

POLITEKNIK ELEKTRONIKA NEGERI SURABAYA
JURUSAN TEKNIK INFORMATIKA
2015



JUDUL PROYEK AKHIR

1. Nama Mahasiswa : Handaru Dwiki Yuntara
2. NRP : 3122500017
3. Kelas : 2 D3 IT A
4. Jurusan/Program Studi : D3 Teknik Informatika
5. Judul Proyek Akhir : Sentiment Analysis Di Media Sosial X Terhadap Pembangunan IKN Menggunakan Algoritma Naïve Bayes dan Algoritma XGBOST
6. Deskripsi Proyek Akhir :
 - Latar belakang Proyek akhir
Sentimeny analysis adalah proses menganalisis teks digital untuk menentukan apakah nada emosional pesan tersebut positif, negatif atau netral. Dalam konteks pembangunan IKN atau ibu kota nusantara, sentiment analysis dapat menjadi alat yang sangat berguna untuk memahami pandangan dan persepsi masyarakat terhadap proyek tersebut.
"IKN" adalah singkatan dari "Ibu Kota Negara". Dalam konteks Indonesia, "IKN" atau "Ibu Kota Nusantara" merujuk pada rencana pembangunan ibu kota baru negara Indonesia. Rencana ini mencakup pembangunan ibu kota baru yang direncanakan untuk menggantikan Jakarta sebagai ibu kota negara.
Beberapa alasan utama di balik rencana pembangunan Ibu Kota Nusantara termasuk kondisi Jakarta yang semakin terbebani oleh masalah-masalah seperti kemacetan lalu lintas, banjir, dan penurunan tanah akibat eksploitasi air tanah yang berlebihan. Selain itu, rencana ini juga dimaksudkan untuk meratakan pembangunan dan pertumbuhan ekonomi di Indonesia dengan memindahkan pusat administrasi ke wilayah timur Indonesia.
Pembangunan Ibu Kota Nusantara menjadi salah satu proyek strategis nasional yang dicanangkan oleh pemerintah Indonesia. Proyek ini melibatkan pengembangan infrastruktur, perumahan, fasilitas publik, serta pembangunan kota yang berkelanjutan.
Rencana ini tidak hanya melibatkan pemindahan kantor-kantor pemerintahan dan institusi negara, tetapi juga pembangunan kota baru yang didesain untuk

POLITEKNIK ELEKTRONIKA NEGERI SURABAYA
JURUSAN TEKNIK INFORMATIKA
2015

menjadi pusat administrasi, ekonomi, dan budaya yang modern. Diharapkan bahwa pembangunan Ibu Kota Nusantara akan membawa dampak positif dalam memperbaiki infrastruktur, meratakan pembangunan, dan meningkatkan kualitas hidup masyarakat di wilayah tersebut.

Pembangunan ibu kota baru adalah proyek besar yang akan memiliki dampak luas, baik secara ekonomi, sosial, maupun lingkungan. Pemerintah dan pemangku kepentingan lainnya perlu memahami bagaimana pendapat publik berkembang seiring waktu terhadap proyek ini. Sentiment analysis dapat membantu mengidentifikasi tren dan pola dalam opini publik.

Pembangunan infrastruktur besar seperti ibu kota baru seringkali menjadi subjek kontroversi dan polarisasi dalam masyarakat. Beberapa orang mungkin mendukung proyek ini karena potensi pembangunan ekonomi dan peluang baru, sementara yang lain mungkin mengkhawatirkan dampak lingkungan, biaya, atau efek sosialnya. Sentiment analysis dapat membantu mengidentifikasi klaster atau kelompok opini yang berbeda.

Dalam mengelola proyek semacam ini, penting bagi para pemangku kepentingan untuk membuat keputusan berdasarkan data yang akurat. Sentiment analysis dapat menyediakan wawasan tambahan yang berguna untuk memandu keputusan strategis terkait dengan pembangunan ibu kota baru.

Sentiment analysis juga dapat digunakan untuk memantau potensi krisis atau penolakan masyarakat terhadap proyek. Dengan mendeteksi sentimen negatif secara dini, pemerintah dan pihak terkait dapat mengambil langkah-langkah mitigasi yang diperlukan untuk meredakan ketegangan atau memperbaiki persepsi masyarakat.

Analisis sentimen juga dapat digunakan sebagai alat evaluasi untuk melacak bagaimana pendapat publik berubah seiring waktu sebagai respons terhadap langkah-langkah yang diambil oleh pemerintah atau pemangku kepentingan lainnya. Hal ini memungkinkan penyesuaian strategi komunikasi dan kebijakan secara lebih responsif.

- **Permasalahan Proyek akhir**

1. **Dampak Ekonomi dari pembangunan Ibu Kota Nusantara**

Analisis sentimen dapat membantu dalam memahami bagaimana perubahan ibu kota baru mempengaruhi persepsi dan harapan

POLITEKNIK ELEKTRONIKA NEGERI SURABAYA
JURUSAN TEKNIK INFORMATIKA
2015

masyarakat terhadap pertumbuhan ekonomi, peluang pekerjaan, dan investasi di daerah baru tersebut.

2. Dampak Lingkungan dari pembangunan Ibu Kota Nusantara

Perubahan ibu kota dapat memiliki dampak signifikan pada lingkungan dan ekosistem. Sentiment analysis dapat membantu dalam memahami perasaan masyarakat terkait dengan pelestarian lingkungan, keberlanjutan, dan langkah-langkah mitigasi yang diambil.

3. Ketidakpastian dan ketegangan

Proyek perubahan ibu kota sering kali memunculkan ketidakpastian dan ketegangan dalam masyarakat, terutama terkait dengan pemindahan penduduk, redistribusi sumber daya, dan pembagian keuntungan. Sentiment analysis dapat membantu dalam memantau dan mengelola ketegangan ini.

4. Keterlibatan Rakyat

Analisis sentimen dapat membantu dalam mengevaluasi tingkat keterlibatan masyarakat dalam proses perubahan ibu kota, termasuk tingkat kepercayaan, partisipasi, dan pemahaman terhadap proses pengambilan keputusan.

- Tujuan Proyek akhir

1. Kebutuhan Akan Pemahaman Opini Publik
2. Polarisasi Opini
3. Manajemen Resiko dan Krisis
4. Pengambilan Keputusan Berbasis data
5. Evaluasi Kinerja dan Responsivitas
6. Membandingkan algoritma Naive Bayes dan Xgbost untuk data ini

- Metode yang digunakan (Tinjauan pustaka secara singkat)

1. Naive Bayes

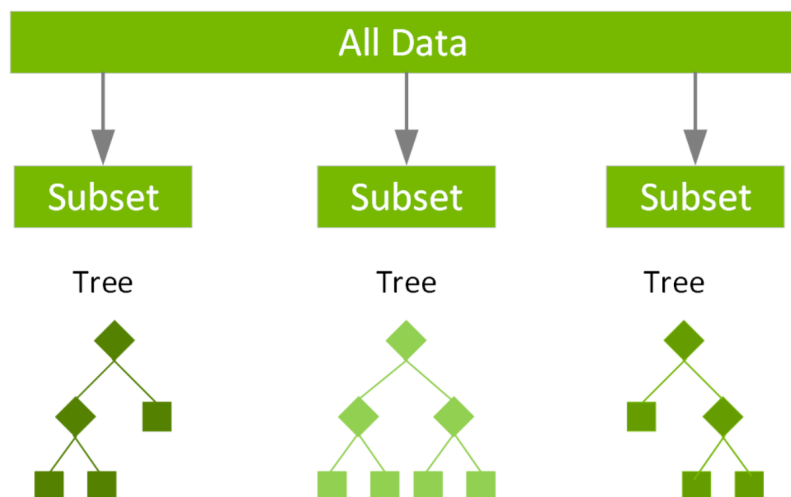
Naive bayes merupakan metode pengklasifikasian paling populer digunakan dengan tingkat keakuratan yang baik. Banyak penelitian tentang pengklasifikasian yang telah dilakukan dengan menggunakan algoritma ini. Berbeda dengan metode pengklasifikasian dengan logistic regression ordinal maupun nominal, pada algoritma naive bayes pengklasifikasian tidak membutuhkan adanya pemodelan maupun uji statistik.

POLITEKNIK ELEKTRONIKA NEGERI SURABAYA
JURUSAN TEKNIK INFORMATIKA
2015

Naive bayes merupakan metode pengklasifikasian berdasarkan probabilitas sederhana dan dirancang agar dapat dipergunakan dengan asumsi antar variabel penjelas saling bebas (independen). Pada algoritma ini pembelajaran lebih ditekankan pada pengestimasian probabilitas. Keuntungan algoritma naive bayes adalah tingkat nilai error yang didapat lebih rendah ketika dataset berjumlah besar, selain itu akurasi naive bayes dan kecepatannya lebih tinggi pada saat diaplikasikan ke dalam dataset yang jumlahnya lebih besar.

2. XGBoost

XGBoost (Extreme Gradient Boosting) adalah sebuah algoritma machine learning yang sangat populer dan efektif dalam analisis data, khususnya dalam tugas-tugas seperti klasifikasi, regresi, dan peringkat. Algoritma ini dikembangkan oleh Tianqi Chen pada tahun 2014 dan menjadi sangat populer karena kecepatan, efisiensi, dan kemampuannya menghasilkan prediksi yang akurat.



umber Gambar: NVIDIA

XGBoost menggunakan teknik ensemble learning, yang menggabungkan prediksi dari beberapa model untuk meningkatkan kinerja keseluruhan. Pendekatan ini berbasis pada teknik boosting, di mana model-model lemah (weak learners), biasanya berupa pohon keputusan (decision trees) yang sederhana, digabungkan menjadi model yang kuat (strong learner).

POLITEKNIK ELEKTRONIKA NEGERI SURABAYA
JURUSAN TEKNIK INFORMATIKA
2015

- Rancangan Diagram sistem

Proses untuk melakukan sentiment analysis terdiri dari beberapa proses yaitu sebagai berikut :

1. Crawling Data

Crawling data adalah proses pengumpulan data dari internet menggunakan kode program. Dalam kasus ini saya menggunakan bahasa pemrograman python. Dan berikut data hasil dari crawling melalui media sosial X

Proses Crawling

```
✓ 19m data = 'data_ikn.csv'
search_keyword = 'ikn until:2024-3-1 since:2023-11-01'
limit = 1000

!npx --yes tweet-harvest@2.2.8 -o "{data}" -s "{search_keyword}" -l {limit} --token ""

Got some tweets, saving to file...
Error parsing response json: {"_type":"Response","_guid":"response@4069cb575b6ca107ef48b3b9cd8605d0"}
Most likely, you have already exceeded the Twitter rate limit. Read more on https://twitter.com/elonmusk/status/1675187969420828

Got some tweets, saving to file...
Your tweets saved to: /content/tweets-data/data_ikn.csv
Total tweets saved: 843

Got some tweets, saving to file...
Your tweets saved to: /content/tweets-data/data_ikn.csv
Total tweets saved: 861

--Taking a break, waiting for 10 seconds...

Got some tweets, saving to file...
Your tweets saved to: /content/tweets-data/data_ikn.csv
Total tweets saved: 879

Got some tweets, saving to file...
Your tweets saved to: /content/tweets-data/data_ikn.csv
Total tweets saved: 897

Got some tweets, saving to file...
Your tweets saved to: /content/tweets-data/data_ikn.csv
Total tweets saved: 914

Got some tweets, saving to file...
Your tweets saved to: /content/tweets-data/data_ikn.csv
Total tweets saved: 929

Got some tweets, saving to file...
Your tweets saved to: /content/tweets-data/data_ikn.csv
```

POLITEKNIK ELEKTRONIKA NEGERI SURABAYA

JURUSAN TEKNIK INFORMATIKA

2015

Data hasil dari crawling

created_at	id_str	full_text
Thu Feb 29 23:59:33 +0000 2024	1763353346315014285	@MARQUEZ_93 Gw setuju dg kritikan yg dilontarkan ahy utk jkw...dan yg gw gk hbs pikir knpa skrg tiba2 ahy muji2 jkw ttg IKN..jd selama ini mengkritik krn gk sebagian jajah kekuasaan?
Thu Feb 29 23:58:56 +0000 2024	1763353188588241332	@bukancowokgalau @opposite6892 Disuruh pindah ikn guys.. Makan ditenggelman 🤔👎
Thu Feb 29 23:58:41 +0000 2024	1763353125778604355	Bongkar semua ketimpangan dan ketidakberasan di lahan IKN
Thu Feb 29 23:58:22 +0000 2024	1763353045063409728	@VIVAacid Bakalan banyak proyek nanti nih di ikn 🤔
Thu Feb 29 23:57:34 +0000 2024	1763352843988455902	Karakter aslinya kekhatan.. Sch dpt tulang... anteng.. AHY Kunjungan Perdana ke IKN Nusantara: Saya Terpukau! https://t.co/oDaPbuytk6
Thu Feb 29 23:56:58 +0000 2024	1763352693744320938	@WagimanDeep212_ Pingin buru2 Istana pindah di IKN
Thu Feb 29 23:56:41 +0000 2024	1763352623099662770	@herculep638 Yang diomongin lg makan² di IKN 🤔
Thu Feb 29 23:56:19 +0000 2024	1763352552519436432	Ya sudah klo gitu presiden yg dipaksakan biar berkantor di ikn dan presiden sesuai konstitusi berkantor di dki jakarta, berani gak? Kalian yg bangga silakan dukung presiden kalian berkantor
Thu Feb 29 23:55:42 +0000 2024	1763352374507393168	Bersyukur dengan kemajuan pembangunan IKN. Harap berhati-hati dalam pemasangan listrik, air serta internet, serahkan saja kepada BUMN yang mengelolanya. Tapi saya yakin Pak @jokowi
Thu Feb 29 23:53:29 +0000 2024	1763351819345178662	@brother_djon Pantess ngebet bgt bikin IKN
Thu Feb 29 23:53:29 +0000 2024	1763351816904057026	@liputan6dotcom Iya terpukau gebetannya banyak di ikn 🤔
Thu Feb 29 23:52:46 +0000 2024	1763351637224284423	Jumat Indonesia 1 Maret #DemilIndonesia PEMILU Presiden RI APBN BUMN IKN Nusantara 0 KM 2024 Reformasi Belum Selesai SHP PIN Merah Putih #stamps #philately #gallery #art Rp 1
Thu Feb 29 23:48:05 +0000 2024	1763350456733250007	友達との約束の時間に起きませんでした!!
Thu Feb 29 23:45:30 +0000 2024	1763349808344101175	@geloraco Ign sombong, inget harimau mu mulut mu, sperti yg Udeh diucapkan KA Nasdem kritik IKN sekarang menjilat ludah sendiri
Thu Feb 29 23:45:00 +0000 2024	1763349681667543087	Jokowi bakal berkantor di IKN setelah Bandara dan Jalan Tol Rampung https://t.co/EquuZCH1eu
Thu Feb 29 23:35:45 +0000 2024	1763347355015975274	Rocky Gerung Sindir AHY yang Sekarang Memuji IKN: Kapasitas Seorang Leader yang jadi Dealer https://t.co/dauRfmTHmb
Thu Feb 29 23:35:37 +0000 2024	1763347319859327054	@tvOneNews IKN ibarat cuman masukin keranjang doang, ga di checkout
Thu Feb 29 23:35:22 +0000 2024	1763347258253492538	@ZeroBot @99propaganda @tham878 @are_inismyname @BANGSAygSUJUD @Bank_Joeie @bengkeldodo @Bul_winner @Lembayung071 @Ndons_Back Kalau kesampingkan ego pribadi
Thu Feb 29 23:34:33 +0000 2024	1763347054448062562	PNS yang Pindah ke IKN Wajib 'Melek' Digital Menpan-RB memastikan ASN yang pindah ditetapkan kriteria kompetensi diantaranya harus menguasai literasi (digital literacy), multitasking,
Thu Feb 29 23:33:54 +0000 2024	1763346887594361329	@CNNIndonesia Kan adanya ikn biar gak Jawa sentris, nah ayo dong mulai sekarang kalo ada simulasi program apapun jangan di Jawa duluan. Bisa kan ya ke pulau-pulau yang lainnya.
Thu Feb 29 23:31:49 +0000 2024	1763346364929651001	@Jayabay19479190 Dulu nyinyir dgn ikn...
Thu Feb 29 23:30:35 +0000 2024	1763346056396632549	Kemarin, Kemensos inovasi kompor hingga Jokowi bayangkan konser di IKN https://t.co/Tj5lG54EN
Thu Feb 29 23:30:21 +0000 2024	1763345995507945500	@benbela Semoga kalimantan ekonominya kembali karna dorongan ekonomi dari ikn, agar ketika menangkap gambar kuyang bisa memakai kamera handphone dengan resolusi tinggi
Thu Feb 29 23:30:06 +0000 2024	1763345934061306236	Pembangunan IKN Makin Padat, Jokowi: Minat Investor Terus Berkembang https://t.co/PumecYAmHo
Thu Feb 29 23:29:31 +0000 2024	1763345788267274319	Sempat mengkritik IKN...Menteri ATR AHY...kagum dg IKN!!!!!!! makaaaa...sebelum mengeritik...lihat..pahami...tinjau... dan pikirkan dulu!!!!!!ini utk yg masih keukeuh nolak IKN 🤔🤔🤔🤔
Thu Feb 29 23:29:30 +0000 2024	1763345782684709248	Biar ada alasan untuk memindahkan Ibukota Negara ke IKN ya Pak @jokowi ... 🤔🤔
Thu Feb 29 23:29:13 +0000 2024	1763345711616471509	Jokowi Klaim Banyak Investor Antre Tanam Modal di IKN https://t.co/mvPxxKc89v
Thu Feb 29 23:28:05 +0000 2024	1763345427263631653	Apa salahnya obrolan begini? Gak harus selalu ngomongin amdal IKN, kan?
Thu Feb 29 23:24:11 +0000 2024	1763344444819226761	@tempodotco Butuh ratusan tahun ikn jadi kota modern 🤔
Thu Feb 29 23:23:36 +0000 2024	1763344295325794437	Jokowi kembali menginap di IKN didampingi sejumlah menteri. Di antaranya Menteri ATR AHY, Menteri BUMN Erick Thohir sampai Panglima TNI https://t.co/rSwbWf33xM
Thu Feb 29 23:22:18 +0000 2024	1763343968996319575	Daripada utakatik dana yg udah mapan, mending pake dana ikn aja.....
Thu Feb 29 23:20:31 +0000 2024	1763343521166361037	Tapi tapi banyak bgt masyarakat bukan sekitar IKN malah pengen oke gas melanjutkan 🤔
Thu Feb 29 23:19:08 +0000 2024	1763343172921594301	Menteri PUPR Pindah Kantor ke IKN pada Juli, Jokowi Tunggu Bandara Selesai https://t.co/t2FkoDKetK #TempoNasional
Thu Feb 29 23:18:56 +0000 2024	1763343123391090990	@wielj4k @doelpaten @12amadi @vanc1Bozz @abu_waras @CutSarina5 @OposisiCerdas @ch_chotimah2 @xix0_3 @pendi138 @Anak2Di @RendangOtentik @Zahrati443738 @papa_joren €

2. Cleaning Data

Karena data yang kita dapatkan masih data mentah dan masih kotor maka perlu kita lakukan adalah melakukan cleaning data. Ada beberapa hal yang harus dilakukan di cleaning data ini yaitu yang pertama bisa menghapus kolom yang dirasa kurang penting . Lalu menghapus data yang duplicat karena data yang duplicat tidak menambah train terhadap model. Dan yang terakhir adalah menghapus karakter seperti karakter @,#,dll yang tidak memiliki arti

3. Preprocessing Data

Setelah data sudah bersih, langkah berikutnya adalah melakukan preprocessing data. Ini mencakup langkah-langkah seperti menghapus karakter khusus, konversi teks menjadi huruf kecil, penghapusan stopwords (kata-kata umum yang tidak memiliki makna penting), tokenisasi (pemisahan teks menjadi kata-kata), dan normalisasi (misalnya, mengubah kata-kata yang bermacam-macam menjadi bentuk dasarnya).

POLITEKNIK ELEKTRONIKA NEGERI SURABAYA
JURUSAN TEKNIK INFORMATIKA
2015

Preprocessing terdiri dari beberapa tahap antara lain :

a. Normalisasi

Pada tahap normalisasi yang dilakukan mengubah kata menjadi bentuk dasar seperti mengubah semua huruf menjadi huruf kecil

b. Stopword

Pada tahap ini yang dilakukan adalah menghilangkan kata yang sering muncul dan tidak memiliki kontribusi contoh : di,dan, ke, dll

c. Tokeniz

Di tahap ini yang dilakukan adalah memecah teks menjadi kata atau frasa yang kemudian dimasukkan ke dalam list.

d. Stemming

Stemming adalah proses menghilangkan imbuhan dari kata untuk mengembalikan kata ke bentuk dasarnya yang disebut "stem". Tujuannya adalah untuk mengurangi variasi kata yang mungkin terjadi dalam teks, sehingga kata-kata dengan akar yang sama dianggap sama. Contoh dari stemming adalah mengubah kata-kata seperti "berjalan", "berjalanlah", dan "berjalanannya" menjadi kata dasar "jalan". Ini membantu mengurangi kompleksitas data teks dan meningkatkan konsistensi dalam analisis.

4. Pemodelan

Setelah ekstraksi fitur selesai, langkah selanjutnya adalah memilih dan melatih model untuk melakukan klasifikasi sentimen. Algoritma yang saya gunakan untuk kasus ini adalah Naive Bayes, dan XGBoost. Model-model ini dilatih menggunakan data yang sudah diproses untuk memprediksi sentimen dari teks baru.

5. Evaluasi Model

Setelah model dilatih, langkah selanjutnya adalah mengevaluasi performanya. Evaluasi dilakukan menggunakan metrik-metrik seperti akurasi, presisi, recall, dan F1-score. Selain itu, validasi silang (cross-validation) juga sering digunakan untuk memastikan bahwa model tidak overfitting.

POLITEKNIK ELEKTRONIKA NEGERI SURABAYA
JURUSAN TEKNIK INFORMATIKA
2015

6. Interpretasi Hasil

Tahap terakhir dalam analisis sentimen adalah menginterpretasi hasil yang diperoleh dari model. Ini melibatkan pemahaman tentang bagaimana model mengklasifikasikan teks menjadi kategori sentimen yang berbeda (positif, negatif, atau netral), serta penyelidikan lebih lanjut terhadap contoh-contoh teks yang diklasifikasikan salah atau ambigu.

7. Nama Dosen Pembimbing :
Pembimbing 1 : Rengga Asmara, S.Kom., M.T.
Pembimbing 2 : Ira Prasetyaningrum, S.Si, M.T.

Surabaya, 12 oktober 2015



(Handaru Dwiki Yuntara)
NRP3122500017