

**STUDI EKSPERIMEN METODE NAIVE BAYES CLASSIFIER PADA
ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP OPINI
PERPINDAHAN IBU KOTA**

SKRIPSI

Untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana S-1
Program Studi Teknik Informatika



Diajukan Oleh:

Syafa'at Adi Nugraha

16650034

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA YOGYAKARTA

2019

HALAMAN PENGESAHAN



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-2054/Un.02/DST/PP.00.9/09/2020

Tugas Akhir dengan judul : STUDI EKSPERIMEN METODE NAIVE BAYES CLASSIFIER PADA ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP OPIN PERPINDAHAN IBU KOTA

yang dipersiapkan dan disusun oleh:

Nama : SYAFA'AT ADI NUGRAHA
Nomor Induk Mahasiswa : 16650034
Telah diujikan pada : Senin, 31 Agustus 2020
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang/Penguji I

Maria Ulfah Siregar, S.Kom. MIT., Ph.D.
SIGNED

Valid ID: 5f61bdc3e9378



Penguji II

Dr. Ir. Bambang Sugiantoro, S.Si., M.T.
SIGNED

Valid ID: 5f517d561b2d6



Penguji III

Agus Mulyanto, S.Si., M.Kom.
SIGNED

Valid ID: 5f570f33522a1

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA



Yogyakarta, 31 Agustus 2020
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi
Dr. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 5f64c5a9e9f74

SURAT PERSETUJUAN SKRIPSI



Universitas Islam Negeri Sunan Kalijaga



FM-UINSK-BM-05-03/R0

SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Syafa'at Adi Nugraha

NIM : 16650034

Judul Skripsi : Studi Eksperimen Metode Naive Bayes Classifier pada Analisis Sentimen Pengguna Twitter Terhadap Opini Perpindahan Ibu Kota

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Teknik Informatika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 25 Agustus 2020

Pembimbing

Maria Ulfah Siregar, S.Kom. MIT., Ph.D.

NIP. 19780106 200212 2 001

PERNYATAAN KEASLIAN SKRIPSI

PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini :

Nama : Syafa'at Adi Nugraha

NIM : 16650034

Jurusan : Teknik Informatika

Fakultas : Sains dan Teknologi

Menyatakan bahwa skripsi saya yang berjudul **“Studi Eksperimen Metode Naive Bayes Classifier Pada Analisis Sentimen Pengguna Twitter Terhadap Opini Perpindahan Ibu Kota”** merupakan hasil penelitian saya sendiri, tidak terdapat pada karya yang pernah di ajukan untuk memperoleh gelar kesarjana di suatu perguruan tinggi, dan bukan plagiasi karya orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Yogyakarta, 25 Agustus 2020



Syafa'at Adi Nugraha
NIM. 16650034

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah dan petunjuk-Nya kepada penulis sehingga dapat menyelesaikan penelitian dengan judul **Studi Eksperimen Metode Naive Bayes Classifier Pada Analisis Sentimen Pengguna Twitter Terhadap Opini Perpindahan Ibu Kota** dengan baik dan tepat waktu, Tujuan penelitian ini adalah sebagai salah satu syarat untuk mencapai gelar sarjana Program Studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta.

Penulis menyadari bahwa dalam proses penelitian ini masih jauh dari kata sempurna maka dari itu diharapkan ada kritik dan saran untuk memperbaiki penelitian yang akan datang.

Penyusunan laporan ini tidak terlepas dari bantuan berbagai pihak, maka dari itu penulis mengucapkan terimakasih kepada:

1. Allah SWT yang telah memberikan kekuatan dan kelancaran.
2. Bapak Prof. Dr. Phil Sahiron, M.A, selaku Rektor UIN Sunan Kalijaga Yogyakarta.
3. Bapak Dr. H. Waryono, M.Ag., selaku Wakil Rektor Bidang Kemahasiswaan dan Kerjasama UIN Sunan Kalijaga Yogyakarta.
4. Ibu Dr. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
5. Ibu Maria Ulfah Siregar, S.Kom. MIT., Ph.D. selaku Kepala Program Studi Teknik Informatika dan Dosen Pembimbing Skripsi yang telah membimbing, meluangkan waktu, motivasi, koreksi dan kritik saran kepada penulis sehingga skripsi ini dapat terselesaikan.

6. Bapak Dr. Agung Fatwanto, S.Si., M.Kom., selaku Dosen Pembimbing Akademik.
7. Bapak Agus Mulyanto, S.Si., M.Kom., Ph.D., Bapak Sumarsono, S.T., M.Kom, Bapak Aulia Faqih Rifa'i, M.Kom., Bapak Muhammad Didik Rohmad Wahyudi, S.T., M.T., Bapak Nurochman, S.Kom., M.Kom., Bapak M. Taufiq Nuruzzaman, S.T., Bapak Rahmat Hidayat, S.Kom., M.Cs., Bapak Dr. Ir. Bambang Sugiantoro, M.T., Ibu Dr. Shofwatul 'Uyun, S.T., M.Kom., selaku dosen pengampu mata kuliah program studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta
8. Seluruh staf dan karyawan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
9. Kedua orang tua, Bapak Daryanto dan Ibu Asmaul Husna yang telah memberikan dukungan, semangat dan do'a sehingga penulis dapat menyelesaikan penelitian ini.
10. Kedua adikku Muhammad Alfandy dan Aulia Zahra Nugraheni.
11. Keluarga besar dari kedua orang tua.
12. Teman-teman yang telah membantu dan memberi semangat Lina, Sekar, Ulfa, Nadia, Fauzi, Nur, Rafi, Ardi, Mbak Nafi.
13. Teman-teman yang telah berpartisipasi dalam proses penelitian Bagus, Dana, Rani.
14. Teman-teman angkatan 2016 semuanya yang telah berproses bersama.
15. Teman-teman KKN Jamban, Salaman Angkatan 99 yang telah menjadi keluarga selama kurang lebih 2 bulan dan seterusnya.

16. Serta semua pihak yang tidak dapat penulis sebutkan satu persatu dan telah memberikan banyak campur tangan, doa, support sehingga penelitian ini dapat terselesaikan.

Yogyakarta, 14 Agustus 2020



HALAMAN PERSEMBAHAN

Skripsi ini dipersembahkan untuk keluarga yang selalu mendukung dan mendoakan.

Terimakasih selalu mendukung dan mendoakan.

Serta kepada selurung orang yang telah membantu dan mendoakan dalam penelitian ini.

Tidak lupa untuk diri sendiri yang sudah berjuang.



HALAMAN MOTTO

“Man Jadda Wajada”

“Man Shobaro Dzofiro”

“Man Yazro Yahshud”

“Bantulah ketika kamu bisa”

“Berusahalah selalu dan dibantu dengan doa”

“Berproseslah walaupun sedikit, tetapi jangan diam”

- فَبِأَيِّ آلَاءِ رَبِّكُمَا تُكَذِّبَانِ - ٥٥

“Maka nikmat Tuhan-mu yang manakah yang kamu dustakan? (QS. Ar Rahman:6)”

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

“Bersyukurlah”

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN.....	ii
SURAT PERSETUJUAN SKRIPSI.....	iii
PERNYATAAN KEASLIAN SKRIPSI.....	iv
KATA PENGANTAR.....	v
HALAMAN PERSEMBAHAN	viii
HALAMAN MOTTO	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
INTISARI	xv
ABSTRACT	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Penelitian	5
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
1.6 Keaslian Penelitian.....	6
1.7 Sistematika Penulisan.....	7
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	9
2.1 Tinjauan Pustaka	9
2.2 Landasan Teori.....	21
2.2.1 Text Mining.....	21
2.2.2 Text Preprocessing	22
2.2.3 Analisis Sentimen.....	23
2.2.4 Twitter	24

2.2.5	Klasifikasi.....	24
2.2.6	Naïve Bayes Classifier	25
2.2.7	TF-IDF	29
2.2.8	Confusion Matrix	30
2.2.9	<i>Split Validation</i>	31
2.2.10	Python.....	32
2.2.11	Ibu Kota	32
BAB III METODE PENELITIAN		33
3.1	Metode Penelitian.....	33
3.2	Alur Penelitian.....	33
3.2.1	Studi Pustaka.....	35
3.2.2	Pengumpulan Data	35
3.2.3	Seleksi Data.....	35
3.2.4	Preprocessing Data.....	36
3.2.5.	Pengolahan Data	37
3.2.6.	Implementasi	37
3.3	Perangkat Penelitian.....	37
3.3.1	Perangkat Keras.....	38
3.3.2	Perangkat Lunak.....	38
BAB IV HASIL DAN PEMBAHASAN		39
4.1	Pengumpulan Data	40
4.2	Seleksi dan Pelabelan Data	41
4.3	<i>Presprocessing</i>	47
4.3.1	<i>Cleansing dan Casefolding</i>	47
4.3.2	<i>Convert Slangword</i>	48
4.3.3	<i>Stopword Removal</i>	49
4.3.4	<i>Stemming</i>	50
4.4	Analisa dan Evaluasi	51
4.4.1	Analisa.....	51
4.4.1.1	<i>Term Frequency – Invers Document Frequency (TF-IDF)</i>	52

4.4.1.2	Algoritma <i>Naive Bayes</i>	54
4.4.2	Evaluasi	59
4.5	Implementasi	62
BAB V	PENUTUP	67
5.1	Kesimpulan.....	67
5.2	Saran.....	68
DAFTAR PUSTAKA.....		69
LAMPIRAN.....		72



DAFTAR GAMBAR

Gambar 3.1 Flowchart.....	34
Gambar 4.1 Flowchart penelitian	39
Gambar 4.2 Presentase Data Latih	44
Gambar 4.3 Visualisasi data dengan kata paling sering muncul sentimen negatif	44
Gambar 4.4 Visualisasi data dengan kata paling sering muncul sentimen positif	45
Gambar 4.5 Visualisasi data wordcloud sering muncul sentimen negative	46
Gambar 4.6 Visualisasi data wordcloud sering muncul sentimen positif	46
Gambar 4.7 Confusion matrix dengan split data 90:10.....	60
Gambar 4.8 Confusion matrix dengan split data 80:20.....	61
Gambar 4.9 Confusion matrix dengan split data 70:30.....	62
Gambar 4.10 Presentase data uji	63
Gambar 4.11 Visualisasi data dengan kata paling sering muncul sentimen negatif ..	64
Gambar 4.12 Visualisasi data dengan kata paling sering muncul sentimen positif ...	65
Gambar 4.13 Visualisasi data uji wordcloud sering muncul sentimen negatif	66
Gambar 4.14 Visualisasi data uji wordcloud sering muncul sentimen positif	66

DAFTAR TABEL

Tabel 2.1 Tabel Tinjauan Pustaka	16
Tabel 2.2 Tabel Tinjauan Pustaka (Lanjutan)	16
Tabel 2.3 Tabel Tinjauan Pustaka (Lanjutan)	17
Tabel 2.4 Tabel Tinjauan Pustaka (Lanjutan)	18
Tabel 2.5 Tabel Tinjauan Pustaka (Lanjutan)	19
Tabel 2.6 Tabel Tinjauan Pustaka (Lanjutan)	20
Tabel 2.7 Tabel Tinjauan Pustaka (Lanjutan)	21
Tabel 2.8 Tabel Confussion Matrix.....	30
Tabel 4.1 Tabel Contoh Pengumpulan Data tweet.....	40
Tabel 4.2 Tabel Trainer.....	42
Tabel 4.3 Contoh Pelabelan Data.....	43
Tabel 4.4 Contoh Hasil Cleansing.....	47
Tabel 4.5 Contoh Hasil Cleansing (Lanjutan).....	48
Tabel 4.6 Contoh Proses Convert Slangword	48
Tabel 4.7 Contoh Proses Convert Slangword (Lanjutan)	49
Tabel 4.8 Contoh Proses Stopword Removal.....	50
Tabel 4.9 Contoh Proses Stemming	51
Tabel 4.10 Contoh Data setelah Pelabelan Data dan Preprocessing	52
Tabel 4.11 Contoh perhitungan TF-IDF	52
Tabel 4.12 Contoh perhitungan TF-IDF (Lanjutan).....	53
Tabel 4.13 Contoh perhitungan TF-IDF	55
Tabel 4.14 Contoh perhitungan TF-IDF (Lanjutan).....	56
Tabel 4.15 Contoh perhitungan laplace smoothing.....	56
Tabel 4.16 Contoh perhitungan laplace smoothing (Lanjutan).....	57
Tabel 4.17 Contoh perhitungan klasifikasi kelas negatif	58
Tabel 4.18 Contoh perhitungan klasifikasi kelas positif	58
Tabel 4.19 Pembagian split data pengujian.....	60

Studi Eksperimen Metode Naive Bayes Classifier Pada Analisis Sentimen Pengguna Twitter Terhadap Opini Perpindahan Ibu Kota

Syafa'at Adi Nugraha

16650034

INTISARI

Keputusan resmi presiden terhadap pemindahan ibu kota yang sebelumnya di Jakarta akan dipindah ke Kalimantan Timur akan menjadi sebuah hal yang besar dan menimbulkan banyak perdebatan dari pihak yang setuju dan tidak setuju karena proyek ini akan membutuhkan dana yang banyak dan akan memberikan dampak terhadap berbagai sektor. Maka dengan adanya penelitian ini bisa dijadikan sebuah informasi terhadap program pemindahan ibu kota baru, dengan sentimen analisis bisa didapatkan informasi topik yang dibicarakan masyarakat beberapa pengguna Twitter.

Penelitian ini menggunakan 1290 data tweet yang terdiri dari 686 sentimen negatif dan 603 sentimen positif yang dijadikan sebagai data latih untuk membuat model yang dievaluasi menggunakan confusion matrix dan split validation. Penelitian ini menggunakan ekstraksi fitur pembobotan kata TF-IDF dengan metode Naive Bayes.

Hasil dari percobaan yang dilakukan menghasilkan akurasi terbaik pada split data 90:10 dengan 76,74%. Model yang telah dibuat diimplementasikan pada data uji sebanyak 1115 data menghasilkan 799 sentimen negatif dan 316 sentimen positif. dari data uji menghasilkan beberapa kata yang muncul diantaranya pada sentimen negatif adalah negara, rakyat, proyek, covid, corona, dan pandemi, dari kata-kata tersebut bisa disimpulkan bahwa masyarakat yang memberikan sentimen negatif ingin bahwa pemerintah lebih fokus terhadap penanggulangan bencana seperti banjir yang sering melanda pada musim hujan dan bencana yang sedang melanda seluruh dunia yaitu corona atau virus covid 19. sedangkan sentimen positif menghasilkan kata yang sering muncul diantaranya kalimantan, jakarta, bangun, mulai, dan jokowi. kata-kata tersebut menggambarkan bahwa dari masyarakat yang memberikan sentimen positif mendukung dengan proses pemindahan ibu kota negara baru ke Kalimantan yang dilakukan oleh Jokowi. Dapat disimpulkan bahwa hasilnya cenderung lebih banyak yang memberikan sentimen negatif.

Kata kunci: *analisis sentimen, twitter, naive bayes, TF-IDF, split validation, confusion matrix, ibu kota negara baru, pindah ibu kota*

Experimental Study of the Naive Bayes Classifier Method on the Sentiment Analysis of Twitter User on Opinion on Capital Moving

Syafa'at Adi Nugraha

16650034

ABSTRACT

The president's official decision regarding the relocation of the capital city, which previously was in Jakarta, will be moved to East Kalimantan, will be a big thing and cause a lot of debate from both parties who agree and disagree because this project will require a lot of funds and will have an impact on various sectors. So with this research, it can be used as an information on the new capital relocation program, with sentiment analysis, information on topics discussed by the community of several Twitter users can be obtained.

This study uses 1290 tweet data consisting of 686 negative sentiments and 603 positive sentiments which are used as training data to create evaluation models using confusion matrix and split validation. In this study, using TF-IDF word-weighted feature extraction with the Naive Bayes method.

The result from the experiments carried out the best accuracy in split data 90:10 with 76.74%. The model that has been made is implemented in 1115 test data resulting in 799 negative sentiments and 316 positive sentiments. From the test data, it produces several words that appear, including negative sentiment, namely negara, rakyat, proyek, covid, corona, and pandemi. From these words can be described that its better for the goverment to take care more during the rainy season like a flood and disasters that are hitting the whole world, namely the corona or covid 19 virus, while positive sentiment results in words that often appear including Kalimantan, Jakarta, bangun, mulai, and Jokowi. These words illustrate that from the people who provide positive sentiment support the process of moving the new nation's capital to Kalimantan by Jokowi. It can be concluded that the results tend to be more of a negative sentiment.

Keywords: *analisis sentimen, twitter, naive bayes, TF-IDF, split validation, confusion matrix, ibu kota negara baru, pindah ibu kota*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Media sosial seringkali sebagai tempat masyarakat untuk memberikan pendapat atau cuitan untuk kritik dan saran suatu hal. Terutama Twitter, yang sudah tidak asing lagi bagi pengguna media sosial, digunakan sebagai sarana bertukar dan berbagi informasi. Dalam penelitian ini menggunakan sampel data tweet dari beberapa pengguna Twitter.

Ibu kota sekarang yaitu Jakarta sudah padat penduduk, sebagai pusat pemerintahan, perdagangan berbagai macam, banyak terdapat perindustrian, pariwisata dan tata ruangnya semrawut dan banyaknya gedung-gedung, serta pemanfaatan lahan yang saling kontradiktif banyak terjadi (Utama FR., 2016). Meski demikian ada beberapa hal yang masih pro dan kontra terhadap wacana tersebut, gubernur Kalimantan Tengah mengakui bahwa terdapat beberapa kekhawatiran dari beberapa tokoh masyarakat, terutama dampak perubahan sosiologis dari masyarakat Kalimantan Tengah yang hidup memiliki kebiasaan sub-urban menjadi masyarakat ibu kota yang metropolitan. Sehingga memunculkan ditakutkan menimbulkan potensi terpinggirnnya masyarakat lokal tersebut menuntut adanya perencanaan yang lebih serius dari pemerintah pusat untuk mengantisipasi perubahan-perubahan yang akan terjadi (Toun NR, 2018). Kalimantan dipilih karena masih memiliki lahan masih sangat luas, sehingga dapat menyusun tata

ruang ibu kota Negara yang sangat ideal dibanding Jakarta yang sudah sempit karena bangunan sudah banyak. Kelemahannya adalah prasarana dan sarana belum memadai, sebagian besar harus membangun yang baru, sehingga membutuhkan anggaran yang besar. Kelemahan lainnya adalah penyediaan air bersih, kebakaran hutan, bencana banjir dan longsor merupakan bahaya yang perlu dipertimbangkan (Yahya, 2018). Selain itu jika program ini dilaksanakan maka akan menimbulkan perubahan di sektor lain, seperti pada tahun 2020 ini adanya *Covid 19* jika pemerintah terlalu fokus dengan pemindahan ibu kota maka dana untuk penanggulangan *Covid 19* bisa terganggu, bahkan pembangunan untuk sektor yang lain akan dikurangi karena fokus terhadap pembangunan ibu kota baru yang membutuhkan dana sangat besar.

Pada tahun 2019 menjadi tahun yang spekulatif karena presiden memberikan wacana terhadap pemindahan ibu kota dan pada tahun yang sama pula diumumkan secara resmi bahwa ibu kota yang sebelumnya di Jakarta akan dipindah ke Kalimantan Timur. Keputusan resmi presiden yang telah diumumkan akan menjadi sebuah hal yang besar dan menimbulkan banyak perdebatan. Hasil keputusan yang kadang tidak sesuai harapan atau ekspektasi masyarakat lalu akan menimbulkan cuitan di twitter mulai dari hal yang baik hingga buruk. Maka pada tahun 2019-selesai proses perpindahan ibu kota akan menjadi keluarnya suara masyarakat terkait pemindahan ibu kota dikarenakan jelas biaya yang dikeluarkan tidaklah sedikit dan akan mempengaruhi berbagai sektor seperti proyek pembangunan atau anggaran untuk berbagai proyek lain di lain daerah. Belum juga semisal ada bencana yang sifatnya mendadak dan tidak terduga maka pengalokasian dana ini

akan sangat penting dipertimbangkan agar satu dengan lainnya bisa berjalan dengan baik dan tidak menimbulkan efek yang buruk bagi salah satu pihak.

Pemindahan ibu kota negara membutuhkan alasan yang kuat dan mendasar tentang efektifitas fungsinya. Beberapa orang berpendapat jika ibu kota negara dipindahkan dari Jakarta, maka negara Indonesia akan menjadi negara serikat seperti Singapura. Jika ibu kota dipindahkan, akan menimbulkan permasalahan terkait permasalahan infrastruktur dan bagaimana jenis pengaturan nantinya, termasuk apabila pemindahan ibu kota jadi dilaksanakan maka akan terjadi perubahan tatanan yang luar biasa (Toun NR, 2018).

Analisis sentimen digunakan untuk penelitian mengetahui sentimen masyarakat terkait pemindahan ibu kota, dengan menggunakan tweet dari Twitter memiliki arti sebagai ungkapan atau pendapat suatu individu. Dalam kasus ini menggunakan beberapa sampel tweet dari pengguna Twitter bisa digunakan sebagai gambaran suatu program atau kebijakan pemerintah tersebut mendapat persetujuan atau penolakan dari masyarakat. Sentimen negatif menandakan bahwa terjadi penolakan dan sentimen positif menandakan persetujuan atau dukungan.

Metode yang digunakan adalah Naive Bayes Classifier (NBC) yang memiliki beberapa kelebihan antara lain, sederhana, cepat dan berakurasi tinggi. Metode *Naive Bayes Classifier* untuk klasifikasi atau kategorisasi teks menggunakan atribut kata yang muncul dalam satu dokumen sebagai dasar klasifikasinya. Algoritma klasifikasi *Naive Bayes* memanfaatkan teori probabilitas yang dikemukakan oleh

ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya (D. Yanti, 2013).

Dengan adanya penelitian “Analisis Sentiment Suara Masyarakat Kepada Pemerintah Terkait Perpindahan Ibu Kota Menggunakan Metode Naive Bayes Classifier” dapat diketahui sejauh mana masyarakat akan memberikan pendapatnya melalui beberapa sampel pengguna Twitter dengan tweet tentang perpindahan ibu kota dan suara masyarakat yang mendukung dan tidak mendukung akan diketahui. Sehingga akan diketahui bahwa langkah presiden tersebut tepat atau justru langkah tersebut salah karena masyarakat menginginkan sektor lain yang sebaiknya perlu diperbaiki dari pada pemindahan ibu kota ini. Karena dengan proyek ini dana yang dikeluarkan jelas tidaklah sedikit. Dari hasil penelitian maka bisa didapatkan informasi topik yang sering diungkapkan masyarakat sehingga pihak pemerintah bisa menyikapi sesuai harapan masyarakat dan dapat mengkajinya dengan tepat sehingga menghasilkan perubahan yang baik bagi masyarakat Indonesia terutama daerah yang terkena imbas dari proyek ini. Selain itu dengan adanya contoh penelitian yang nyata bisa digunakan sebagai informasi bahwa program pemindahan ibu kota ini memiliki hasil yang pro kontra pada masyarakat yang diambil dari data tweet beberapa pengguna Twitter.

Dari data yang didapat akan diolah menggunakan bahasa pemrograman Python dan akan memuat informasi dengan kecenderungan positif dan negatif. Dengan adanya penelitian ini yang berjudul “*Studi Eksperimen Metode Naive Bayes Classifier Pada Analisis Sentimen Pengguna Twitter Terhadap Opini Perpindahan*

Ibu Kota”, diharapkan dari penelitian ini dapat dihasilkan informasi yang dapat berguna bagi pihak terkait dalam pelaksanaan perpindahan ibu kota ini.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah yang akan diselesaikan dalam penelitian ini yaitu:

“Belum adanya analisis sentimen untuk klasifikasi opini masyarakat pada perpindahan ibu kota baru menggunakan metode Naive Bayes Classifier”.

1.3 Batasan Penelitian

Dalam penelitian ini terdapat beberapa batasan masalah yang dibahas agar penyusunan dan pembahasan penelitian dapat dilakukan secara terarah dan tercapai sesuai dengan yang diharapkan. Antara lain yaitu sebagai berikut:

1. Data yang digunakan berasal dari sampel pengguna Twitter
2. Data yang diambil adalah data mulai tahun 2019-2020
3. Menggunakan metode TF-IDF untuk menghitung bobot kata.
4. Pengolah data menggunakan bahasa pemrograman Python.
5. Algoritma yang digunakan dalam pengklasifikasian data adalah metode *Naive Bayes Classifier* .
6. Data diklasifikasi sebagai positif dan negatif
7. Pengambilan data menggunakan API Twitter yang mengandung kalimat ibu kota.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai oleh penulis dari penelitian dan penyusunan tugas akhir ini adalah:

1. Menganalisis sentimen masyarakat pengguna Twitter dengan proses Text Mining.
2. Mengetahui penilaian masyarakat atau topik terhadap pemerintah terkait perpindahan ibu kota yang dapat digunakan sebagai informasi terhadap proses perpindahan ibu kota baru.

1.5 Manfaat Penelitian

Penelitian ini akan memberi manfaat bagi peneliti dan pembaca untuk:

1. Mengklasifikasi cuitan Twitter ke dalam klasifikasi positif dan negatif dalam jumlah besar.
2. Dapat mengetahui mengenai cara sentimen analisis menggunakan metode *Naive Bayes* dan pembobotan kata TF-IDF.
3. Dapat mengetahui tentang pandangan masyarakat terhadap pemerintah yang melakukan perpindahan ibu kota.

1.6 Keaslian Penelitian

Penelitian sentimen menggunakan *Metode Naive Bayes Classifier* memanglah sudah banyak, tetapi peneliti belum menemukan judul yang sama untuk metode yang digunakan dalam sentimen analisis dengan judul tersebut.

1.7 Sistematika Penulisan

Sebagai gambaran dan kerangka yang jelas mengenai pokok bahasan setiap bab dalam penelitian ini, maka diperlukan sistematika penulisan. Penyusunan laporan tugas akhir ini memiliki sistematika penulisan yang diawali dari BAB I dan diakhiri BAB V. Berikut adalah penjelasan pada tiap-tiap bab dalam laporan penelitian ini :

BAB I PENDAHULUAN

Bab pendahuluan berisikan penjelasan mengenai latar belakang dilakukannya penelitian, rumusan masalah penelitian, batasan masalah, tujuan penelitian, manfaat penelitian, keaslian penelitian, dan sistematika penulisan penelitian.

BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI

Bab tinjauan pustaka dan landasan teori berisikan mengenai penelitian terdahulu dan teori-teori dasar yang terkait dengan penelitian ini. Teori yang digunakan terdiri dari analisis sentimen, *TF-IDF*, metode *Naïve Bayes Classifier*, *text preprocessing*, *Feature Extraxtion*, *Confussion Matrix* dan *Python*.

BAB III METODE PENELITIAN

Bab metode penelitian berisi tentang penjelasan mengenai metode ataupun algoritma yang digunakan serta tahapan-tahapan yang dilakukan untuk mencapai tujuan dan kesimpulan tugas akhir.

BAB IV HASIL DAN PEMBAHASAN

Bab hasil dan pembahasan membahas analisis data dan hasil dari penelitian yang telah dilakukan.

BAB V PENUTUP

Bab penutup berisi tentang kesimpulan dari hasil penelitian yang telah dilakukan. Selanjutnya, kekurangan yang ada pada penelitian dituliskan pada saran untuk pengembangan penelitian di masa yang akan datang.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang dilakukan dari data latih sebanyak 1290 menggunakan *split validation* menghasilkan *split data* dengan hasil tertinggi yaitu 90:10 yang dihitung menggunakan *confusion matrix* metode *Naive Bayes* dan pembobotan kata TF-IDF menghasilkan akurasi 76,74%. Hasil dari implementasi pada data uji sebanyak 1115 data menghasilkan sentimen dengan label negatif sebanyak 799 dan sentimen positif sebanyak 316. Dari implementasi yang dilakukan menghasilkan sentimen negatif sebanyak 71,7% dan sentimen positif sebanyak 28,3%. Jadi hasilnya mayoritas memberikan sentimen negatif atau tidak setuju untuk proses pemindahan ibu kota baru.

Hasil dari penelitian memunculkan beberapa kata yang menjadi fokus diperbincangkan diantaranya adalah negara, bangun, kalimantan, tetap dan jokowi, dari kata-kata tersebut bisa digambarkan bahwa tweet yang memiliki sentimen positif mendukung dengan adanya proses perpindahan ibu kota negara ke Kalimantan yang dilakukan oleh Jokowi sedangkan untuk sentimen negatif kata yang sering muncul diantaranya adalah enggak, covid, corona, banjir dan pandemi, dari kata-kata tersebut tergambar bahwa masyarakat yang memberikan sentimen negatif ingin bahwa pemerintah lebih fokus terhadap penanggulangan bencana seperti banjir yang sering melanda pada musim hujan dan bencana yang sedang melanda seluruh dunia yaitu corona atau virus covid 19 dari pada proyek pindah ibu kota negara yang dilakukan pemerintah. Sedangkan pada data uji dengan sentimen positif juga menghasilkan kata yang sering muncul diantaranya kalimantan, jakarta, bangun, mulai, dan jokowi, hasil tersebut bahwasanya hampir sama dengan data latih. Untuk sentimen negatif muncul

kata-kata diantaranya negara, rakyat, proyek, *covid*, *corona*, dan pandemi. Sama bahwa hasil ini juga hampir sama dengan yang dihasilkan pada data latih. Maka dapat disimpulkan bahwa masyarakat lebih dominan memberikan sentimen negatif karena dari pada membangun atau memindahkan ibu kota baru ke Kalimantan alangkah lebih baiknya dana yang digunakan untuk memindahkan ibu kota sebaiknya digunakan untuk menyelesaikan masalah bencana seperti banjir, virus *covid 19* dan untuk membantu rakyat yang mengalami kesusahan.

5.2 Saran

Pada penelitian yang dilakukan pasti ada kekurangan. Maka penulisan menyarankan beberapa hal yang dapat dilakukan agar penelitian yang dilakukan selanjutnya dapat lebih baik, diantaranya adalah :

1. Membandingkan kembali menggunakan metode pembelajaran mesin lainnya untuk mendapatkan hasil yang terbaik.
2. Menggunakan data latih yang lebih banyak supaya bisa mendapatkan hasil yang lebih akurat.
3. Menggunakan pakar pada bidang terkait dalam proses pelabelan data supaya data latih yang digunakan benar-benar teruji kebenarannya.

DAFTAR PUSTAKA

- Abpama Sevsa Bama. (2019). *Analisis Sentimen pada Indeks Kinerja Dosen Fakultas SAINTEK UIN Sunan Kalijaga Menggunakan Naive Bayes Classifier*. 10, 112–123.
- Bestari, D. P. B., Saptono, R., & Anggrainingsih, R. (2019). ACADEMIC ARTICLES CLASSIFICATION USING NAIVE BAYES CLASSIFIER (NBC) METHOD. *ITSMART: Jurnal Teknologi Dan Informasi*, 7(2), 74–81.
- Buntoro, G. A. (2017). Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter. *INTEGER: Journal of Information Technology*, 2(1).
- Daniel Jurafsky & James H. Martin. (2019). Naive Bayes and Sentiment Classification. In *Speech and Language Processing*.
- Dehaff, M. (2010). Sentiment Analysis, Hard But Worth It! Tersedia Di: [Http://Www.Customerthink.Com/Blog/Sentiment_analysis_hard_but_worth_it](http://www.Customerthink.Com/Blog/Sentiment_analysis_hard_but_worth_it).
- Fachrurrozi, M., & Yusliani, N. (2015). *Analisis Sentimen Pengguna Jejaring Sosial Menggunakan Metode Support Vector Machine*.
- Hadna, N. M. S., Santosa, P. I., & Winarno, W. W. (2016). *STUDI LITERATUR TENTANG PERBANDINGAN METODE UNTUK PROSES*. 2016(Sentika), 18–19.
- Hall, M. (2006). A decision tree-based attribute weighting filter for naive Bayes. *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, 59–70.

- Hamidah, N. N. (2019). *Analisis Sentiment Masyarakat Terhadap Fintech Menggunakan Metode Naive Bayes Classifier*.
- Hidayatullah, A. F., Azhari, S. N., & others. (2014). *Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Data Twitter Menggunakan Naive Bayes Classifier*. [Yogyakarta]: Universitas Gadjah Mada.
- Kalokasari, D. H., Shofi, I. M., & Setyaningrum, A. H. (2017). *IMPLEMENTASI ALGORITMA MULTINOMIAL NAIVE BAYES CLASSIFIER PADA SISTEM KLASIFIKASI SURAT KELUAR (Studi Kasus : DISKOMINFO Kabupaten Tangerang)*. 10(2). <https://doi.org/10.15408/jti.v10i2.6822>
- Manalu, B. U. (2014). *Analisis Sentimen Pada Twitter Menggunakan Text Mining*.
- Novantirani, A., Sabariah, M. K., & Effendy, V. (2015). Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota dengan Metode Support Vector Machine. *EProceedings of Engineering*, 2(1).
- Nurhuda, F., Sihwi, S. W., & Doewes, A. (2014). Analisis sentimen masyarakat terhadap calon Presiden Indonesia 2014 berdasarkan opini dari Twitter menggunakan metode Naive Bayes Classifier. *ITSmart: Jurnal Teknologi Dan Informasi*, 2(2), 35–42.
- Prasetyo, E. (2019). *Data mining mengolah data menjadi informasi menggunakan matlab*.
- Rish, I., & others. (2001). An empirical study of the naive Bayes classifier. *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, 3(22), 41–46.
- Rocha, A. de R., 2006. Naive Bayes Classifier Teaching Material. , pp.1–9. Available at: <http://www.ic.unicamp.br/~rocha/teaching/2011s2/mc906/aulas/naive-bayes->

classifier.pdf.

- Rustiana, D., & Rahayu, N. (2017). Analisis Sentimen Pasar Otomotif Mobil: Tweet Twitter Menggunakan Naïve Bayes. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 8(1), 113–120.
- Suryono, S., Utami, E., & Luthfi, E. T. (2018). Klasifikasi Sentimen Pada Twitter Dengan Naive Bayes Classifier. *Angkasa: Jurnal Ilmiah Bidang Teknologi*, 10(1), 89. <https://doi.org/10.28989/angkasa.v10i1.218>
- Taheri, S., & Mammadov, M. (2013). Learning the naive bayes classifier with optimization models. *International Journal of Applied Mathematics and Computer Science*, 23(4), 787–795. <https://doi.org/10.2478/amcs-2013-0059>
- Triasanti, D. (2000). Konsep dasar python. *Python*, 1–6. <https://doi.org/aa>
- Untari, D. (2010). Data Mining untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode Decision Tree C4.5. *Fakultas Ilmu Komputer Universitas Dian Nuswantoro*.
- Wahyunita, L. (2017). Klasifikasi Penyebab Penyalahgunaan Narkoba Dari Berita Online Dengan Menggunakan Naive Bayes. *Jurnal ELTIKOM*, 1(1), 23–30. <https://doi.org/10.31961/eltikom.v1i1.12>
- Wu, X., & Kumar, V. (2009). *The top ten algorithms in data mining*. CRC press.
- Yahya, H. M. (2018). *Pemindahan Ibu Kota Negara Maju dan Sejahtera*. 14(01), 21–30. <https://doi.org/10.23971/jsam.v14i1.kemerdekaan>