebesar 86.673%.

Kata kunci : Institusi Polri, *Data Mining*, *Support Vector Machine*

KATA PENGANTAR

Puji dan Syukur Penulis panjatkan kepada Allah *Subhanahuwa Ta'ala* dan salam atas Nabi Muhammad *Shalallahu 'alaihiwasallam*, pembawa risalah pencerahan bagi manusia.

Penulisan skripsi ini dimaksudkan untuk memenuhi salah satu syarat akademik guna mendapatkan gelar Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Sains dan Teknologi UIN Sulthan Thaha Saifuddin Jambi.

Dalam penyusunan skripsi ini banyak ditemukan kesulitan-kesulitan dan hambatan-hambatan, namun berkat inayah Allah SWT, serta bantuan dari berbagai pihak segala kesulitan dan hambatan tersebut dapat diatasi, sehingga skripsi ini dapat diselesaikan. Untuk itu, penulis mengucapkan terima kasih kepada yang terhormat:

1. Bapak Prof. DR. H. Su'aidi, MA.Ph.D. selaku Rektor UIN Sulthan Thaha Saifuddin Jambi
2. Bapak Iskandar, S.Ag., M.Pd., M.S.I., M.H., Ph.D. selaku Dekan Fakultas Sains dan Teknologi UIN Sulthan Thaha Saifuddin Jambi.
3. Ibu Dr. Try Susanti, S.Si., M.Si. selaku Ketua Program Studi Sistem Informasi Fakultas Sains dan Teknologi.
4. Ibu Dr. Try Susanti, S.Si., M.Si. selaku Pembimbing I dan Bapak Mhd. Theo Ari Bansa, M.Cs selaku Pembimbing II.
5. Bapak-bapak dan ibu-ibu dosen serta staf Fakultas Sains dan Teknologi UIN Sulthan Thaha Saifuddin Jambi.
6. Ayah, Ibu, dan saudara-saudaraku yang telah mendukung dan memberikan motivasi.
7. Rekan-rekan seperjuangan Angkatan 2019 di Fakultas sains dan teknologi Sulthan Thaha Saifuddin Jambi.
8. Almamater

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

Akhirnya kepada semua pihak, penulis sertakan do'a semoga Allah SWT membalas pahala kebaikan yang telah diberikan agar berlipat ganda dan berkenan untuk mengabulkannya, amin ya rabbal alamin. Semoga skripsi ini akan memberikan banyak manfaat bagi yang membacanya.

Jambi, Agustus 2023

Said Ahmad Sirojudin
Nim. 701190051

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sunthha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sunthha Jambi



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iv
PERSEMBAHAN	v
MOTO	vi
PERNYATAAN KEASLIAN NASKA	vii
ABSTRACT	viii
ABSTRAK	ix
KATA PENGANTAR	x
DAFTAR ISI	xii
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
LAMPIRAN	xvi
DAFTAR RIWAYAT HIDUP	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
BAB II LANDASAN TEORI	5
2.1 Kajian Teoritis	5
2.2 Penelitian Terdahulu	14
BAB III METODOLOGI PENELITIAN	17
3.1 Tahapan Penelitian	17
3.2 Data Crawling	17
3.3 Preprocessing	21
3.4 Ekstrasi Fitur	27
3.5 Klasifikasi SVM	30
3.6 Evaluasi	34

BAB IV HASIL DAN PEMBAHASAN	36
4.1 Hasil Penelitian	36
4.2 Pembahasan Penelitian	42
BAB V KESIMPULAN DAN SARAN	44
5.1 Kesimpulan	44
5.2 Saran	44
DAFTAR PUSTAKA	46
LAMPIRAN	47

@ Hak cipta milik IAIN Sutha Jambi

State Islamic University of Sulthan Thaha Saifuddin Jambi



UNIVERSITAS ISLAM NEGERI
SULTHAN THAHA SAIFUDDIN
J A M B I

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

DAFTAR TABEL

Tabel 2.1 Penelitian terdahulu	15
Tabel 2.1 (Lanjutan)	16
Tabel 3.1 Contoh text data mentah	20
Tabel 3.1 (Lanjutan)	21
Tabel 3.2 Contoh text <i>cleansing</i> replace rt	22
Tabel 3.3 Contoh text <i>cleansing</i> replace url	23
Tabel 3.4 Contoh text <i>cleansing</i> replace hastag	23
Tabel 3.5 Contoh text <i>cleansing</i> replace mention	24
Tabel 3.6 Contoh text <i>cleansing</i> replace simbol	24
Tabel 3.7 Contoh text <i>case folding</i>	25
Tabel 3.8 Pelabelan manual	29
Tabel 3.8 (Lanjutan)	30
Tabel 3.9 Ilustrasi skema <i>5 fold validation</i>	35
Tabel 4.1 Keterangan <i>confusion matrix</i>	37
Tabel 4.2 Hasil pengujian data <i>testing 600</i>	38
Tabel 4.3 Hasil pengujian data <i>testing 400</i>	38
Tabel 4.4 Hasil percobaan data baru	41
Tabel 4.1 Hasil prediksi	42

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sunha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sunha Jambi



DAFTAR GAMBAR

Gambar 2.1 <i>SVM Linear</i>	8
Gambar 2.2 <i>SVM Non-Linear</i>	8
Gambar 2.3 Ilustrasi <i>Hyperplane</i>	10
Gambar 2.4 <i>Support Vector</i>	11
Gambar 2.5 <i>Max Margin</i>	12
Gambar 2.6 Dua kategori	12
Gambar 2.7 Dua kategori terpisah oleh kurva	13
Gambar 2.8 Dua kategori terpisah oleh <i>Hyperplane</i>	13
Gambar 3.1 Tahapan Penelitian	17
Gambar 3.2 Pengambilan Token	18
Gambar 3.3 Tahapan <i>crawling</i>	19
Gambar 3.4 Tahapan penggabungan data kata kunci	20
Gambar 3.5 Contoh <i>cleansing</i>	21
Gambar 3.6 Tahapan pembersihan data	22
Gambar 3.7 Contoh <i>case folding</i>	25
Gambar 3.8 Contoh <i>tokenizing</i>	26
Gambar 3.9 Contoh <i>filtering</i>	26
Gambar 3.10 Contoh <i>TF-IDF</i>	27
Gambar 3.11 Proses <i>TF-IDF</i>	28
Gambar 3.12 Membuat model <i>SVM</i> pada Rapid Miner	31
Gambar 3.13 <i>Sentiment is not missing</i>	32
Gambar 3.14 Proses pelabelan dengan model data latih	32
Gambar 3.15 <i>Sentiment no missing</i>	33
Gambar 4.1 Hasil <i>TD-IDF</i>	36
Gambar 4.2 Hasil <i>accuracy split validation</i> 40%	39
Gambar 4.3 Hasil <i>accuracy split validation</i> 20%	40
Gambar 4.4 Hasil <i>accuracy crossvalidation</i>	40
Gambar 4.5 Jumlah positif negatif	41

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sunha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sunha Jambi

DAFTAR LAMPIRAN

Lampiran 1 Aplikasi rapid miner	47
Lampiran 2 Tampilan <i>recent</i>	47
Lampiran 3 Tampilan awal rapid miner	48
Lampiran 4 Tampilan <i>repository</i> data	49
Lampiran 5 Tampilan <i>repository proces</i>	49
Lampiran 6 Tampilan <i>operators</i>	50
Lampiran 7 Tampilan <i>parameters</i>	50

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

DAFTAR RIWAYAT HIDUP

A. Identitas

Nama : Said Ahmad Sirojudin

Jenis Kelamin : Laki-Laki

Tempat / Tanggal Lahir : Pauh 20 Agustus 2001

Alamat : Jl. Tembesi-Jambi, Mendalo Asri, Lorong Baruga
Blok C, RT 04, Kec. Jambi Luar Kota, Jambi.

Email : ahmadsirojudin03082019@gmail.com

HP : 082210871091



B. Riwayat Pendidikan

1. SD Negeri 50 Pauh Tahun 2013
2. SMP Negeri 7 Sarolangun Tahun 2016
3. SMA Negeri 3 Sarolangun Tahun 2019

Jambi, Agustus 2023

Said Ahmad Sirojudin
Nim. 701190051

Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kepolisian Negara Republik Indonesia (Polri) salah satu aparaturnegara yang menjaga keamanan dan ketertiban masyarakat, Misinya adalah melindungi, memberikan pelayanan, dan menegakkan hokum bagi masyarakat indonesia(Ritonga et al., 2022). Negara Indonesia didasarkan pada aturan hukum, bukan hanya kekuasaan. Indonesia mewujudkan negara hukum sebagai ideologi untuk keamanan, ketertiban, keadilan dan kesejahteraan bagi warga negaranya.

Polisi merupakan dua unsur yang saling berhubungan dengan masyarakat. Tidak akan berjalan dengan lancer dan produktif tanpa polisi, tanpa masyarakat tidak ada polisi. Fakta ini menempatkan Polri dalam peran ganda dengan memenuhi tugasnya, sebagai pekerja sosial (*social worker*) dan penegak hukum dalam bidang kemasyarakatan (pelayanan dan pengabdian) dan sosial(Edi Saputra Hasibuan, 2021).

Pandangan masyarakat terhadap kinerja polisi dapat berupa penilaian positif maupun negatif. Masyarakat percaya pada penegakan hukum di negeri ini semakin hari semakin berkurang. Ketidakpuasan masyarakat pada penegak hukum semakin meningkat di negeri ini, terlihat jelas dari hasil survei lembaga di Indonesia, di mana ketidakpercayaan masyarakat terhadap penegakan hukum saat ini mencapai 56%, dan hanya 29.8% yang

melaporkannya. Mereka puas, dibandingkan dengan pemerintahan sebelumnya, hanya 22,6% yang menyatakan polisi lebih baik dibawah pemerintahan sekarang daripada di pemerintahan sebelumnya(Alfian, 2020)

Media sosial adalah platform komunitas yang berfokus pada kehadiran pengguna dan mendorong aktivitas dan kolaborasi pengguna. Oleh karena itu, dapat dikatakan bahwa social media merupakan media online yang memiliki kemampuan untuk mempererat hubungan sosial.

Twitter adalah salah satu media sosial yang selalu menjadi trending di Indonesia, kini banyak sekali informasi yang dapat dijawab dengan cepat dan akurat dari berbagai sudut pandang. Oleh karena itu, Twitter tidak hanya memberikan dampak positif tetapi juga berdampak negatif bagi pengguna maupun non pengguna Twitter, salah satunya adalah opini publik terhadap lembaga kepolisian nasional(Rizki et al., 2021). Untuk melihat informasi data yang besar pada media social twitter masih menggunakan cara yang manual, kesulitan dari cara manual ini tidak bisa mengumpulkan data yang begitu banyak dengan otomatis. Maka dari itu perlu adanya analisis sentimen sehingga untuk mempermudah dalam mengklasifikasi opini masyarakat terhadap tweet dari pengguna twitter dengan cepat dan secara otomatis.

Analisis sentiment atau penelitian opini termasuk data berbasis data yang tujuannya untuk menganalisis, memahami, mengelola dan mengekstrak data tekstual sebagai data pendapat.

Beberapa algoritma termasuk Support Vector Machine (SVM), Naïve Bayes Classifier, K-Nearest Neighbor (KNN), Artificial Network (ANN) dan



Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

Decision Tree. Algoritma yang digunakan Support Vector Machine terbukti mempunyai akurasi klasifikasi yang lebih stabil dibandingkan algoritma lainnya serta lebih kuat dan lebih dapat digeneralisasikan (Ananda & Pristyanto, 2021).

Berdasarkan penelitian tersebut, maka penulis memilih metode *Support Vector Machine* (SVM), karena metode tersebut untuk mencari garis pemisah optimal (hyperplane) dua dataset terpisah yang termasuk dalam dua kelas berbeda dan memiliki keunggulan seperti menentukan jarak menggunakan vector pendukung untuk perhitungan atau kompilasi yang cepat. Berdasarkan uraian latar belakang tersebut, maka penulis ingin meneliti “ANALISIS SENTIMEN OPINI MASYARAKAT TENTANG INSTITUSI POLRI BERDASARKAN TWITTER MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM).

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang yang dijelaskan diatas maka dapat dirumuskan masalah dalam penelitian ini adalah “bagaimana mengklasifikasikan sentimen publik terhadap pengguna twitter dengan metode Support Vector Machine (SVM)?”

1.3 Batasan Masalah

Adapun batasan permasalahan penelitian ini ialah sebagai berikut:

1. Proses klasifikasi hanya data teks berbahasa Indonesia.
2. Data media center yang diteliti adalah twitter.



Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

3. Metode algoritma yang dipakai dalam pengklasifikasian ialah metode Support Vector Machine (SVM).
4. Analisis sentimen opini publik yang dibuat diambil dari opini public terhadap institusi polri.
5. Sentimen positif, negatif adalah hasil keluaran dari analisis ini.
6. Khusus wilayah Indonesia sumber opini digunakan.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini ialah mengklasifikasikan sentiment terhadap opini masyarakat berdasarkan tweet dari pengguna twitter menggunakan metode Support Vector Machine (SVM).

1.5 Manfaat Penelitian

Adapun manfaat penelitian ini adalah dari hasil penelitian dapat dijadikan sebagai informasi mengenai opini masyarakat terkait institusi polri, dan juga untuk memberikan informasi metode pengklasifikasi *Support Vector Machine* (SVM) dapat digunakan untuk menganalisis sentimen.

BAB II

LANDASAN TEORI

2.1 Kajian Teoritis

2.1.1 Analisis Sentimen

Analisis sentimen adalah teknik untuk menemukan, mengekstrak, dan menjalankan informasi secara otomatis melakukan informasi tekstual untuk menemui informasi sentimental dalam mengungkapkan pemikiran. Analisis sentimen digunakan untuk melihat pendapat seseorang atau pandangan serupa tentang suatu masalah atau topic secara positif atau negatif(Gifari et al., 2022)

2.1.2 Metode Support Vector Machine (SVM)

Kelebihan metode adalah dapat diklasifikasikan dan mengatasi regresi linier dan nonlinier. Metode Support Vector Machine, teknik pemisah (hyperplane) dua kumpulan data kelas yang berbeda dengan memaksimalkan batas fungsi pemisah yaitu (hyperplane). Metode Support Vector Machine mempunyai akurasi pengklasifikasi lebih baik dibandingkan dengan metode klasifikasi lainnya(Dasmasela et al., 2021).

Data pada dataset diberi variabel x , sedangkan kelas dari dataset diberi variabel y . Metode SVM membagi kumpulan data menjadi 2 kelas. Kelas pertama yang dipisahkan oleh hyperplane memiliki nilai 1, sedangkan kelas lainnya memiliki nilai -1.

$$X_i \cdot W + b \geq 1 \text{ untuk } Y_i = 1 \quad (1)$$

$$X_i \cdot W + b \leq -1 \text{ untuk } Y_i = -1 \quad (2)$$

Keterangan :

X_i = data ke -i

W = nilai bobot support vector yang tegak lurus dengan hyperplane

b = nilai bias Y_i = kelas data ke - i

2.1.3 Twitter

Twitter merupakan salah satu media sosial yang sangat populer dengan lebih dari 300 juta pengguna di seluruh dunia pada tahun 2018. Melakukan penelitian dan analisis dengan data yang bersumber dari Twitter telah menjadi salah satu topik yang sangat digemari akhir-akhir ini. Salah satu topik yang banyak dipelajari dalam menganalisis data dari Twitter merupakan analisis sentimen. Selain itu, pada Twitter terdapat banyak opini, sehingga dapat memudahkan penelitian yang menggunakan algoritma klasifikasi (Aldisa & Maulana, 2022)

2.1.4 Teks Mining

Teks mining adalah teknik untuk mengekstraksi informasi dari kumpulan data tidak terstruktur berkualitas tinggi dan mengumpulkan data tentang masalah dalam teks tematik tertentu, dengan secara otomatis mengekstraksi informasi dari sumber yang berbeda (Mustofa & Novita, 2022).

2.1.5 Split Validation

Validasi split adalah model klasifikasi mesin support vector sesuai dengan pembagian dua dataset dijadikan dua dataset, data pelatihan dan

Hak Cipta Dilindungi Undang-Undang:

data uji (testing) data training atau data latih digunakan untuk training sedangkan data test (uji) digunakan untuk testing (Residencia, 2019).

2.1.6 Confusion Matrix

Confusion matrix digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada dasarnya confusion matrix mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya. Pada pengukuran kinerja menggunakan confusion matrix, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN). Nilai True Negative (TN) merupakan jumlah data negatif yang terdeteksi dengan benar, sedangkan False Positive (FP) merupakan data negatif namun terdeteksi sebagai data positif. (Residencia, 2019)

2.1.7 Algoritma Klasifikasi Support Vector Machine

Kinerja metode Support Vector Machine, terutama untuk masalah nonlinear, menambahkan konsep inti ke ruang dimensi tinggi. Salah satu fungsi mendefinisikan hyperplane (pemisah) atau maksimal jarak pemisah (margin) antar kelas data. Berikut ini metode Support Vector Machine dapat dibedakan menjadi dua jenis, yaitu:

1. Support Vector Machine Linear

Support Vector Machine (SVM) linear digunakan untuk data yang dipisahkan secara linear. Maksudnya, jika dataset diklasifikasikan

Hak Cipta Dilindungi Undang-Undang:

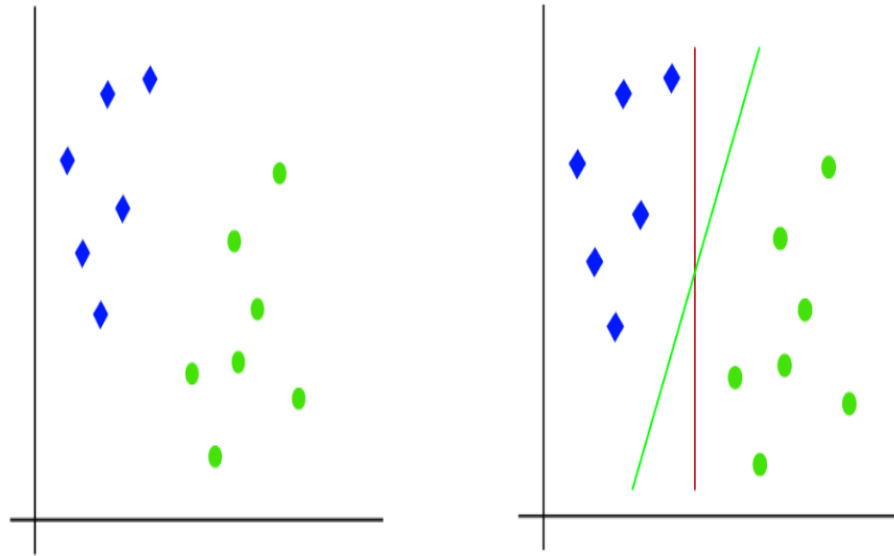
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suntha Jambi

2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suntha Jambi

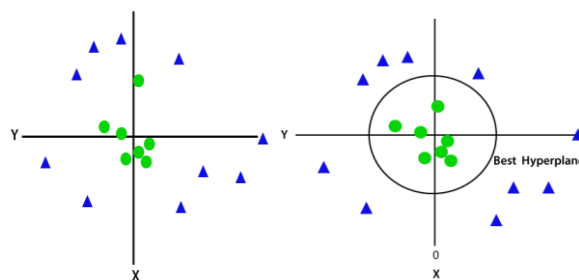
dengan satu baris menjadi dua kelas, data tersebut dapat dipisahkan secara linear dan pengklasifikasi yang digunakan disebut pengklasifikasi linear.



Gambar 2.1 SVM Linear

2. Support Vector Machine Non-linear

Support Vector Machine nonlinear digunakan untuk data yang dapat didiskritisasi secara nonlinear, Artinya, jika dataset tidak dapat diklasifikasikan secara linear, jadi data tersebut sebagai data classifier dan nonlinear yang digunakan disebut Support Vector Machine nonlinear classifier.



Gambar 2.2 SVM No-Linear

- Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
 2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

Algoritma Support Vector Machine berikut kelebihan dan kelemahan:

1. Kelebihan

- a. Bisa digunakan untuk ruang yang berdimensi tinggi.
- b. Lebih efektif jika jumlah pengukuran lebih besar dari jumlah sampel.
- c. penghemat memori karena titik pembelajaran fungsi keputusan (vector pendukung) digunakan.
- d. Bekerja relatif baik bila ada perbedaan yang jelas antara lapisan.

2. Kelemahan

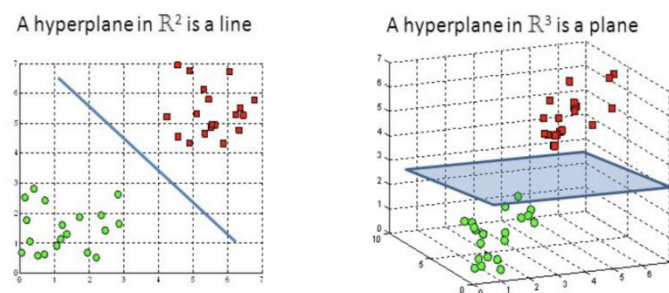
- a. Algoritma Support Vector Machine tidak bisa digunakan untuk kumpulan data besar karena membutuhkan pelatihan yang sangat lama.
- b. Mesin vektor pendukung kinerja tidak baik saat dataset memiliki banyak noise, seperti kelas target yang tumpang tindih.
- c. Ketika jumlah fitur per titik data melebihi jumlah sampel data pelatihan, mesin vektor pendukung akan mengalami kinerja buruk.
- d. Karena pengklasifikasi vektor dukungan bekerja dengan menempatkan titik data di atas dan di bawah hyperplane, klasifikasi tidak memiliki probabilitas yang jelas, yang dapat menyebabkan beban kerja komputasi yang besar.

Berikut ada beberapa uraian dari metode *Support Vector Machine*

Sebagai berikut:

1. Hyperplane

Mungkin ada beberapa garis keputusan atau batasan saat memisahkan kelas dalam ruang n -dimensi. Tetapi, untuk mengklasifikasikan Support Vector Machine kita perlu menemukan batas keputusan terbaik. Hyperplane adalah batas penting yang memisahkan dua kelas di Support Vector Machine. Dapat ditetapkan ke kategori yang berbeda pada poin data yang terletak di kedua sisi hyperplane. Ukuran hyperplane tergantung pada karakteristik kumpulan data. Maksudnya, jika ada dua fitur, hyperplane berbentuk persegi panjang, jika hyperplane tersebut merupakan bidang dua dimensi maka ada tiga fitur. Berikut ini contoh ilustrasi dari hyperplane di tunjukkan Gambar 2.3 di bawah ini.



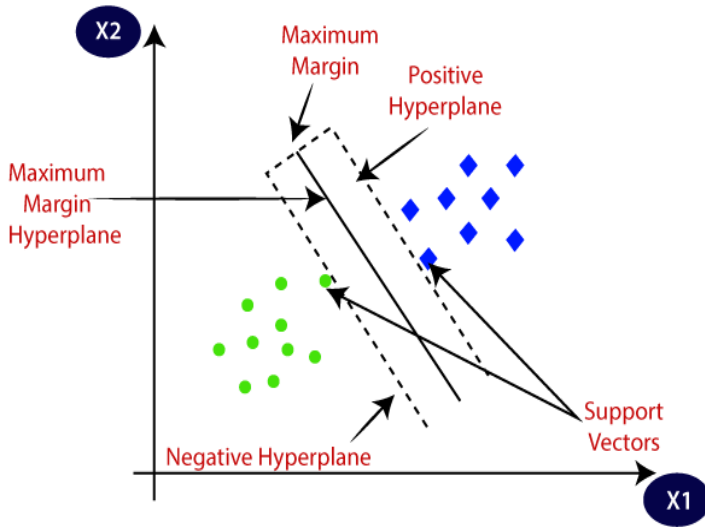
Gambar 2.3 Ilustrasi *Hyperplane*

2. Support Vector

Support vector atau pendukung vector adalah vektor terdekat dengan hyperplane (pemisah) dan mempengaruhi posisi titik data. Disebut vector dukungan, karena vektor ini mendukung hyperplane.

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi



Gambar 2.4 Support Vector

3. Max Margin

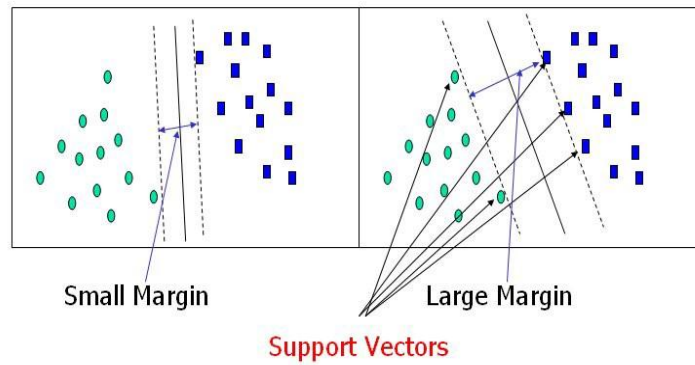
Margin adalah jarak support vector dari setiap kelas di sekitar hyperplane (pemisah). Pada gambar 2.5 di bawah, margin diilustrasikan dengan jarak antara 2 yang garis putus-putus. Batas maksimum (maximum margin) dapat dicari dengan cara memaksimalkan nilai jarak antara hyperplane dengan titik terdekatnya. Anda bias melihat pada gambar 2.5 di sebelah kanan memiliki margin yang lebih besar dari pada gambar di sebelah kiri. Secara intuitif, pengambilan yang lebih besar menghasilkan kinerja penyortiran yang lebih baik.

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

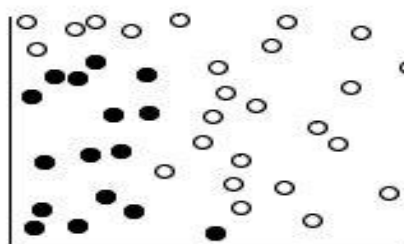


Gambar 2.5 Max margin

2.1.8 Cara Kerja Metode Algoritma *Support Vector Machine*

Algoritma Support Vector Machine bekerja dengan menggabungkan data dalam ruang fitur berdimensi tinggi untuk mengklasifikasikan titik data meskipun data tidak dapat dipisahkan secara linear. Dapat digunakan karakter data baru untuk memprediksi kelompok mana yang harus dimiliki kumpulan data baru. Jika pembatas ditemukan di antara kelas, data dapat diubah sehingga pembatas dapat digambarkan sebagai pembatas dapat digambarkan sebagai hyperplane.

perhatikan gambar 2.6 di bawah, yang berisi titik data terpisah dalam dua kategori, lingkaran hitam dan lingkaran putih.

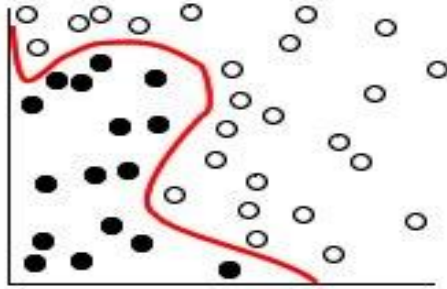


Gambar 2.6 Dua kategori

Hak Cipta Dilindungi Undang-Undang:

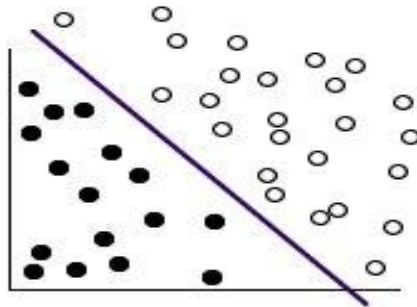
1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

Kedua kelas selanjutnya dipisahkan oleh kurva, ditunjukkan pada gambar 2.7 di bawah ini.



Gambar 2.7 Dua kategori terpisah oleh kurva

Setelah transformasi selesai, batas antara kedua kelas dapat ditentukan pada hyperplane, ditunjukkan pada gambar 2.8



Gambar 2.8 Dua kategori terpisah oleh *Hyperplane*

Fungsi matematis yang digunakan dalam proses transformasi algoritma Support Vector Machine disebut fungsi kernel. Untuk mengklasifikasikan data nonlinear menggunakan fungsi kernel. Caranya dengan merubah data nonlinear menjadi linear selanjutnya membuat hyperplane.

Fungsi kernel linier merekomendasikan saat melakukan pembagian linear langsung. Dalam kasus lain, fungsi tambahan mungkin diperlukan. Untuk menemukan model terbaik kami harus mencoba fungsi berbeda di setiap kasus, karena hal ini sering membutuhkan algoritma dan parameter yang berbeda.

2.2 Penelitian Terdahulu

Membedakan penelitian sebelumnya dengan penelitian yang dilakukan maka penulis melakukan tinjauan pustaka yang telah dilakukan peneliti sebelumnya, mirip dengan tinjauan pustaka sebelumnya.

Irfangi (2019) Hasil percobaan akurasi dari 109 data percobaan memberikan akurasi sebesar 84%. Melakukan penelitian dengan mengimplementasikan metode Naïve Bayes Classifier untuk melakukan analisis sentiment transfortasi online di Indonesia pada Twitter.

Septian (2019) Hasil percobaan akurasi cluster pada sistem ini adalah sebesar 86%. Melakukan penelitian mengimplementasikan metode Naïve Bayes Classification untuk melakukan analisis sentiment pada media Twitter milik Departemen Humas Polri dimana tweet yang dianalisis diklasifikasikan menjadi tiga tema: kebijakan, layanan masyarakat dan komentar masyarakat.

Bety Wulan Sari (2019) Hasil pengujian menggunakan metode *K-Fold* dan cross-validation matriks pada model yang dibuat dengan algoritma *Support Vector Machine* yang menghasilkan hasil dengan akurasi 83,2%,

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

presisi 78,8%, recall 71,6%, dan F1-score 75% untuk Biznet serta akurasi 79,6%, presisi 76,5%, recall 72,8% , dan F1-skot 74,6% untuk Telkom.

Ira Diana Sholihati (2020) penulis menggunakan metode *Support Vector Machine* (SVM) dalam analisis sentimen terhadap opini pengguna layanan Indihome di Twitter, tujuan untuk mendapatkan model klasifikasi sentimen menggunakan SVM dan menentukan tingkat akurasi yang akurat karena metode SVM yang diterapkan pada menganalisis sentimen, serta mengetahui tingkat kepuasan pengguna layanan Indihome berbasis Twitter. Setelah dilakukan pengujian dengan metode SVM didapatkan hasil accuracy 87%, precision 86%, recall 95%, error rate 13%, dan f1-score 90%.

Safitri Juanita (2021) Berdasarkan hasil yang diperoleh dari taghap pemodelan dengan menggunakan algoritma knearest neighbors dan membandingkan 60:40 untuk data latih dan uji, nilai akurasi dan presisi recall yang dihasilkan dari setiap aplikasi yaitu untuk bibit 85,14% , 91,91%, dan 76,44%, sedangkan untuk bareksa yaitu 81,70%, 87,15%, dan 75,73%. Penelitian ini menggunakan model CRISP-DM (*Cross Industry Standard Process for Data Mining*) dan algoritma yang digunakan dalam penelitian ini adalah K-Nearest Neighbors.

Tabel 2.1 Penelitian terdahulu

No	Peneliti	Objek	Metode	Hasil
1	Irfangi (2019)	transportasi online mengenai tweet	Naïve Bayes Classifier sebagai metode penelitian Tweet mengenai transportasi online	Hasil uji akurasi pengujian 109 data, dihasilkan akurasi sebesar 84%.

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

Tabel 2.1 (Lanjutan)

2	Septian Narsa Putra (2019)	Tweet mengenai Divisi Humas Polri	Naïve Bayes Classifier	Hasil akurasi pengujian adalah 86%. Tweet terbagi menjadi tiga topik: kegiatan polisi, layanan masyarakat dan komentar masyarakat.
3	Bety Wulan Sari (2019)	Objek berupa 500 data Tweet. Sebanyak 250 adalah Tweet mengenai Telkom dan 250 data teks lainnya mengenai Biznet.	Penelitian menggunakan metode SVM untuk melakukan analisis sentiment	Hasil penelitian menyimpulkan bahwa nilai accuracy sebesar 79,6% untuk Telkom dan accuracy 83,2% untuk Biznet.
4	Ira Diana Sholihati (2020)	Objek penelitian sebanyak 10.000 Tweet dengan rasio penggunaan data latih (training) dan data uji (testing) adalah perbandingan data 80:20	Metode menggunakan SVM dalam analisis sentimen.	Hasil pencarian dengan metode SVM diperoleh nilai accuracy 87%, precision 86%, recall 95%, error rate 13%, dan F1-score 90%.
5	Safitri Juanita (2021)	Objek penelitian adalah berupa review aplikasi pada Google Play Store. Dataset yang digunakan berjumlah 4.362 dengan 3.489 data latih 3.489 dan 873 data uji.	KNN sebagai metode dalam penelitian ini	Hasil dari penelitian dengan metode KNN ini mendapat nilai accuracy sebesar 81,70%.

Hak Cipta Dilindungi Undang-Undang:

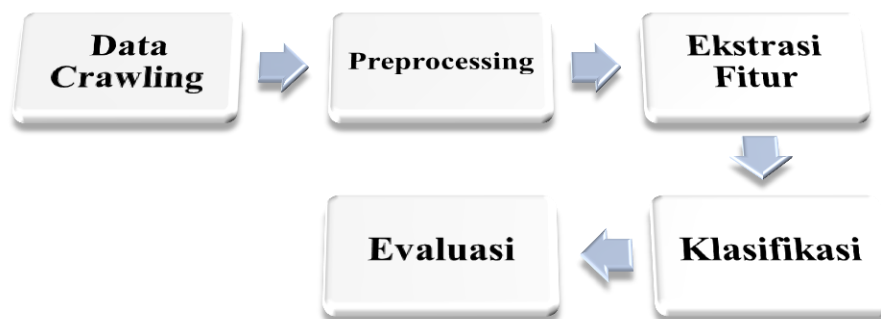
1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

BAB III

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Tahapan penelitian sebagai kumpulan fakta atau data yang diperoleh dari suatu objek penelitian. Objek penelitian dalam hal ini yang dimaksud data tweet pada media sosial Twitter selanjutnya data akan dijadikan sebagai sumber data. Dalam melakukan penelitian ini, diperlukan proses penelitian, agar proses penelitian tetap berjalan sesuai rencana. Gambar 3.1 menunjukkan proses alur pada penelitian ini.



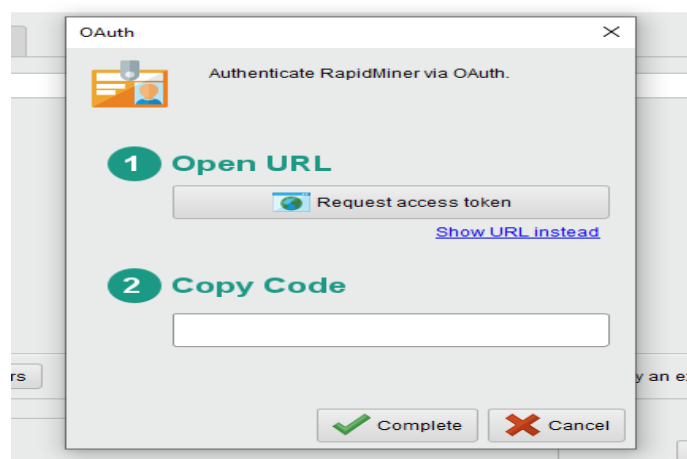
Gambar 3.1 Tahapan Penelitian

3.2 Data Crawling

Tahap pengumpulan data dengan melakukan crawling melalui RapidMiner Data yang dicrawl adalah berdasarkan kumpulan tweet pada media social Twitter. Kemudian, dataset tweet mentah yang disimpan sebagai CSV (Comma Separated Values) akan diproses lebih lanjut.

Data crawling adalah proses otomatis untuk mengumpulkan dari berbagai sumber seperti situs web, database, atau dokumen, pengumpulan data dilakukan dengan crawling dari data *tweet* menggunakan aplikasi RapidMiner dengan menentukan Keyword yang sedang ramai dibicarakan oleh banyak orang yaitu *#intitusi polri*, *#humas polri*, dan *#polda*. Setelah pengambilan data mentah dari kata kunci tersebut di tarik lalu data tersebut di simpan dalam format file CSV (Comma Separated Values) kemudian data tersebut digunakan dan di proses lebih lanjut.

RapidMiner adalah perangkat lunak yang berfungsi sebagai alat pembelajaran penambangan data. Tahapan yang awal dalam penelitian ini adalah Crawling Data. Sebelum melakukan penarikan data pada Twitter menggunakan aplikasi Rapid Miner perlu mengkoneksikan antara kedua aplikasi tersebut Rapid miner dan Twitter, pengkoneksian ditunjukkan pada Gambar 3.2 di bawah ini.



Gambar 3.2 Pengambilan Token

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

Pada gambar di atas proses pengkoneksian Rapid Miner dan Twitter, nantinya Twitter akan memberikan token akses (kode angka) di gunakan dalam proses pengkoneksian kedua aplikasi tersebut. Adapun tahapan Crawling data. Lihat Gambar 3.3 di bawah ini.



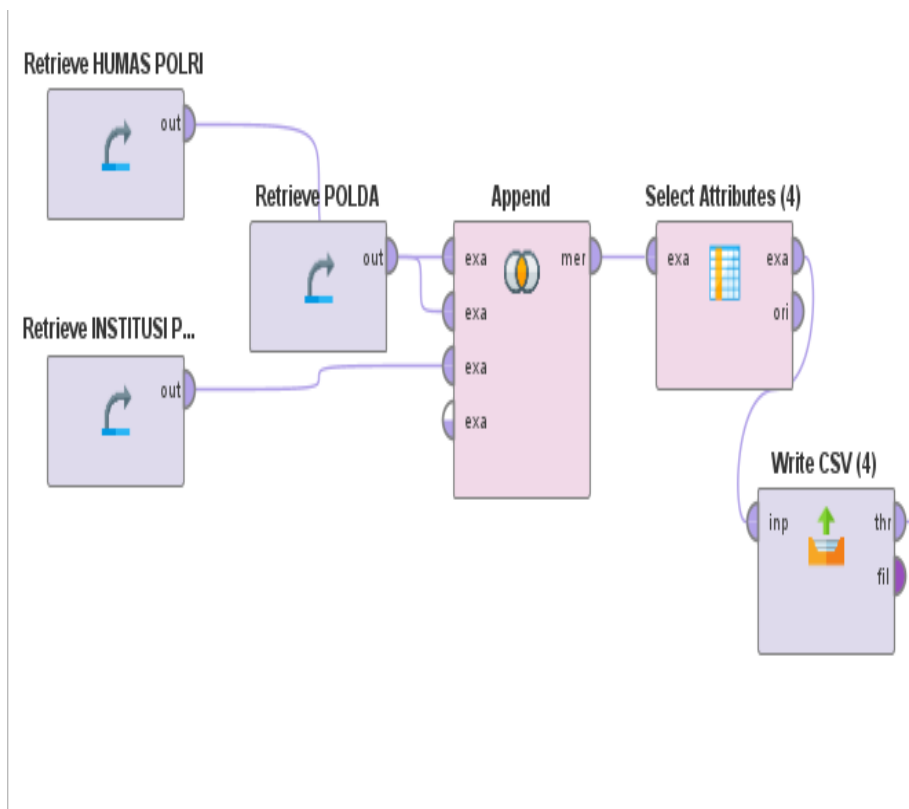
Gambar 3.3 Tahapan *crawling*

Tahapan *crawling* terdapat pada gambar 3.3 tahapan ini menggunakan tiga kata kunci yaitu (institusi polri), (humas polri) dan (polda) terdapat pada kata kunci (institusi polri) sebanyak 756 data, (humas polri) sebanyak 1.255 data dan kata kunci (polda) sebanyak 4.914 data.

Terdapat id didalam data maka diperlukan operator select attribute fungsinya hanya menampilkan text saja. Selanjutnya akan disimpan kedalam bentuk file write CSV.

Pada gambar 3.4 menjelaskan setiap kata kunci yang didapatkan pada data base twitter akan digabungkan menggunakan operator append fungsinya untuk menambah data dari setiap kata kunci, ketika masih terdapat id didalam

data akan diproses menggunakan select attribute untuk menampilkan sebuah text. lalu disimpan dalam bentuk file write CSV.



Gambar 3.4 Tahapan penggabungan data kata kunci

Table 3.1 Contoh text data mentah

No	Text
1	Ciptakan kamtibmas yang kondusif jelang Pemilu 2024, Divisi Humas Polri menggandeng sejumlah media massa. Hal tersebut disampaikan Kadiv Humas Polri Irjen Pol. Prof. Dr. Dedi Prasetyo, M.Hum., https://t.co/4DihjLHgDu.. M.M., saat Rakernis Humas Polri T.A 2023 di Bali, Jumat (3/3) https://t.co/RcHN6TyE8e
2	Kanit Binmas Polsek Rembang dan Anggota menerima kunjungan dan memberikan materi pengenalan profesi Polri kepada kelompok bermain (KB) Barokah desa Panusupan, KB Alkhoir Desa Sumampir
3	RT @antaranews: Karopenmas Divisi Humas Polri Brigjen Pol. Ahmad Ramadhan menegaskan bahwa penerimaan anggota kepolisian tidak dipungut bia...
4	Gerimis tak menjadi halangan bagi Personel Polres Solok Selatan dalam melaksanakan salah satu kewajibannya yakni apel pagi.

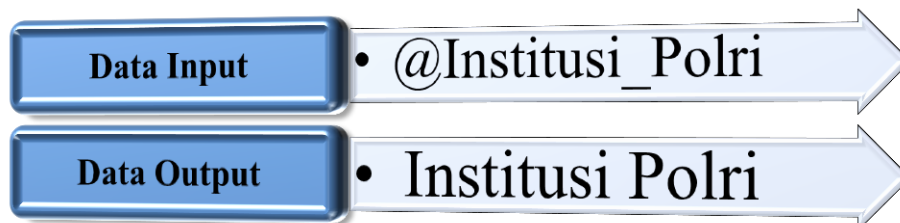
Table 3.1 (Lanjutan)

5	RT @DivHumas_Polri: Ciptakan kamtibmas yang kondusif jelang Pemilu 2024, Divisi Humas Polri menggandeng sejumlah media massa. Hal tersebut...
6	RT @Polsek_Denbar: Menghadapi Pemilu 2024. Kadiv Humas Polri Irjen Pol. Prof. Dr. Dedi Prasetyo, M.Hum., https://t.co/Ey702i4bns , M.M., me...
7	Mabes Polri memeriksa 24 saksi dalam rangka mengusut kebakaran hebat Depo Pertamina Plumpang yang terjadi pada Jumat (3/3/2023) silam. Kepala Biro Penerangan Masyarakat (Karo Penmas) Humas Polri Brigjen Ahmad Ramadhan menyebut, pengambilan keterangan... https://t.co/5dvVn1n83x

3.3 Preprocessing

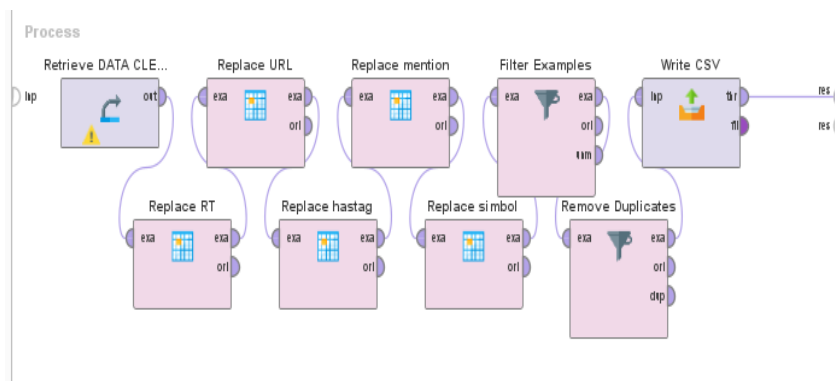
3.3.1 Cleansing

Selama proses ini, semua karakter di dalam post no-alfabet dihapus untuk mengurangi karakter yang tidak diinginkan dan tidak masuk akal dalam analisis sentiment. Karakter tersebut seperti angka, #, @, emoji, hingga link pada halaman web dalam sebuah artikel..



Gambar 3.5 Contoh *Cleansing*

Data yang dikumpulkan merupakan data yang tidak terstruktur sehingga membutuhkan Pre-Processing. Maka dibutuhkan proses tahapan Replace RT, URL, Hastag, mention, Simbol fungsinya untuk membersihkan data data yang dianggap tidak penting



Gambar 3.6 Tahapan pembersihan data

Berikut ada beberapa proses tahapan *cleansing*:

a. *Cleansing* Replace RT

Pada operator Replace RT digunakan sebagaimana fungsinya hanya menghilangkan retweet seperti RT @antaranews: terdapat contoh pada table 3.2

Table 3.2 Contoh text *cleansing* replace rt

No	Sebelum	Sesudah
1	RT @antaranews: Karopenmas Divisi Humas Polri Brigjen Pol. Ahmad Ramadhan menegaskan bahwa penerimaan anggota kepolisian tidak dipungut bia...	Karopenmas Divisi Humas Polri Brigjen Pol. Ahmad Ramadhan menegaskan bahwa penerimaan anggota kepolisian tidak dipungut bia...
2	RT @DivHumas_Polri: Ciptakan kamtibmas yang kondusif jelang Pemilu 2024, Divisi Humas Polri menggandeng sejumlah media massa. Hal tersebut...	Ciptakan kamtibmas yang kondusif jelang Pemilu 2024, Divisi Humas Polri menggandeng sejumlah media massa. Hal tersebut...
3	RT @polsekbandungTA: Divisi Humas Polri menggelar diskusi publik pada kegiatan Rakernis Humas Polri T.A 2023 di Pulau Dewata pada Kamis (2/...	Divisi Humas Polri menggelar diskusi publik pada kegiatan Rakernis Humas Polri T.A 2023 di Pulau Dewata pada Kamis (2/...

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:

 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi

2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

Hak Cipta Dilindungi Undang-Undang:
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
 2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

b. *Cleansing* Replace URL

Pada proses ini, semua karakter non-alfabet dalam pesan dihapus untuk mengurangi karakter yang tidak diinginkan dan tidak masuk akal dalam analisis sentiment. Karakternya seperti (tautan) dalam sebuah postingan. Contohnya seperti table 3.3

Table 3.3 Contoh text *cleansing* replace url

No	Sebelum	Sesudah
1	Perawatan Personel, Polres Sijunjung Gelar Uji Kesamaptaan Jasmani Semester 1 2023 https://t.co/U7qLW62az	Perawatan Personel, Polres Sijunjung Gelar Uji Kesamaptaan Jasmani Semester 1 2023
2	RT @pendim_tegal: Dandim 0712 Tegal Apel Bersama Sinergitas TNI-Polri Di Wilayah Tegal https://t.co/OXpCmsuZ9d	RT @pendim_tegal: Dandim 0712 Tegal Apel Bersama Sinergitas TNI-Polri Di Wilayah Tegal
3	RT @Polsek_Denbar: Menghadapi Pemilu 2024. Kadiv Humas Polri Irjen Pol. Prof. Dr. Dedi Prasetyo, M.Hum., https://t.co/Ey702i4bns. , M.M., me...	RT @Polsek_Denbar: Menghadapi Pemilu 2024. Kadiv Humas Polri Irjen Pol. Prof. Dr. Dedi Prasetyo, M.Hum.,

c. *Cleansing* Replace Hastag

Pada proses ini, semua karakter no-alfabet dalam pesan dihapus untuk mengurangi karakter yang tidak diinginkan dan tidak masuk akal dalam analisis sentiment. Karakter tersebut sama dengan symbol (#) pada sebuah pesan. Contohnya seperti table 3.4

Table 3.4 Contoh text *cleansing* replace hastag

No	Sebelum	Sesudah
1	Hati-hati ya sobat Polri dalam berkendara, tetap utamakan keselamatan. #polri	Hati-hati ya sobat Polri dalam berkendara, tetap utamakan keselamatan.
2	Divisihumaspolri #divhupolri	Divisihumaspolri
3	Kadiv Humas Polri #humaspolri #divhupolri	Kadiv Humas Polri

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

d. *Cleansing* Replace Mention

Pada proses ini, semua karakter non-alfabet dalam pesan dihapus untuk mengurangi karakter yang tidak diinginkan dan tidak masuk akal dalam analisis sentiment. Karakter seperti symbol (@) yang ada didalam sebuah postingan. Contohnya seperti table 3.5

Table 3.5 Contoh text *cleansing* replace mention

No	Sebelum	Sesudah
1	@polisi_indonesia	polisi_indonesia
2	@polisiindonesia	Polisiindonesia
3	@polantasindonesia	Polantasindonesia

e. *Cleansing* Replace Symbol

Pada proses ini, semua karakter non-alfabet dalam pesan dihapus untuk mengurangi karakter yang tidak diinginkan dan tidak masuk akal dalam analisis sentiment. Karakter tersebut seperti symbol (!@#\$\$%^&*_) yang ada didalam sebuah postingan. Contohnya seperti table 3.6

Table 3.6 Contoh text *cleansing* replace simbol

No	Sebelum	Sesudah
1	@berita_polisi_terkini	Beritapolisiterkini
2	*Disiplin, Militan, Profesional dan Dicintai Rakyat*	Disiplin Militan Profesional dan Dicintai Rakyat
3	Twitter : @humas_korbrimob	Twitter humaskorbrimob

f. *Cleansing* Filter Examples

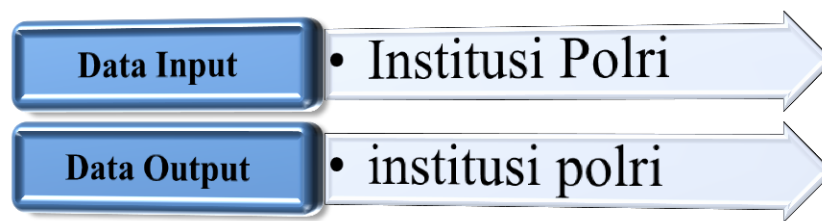
Saat menggunakan filter examples semua data yang kosong tidak mempunyai makna seperti ? akan akan dihapus permanen, sehingga tidak ada data text yang kosong.

g. *Cleansing* remove duplicates

Saat menggunakan remove duplilat semua data yang terduplilat akan dihapus permanen, sehingga tidak ada data yang secara tidak sengaja terhapus.

3.3.2 Case Folding

Proses mengubah setiap huruf didalam posts menjadi lowercase atau semuanya dijadikan huruf kecil.



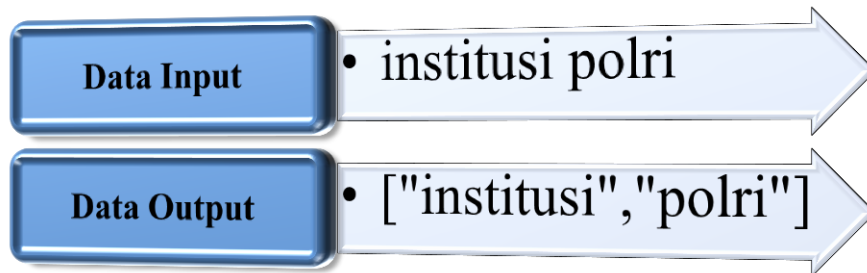
Gambar 3.7 Contoh case folding

Tabel 3.7 Contoh text case folding

No	Sebelum	Sesudah
1	Mabes Polri memeriksa saksi dalam rangka mengusut kebakaran hebat Depo Pertamina Plumpang yang terjadi pada Jumat silam Kepala Biro Penerangan Masyarakat Karo Penmas Humas Polri Brigjen Ahmad Ramadhan menyebut pengambilan keterangan...	mabes polri memeriksa saksi dalam rangka mengusut kebakaran hebat depo pertamina plumpang yang terjadi pada jumat silam kepala biro penerangan masyarakat karo penmas humas polri brigjen ahmad ramadhan menyebut pengambilan keterangan...
2	GERIMIS TAK MENJADI HALANGAN BAGI PERSONEL POLRES SOLOK SELATAN DALAM MELAKSANAKAN SALAH SATU KEWAJIBANNYA YAKNI APEL PAGI	gerimis tak menjadi halangan bagi personel polres solok selatan dalam melaksanakan salah satu kewajibannya yakni apel pagi
3	Mabes Polri memeriksa saksi dalam rangka mengusut kebakaran hebat Depo Pertamina Plumpang yang terjadi pada Jumat silam Kepala Biro Penerangan Masyarakat Karo Penmas Humas Polri Brigjen Ahmad Ramadhan menyebut pengambilan keterangan...	mabes polri memeriksa saksi dalam rangka mengusut kebakaran hebat depo pertamina plumpang yang terjadi pada jumat silam kepala biro penerangan masyarakat karo penmas humas polri brigjen ahmad ramadhan menyebut pengambilan keterangan...

3.3.3 Tokenizing

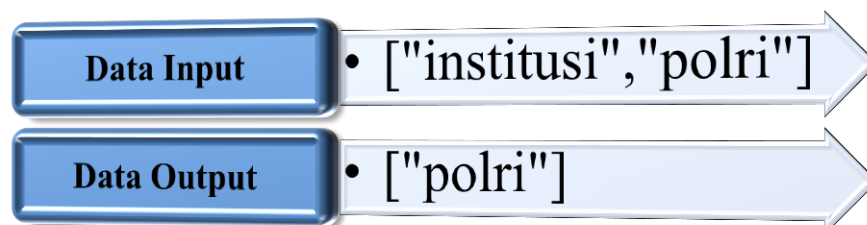
Pada tahap ini, sebuah tweet atau komentar dari pengguna yang sudah melalui tahap cleansing dan case folding dipisahkan dari kalimatnya menjadi sebuah kata.



Gambar 3.8 Contoh *Tokenizing*

3.3.4 Filtering

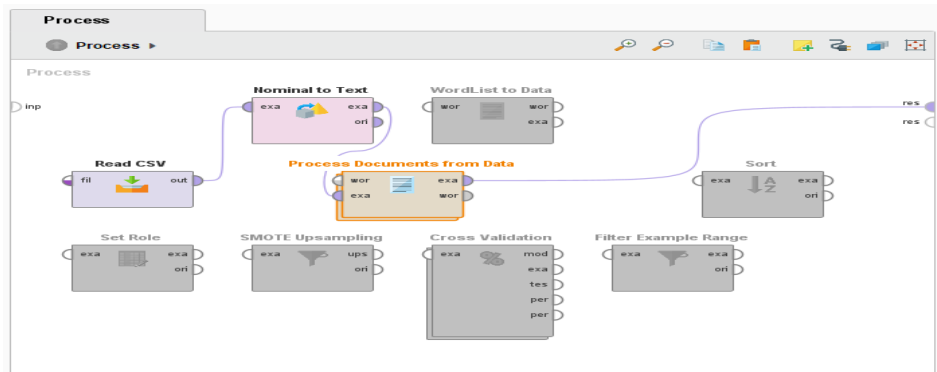
Langkah menghilangkan kata yang muncul dalam jumlah banyak tetapi memiliki arti. Alasan untuk menghapus kata-kata berhenti karena terlalu umum, sehingga pengguna dapat fokus pada kata-kata yang lebih penting. Pada dasarnya stopwords adalah kelompok kata yang banyak digunakan dalam berbagai bahasa.



Gambar 3.9 Contoh *filtering*

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi



Gambar 3.11 Proses TF-IDF

Tahap pertama nominal ke teks, untuk merubah jenis atribut nominal yang dipilih menjadi teks menggunakan operator nominal to text. Operator ini juga memetakan semua nilai properti ini ke nilai string yang sesuai. Data teks diproses untuk mengekstrak kata dan mengirim vector kata (numerik untuk mewakili teks). Selanjutnya process document from data atau TF-IDF didalam operator TF-IDF ini terdapat beberapa operator (*Tokenize1*, *Tokenize2*, *transform case*, *filter stopwords*, *filter token by length*). Tokenize pertama menghilangkan symbol-symbol, tokenize kedua menghilangkan angka-angka dan juga fungsinya untuk memecah kalimat menjadi kata perkata, transform case fungsinya untuk merubah seluruh suku kata menjadi huruf kecil semua supaya seragam, filter stopwords menghapus kata-kata tidak mempunyai makna contohnya kata sambung, filter by length untuk membatasi minimal character dalam satu kata. Berikut hasil dari TF-IDF terdapat pada Gambar 4.1

3.4.2 Pelabelan Manual

Pada tahap pelabelan manual ini menentukan sentiment dari dua kelas positif dan negative, yang didapatkan dari pendapat pertama, kedua

dan ketiga. Contohnya, Jika pendapat pertama mengatakan positif, kedua positif dan ketiga negatif maka sentiment positif. Maupun sebaliknya jika dua negatif satu positif maka sentiment negatif.

Keterangan:

P1 = pendapat pertama
P2 = pendapat kedua
P3 = pendapat ketiga
P = Positif
N = Negatif

Tabel 3.8 data pelabelan yang diproses secara manual, pada tahap pelabelan peneliti melakukan pelabelan dengan 3 orang untuk melakukan pelabelan tersebut.

Tabel 3.8 Pelabelan manual

No	Text	P1	P2	P3	Sentimen
1	DivHumasPolri kami ingatkanjgn sampai institusi polri dpt malu! sampe nitezen ikut menyelidikikecuali emang udah gak py urat malu	N	N	N	Negatif
2	HusinShihab polisijaksel Poldametrojaya DivHumasPolri ListyoSigitP Kalau blunder lagi berarti memang institusi nya yang salah bukan oknum nya kok tidak ada jera ² nya mereka memperlakukan diri	N	N	N	Negatif
3	Giat Bhabinkamtibmas Bripka Hendriyansah desa karya jaya Polsek Sinar Peninjauan Polres Oku POLDA SUMSEL mesosialisasikan dan memasang stiker Bantuan Polisi di desa karya jaya KecSinar Peninjauan kaboku	P	P	P	Positif
4	Kabid Humas Polda Jateng menyebut lima polisi yang diduga terlibat kasus KKN dalam penerimaan Bintara Polri telah disidang etik Dia menyebut ada dua ASN Polri yang diduga ikut terlibat	P	P	P	Positif
5	poldabali Dikypamanahrassa HumasPolri HumasBali niluhdjelantik ListyoSigitP kok jadi semacam mesin penjawab aja tapi...	N	P	N	Negatif
6	Terbaik Dalam Pengelolaan Media Bidhumas Polda Jateng Raih Penghargaan Bergengsi	P	P	P	Positif
7	Memalukan dan memuakkan Mohon supaya Kapolda Jateng segera di copotListyoSigitP mohmahfudmd jokowi ada kecenderungan berl...	N	N	P	Negatif

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

Tabel 3.8 (Lanjutan)

8	Sat Brimob Polda Metro Jaya menerima kunjungan dari TK Kingdom Kids dalam rangka kegiatan 'BRIMOB SAHABAT ANAK' adapun dalam kegiatan ini anak-anak diperkenalkan seragam Brimob pengenalan alat persenjataan dan kendaraan multifungsi Brimob Rabu	P	P	P	Positif
9	Anggota Polsek Sinar Peninjauan Polres Oku POLDA SUMSEL melaksanakan SISPAM MAKO Polsek sinar peninjauan	P	P	P	Positif
10	Polri yang tergabung dalam Emergency Medical Team EMT melaksanakan bakti sosial untuk masyarakat korban gempa di wilayah Hassa Provinsi Hatay Turki Minggu	P	P	P	Positif

3.5 Klasifikasi SVM

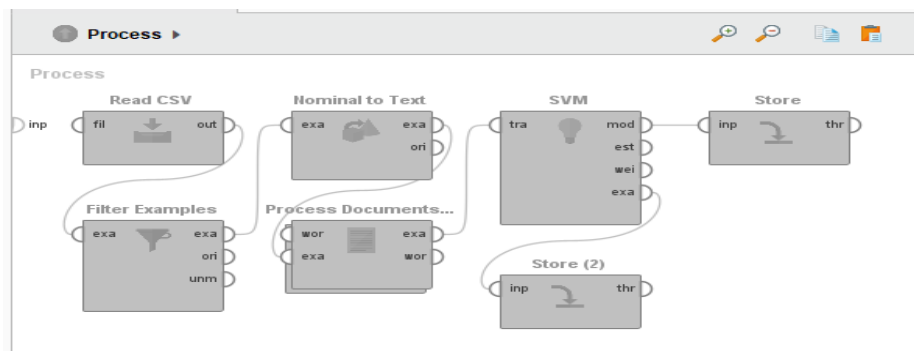
Klasifikasi metode svm berikut gambaran ilustrasi dan cara kerja metode svm terdapat pada landasan teori. Setelah diklasifikasikan menggunakan metode SVM. Metode ini dipilih karena memiliki akurasi yang lebih baik. Data dalam bentuk tweet akan di ambil dari media social Twitter lalu disimpan dalam bentuk file csv. File CSV (Comma Separated Values) adalah file teks sederhana yang berisi daftar data, file ini disebut nilai yang dipisahkan koma. File-file ini sering digunakan untuk bertukar data antara aplikasi yang berbeda. File CSV umumnya, gunakan koma untuk memisahkan atau membatasi data, tetapi terkadang karakter lain seperti titik koma juga digunakan. Kita dapat mentransfer data menggunakan file CSV dari satu sistem ke sistem lain dengan lebih mudah dan tanpa entri manual. Misalnya, Mengekspor data kompleks dari satu aplikasi ke file CSV, yang kemudian dapat diimpor / dimuat ke aplikasi lain.

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

Data dalam bentuk tweet diambil dari Twitter lalu disimpan dalam bentuk file CSV. Data dibagian dua dataset dijadikan dua subset data, data latih dan data uji. Dalam pelabelan akan disediakan untuk membedakan antara tweet negatif dan positif. Metode support vector machine digunakan pada tahap klasifikasi sentimen dan hasil interpretasi analis sentimen, selanjutnya data yang dipisahkan pada perbandingan data percobaan dan data uji kemudian diproses dengan metode *Support Vector Machine*.

Tahapan awal menggunakan metode *Support Vector Machine* pada *Tools Rapid Miner*, Modelkan metode *Support Vector Machine*. dilihat Gambar 3.12 bisa dilihat di bawah ini.



Gambar 3.12 Membuat model SVM pada Rapid Miner

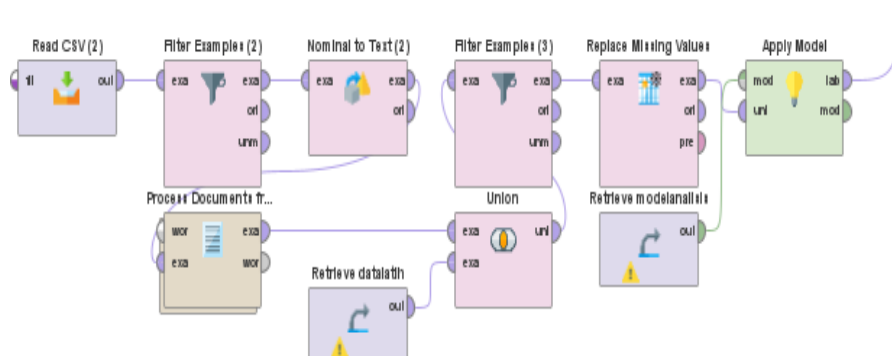
Pada Gambar 3.12 dapat dilihat model *Support Vector Machine* tahap pertama menggunakan operator Read CSV dalam operator tersebut masukkan file data latih yang sudah terlabel secara manual , selanjutnya operator *filter examples* berfungsi untuk menampilkan data yang hanya memiliki sentimen saja dengan pengaturan filter (sentimen, *is not missing*), kemudian untuk mengubah data nominal menjadi data text menggunakan operator nominal to text, selanjutnya operator process document untuk memproses document yang dalamnya berisi tokenize, transform cases filter stopwords, filter token

bylength. Selanjutnya buat model algoritma *Support Vector Machine* (SVM) menggunakan operator *Support Vector Machine* (SVM) setelah model terbentuk dan disimpan model dan data tersebut menggunakan operator store. Dalam membuat model ini yang diambil data yang sudah dilabel sentimennya menggunakan operator filter examples untuk tampilan datanya yang diproses, ditunjukkan Gambar 3.13 di bawah ini.

Row No.	LABEL	text
1	Positif	ciptakan kamtibmas kondusif jelang pemilu divisi humas polri menggandeng media massa kadiv humas polri irjen prof ...
2	Positif	kanit binmas polsek rembang anggota menerima kunjungan materi pengenalan profesi polri kelompok bermain barokah...
3	Positif	karopenmas divisi humas polri brigjen ahmad ramadhan penerimaan anggota kepolisian dipungut
4	Positif	hatihati sobat polri berkendara utamakan keselamatan
5	Positif	gerimis halangan personel polres solok selatan melaksanakan salah kewajibannya apel pagi
6	Positif	ciptakan kamtibmas kondusif jelang pemilu divisi humas polri menggandeng media massa
7	Positif	berani jujur bertanggung
8	Positif	divisi humas polri menggelar diskusi publik kegiatan rakernis humas polri pulau dewata Kamis
9	Positif	polri menerima kantong jenazah identifikasi lanju
10	Positif	kesempatan kadiv humas polri didampingi karo jianstra ssdm polri brigjen sandi nugroho mhum karo multimedia divhu...
11	Positif	tcoexryyris humas sembilan polda laki-laki perempuan ikuti
12	Positif	tcoexryyris humas biro perencanaan anggaran polda menggelar focus group discussion gedu
13	Positif	terima kunjungan danlantamal kupang kapolda polri garda terdepan bangsa Indonesia

Gambar 3.13 *Sentiment is not missing*

Berikut tahap mengisi data yang sentimennya masih kosong menggunakan model *Support Vector Machine* ditunjukkan pada Gambar 3.14 di bawah ini.



Gambar 3.14 Proses pelabelan dengan model data latih

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber aslinya:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

Pada gambar 3.14 proses pelabelan dengan model data data latih *Support Vector Machine* yang pertama pada tahapan ini masukan data menggunakan operator read CSV kemudian data filter menggunakan operator filter examples sama seperti membuat model pada tahap sebelumnya namun, pada filter examples ini kebalikannya hanya menampilkan data yang sentimennya kosong atau yang belum dilabel contohnya pada gambar dibawah.

Row No.	LABEL	prediction(LABEL)	confidence(Positif)	confidence(Negatif)	text
1	?	Positif	0.630	0.370	polisiindonesia
2	?	Positif	0.630	0.370	polantasindonesia
3	?	Positif	0.659	0.341	divisihumaspolri
4	?	Positif	0.630	0.370	ntmcpolri
5	?	Positif	0.630	0.370	beritapolisiterkini
6	?	Positif	0.630	0.370	beritapolantasterkini
7	?	Positif	0.630	0.370	humaspoldalampung
8	?	Positif	0.671	0.329	perawatan personel Polres Sijunjung gelar K...
9	?	Positif	0.681	0.319	jalur perkotaan kota cilegon
10	?	Positif	0.648	0.352	rabu
11	?	Positif	0.630	0.370	polrescilegon
12	?	Positif	0.630	0.370	lantasbanten

Gambar 3.15 *Sentiment no missing*

Selanjutnya mengubah data nominal menjadi text dengan operator nominal to text setelah data menjadi text berikut data diproses menggunakan operator process document untuk memproses document yang didalamnya terdiri tokenize, transform cases, filter stopword, filter tokenize by length sama seperti membuat model kemudian operator union diperlukan untuk proses ini untuk menggabungkan data percobaan dan data pengujian. Data percobaan diambil dari data yang telah memasuki tahap pembentukan model pada gambar 3.12 kemudian filter examples digunakan lagi untuk menampilkan data sentiment yang belum terisi selanjutnya menggunakan operator replace missing values untuk mengubah data yang tidak sama atributnya atau masi tanda Tanya dirubah menjadi angka 0, selanjutnya model

- Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
 2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

yang telah dibuat digunakan menggunakan operator apply model dihubungkan model tersebut dan data percobaan dengan data yang belum mempunyai sentiment kemudian diproses dan dijalankan, kemudian aplikasi rapid miner tersebut akan belajar dari data latih dan pemodelan data tersebut untuk menampilkan prediksi dari data yang belum ada sentimennya.

3.6 Evaluasi

3.6.1 Split validation

Validasi split adalah model klasifikasi mesin support vector sesuai dengan pembagian dua dataset dijadikan dua dataset, data pelatihan dan data uji (testing) data training atau data latih digunakan untuk training sedangkan data test (uji) digunakan untuk testing. Untuk representasi model validasi split yang lebih jelas, lihat gambar di bawah ini.



3.6.2 Cross Validation

Evaluasi kinerja dilakukan untuk memeriksa hasil klasifikasi dari metode tersebut, sehingga berupa hasil dari klasifikasi terdapat beberapa persen opini yang positif dan negative. Evaluasi dan pengujian dilakukan untuk pengujian terhadap model klasifikasi pengujian *k-fold crossvalidation* akan menggunakan $K = 5$ data dibagi lima kali pengujian yang memiliki jumlah sama data tersebut terdapat dari 1 data test dan 4 data training.

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jamb
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jamb

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

Tabel 3.9 Ilustrasi skema *5 fold validation*

Iterasi 1	Iterasi 2	Iterasi 3	Iterasi 4	Iterasi 5
Partisi 1	Partisi 2	Partisi 3	Partisi 4	Partisi 5
Partisi 1	Partisi 2	Partisi 3	Partisi 4	Partisi 5
Partisi 1	Partisi 2	Partisi 3	Partisi 4	Partisi 5
Partisi 1	Partisi 2	Partisi 3	Partisi 4	Partisi 5
Partisi 1	Partisi 2	Partisi 3	Partisi 4	Partisi 5

Data latih

Data uji

3.6.3 Alat yang digunakan

Alat yang digunakan dalam penelitian ini adalah :

1. PC/Laptop dengan spesifikasi processor Intel(R), Core(TM) i5-8250U, CPU @ 1.60GHz 1.80GHz, RAM 8.00 GB, System operasi Windows 10,64-bit.
2. Aplikasi RapidMiner

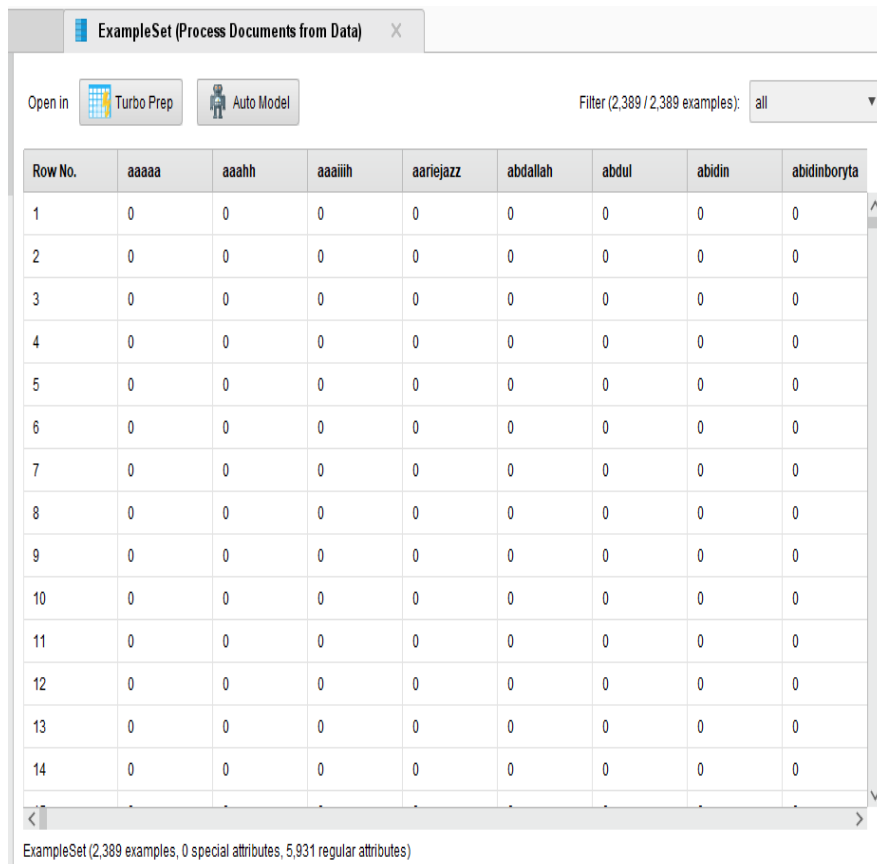
BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

4.1.1 Hasil Etrasi Fitur

Hasil proses tf-idf dalam sebuah dokumen dari jumlah data 2389, kata setiap teks tersebut di pecah menjadi kata perkata yang sebanyak 5931 kata (regular atribut). angka 0 tersebut jumlah data hasil dari tf-idf yang menghitung suatu kata yang muncul dalam dokumen bisa dilihat Gambar 4.1 di bawah ini:



Row No.	aaaaa	aaahh	aaaiih	aariejazz	abdallah	abdul	abidin	abidinboryta
1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0

Gambar 4.1 Hasil TF-IDF

Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli;
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah,
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

4.1.2 Hasil Pemodelan Manual Metode Support Vector Machine

Tabel 4.1 Keterangan *confusion matrix*

No	Kelas	Terklasifikasi Positive	Terklasifikasi Negative
1	Positif	TP (True Positif)	FN (False Negatif)
2	Negatif	FP (False Positif)	TN (True Negatif)

Keterangan:

- a. TP (*True Positive*), yaitu jumlah data positif yang diklasifikasikan dengan benar oleh sistem.
- b. TN (*True Negative*), yaitu jumlah data negatif yang diklasifikasikan dengan benar oleh sistem.
- c. FN (*False Negative*), yaitu jumlah data yang negatif tetapi diklasifikasikan oleh sistem memiliki kesalahan.
- d. FP (*False Positive*), yaitu jumlah data yang positif tetapi diklasifikasikan sebagai salah oleh sistem.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} * 100\%$$

$$\text{Presisi} = \frac{TP}{FP+TP} * 100\%$$

$$\text{Recall} = \frac{TP}{FN+TP} * 100\%$$

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

Tabel 4.2 Hasil pengujian data *testing* 600

	True Negatif	False Positif
Prediksi Negatif	2 Benar Negatif	174 False Positif
Prediksi Positif	1 False Negatif	423 Benar Positif
Total	600	

$$\text{Akurasi} = \frac{423+2}{423+2+174+1} * 100\%$$

$$= 71$$

$$\text{Presisi} = \frac{423}{174+423} * 100\%$$

$$= 70.85$$

$$\text{Recall} = \frac{423}{1+423} * 100\%$$

$$= 99.76$$

Tabel 4.3 Hasil pengujian data *testing* 400

	True Negatif	False Positif
Prediksi Negatif	0 Benar Negatif	55 False Positif
Prediksi Positif	1 False Negatif	344 Benar Positif
Total	400	

$$\text{Akurasi} = \frac{344+0}{344+0+55+1} * 100\%$$

$$= 86$$

$$\text{Presisi} = \frac{344}{55+344} * 100\%$$

$$= 86.21$$

$$\text{Recall} = \frac{344}{1+344} * 100\%$$

$$= 99.71$$

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

Berdasarkan pengujian manual Dari hasil pengujian manual data testing 600 (73:27) didapatkan Accuracy sebanyak 71%, Precision 70.85% dan Recall 99.76%. Pengujian kedua data testing 400 (82:18) didapatkan Accuracy sebanyak 86%, Precision 86.21% dan Recall 99.71%. Bisa kita lihat Tabel di atas 4.2 dan 4.3

4.1.3 Hasil Split Validation Otomatis

Pada gambar 4.2 dihasilkan *confusion matrix* dengan data testing 915 (60:40) Ketika diklasifikasikan menggunakan metode *Support Vector Machine (SVM)* dinilai s[plit ratio = 0.6 adalah $TP = 772$, $FP = 117$, $FN = 3$, $TP = 23$ dan hasil *Precision* dikalkulasikan 86.84% sedangkan *recall* didapat 99.61%. Berdasarkan hasil *confusion matrix support vector machine (SVM)* didapat akurasi sebesar 86.89%.

accuracy: 86.89%

	true Positif	true Negatif	class precision
pred. Positif	772	117	86.84%
pred. Negatif	3	23	88.46%
class recall	99.61%	16.43%	

Gambar 4.2 Hasil *accuracy split validation* 40%

Pada gambar 4.3 dihasilkan *confusion matrix* dengan data testing 458 (80:20) Ketika diklasifikasikan menggunakan metode *Support Vector Machine (SVM)* dinilai split ratio = 0.6 adalah $TP = 386$, $FP = 55$, $FN = 2$, $TP = 15$ dan hasil *Precision* dikalkulasikan 87.53% sedangkan *recall*

didapat 99.48%. Berdasarkan hasil *confusion matrix support vector machine (SVM)* didapat akurasi sebesar 87.55%.

accuracy: 87.55%

	true Positif	true Negatif	class precision
pred. Positif	386	55	87.53%
pred. Negatif	2	15	88.24%
class recall	99.48%	21.43%	

Gambar 4.3 Hasil *accuracy split validation* 20%

4.1.4 Hasil Crossvalidation Otomatis

Pada gambar 4.4 dihasilkan *confusion matrix* pada klasifikasi dengan menggunakan metode *Support vector machine* dinilai $k=5$ adalah $TP = 1922$, $FP = 290$, $FN = 16$, $TP = 61$ dan hasil *precision* dikalkulasikan 86,89% sedangkan *recall* didapat 99,17%. Berdasarkan hasil *confusion matrix Support vector machine* didapat akurasi sebesar 86,63%.

accuracy: 86.63% +/- 0.59% (micro average: 86.63%)

	true Positif	true Negatif	class precision
pred. Positif	1922	290	86.89%
pred. Negatif	16	61	79.22%
class recall	99.17%	17.38%	

Gambar 4.4 Hasil *accuracy crossvalidation*

4.1.5 Hasil Percobaan Data Baru

Percobaan data baru peneliti mencoba menambahkan data baru dalam klasifikasi ini untuk melihat hasil prediksi. Untuk hasil percobaan data baru lihat Tabel 4.4

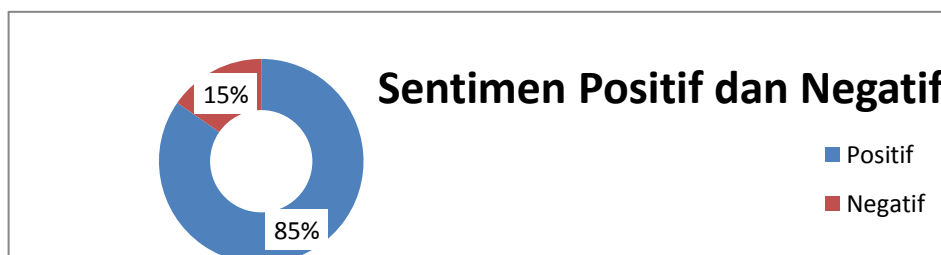
Tabel 4.4 Hasil percobaan data baru

No	Text	Prediksi Sentimen	Hasil
1	Saya bangga adanya institusi polri	Positif	Benar
2	polri the best	Positif	Benar
3	Polri Maksimalkan Penanganan Medis Korban Kebakaran	Positif	Benar
4	institusi polri bobrok	Positif	Salah
5	semogah saja institusi polri bisa beruba supaya benar benar mengayomi dan melindungi masyarakat bukan karena adanya uang	Positif	Benar
6	sehat selalu pak kapolri	Positif	Benar
7	luar biasa polri semogah langkah ini kedepannya membawa perubahan terhadap institusi polri	Positif	Benar
8	semogah polri bisa jauh lebih profesional dan lebih baik	Positif	Benar
9	institusi polri tidak tegas	Positif	Salah

Berdasarkan data di atas terprediksi 7 data benar dan 2 data salah.

4.1.6 Jumlah Sentimen Positif Dan Negatif

Pada gambar 4.5 menampilkan diagram lingkaran dengan jumlah data 85% positif dan 15% data negatif. Sentiment positif berjumlah 1938 data dan sentiment negatif berjumlah 351 data.



Gambar 4.5 Jumlah positif negatif

4.1.7 Evaluasi

Dari hasil pengujian diatas, pengukuran akurasi menggunakan *confusion matrix* membuktikan bahwa hasil uji algoritma Support Vector Machine memiliki skor akurasi baik. Setelah dilakukan pengujian manual data testing 600 akurasi sebesar 71%, pengujian manual data testing 400 akurasi sebesar 86%, pengujian sistem split validation data (60:40) akurasi sebesar 86.89%, pengujian sistem split validation data (80:20) akurasi sebesar 87.55% dan pengujian sistem cross validation data keseluruhan menggunakan k-fold 5 akurasi sebesar 86.673%. Perbandingan hasil pengujian terbaik ditunjukkan Tabel 4.5 di bawah ini.

Tabel 4.5 Hasil prediksi

Perbandingan Hasil Tes	Accuracy	Precision	Recall
pengujian manual data testing 600	71%	70.85%	99.76%
Pengujian manual data testing 400	86%	86.21%	99.71%
Split validation data (60:40)	86.89%	86.84%	99.61%
Split validation data (80:20)	87.55%	87.53%	99.48%
Cross validation data keseluruhan	86.63%	86.89%	99.17%

4.2 Pembahasan Penelitian

Penelitian analisis sentimen dalam mengklasifikasi *sentiment Public* opini masyarakat tentang institusi polri berdasarkan Twitter. Adapun keunggulan yang dimiliki yaitu menggunakan pengujian manual untuk memastikan *accuracy* yang didapatkan apakah sesuai dengan *accuracy* pengujian sistem.

Melakukan proses pengujian menggunakan metode *Support Vector Machine*. Berdasarkan hasil untuk mengetahui *Accuracy* menggunakan

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

aplikasi Rapid Miner dengan pengujian *Split Validation* data (60:40) dan (80:20) maka diketahui semakin besar data training semakin bagus *Accuracy* yang didapatkan, dengan hasil tersebut maka dapat dikatakan bahwa pengujian manual dan pengujian sistem tidak ada perbedaan.

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Hasil penelitiandan pembahasan dapat disimpulkan bahwa:

Metode support vector machine dapat digunakan sebagai mengklasifikasi sentiment positif dan negative. Analisis sentiment opini masyarakat tentang institusi polri berdasarkan twitter dengan metode support vector machine (SVM) dimulai tahapan pengumpulan data (Crawling) sebanyak 6.925 data mentah dari 3 kata kunci *Institusi polri*, *Humas polri* dan *Polda*. Hasil kesimpulan dari pengujian pelabelan manual data testing 600 (73:27) didapatkan *accuracy* sebanyak 71% dan data testing 400 (82:18) didapatkan *accuracy* 86%. Pengujian sistem data (80:20) didapatkan *accuracy* sebanyak 87.55% dan data (60:40) didapatkan *accuracy* sebanyak 86.89%. Semakin banyak data training semakin besar *accuracy* yang didapatkan

5.2 Saran

Berdasarkan kajian yang telah dilakukan, penulis dapat memberikan rekomendasi untuk pengembangan selanjutnya, antara lain:

1. Menambah jumlah data latih yang digunakna untuk melatih model klasifikasi agar akurasi yang diperoleh lebih tinggi.

2. Peningkatan pembersihan dokumen tweet pada prapemrosesan untuk hasil terbaik
3. Tambahkan sentiment netral
4. Menambahkan beberapa metode ekstrasi fitur lainnya agar mendapatkan bobot yang maksimal dan fitur yang tepat.
5. Meningkatkan kinerja algoritma Support Vector Machine pada sistem, mempercepat waktu yang diperlukan untuk proses klasifikasi.
6. Sistem dapat dikembangkan untuk analisis sentiment dengan topic berbeda lainnya.
7. Menambahkan metode untuk menentukan opini dan bukan opini.

@ Hak cipta milik UIN Sutha Jambi

State Islamic University of Sulthan Thaha Saifuddin Jambi



UNIVERSITAS ISLAM NEGERI
SULTHAN THAHA SAIFUDDIN
J A M B I

- Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
 2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi

DAFTAR PUSTAKA

- Aldisa, R. T., & Maulana, P. (2022). Analisis Sentimen Opini Masyarakat Terhadap Vaksinasi Booster COVID-19 Dengan Perbandingan Metode Naive Bayes, Decision Tree dan SVM. *Technology and Science (BITS)*, 4(1), 106–109. <https://doi.org/10.47065/bits.v4i1.1581>
- Alfian, E. (2020). Tugas dan Fungsi Kepolisian Untuk Meningkatkan Kepercayaan Publik terhadap Penegak Hukum. *Legalitas: Jurnal Hukum*, 12(1), 27. <https://doi.org/10.33087/legalitas.v12i1.192>
- Ananda, F. D., & Pristyanto, Y. (2021). Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine. *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 20(2), 407–416. <https://doi.org/10.30812/matrik.v20i2.1130>
- Dasmasela, R., Tomasouw, B. P., & Leleury, Z. A. (2021). Penerapan Metode Support Vector Machine (SVM) untuk Mendeteksi Penyalahgunaan Narkoba. *Matematika, Statistik dan Terapannya*, 1(02), 93–101. <https://ojs3.unpatti.ac.id/index.php/parameter/article/view/3630>
- Edi Saputra Hasibuan. (2021). Persepsi Masyarakat Terhadap Penerimaan Anggota Polri. *Jurnal Hukum Sasana*, 7(1), 33–50. <https://doi.org/10.31599/sasana.v7i1.526>
- Gifari, O. I., Adha, M., Freddy, F., & Durrand, F. F. S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. *Journal of Information Technology*, 2(1), 36–40. <https://doi.org/10.46229/jifotech.v2i1.330>
- Mustofa, A., & Novita, R. (2022). Klasifikasi Sentimen Masyarakat Terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan Text Mining Pada Twitter. *Jurnal Building of Informatics, Technology and Science (BITS)*, 4(1), 200–208. <https://doi.org/10.47065/bits.v4i1.1628>
- Residencia, A. (2019). *KLASIFIKASI KELAYAKAN PESERTA PENGAJUAN KREDIT RUMAH DENGAN ALGORITMA NAÏVE BAYES DI PERUMAHAN AZZURA RESIDENCIA*. 9, 43–48.
- Ritonga, H., Marlina, M., & Mustamam, M. (2022). PENINDAKAN PROPAM (POLRI TERHADAP ANGGOTA POLISI YANG MELAKUKAN PENGANIAYAAN (Studi Di Bidang Propam Kepolisian Resor Nias Selatan). *Jurnal Ilmiah METADATA*, 4(3), 215–227.
- Rizki, M. F., Auliasari, K., & Primaswara Prasetya, R. (2021). Analisis Sentiment Cyberbullying Pada Sosial Media Twitter Menggunakan Metode Support Vector Machine. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 5(2), 548–556. <https://doi.org/10.36040/jati.v5i2.3808>

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:

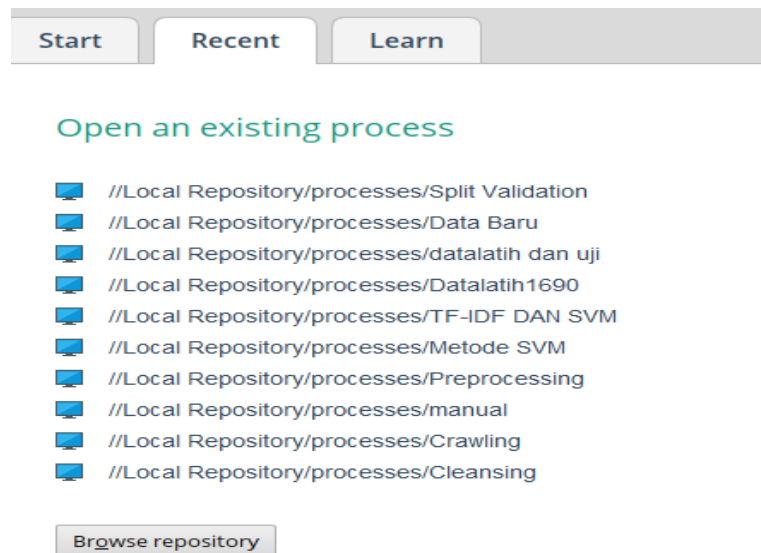
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suftha Jambi

LAMPIRAN



Lampiran 1 Aplikasi rapid miner



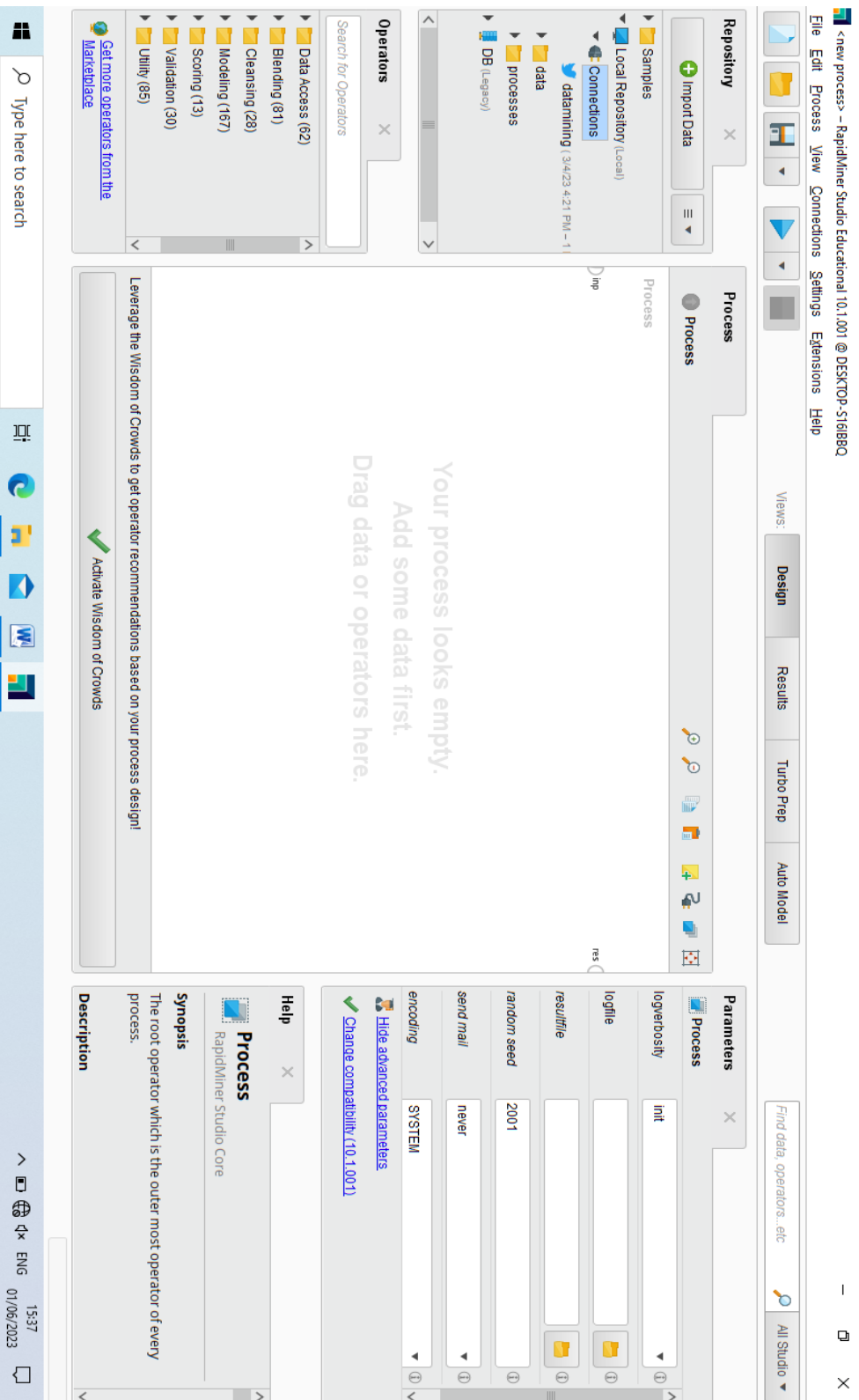
Lampiran 2 Tampilan recent

@ Hak cipta milik UIN Sutha Jambi

State Islamic University of Sulthan Thaha Saifuddin Jambi



Hak Cipta Dilindungi Undang-Undang:
1. Dilarang mengutip sebagian dan atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah,
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian dan atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi



Lampiran 3 Tampilan awal rapid miner

@ Hak cipta milik UIN Sultha Jambi

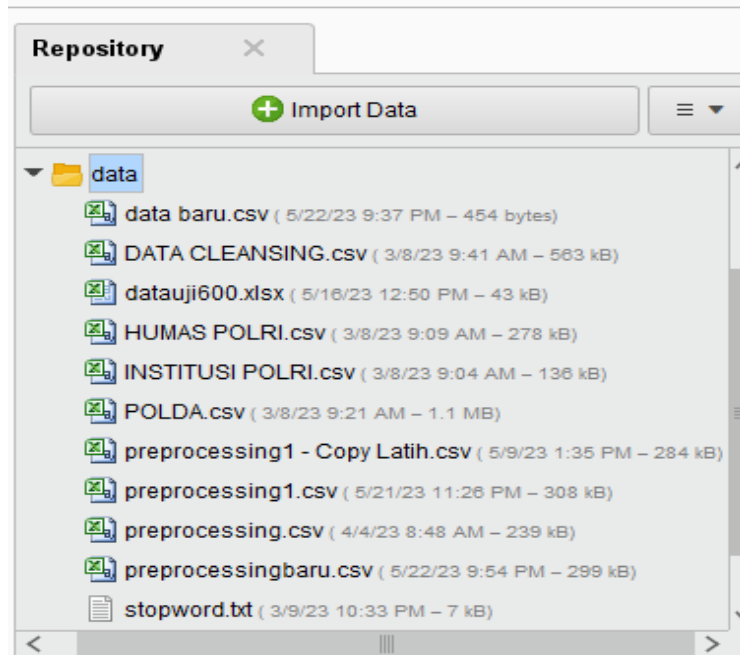
State Islamic University of Sulthan Thaha Saifuddin Jambi

Hak Cipta Dilindungi Undang-Undang:

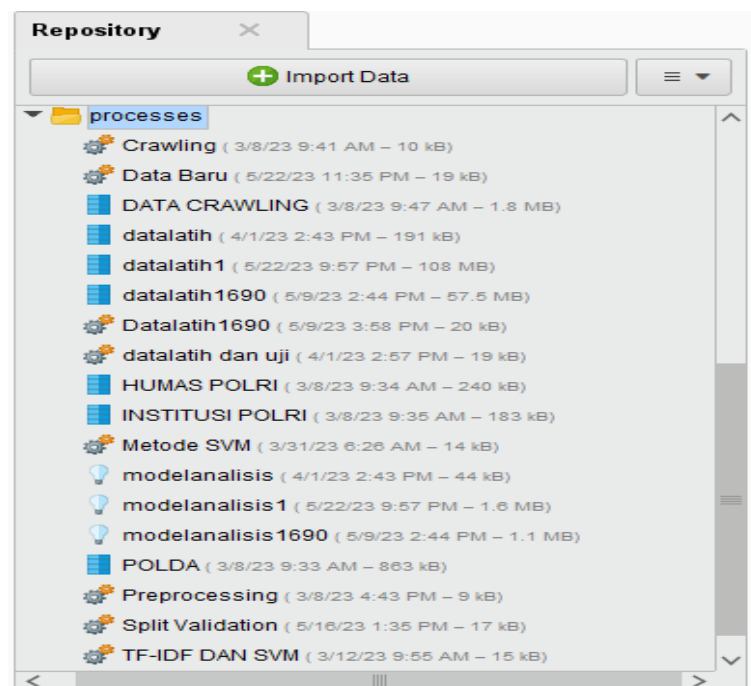
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber aslinya:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sultha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sultha Jambi

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber aslinya:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi



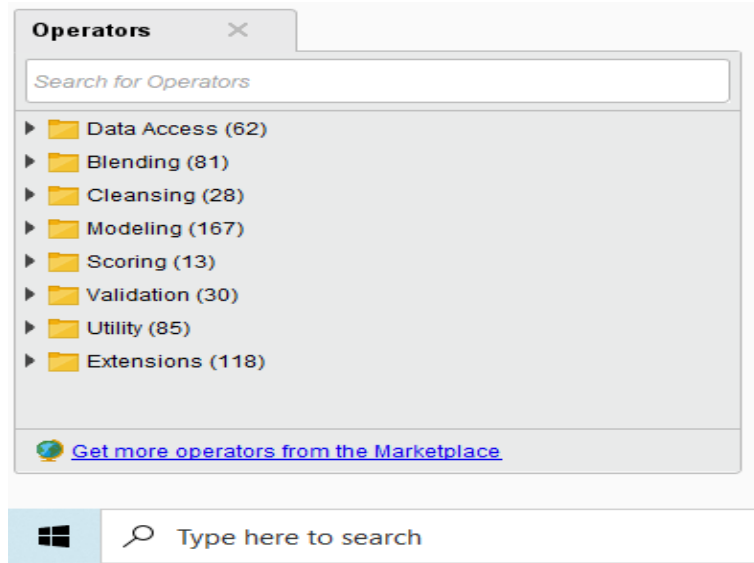
Lampiran 4 Tampilan *repository* data



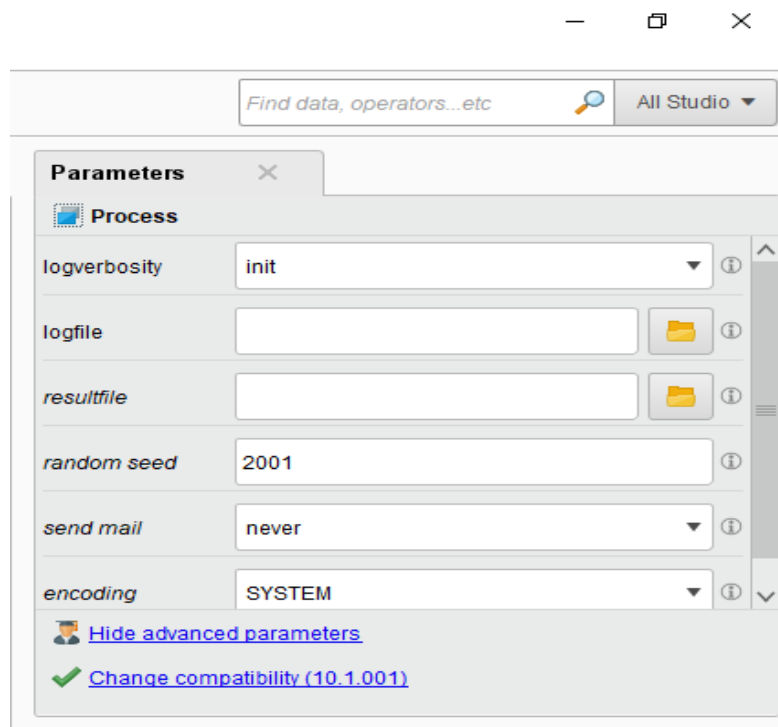
Lampiran 5 Tampilan *repository* process

Hak Cipta Dilindungi Undang-Undang:

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan da menyebutkan sumber asli:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Sutha Jambi
2. Dilarang memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Sutha Jambi



Lampiran 6 Tampilan operators



Lampiran 7 Tampilan parameters