



## Analisis Sentimen Masyarakat Terhadap Perekutan PPPK Pada Twitter Dengan Metode Naive Bayes Dan Support Vector Machine

**Fajar Nurrachmat Hidayat<sup>1</sup>, Sugiyono<sup>2</sup>**<sup>1,2</sup> Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika<sup>1</sup>fajarnurrachmat@gmail.com, <sup>2</sup>inosoguy007@gmail.com

**Abstrak**— Peneliti mengidentifikasi beberapa masalah yang sering disorot oleh masyarakat terkait Isu Perekutan PPPK melalui marketplace guru oleh MENDIKBUD Perekutan melalui marketplace guru oleh MENDIKBUD menjadi isu terhangat dan kontroversial baru-baru ini. Isu ini dilontarkan pertama kali oleh Mendikbud Ristek Nadiem Makarim, Postingan akun media sosial twitter @cnbcindonesia dengan Tweet Jelang perekutan 1 juta guru PPPK, Mendikbud Ristek Nadiem Makarim sudah menyiapkan strategi. Penelitian ini dilakukan untuk mengetahui pendapat masyarakat terhadap Perekutan PPPK pada media sosial Twitter. Salah satu algoritma yang dapat digunakan untuk melakukan analisis sentimen adalah Naïve Bayes dan Support Vector Machine (SVM), bertujuan sebagai variable tambahan untuk menganalisis pendapat para pengguna aplikasi Perekutan PPPK pada media sosial Twitter. Hasil akhir dari Perbandingan dengan dua metode pengujian ini, yaitu hasil prediksi Sentimen Masyarakat Terhadap Isu Perekutan PPPK berdasarkan data yang didapat dari Twitter dan diimplementasikan dengan metode SVM (Support Vector Machine) menunjukkan nilai akurasi sebesar 94,80 %. Dari 482 data uji, terprediksi sebesar 335 data sebagai Sentimen Negatif dan 147 data sebagai Sentimen Positif dan Metode Naïve bayes menunjukkan nilai akurasi sebesar 96,14 %. Dari 482 data uji, terprediksi sebesar 335 data sebagai Sentimen Negatif dan 147 data sebagai Sentimen Positif.

**Kata Kunci:** Sentimen, Twitter, PPPK, Naïve Bayes, Support Vector Machine

**Abstract**— The researcher identified several problems that are often highlighted by the community related to the Issue of PPPK Recruitment through the teacher marketplace by MENDIKBUD Recruitment through the teacher marketplace by MENDIKBUD has become the hottest and most controversial issue recently. The Minister of Education and Culture and Research and Technology Nadiem Makarim first raised this issue, posting a Twitter social media account @cnbcindonesia with a Tweet Ahead of the recruitment of 1 million PPPK teachers, the Minister of Education and Culture and Research and Technology Nadiem Makarim has prepared a strategy. This research was conducted to find out the public's opinion of PPPK recruitment on Twitter social media. One of the algorithms that can be used to carry out sentiment analysis is Naïve Bayes and Support Vector Machine (SVM), intended as an additional variable to analyze the opinions of users of the PPPK Recruitment application on Twitter social media. The final result of the comparison with these two test methods, namely the prediction of Community Sentiment on PPPK Recruitment Issues based on data obtained from Twitter and implemented with the SVM (Support Vector Machine) method shows an accuracy value of 94.80%. Of the 482 test data, it is predicted that 335 data are Negative Sentiment and 147 data are Positive Sentiment and the Naïve Bayes Method shows an accuracy value of 96.14%. Of the 482 test data, it is predicted that 335 data are Negative Sentiment and 147 data are Positive Sentiment.

**Keywords:** Sentiment, Twitter, PPPK, Naïve Bayes, Support Vector Machine

### I. PENDAHULUAN

Perekutan Pegawai Pemerintah dengan Perjanjian Kerja (PPPK) melalui marketplace guru adalah sebuah konteks yang tidak masuk akal dan tidak sesuai dengan prosedur rekrutmen resmi yang diterapkan oleh pemerintah. Namun, secara umum, PPPK adalah program rekrutmen pegawai pemerintah di Indonesia yang diperkenalkan untuk mengisi kekurangan tenaga kerja di sektor pendidikan, kesehatan, dan bidang lainnya. Proses rekrutmennya dilakukan secara terbuka, transparan, dan sesuai dengan peraturan yang telah ditetapkan oleh pemerintah.

Pemerintah biasanya mengumumkan lowongan PPPK melalui situs web resmi, media sosial, atau surat kabar resmi. Prosedur rekrutmen dilakukan melalui seleksi administrasi, tes tertulis atau ujian kompetensi, serta wawancara dan evaluasi akhir. Semua tahap ini dilakukan untuk memastikan calon pegawai yang terpilih memenuhi persyaratan kualifikasi dan memiliki kompetensi yang dibutuhkan.

Perekutan melalui marketplace guru tidak sesuai dengan prosedur ini dan dapat menimbulkan keraguan mengenai keabsahan rekrutmen tersebut. Jika ada informasi tentang perekutan PPPK melalui marketplace guru, penting untuk memverifikasinya melalui sumber resmi, seperti situs web resmi pemerintah atau instansi terkait, sebelum mengambil langkah lebih lanjut.

Dalam situasi seperti ini, disarankan untuk menghubungi instansi terkait, seperti Kementerian/Lembaga terkait atau Badan Kepegawaian Negara (BKN), untuk mendapatkan klarifikasi dan informasi yang akurat mengenai proses rekrutmen PPPK.

Berdasarkan pencarian di media sosial Twitter, Peneliti mengidentifikasi beberapa masalah yang sering disorot oleh masyarakat terkait Isu Perekutan melalui marketplace guru oleh MENDIKBUD Perekutan melalui marketplace guru oleh MENDIKBUD menjadi isu terhangat dan kontroversial baru-baru ini. Isu ini dilontarkan pertama kali oleh Mendikbud Ristek Nadiem Makarim, Postingan akun media sosial twitter @cnbcindonesia dengan Tweet Jelang perekutan 1 juta



guru PPPK, Mendikbud Ristek Nadiem Makarim sudah menyiapkan strategi.

Analisis sentimen merupakan kegiatan untuk mencari opini masyarakat tentang sebuah objek yang ingin di ketahui. Penelitian ini dilakukan untuk mengetahui pendapat masyarakat terhadap Perekutan PPPK pada media sosial Twitter. Salah satu algoritma yang dapat digunakan untuk melakukan analisis sentimen adalah Naive Bayes dan Support Vector Machine (SVM), bertujuan sebagai variable tambahan untuk menganalisis pendapat para pengguna aplikasi Perekutan PPPK pada media sosial Twitter.

Berdasarkan Hasil akhir dari Perbandingan dengan dua metode pengujian ini, yaitu hasil prediksi Sentimen Masyarakat Terhadap Isu Perekutan PPPK berdasarkan data yang didapat dari Twitter dan diimplementasikan dengan metode SVM (Support Vector Machine) menunjukkan nilai akurasi sebesar 94,80 %. Dari 482 data uji, terprediksi sebesar 335 data sebagai Sentimen Negatif dan 147 data sebagai Sentimen Positif dan Metode Naïve bayes menunjukkan nilai akurasi sebesar 96,14 %. Dari 482 data uji, terprediksi sebesar 335 data sebagai Sentimen Negatif dan 147 data sebagai Sentimen Positif[1].

Penyelenggaraan pendidikan di Indonesia masih diwarnai dengan persoalan manajemen guru, khususnya di kalangan pengajar honorer, serta reformasi birokrasi. Aparatur Sipil Negara berupaya meningkatkan kualitas pelayanan. Kementerian Pendidikan dan Kebudayaan bekerjasama dengan Kementerian Pendayagunaan Aparatur Negara, Reformasi Birokrasi, dan Keuangan menyepakati perubahan sistem rekrutmen pegawai pemerintah dari penerimaan Calon Pegawai Negeri Sipil (CPNS) menjadi Pegawai Pemerintah dengan Perjanjian Kerja (PPPK), yang masih terjadi pro dan kontra, oleh karena itu, peneliti akan menggunakan data mining untuk melakukan analisis sentimen terhadap pelaksanaan Pegawai Pemerintah dengan Perjanjian Kontrak (PPPK), di salah satu media sosial yaitu Twitter, menggunakan jumlah data sebanyak 871 data, yg selanjutnya difiltering dan cleaning sehingga tersisa data yg sudah bersih sebanyak 519 data, sesudah menghapus data yan ganda atau terduplikat, data yg kosong serta membersihkan data[1].

Pendidikan merupakan salah satu sektor yang dipengaruhi oleh perkembangan teknologi. Saat ini pesatnya kemajuan teknologi menjadi salah satu unsur penting dalam dunia pendidikan[2].

Undang-Undang Republik Indonesia Nomor 25 Tahun 2009 tentang Pelayanan Publik mengamanatkan agar semua penyelenggara pelayanan publik dapat menyediakan pelayanan yang berkualitas bagi pengguna layanan atau yang disebut dengan pelayanan prima. Salah satu asas dalam penyelenggaraan pelayanan publik adalah tersedianya SDM yang profesional. Untuk mewujudkan kinerja pemerintah yang prima, dibutuhkan SDM (ASN) yang berkualitas. SDM merupakan sumber daya yang digunakan untuk menggerakkan dan menyinergikan sumber daya lainnya untuk mencapai tujuan organisasi[3].

Indonesia merupakan negara berkembang yang sedang melaksanakan pembangunan nasional salah satunya berupaya menyelenggarakan pemerintahan yang baik (good governance). Proses penyelenggaraan pembangunan nasional dapat terwujud dengan sumber daya manusia yang berkualitas dan disiplin tinggi. Dengan adanya pembangunan nasional yang baik untuk mewujudkan suatu masyarakat adil dan makmur yang merata materil dan spiritual berdasarkan Pancasila di dalam wadah Negara Kesatuan Republik Indonesia, maka untuk mewujudkannya adalah dengan melakukan rekrutmen[4].

Kebutuhan akan Aparatur Sipil Negara (ASN) yang kompeten terus meningkat sesuai dengan tujuan pemerintah dalam menyelenggarakan roda pemerintahan yang efektif dan efisien.

Rekrutmen CASN (Calon Aparatur Sipil Negara) menjadi ujung tombak dalam mendapatkan pegawai yang kompeten dan profesional. Melalui proses rekrutmen CASN yang berlandaskan pada meritokrasi diharapkan dapat memperbaiki tatanan birokrasi yang dulu dikenal sebagai pegawai yang bekerja lamban, tidak profesional, rawan KKN, dan cenderung tidak netral dapat dihapuskan serta berangsur-angsur hilang. ASN yang terdiri dari PNS dan PPPK tentu harus diisi oleh putera-puteri terbaik negeri ini, yang kelak akan menghadapi berbagai disruptif di ranah birokrasi[5].

Perjalanan panjang birokrasi yang penuh rintangan dan berbagai kelemahan sejatinya sangat menuai banyak kritikan sehingga menjadi sebuah tantangan besar bagi birokrasi untuk melakukan upaya mereformasinya. Salah satu upaya mendasar yang perlu disasar dalam reformasi birokrasi adalah penataan sistem manajemen SDM Aparatur Negara dalam peraturan perundang-undangan. Hadirnya Undang-Undang Aparatur Sipil Negara yang secara sekuisial menghadirkan PPPK dalam cakrawala kepegawaian di Indonesia merupakan langkah Pemerintah dalam upaya mereformasi birokrasi. Pasalnya, kehadiran PPPK yang secara mutatis mutandis menghapuskan eksistensi tenaga honorer disinyalir mampu untuk lebih memberikan perlindungan hukum dan menjawab carut marut perihal kompetensi, integritas dan profesionalitas Aparatur Negara[6].

Tidak dapat dipungkiri bahwa untuk kelancaran tugas administrasi pemerintahan dibutuhkan adanya SDM aparatur/Pegawai ASN yang bertanggung jawab dalam melaksanakan tugasnya sebagai penyelenggara pelayanan publik bagi masyarakat, berupa pemberian pelayanan atas barang, jasa, dan/atau pelayanan administratif. Pegawai ASN memiliki kewajiban melayani masyarakat yang berkaitan dengan hak dasar, seperti akses pendidikan, kesehatan, dan perizinan, secara efektif dan efisien. Sayangnya, tingkat kepuasan masyarakat sebagai pengguna pelayanan masih rendah. Selain masih banyaknya keluhan masyarakat, kompetensi dan profesionalitas SDM aparatur sebagai penyedia layanan juga masih dipertanyakan[7].

Pegawai aparatur negara merupakan suatu unsur pelaksana penyelenggaraan pemerintah negara yang



mempunyai peran sentral dan strategis terhadap keberhasilan pembangunan nasional. Kinerja aparatur negara dari waktu ke waktu terus mengalami penyempurnaan dan peningkatan seirama dengan tuntutan dan perubahan lingkungan strategis yang berkembang begitu cepat, baik secara nasional, regional maupun global. Hal ini sejalan dengan semangat reformasi birokrasi yang didengungkan di berbagai kementerian/lembaga dan pemerintah daerah mencakup 3 elemen dasar, yaitu kelembagaan, ketatalaksanaan dan sumber aparaturnya[8].

Di sisi lain dengan disetujuinya Permenpan-RB Nomor 27 Tahun 2021 mengatur tentang Pengadaan Pegawai Negeri Sipil (PNS) serta Permenpan-RB Nomor 28 Tahun 2021 mengatur tentang Pengadaan Pegawai Pemerintah dengan Perjanjian Kerja (PPPK) untuk Jabatan Fungsional Guru pada Instansi Daerah mengindikasikan bahwa tes CASN tahun 2021 tidak ada pengangkatan PNS baik bagi sarjana pendidikan maupun guru honorer. Menteri Kemendikbud Nadiem Makarim mengindikasikan bahwa seleksi CPNS bagi guru akan ditiadakan untuk beberapa tahun kedepan dengan berbagai pertimbangan. Selain itu seleksi PPPK menjadi perdebatan baru dalam pelaksanaannya, dimana guru honorer dengan usia lebih dari 40 tahun diharuskan berkompetisi dengan guru honorer yang berusia muda. Serentetan permasalahan tersebut mengindikasikan bahwa perlakuan pemerintah terhadap guru honorer jauh dari nilai keadilan dan Hak Asasi Manusia[9].

Negara merupakan hasil konstruksi suatu organisasi sosial politik yang diakui memiliki kekuasaan (match) atau kedaulatan (sovereignty) atas suatu wilayah, penduduk dan pemeritahan. Pengertian tentang negara mengalami pertumbuhan dan perkembangan mengikuti irama peradaban umat manusia. Karena itu, pemaknaan tentang negara dirumuskan sesuai dengan irama zamannya. Pertumbuhan[10].

Salah satu persoalan mendasar dalam sistem kepegawaian adalah pekerjaan tempat Pegawai Negeri Sipil (PNS) mengabdi saat ini belum dipandang sebagai sebuah profesi yang mulia, harus dihormati, dijaga, dan dijadikan dasar dalam berbagai kebijakan dan manajemen Sumber Daya Manusia (SDM). PNS sebagai abdi negara dan abdi masyarakat tidak dianggap sebagai aset negara, bahkan kadang-kadang di- pandang menjadi beban negara[11].

## II. METODE PENELITIAN

Pada bagian ini akan dijelaskan mengenai perancangan sistem dalam penelitian ini. Metode dalam penelitian ini dilakukan melalui beberapa tahap. Secara garis besar, alur penelitian dapat dijelaskan dibawah ini:

### A. Analisa Sentimen

Analisa Sentimen atau opinion mining mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistik dan text mining yang bertujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenaan dengan suatu topik, produk layanan, organisasi, individu,

ataupun kegiatan tertentu Tujuan dari analisa sentimen adalah untuk menentukan perilaku atau opini dari seorang penulis dengan memperhatikan suatu topik tertentu. Perilaku bisa mengindikasikan alasan, opini atau penilaian, kondisi kecenderungan Analisa sentimen dapat juga berupa bentuk perasaan marah, gembira maupun sedih. Langkah-langkah yang umumnya ditemukan pada klasifikasi teks analisa sentimen adalah Definisikan domain dataset Pengumpulan dataset yang melingkupi suatu domain, misalnya dataset review film, dataset review produk, dan lain sebagainya.

- a) Pre-processing Tahap pemrosesan awal yang umumnya dilakukan dengan proses Tokenization, Stopwords removal, dan Stemming.
- b) Transformation Proses representasi angka yang dihitung dari data tekstual. Binary representation yang umumnya digunakan dan hanya menghitung kehadiran atau ketidakhadiran sebuah kata di dalam dokumen. Berapa kali sebuah kata muncul di dalam suatu dokumen juga digunakan sebagai skema pembobotan dari data tekstual. Proses yang umumnya digunakan yaitu TF-IDF, Binary transformation, dan Frequency transformation.
- c) Feature Selection Pemilihan fitur (feature selection) bisa membuat pengklasifikasi lebih efisien/efektif dengan mengurangi jumlah data untuk dianalisa dengan mengidentifikasi fitur yang relevan yang selanjutnya akan diproses. Metode pemilihan fitur yang biasanya digunakan adalah Expert. Knowledge, Minimum Frequency, Information gain, Chi-Square, dan lain sebagainya.
- d) Classification Proses klasifikasi umumnya menggunakan [12].

### B. Text Mining

Text mining atau text analytics adalah istilah yang mendeskripsikan sebuah teknologi yang mampu menganalisis data teks semi-terstruktur maupun tidak terstruktur, hal inilah yang membedakannya dengan data mining dimana data mining mengolah data yang sifatnya terstruktur. Pada dasarnya, text mining merupakan bidang interdisiplin yang mengacu pada perolehan informasi (information retrieval), data mining, pembelajaran mesin (machine learning), statistik, dan komputasi linguistic[12].

### C. Naïve Bayes

Salah satu metode klasifikasi yang dapat digunakan adalah metode Naïve Bayes yang sering disebut dengan Naive Bayes Classifier (NBC). Kelebihan metode NBC adalah sederhana tetapi memiliki akurasi yang tinggi. Berdasarkan penelitian yang berjudul sentimen analysis untuk memanfaatkan saran kuesioner dalam evaluasi pembelajaran dengan menggunakan naïve bayes classifier menyebutkan metode NBC dapat memperoleh akurasi mencapai 85,95% [13].

Support Vector Machine (SVM) merupakan salah satu bahasa pesan ke bahasa alami dengan menghilangkan ekspresi teknik pembelajaran supervised dengan kualitas dan akurasi atipikal sehingga dapat meminimalkan noise pada tahapan baik yang membuatnya menjadi algoritma



populer diantara selanjutnya[14]. SVM adalah algoritma supervised learning yang banyak digunakan pada masalah klasifikasi yang memiliki kehandalan[15].

#### D. Metode Pengumpulan Data

Proses pengumpulan data pada penelitian ini akan menggunakan data set publik. Data set publik didapat dengan Crawling Data *Tweet* para pengguna *Twitter*.

#### E. Sampel Data Traning

Dataset yang digunakan terdiri dari 2000 baris record tweets. Tweet's telah dilabeli, 1330 tweets positif dan 439 tweets negatif. Record terdiri dari 2 kolom, yang pertama adalah "Text" memuat isi tweets opini-opini Perekurban PPPK adalah sentiment yang merupakan label dari setiap tweets tersebut, dengan nilai positive dan negatif. Sampel dari data training yang digunakan 2 data terlihat seperti tabel 1 berikut:

Tabel 1. Sampel Data Training

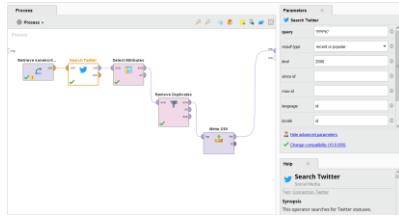
Text	Sentimen
 CNBC Indonesia @cnbcindonesia · May 25 Jelang perekrutan 1 juta guru PPPK, Mendikbud Ristek Nadiem Makarim sudah menyiapkan strategi. Salah satunya dengan menciptakan marketplace guru. Tentunya gimana ya?	Positive (1)
 Qfr @Unusualfr · May 27 Nih yang guru-guru ceritanya mau masukkan ke marketplace atau gmn ricnhh+ Apakah nantinya akan solidout??  0 12 73 5	Negative (2)

#### F. Penerapan Metodologi

Metode Naive Bayes dan SVM (Support Vector Machine) merupakan dua algoritma yang sering digunakan dalam bidang Machine Learning. Berikut adalah penerapan metodologi menggunakan kedua metode tersebut.

#### G. Proses Pengumpulan Dataset

Data dikumpulkan dari platform media sosial Twitter dengan menggunakan alat pengumpul data yang disebut API. Data yang dikumpulkan meliputi komentar pengguna pada Tweet pppk yang berbeda. Data yang diambil adalah 2000 komentar yang diambil dari Tweet Perekurban PPPK populer. Proses pengumpulan data pada penelitian ini akan menggunakan data set publik. Data set publik didapat dari Tweet para pengguna Twitter dengan menggunakan Operator "Search Twitter" yang ada pada RapidMiner Studio dengan kata kunci "PPPK" sebanyak 2000 data yang selanjutnya di filter untuk membuang tweet duplikasi, lalu data di ubah kedalam bentuk CSV. Seperti contoh gambar dibawah ini:



Gambar 1. Proses Crawling Data

- a) **Retrive koneksi Twitter** pada operator RapidMiner adalah salah satu fitur yang memungkinkan pengguna untuk mengambil data dari Twitter dan melakukan analisis data menggunakan platform RapidMiner. Fitur ini memungkinkan pengguna untuk mengakses data

publik di Twitter, seperti tweet, pengguna, dan topik terkait, dan menggunakan data tersebut untuk melakukan analisis data. Untuk menggunakan fitur retrive koneksi Twitter pada RapidMiner, pengguna perlu menghubungkan akun Twitter mereka dengan platform RapidMiner dan mendapatkan akses token untuk mengakses API Twitter. Setelah itu, pengguna dapat mengambil data yang diinginkan dari Twitter menggunakan operator Retrive Twitter. Operator ini memungkinkan pengguna untuk mengatur parameter pencarian seperti kata kunci, tanggal, lokasi, pengguna tertentu, dan banyak lagi, sehingga pengguna dapat memilih data yang ingin diambil dengan lebih spesifik.

- b) **Operator RapidMiner Search Twitter** juga memungkinkan pengguna untuk melakukan filter terhadap tweet berdasarkan tanggal, lokasi, jumlah retweet, dan follower. Setelah tweet diambil, pengguna dapat melakukan preprocessing pada data, seperti membersihkan tweet dari karakter yang tidak diinginkan, mengubah huruf kecil menjadi huruf besar, dan menghilangkan stopword.

- c) **Operator RapidMiner Select Attributes** adalah operator yang digunakan untuk memilih atribut atau variabel tertentu dari dataset. Dalam pengolahan data, terkadang tidak semua atribut atau variabel pada dataset diperlukan dalam analisis atau pemodelan. Operator Select Attributes membantu dalam mengurangi dimensi data atau feature selection, sehingga dapat meningkatkan performa analisis dan pemodelan data.

- d) **Operator RapdiMiner "Remove Duplicates"** digunakan untuk menghapus baris data duplikat dari dataset yang diberikan. Operator ini sangat berguna ketika peneliti memiliki data yang besar dan ingin memastikan bahwa tidak ada duplikasi dalam dataset Anda. Dengan menggunakan operator "Remove Duplicates", Peneliti dapat menghapus baris data yang identik dari dataset dengan mudah dan cepat.

- e) **Operator RapdiMiner "Write CSV"** adalah operator yang digunakan untuk menulis data hasil pengolahan RapidMiner ke dalam format file CSV (Comma Separated Values). CSV adalah format file yang umum digunakan untuk menyimpan data tabel dalam bentuk teks, di mana setiap baris dalam file mewakili sebuah baris dalam tabel, dan kolom-kolom dipisahkan oleh tanda koma.

#### H. Pelabelan Dataset

Sebanyak 482 data akan dilabeli secara manual yang nantinya akan digunakan sebagai data latih. Untuk mengurangi penilaian secara subyektif, proses Pelabelan data dibagi menjadi 2 label, yaitu label "Positif" dan label "Negatif".



A	B
text	SENIMEN
tenaga kesehatan sidoarjo terima pppk https://dijijim	POSITIVE
bupati ahmad muhdior menerahkan pppk tenaga kesehatan sidoarjo https://gfvvq	POSITIVE
diangkat pppk pesan khusus bupati gresik fandi akhmad yani yani hadapan tenaga pp	POSITIVE
kempanrb	POSITIVE
rekrutmen calon pegawai negeri cpns pegawai pemerintah perjanjian kerja pppk	NEGATIVE
kabangunkid bkppd gunungkidul usulkan rekrutmen pppk formasi https://kglopywy	POSITIVE
bupati sidoarjo ahmad muhdior menerahkan pppk pegawai pemerintah pe	NEGATIVE
bupati mubar upayakan tenaga pppk https://jscaojje/zonasultra.com	NEGATIVE
surabay update penetapan pppk	NEGATIVE
lantik pppk bupati pamekasan harapan kontribusi penuh	NEGATIVE
program market place nadien	NEGATIVE
beasiswa prajabatan objek pengensi kesosongan guru guru daerah	NEGATIVE
mindos kedepan syarat guru wajib	NEGATIVE
bkppd gunungkidul usulkan rekrutmen pppk formasi https://kglopywy	POSITIVE
serahkann pppk kesehatan gresik yani berharap membawa manfaat masyarakat gresik	POSITIVE
collegemenfess setau skrg guru cpns seleskiannya namanya pppk nder	NEGATIVE
bupati ponorogo sugiri sancoko pesan khusus tenaga kesehatan nakes diangkat pega	POSITIVE
pemerintah membuka lowongan calon pegawai negeri cpns pegawai pemerintah NEGATIVE	NEGATIVE
serahkann pengangkatan pppk bupati gresik warga berobat dicemberuti https://fifaa bu	NEGATIVE
pmekot serang usulkan formasi pppk https://jyhh.aakhm	POSITIVE
	NEGATIVE

Gambar 2. Tahap Pelabelan data

### I. Tahap Cleaning Data

Tahap ini bertujuan untuk membersihkan Tweet dari kata-kata yang tidak diperlukan seperti karakter hastag “#”, mention “@”, membuang url-url atau simbol simbol yang tidak dibutuhkan dalam proses sentimen analisis.

#	Text
1	text
2	@sidoarjokota memberikan pesan mengenai seluruh bantuan, <a href="https://tinyurl.com/28202023">https://tinyurl.com/28202023</a>
3	RT @sidoarjokota: Bagaimana yg kita lakukan yg benar? Maka yg benar adalah yg benar
4	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47

Gambar 3. Sebelum Cleaning data

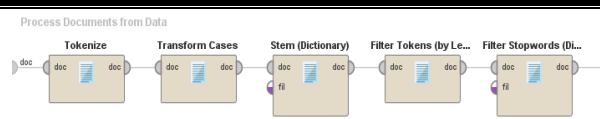
Row No.	SENIMEN	Text
1	Positif	tenaga kesehatan salinan teman sopan doyan
2	Positif	bupati ahmad muhdior menerahkan pppk tenaga kesehatan sidoarjo
3	Positif	orangtua anak jenius bupati pppk tenaga kesehatan sidoarjo
4	Positif	kesenangan
5	Negatif	memohon calon pegawai negri cpns pegawai pemerintah perjanjian kerja pppk obatka
6	Positif	ketarung ketiup gunungan cikulan melencong pppk formasi kipas
7	Negatif	bupati salatiga ahmad mulyadi mendorong masyarakat pppk formasi kipas
8	Negatif	bupati menteri upayaan teman pppk birokrasi imanachan
9	Negatif	suratnya update pencapaian pppk
10	Negatif	teman pppk birokrasi pencapaian harapan kontribusi pppk
11	Negatif	program market place nadien
12	Negatif	beasiswa prajabatan objek pengensi kesosongan guru daerah
13	Negatif	minangs kebutuhan cari guru wajib
14	Positif	kegiatan pengangkatan pppk tenaga kesehatan
15	Positif	suratnya pppk kesehatan gresik warga berobat dicemberuti
16	Negatif	collegemenfess setau skrg guru cpns seleskiannya namanya pppk nder
17	Positif	tenaga kesehatan pesan khusus tenaga kesehatan nakes diangkat pega
18	Negatif	pemerintah membuka lowongan calon pegawai negeri cpns pegawai pemerintah NEGATIVE
19	Positif	serahkann pengangkatan pppk bupati gresik warga berobat dicemberuti https://fifaa bu
20	Negatif	pmekot serang usulkan formasi pppk https://jyhh.aakhm

Gambar 4. Setelah Proses Cleaning data

### J. Tahap Preprocessing Data

Tahap Preprocessing data bertujuan untuk mengubah data mentah menjadi data yang siap diolah, adapula tahap ini dibagi menjadi beberapa bagian, yaitu :

- Tokenizing**, yaitu tahapan untuk membagi teks menjadi kata, seperti teks “sedang mengerjakan skripsi” setelah melewati tahap Tokenizing, akan menjadi 4 kata, yaitu “sedang”, “mengerjakan”, dan “skripsi”. Case Folding, yaitu tahapan untuk mengubah data tweet menjadi lower case (huruf kecil).
- Case Folding**, yaitu tahapan untuk mengubah data tweet menjadi lower case (huruf kecil).
- Stemming**, yaitu tahapan untuk membersihkan kata-kata imbuhan awalan dan akhirnya yang terdapat dalam teks seperti “mengerjakan” menjadi “kerja”.
- Filter Tokens (by Length)**, yaitu tahapan untuk membuang kata-kata yang kurang dari 2 huruf dan kata-kata yang melebihi 25 huruf.
- Filter Stopwords**, yaitu tahapan untuk menghapus kata bantu seperti “saya”, “dia”, “aku”, dan “mereka”.



Gambar 5. Proses Preprocessing

### K. Tahap Pembobotan Kata

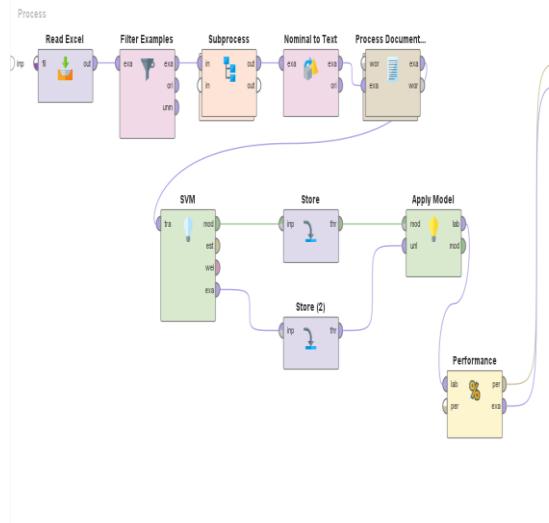
Pada tahap ini, hasil preprocessing akan diolah agar setiap kata memiliki bobot (nilai). Pembobotan kata yang penulis gunakan adalah algoritma TF-IDF. Term Frequency-Inverse Document Frequency atau TF-IDF adalah suatu metode algoritma yang berguna untuk menghitung bobot setiap kata yang umum digunakan. Metode ini juga terkenal efisien, mudah dan memiliki hasil yang akurat. Metode ini akan menghitung nilai Term Frequency (TF) dan Inverse Document Frequency (IDF) pada setiap token (kata) di setiap dokumen dalam korpus. Secara sederhana, metode TF-IDF digunakan untuk mengetahui berapa sering suatu kata muncul di dalam dokumen.

Row No.	Sentimen	akal	allah	amin	bacot	banding	bicara	dungu	fokus
1	Positif	0.335	0	0	0	0.335	0.335	0	0
2	Negatif	0	0	0	0	0	0	0	0.219
3	Negatif	0	0	0	0	0	0	0	0
4	Positif	0	0.553	0.396	0	0	0	0	0
5	Negatif	0	0	0	0	0	0	0	0
6	Positif	0	0.215	0	0	0	0	0	0
7	Negatif	0	0	0	0.316	0	0	0	0
8	Negatif	0	0	0	0	0	0	0.472	0
9	Negatif	0	0	0	0	0	0	0	0
10	Positif	0	0	0	0	0	0	0	0

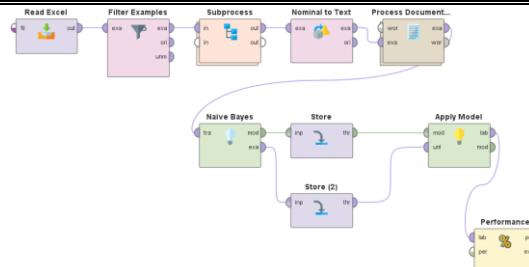
Gambar 6. Hasil Pembobotan Kata dengan TF – IDF

### L. Tahap Pembuatan Model

Output dari tahapan ini adalah sebuah model klasifikasi dengan metode *Naïve Bayes* dan *Support Vector Machine* dan Data Latih yang nanti akan digunakan dalam proses sentimen analisis.



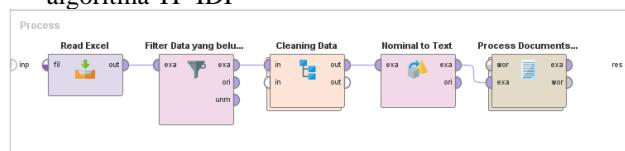
Gambar 7. Pembuatan Model Support Vector Machine

Gambar 8. Pembuatan Model *Naïve Bayes*

#### M. Tahap Mempersiapkan Data Uji

Pada tahap ini data uji akan diolah kembali agar dapat dilakukan proses selanjutnya, tahapan mengolah data uji antara lain :

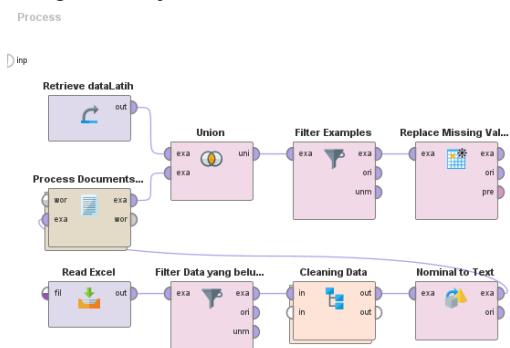
- Filter data yang belum memiliki label
- Cleaning data untuk menghilangkan kata-kata yang tidak dibutuhkan
- Preprocessing data dan pembobotan kata dengan algoritma TF-IDF



Gambar 9. Persiapan data uji

#### N. Tahap Union / Penggabungan kata

Pada tahap ini, data uji dan data latih akan digabungkan menjadi satu data.

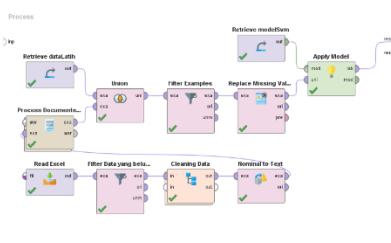


Gambar 10. Proses Penggabungan Data

### III. HASIL DAN PEMBAHASAN

#### A. Tahap Pengujian

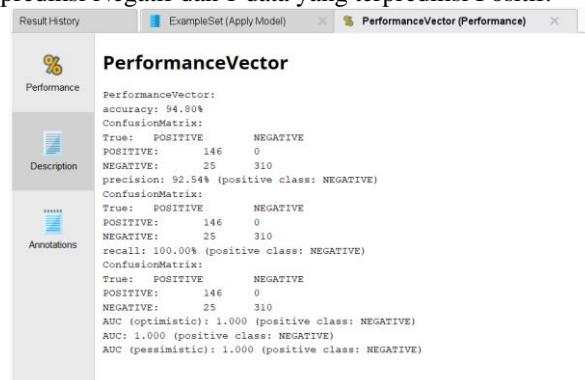
Pada tahap ini, model yang telah dibuat sebelumnya akan diterapkan untuk memprediksi sentimen pada data uji.



Gambar 11. Pengujian Data

#### B. Nilai Akurasi Support Vector Machine dan Naïve Bayes

Hasil perhitungan akurasi data latih dengan menggunakan metode Support Vector Machine, didapatkan nilai Accuracy sebesar 94,80%, nilai Recall Positif sebesar 85,38%, nilai Recall Negatif sebesar 100,00%, nilai Precision Positif sebesar 100,00%, nilai Precision Negatif sebesar 92,54%. Dari 482 data latih, penulis melabeli sebanyak 147 data sebagai Sentimen Positif dan 335 data sebagai Sentimen Negatif. Untuk hasil prediksi dari Sentimen Negatif, terdapat 335 data terprediksi Negatif dan 1 data yang terprediksi Positif.

Gambar 12. akurasi metode *Support Vector Machine*

#### C. Perbandingan Akurasi Metode Support Vector Machine dan Naïve Bayes

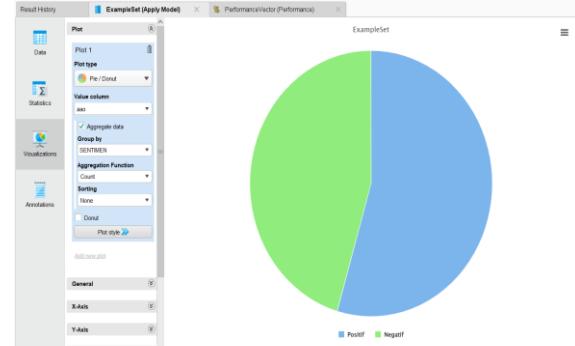
Hasil dari Implementasi yang telah dilakukan, perbandingan tingkat akurasi antara metode Naïve Bayes dan Support Vector Machine.

Tabel 2. Sampel Data Training

Metode	Nilai Akurasi
Naïve Bayes	96,14%
Support Vector Machine	94,80%

#### D. Hasil Sentimen Analisis

Dari proses sentimen analisis dengan 482 data uji, dihasilkan prediksi sebanyak 335 Sentimen Negatif dan 147 Sentimen Positif, berikut penulis sajikan data tersebut kedalam Pie Chart.

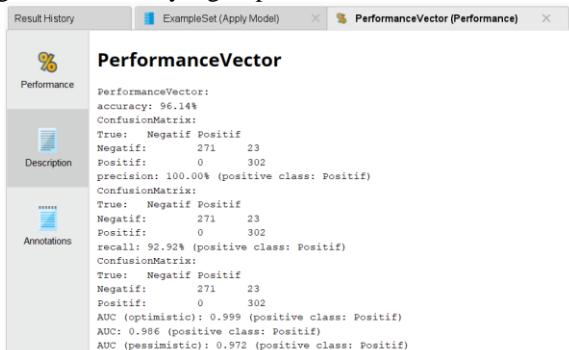


Gambar 13. Pie Chart

#### E. Hasil Akhir Pengujian



Hasil perhitungan akurasi data latih dengan menggunakan metode Naïve Bayes, didapatkan nilai Accuracy sebesar 96,14%, nilai Recall Positif sebesar 92,92%, nilai Recall Negatif sebesar 100,00%, nilai Precision Positif sebesar 100,00%, nilai Precision Negatif sebesar 92,18%. Dari 482 data latih, penulis melabeli sebanyak 147 data sebagai Sentimen Positif dan 335 data sebagai Sentimen Negatif. Hasil prediksi dari Sentimen Positif,. Untuk hasil prediksi dari Sentimen Negatif, terdapat 335 data terprediksi Negatif dan 1 data yang terprediksi Positif.



Gambar 14. akurasi metode *Naïve Bayes*

Hasil akhir dari Perbandingan dengan dua metode pengujian ini, yaitu hasil prediksi Sentimen Masyarakat Terhadap Isu Perekutan PPPK berdasarkan data yang didapat dari Twitter dan diimplementasikan dengan metode SVM (Support Vector Machine) menunjukkan nilai akurasi sebesar 94,80 %. Dari 482 data uji, terprediksi sebesar 335 data sebagai Sentimen Negatif dan 147 data sebagai Sentimen Positif dan Metode Naïve bayes menunjukkan nilai akurasi sebesar 96,14 %. Dari 482 data uji, terprediksi sebesar 335 data sebagai Sentimen Negatif dan 147 data sebagai Sentimen Positif.

#### IV. KESIMPULAN

Berdasarkan pembahasan, hasil, dan pengujian yang telah dilakukan, dalam penerapan dan perbandingan algoritma Support Vector Machine dan naïve bayes untuk melakukan Sentimen Analisis, maka dapat diambil kesimpulan Sentimen analisis dapat dilakukan menggunakan Software RapidMiner Studio menggunakan Metode Support Vector Machine dan naïve bayes dengan mengolah data yang didapat dari Media Sosial Twitter melalui proses Crawling data, Labeling data, Cleaning data, Preprocessing dan Klasifikasi.

Dari proses Sentimen analisis didapatkan hasil perbandingan nilai akurasi metode Support Vector Machine sebesar 94,80% dan naïve bayes 96,14% dari 482 data uji, terprediksi sebanyak 335 data sebagai Sentimen Negatif dan 147 data sebagai Sentimen Positif, sehingga dapat disimpulkan bahwa sebagian besar masyarakat tidak menginginkan adanya perekutan pppk melalui market place.

#### UCAPAN TERIMA KASIH

Saya ingin mengucapkan terima kasih yang sebesar-besarnya atas bimbingan, dukungan, dan dedikasi Bapak

Ir. Sugiyono dalam membantu pembuatan jurnal ini. Tanpa panduan dan dorongan yang berharga dari Anda, pencapaian ini tidak akan mungkin terwujud. Saya sangat berterima kasih atas waktu dan pengetahuan yang Anda luangkan untuk membimbing saya dalam menyusun jurnal ini. Dalam setiap pertemuan dan diskusi, Bapak telah memberikan wawasan yang berharga, saran yang berharga, dan arahan yang tepat. Bimbingan Bapak telah memainkan peran penting dalam membantu saya memahami proses penelitian, mengembangkan metodologi yang tepat, dan menganalisis data dengan benar. Selain itu, terima kasih juga karena telah memberikan koreksi dan umpan balik konstruktif dalam setiap tahap penulisan jurnal. Pengamatan dan penilaian Bapak yang mendalam telah membantu saya untuk meningkatkan kualitas tulisan dan memperbaiki kelemahan yang ada. Saya sangat menghargai kesabaran dan ketelitian yang Anda tunjukkan dalam membimbing saya menuju hasil akhir yang memuaskan. Lebih dari sekadar menjadi pembimbing, Bapak juga menjadi inspirasi bagi saya. Kepedulian dan semangat Bapak terhadap penelitian dan pengembangan ilmu pengetahuan telah mendorong saya untuk terus belajar dan berkembang. Saya merasa sangat beruntung dan bersyukur dapat belajar dari Bapak, seorang akademisi yang berkompeten dan berdedikasi. Terima kasih karena telah memberikan kesempatan kepada saya untuk terlibat dalam penelitian ini dan memperluas pengetahuan serta wawasan saya. Saya meyakini bahwa karya tulis ini tidak hanya akan memberikan manfaat bagi saya secara pribadi, tetapi juga akan memberikan kontribusi yang berarti dalam perkembangan bidang ilmu yang kita geluti. Sekali lagi, terima kasih yang tak terhingga atas semua bimbingan, dorongan, dan kepercayaan yang Bapak berikan kepada saya. Saya sangat beruntung dan berbangga dapat belajar di bawah arahan Bapak. Semoga keberhasilan ini menjadi bukti nyata dari dedikasi dan komitmen Bapak dalam membentuk generasi muda yang berkualitas dan berkontribusi positif dalam dunia akademik.

#### REFERENSI

- [1] Fitriani Fitriani, Ema Utami, and Anggit Dwi Hartanto, “Analisis Sentimen Masyarakat Terhadap Pelaksanaan P3K Guru Dengan Algoritma Naive Bayes Dan Decision Tree,” *Tek. Teknol. Inf. dan Multimed.*, vol. 3, no. 1, pp. 23–30, 2022, doi: 10.46764/teknimedia.v3i1.53.
- [2] C. H. Yutika, A. Adiwijaya, and S. Al Faraby, “Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes,” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 422, 2021, doi: 10.30865/mib.v5i2.2845.
- [3] M. T. Student *et al.*, “PPPK Sebagai Solusi Pemenuhan Kebutuhan SDM Profesional Dalam Birokrasi,” *JSPG J. Soc. Polit. Gov.*, vol. 14, no. 1, pp. 1–13, 2021.
- [4] N. ANNET and J. Naranjo, “ENGANGKATAN TENAGA HONORER EKS KATEGORI 2 (K-II)



- BERDASARKAN SKEMA PEGAWAI PEMERINTAH DENGAN PERJANJIAN KERJA (PPPK) DI KABUPATEN BANYUMAS APPOINTMENT HONORARY WORKER EX CATEGORY 2 (K-II) BASED ON SCHEME NON-ONGOING EMPLOYEE (PPPK) IN KABUPATEN BANYUMAS,” *Appl. Microbiol. Biotechnol.*, vol. 85, no. 1, pp. 2071–2079, 2014.
- [5] R. Eksekutif, “Penguatan Sistem Rekrutmen PPPK,” 2024.
- [6] L. N. Qomarani, “Anomali Kehadiran Pegawai Pemerintah Dengan Perjanjian Kerja (Pppk) Dalam Cakrawala Kepegawaian Di Indonesia,” *Cepalo*, vol. 4, no. 2, pp. 95–110, 2020, doi: 10.25041/cepalov4no2.1979.
- [7] H. Juliani, “Diskresi Dalam Rekrutmen Pegawai Non Pegawai Negeri Sipil Setelah Pemberlakuan Peraturan Pemerintah Nomor 49 Tahun 2018 tentang Manajemen Pegawai Pemerintah Dengan Perjanjian Kerja,” *Adm. Law Gov. J.*, vol. 2, no. 2, pp. 314–325, 2019, doi: 10.14710/alj.v2i2.314-325.
- [8] N. S. Fahrani, “Manajemen pegawai pemerintah dengan perjanjian kerja di pemerintah kota banda aceh,” *Civ. Serv. J.*, pp. 55–70, 2015, [Online]. Available: <https://jurnal.bkn.go.id/index.php/asn/article/view/51%0Ahttps://jurnal.bkn.go.id/index.php/asn/article/download/51/120>
- [9] G. A. Fauzan, “Guru Honorer dalam Lingkaran Ketidakadilan,” *J. Educ.*, vol. 4, no. 1, pp. 197–208, 2021, doi: 10.31004/joe.v4i1.418.
- [10] D. I. K. Takalar, “Pengangkatan dan pemecatan pegawai honorer di kabupaten takalar,” vol. 1, pp. 232–240.
- [11] G. Besar, F. Universitas, and W. Menteri, “UNDANG-UNDANG APARATUR SIPIL NEGARA : MEMBANGUN PROFESIONALISME APARATUR SIPIL NEGARA CIVIL STATE APPARATUSLAW : BUILDING THE PROFESSIONALISM OF CIVIL STATE APPARATUS,” vol. 8, no. 1, pp. 13–31, 2014.
- [12] D. Rusdiaman and D. Rosiyadi, “Analisa Sentimen Terhadap Tokoh Publik Menggunakan Metode Naïve Bayes Classifier Dan Support Vector Machine,” *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 4, no. 2, pp. 230–235, 2019.
- [13] L. Oktasari, Y. H. Chrisnanto, and R. Yuniarti, “Text Mining Dalam Analisis Sentimen Asuransi Menggunakan Metode Niave Bayes Classifier,” *Pros. SNST*, vol. 7, pp. 37–42, 2016, [Online]. Available: [https://www.publikasiilmiah.unwahas.ac.id/index.php/PROSIDING\\_SNST\\_FT/article/view/1506/1589](https://www.publikasiilmiah.unwahas.ac.id/index.php/PROSIDING_SNST_FT/article/view/1506/1589)
- [14] H. Tuhuteru and A. Iriani, “Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon
- [15] Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier,” *J. Inform. J. Pengemb. IT*, vol. 3, no. 3, pp. 394–401, 2018, doi: 10.30591/jpit.v3i3.977.
- E. Indrayuni, A. Nurhadi, and D. A. Kristiyanti, “Implementasi Algoritma Naive Bayes, Support Vector Machine, dan K-Nearest Neighbors untuk Analisa Sentimen Aplikasi Halodoc,” *Fakt. Exacta*, vol. 14, no. 2, p. 64, 2021, doi: 10.30998/faktorexacta.v14i2.9697.