ANALISIS SENTIMENT PADA TWEET DENGAN TAGAR #MAHKAMAKONSTITUSI MENGGUNAKAN METODE MULTI LAYER PERCEPTRON

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Syarat
untuk Memperoleh Gelar sarjana Teknik Pada Fakultas Teknik
Universitas Islam Riau

OLEH :

ASIF UMMATUL KHAIRA
143510567

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS ISLAM RIAU
PEKANBARU
2021

KATA PENGANTAR

Assalamu'alaikum Wr, Wb.

Dengan mengucapkan Puji dan syukur kepada Allah SWT yang senantiasa melimpahkan rahmat dan hidayah-Nya kepada penulis sehingga penulis mampu menyelesaikan laporan penelitian skripsi yang berjudul **"Analisis Sentiment Pada Tweet Dengan Tagar #Mahkamahkonstitusi Menggunakan Metode Multi Layer Perceptron".**

Laporan penelitian skripsi ini adalah salah satu syarat untuk memperoleh gelar sarjana Teknik di Fakultas Teknik Universitas Islam Riau. Penulis menyadari, bahwa penulisan skripsi ini tidak akan terwujud tanpa adanya dukungan, bantuan, saran, dan kritik yang telah penulis terima, oleh karena itu dalam kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Kedua orang tua penulis, Bapak Zulyadi & Ibu Lidiawati serta Adek Nuril Ilma yang selalu mendoakan dan selalu memberikan segala dukungan kepada penulis dalam menyelesaikan skripsi ini.
2. Bapak Dr. Eng. Muslim, ST., MT selaku Dekan Fakultas Teknik Universitas Islam Riau. Beserta seluruh Dekan I, Dekan II dan Dekan III.
3. Kepada Bapak Dr. Arbi Haza Nasution, B.IT.,M.IT. selaku Ketua Program Studi Teknik Informatika Universitas Islam Riau. Sekaligus Pembimbing I yang telah memberikan bimbingan, motivasi, dan arahan kepada penulis sehingga skripsi ini bisa selesai dengan baik.

4. Seluruh dosen Program Studi Teknik Informatika Universitas Islam Riau atas segala ilmu pengetahuan yang diberikan kepada penulis. Beserta seluruh staff akademik Fakultas Teknik Universitas Islam Riau.
5. Kepada seluruh teman-teman Teknik Informatika angkatan 2014, serta rekan-rekan seperjuangan yang saya banggakan.

Akhir kata, penulis mohon maaf atas kekeliruan dan kesalahan yang terdapat dalam skripsi ini, dan penulis berharap semoga skripsi ini dapat bermanfaat dan menambah wawasan serta pengetahuan bagi para pembaca.

Pekanbaru,

Asif Ummatul Khaira

DAFTAR ISI

KATA PENGANTAR..... i

DAFTAR ISI.....iii

DAFTAR GAMBAR.....vi

DAFTAR TABEL.....vii

BAB I PENDAHULUAN..... 1

1.1 Latar Belakang..... 1

1.2 Identifikasi Masalah..... 4

1.3 Rumusan Masalah..... 4

1.4 Batas Masalah..... 4

1.5 Tujuan Penelitian..... 5

1.6 Manfaat Penelitian..... 5

BAB II LANDASAN TEORI..... 6

2.1 Studi Kepustakaan..... 6

2.2 Dasar Teori..... 8

2.2.1 Sentimen Analysis..... 8

2.2.2 Twitter..... 9

2.2.3 K-fold Cross Validation..... 9

2.2.4 Klasifikasi..... 11

2.2.5 Multilayer Perceptron (MLP)..... 11

2.2.6 Term Frequancy-Inverse Document Frequency (TF-IDF)..... 14

2.2.7 Evalution Measure..... 15

2.2.8 Text Mining..... 16

2.2.9 Pre-processing Text..... 17
2.2.10 RapidMiner..... 17
2.2.11 Python..... 18
2.2.12 Jupyter Notebook..... 18
2.2.13 Flowchart..... 19
BAB III PENDAHULUAN..... 20
3.1 Metode Penelitian..... 20
3.1.1 Metode Peneltian..... 20
3.1.2 Spesifikasi Perangkat Keras (Hardware)..... 21
3.1.3 Spesifikasi Perangkat Lunak (Software)..... 21
3.2 Perancangan Sistem..... 22
3.2.1 Gambaran Umum..... 22
3.2.2 Proses Dataset..... 23
3.2.3 Preprocessing..... 24
3.2.4 Term Frequency-Inverse Document Frequency (TF_IDF)..... 26
3.2.5 Multinominal Multi-layer Perceptron..... 29
3.2.6 K-Fold Cross Validation..... 34
BAB IV HASIL DAN PEMBAHASAN.....36
5.1 Model Analisa Sentimen..... 36
5.1.1 Pelatihan..... 36
5.1.2 Pengujian Dengan Data Uji..... 37
5.2 Evaluation Measure..... 38
5.2.1 Accuracy..... 39

5.3 Antar Muka Pada Analisis Sentimen..... 40
5.4 Penguji Kepada User..... 41
5.4.1 Implementasi User..... 41
5.4.2 Kesimpulan Implementasi Sistem..... 43
BAB V KESIPULAN DAN SARAN..... 45
5.1 Kesimpulan..... 45
5.2 Saran..... 45
DAFTAR PUSTAKA..... 46

DAFTAR GAMBAR

Gambar 2.1 3-fold validation..... 10

Gambar 2.2 Arsitektur Multi Layer Perceptron..... 12

Gambar 3.1 Gambaran Umum Analisis Sentimen..... 22

Gambar 3.2 Diagram Alir *Dataset*..... 23

Gambar 3.3 MLP..... 30

Gambar 3.4 Pembagian Data k-fold..... 34

Gambar 4.1 Tampilan Data Latih..... 37

Gambar 4.2 Tampilan *Form* Input Data Uji..... 37

Gambar 4.1 Tampilan Ketepatan Prediksi..... 38

Gambar 4.2 Perhitungan *Accuracy* Pada Sistem..... 40

Gambar 4.5 Form *Input* User..... 41

Gambar 4.6 Form *Output* User..... 41

Gambar 4.7 Grafik Hasil Kuisioner..... 42

DAFTAR TABEL

Tabel 2.1 Simbol *Flowchart*..... 19

Tebal 3.1 Contoh Cleansing..... 24

Tebal 3.2 Contoh Tokenisasi..... 25

Tebal 3.3 Contoh Case Folding..... 25

Tebal 3.4 Contoh stopword..... 26

Tabel 3.5 Contoh Dokumen..... 26

Tabel 3.6 TF-IDF Doc A..... 27

Tabel 3.7 TF-IDF Doc B..... 28

Tabel 3. 8 TF-IDF Doc C..... 28

Tabel 3.9 Rincian pembagian data..... 34

Tabel 3.10 hasil uji coba..... 34

Tabel 4.1 Tabel *Confusion Matrix*..... 39

Tabel 4.2 Hasil Nilai Persentase Tiap Pertanyaan Kuisioner..... 43

ABSTRAK

Twitter adalah salah satu sosial media yang penggunanya paling banyak menulis berbagai opini, komentar serta berita yang terupdate. Banyak pengguna yang menggunakan aplikasi twitter untuk posting status atau menulis pendapat mereka tentang suatu berita atau informasi. Demikian hal ini dilakukan pemanfaatan data dan digunakan sebagai sumber data untuk menilai sentimen analisis pada twitter, salah satu analisis animo masyarakat tentang berita pemilihan presiden yang berujung ke mahkamah konstitusi pada tahun 2019. Teknik yang diterapkan yaitu klasifikasi, untuk mengklasifikasikan data pada tweet, sebelum itu lakukan tahap preprocessing dan pembobotan term requancy pada data tweet. Dari pembobotan ini menghasilkan data training dan data testing yang akan dilakukan klasifikasi menggunakan Algoritma Multilayer Perceptron, pendekatan Multilayer Perceptron itu sendiri untuk mendapatkan hasil Sentimen Analisis Anemo masyarakat terhadap pilpres 2019 yang berujuk kemahkamah konstitusi dengan kategori positif atau negatif . Dari hasil pengujian ini mendapat akursi yang cukup tinggi, sebesar 89% dengan jumlah data 300.

Kata kunci : analisis sentiment pada tweet menggunakan metode multilayer perceptron

ABSTRACT

Twitter is one of the social media whose users write the most various opinions, comments and updated news. Many users use the Twitter application to post status or write their opinion about a news or information. Thus this is done by utilizing data and being used as a data source to assess the sentiment analysis on Twitter, one of the public animo analyzes about the news of the presidential election which led to the constitutional court in 2019. The technique applied is classification, to classify data on tweets, before that. perform the preprocessing stage and term requancy weighting on the tweet data. From this weighting, training data and testing data will be classified using the Multilayer Perceptron Algorithm, the Multilayer Perceptron approach itself to get the results of the Public Anemo Analysis Sentiment towards the 2019 presidential election which refers to the constitutional court with a positive or negative category. From the results of this test, the accuracy is quite high, amounting to 89% with a total of 300 data.

Keyword : Sentiment analysis on tweets using the multilayer perceptron method

BAB I
PENDAHULUAN

1.1 Latar Belakang

Pemilihan presiden dan wakil presiden 2019 telah selesai di selenggarakan pada hari Rabu 17 April 2019 yang di ikuti oleh seluruh masyarakat indonesia, dari hasil perhitungan suara pasangan Joko Widodo-Ma'ruf Amin terpilih menjadi presiden dan wakil presiden 2019 mengalahkan pasangan Prabowo Subianto-Sandiaga Uno. Menurut Badan Pemenangan Nasional Prabowo-Sandi, terdapat sejumlah pelanggaran Pilpres 2019 yang sistematis, terstruktur, dan masif meliputi penyalahgunaan APBN atau Program Kerja Pemerintah, ketidak netralan aparatur negara, penyalahgunaan birokrasi dan BUMN, pembatasan kebebasan media, pers, diskriminasi perlakuan, dan penyalahgunaan penegakan hukum. Disebutkan juga dalam permohonan adanya kekacauan Sistem Informasi Penghitungan Suara (situng) KPU dalam kaitannya dengan daftar pemilihan tetap, seperti banyaknya kesalahan input data pada Situng yang mengakibatkan terjadinya ketidak sesuaian data informasi dengan data yang terdapat pada C1 yang dipindai KPU sendiri di 34 provinsi.

BPN Prabowo-Sandi sudah menemukan perhitungan suara yang tidak sesuai dengan jumlah suara DPT/DPTb/DPK dan kesalahan pada data C1. Prabowo Sandi menilai KPU tidak teliti, tidak profesional serta memiliki Aplikasi sistem Perhitungan yang belum sempurna, dan terdapat kejanggalan yang lain pada data C1. Selain itu, BPN Prabowo-Sandi juga menyampaikan perolehan suara berbanding terbalik dengan KPU. Perolehan suara yang diajukan dalam

permohonan pasangan Jokowi-Ma'ruf mendapat suara 63.573.169 atau 48% sedangkan Prabowo-Sandiaga Uno mendapat total suara 68.650.239 atau 52%.

Tim Hukum Badan Pemenang Nasional (BPN) pasangan calon presiden dan wakil presiden nomor urut 02 Prabowo Subianto-Sandiaga Uno telah resmi mengajukan gugatan sengketa perselisihan hasil pemilihan umum (PHPU) 2019 ke Mahkamah Konstitusi, jumat malam (24/5/2019). Tim hukum Prabowo-Sandiaga mencoba membuktikan bahwa penyelenggara Pilpres 2019 melakukan kecurangan yang terstruktur, sistematis dan masif. Hal ini di ukur dari penyalahgunaan APBN, ketidak netralan petugas, dalam birokrasi, pembaatasan media dan diskriminasi dalam penyalagunaan penegakkan hukum. Mahkamah Konstitusi jumat 14 Juni 2019 menggelar sidang pertama pemilihan presiden dan wakil presiden 2019 yang diajukan oleh pasangan Prabowo-Sandiaga Uno. Sidang pertama dipimpin oleh majelis hakim beranggota sembilan orang dan di ketuai oleh Anwar Usman. Materi gugatan dibacakan secara bergantian oleh mantan wakil Mentri hukum dan Hak Asasi manusia Denny Indrayana, mantan ketua Yayasan Lembaga Hukum Indonesia (YBHI) Bambang Widjojanto, dan Pengacara Tengku Nasrullah.

Proses sidang diadakan beberapa kali hingga mendapatkan hasil keputusan, selama masa persidangan perang antar pendukung tidak juga surut. Kekalahan Prabowo Subianto-Sandiaga Uno dalam pemilihan presiden 2019, tak sejalan dengan keinginan pendukungnya. Mereka tetap menilai jagoannya lebih layak memimpin Indonesia lima tahun ke depan, dari pada pasangan Jokowi-Ma'ruf Amin Pemberitaan mengenai sengketa pemilihan umum Presiden Indonesia 2019

menjadi topik hangat dimasyarakat dan kalangan pengguna sosial media, simpang siur pemberitaan dan argumen masyarat dalam menunggu hasil sidang tertuang dalam cuitan-cuitan yang terangkum dalam bentuk tagar #mahkamahkonstitusi.

Sosial media menjadi wadah populer dalam menampung aspirasi masyarakat terutama dalam menyikapi permasalahan sengketa pemilihan umum presiden 2019 ini, dan sosial media yang cukup di gemari sebagai wadahnya adalah sosial media *twitter* yang terangkum melalui tagar (#) baik bernilai positif maupun bernilai negatif.

Dari beragam cuitan pada tagar tersebut akan dapat menjadi survei dalam bentuk analisa animo masyarakat terhadap kasus sengketa pemilihan umum Presiden Indonesia 2019, maka dari itu untuk melihat antusias dikalangan masyarakat terhadap kasus sengketa pemilihan umum Presiden Indonesia 2019 di nilai perlunya sebuah survei dalam bentuk sentimen analisis yang mampu menganalisa animo masyarakat terhadap gugatan sengketa perselisihan hasil pemilihan umum (PHPU) 2019 ke mahkamah konstitusim (MK) berdasarkan tagar media sosial *twitter*. Penulis tertarik untuk mengangkat topik tersebut dalam tugas akhir dengan judul **"Analisis Sentiment Pada Tweet Dengan Tagar #Mahkamahkonstitusi Menggunakan metode Multi layer Perceptron".**

1.2 Identifikasi Masalah

Indentifikasi masalah yang dapat diambil dari latar belakang yaitu: "memerlukan survei dan analisa mengenai pendapat masyarakat yang bersifat positif maupun negatif dan menganalisa opini yang berkembang di kalangan masyarakat dalam kasus Gugatan Sengketa Perselisihan Hasil Pemilihan Umum (PHPU) 2019 ke Mahkamah Konstitusi (MK)", di media sosial twitter.

1.3 Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini adalah bagaimana mengamati dan menganalisa opini pengguna twitter mengenai kasus gugatan sengketa perselisihan hasil pemilihan umum (PHPU) 2019 ke Mahkamah Konstitusi (MK) dan mengklasifikasikan sentimen data tweet tersebut.

1.4 Batas Masalah

Mengingat keterbatasan kemampuan penulis, maka batasan masalah dalam penelitian ini dibatasi, sebagai berikut :

1. Sosial media yang digunakan adalah sosial media twitter
2. Penelitian menggunakan data yang diperoleh dari trending topik yang ada di kalangan masyarakat pengguna sosial media twitter
3. Tagar yang digunakan adalah #mahkamahkonstitusi
4. Hanya menampilkan sentimen positif dan negatif.
5. Data yang digunakan 300 data
6. Data yang dimasukkan dengan data csv

1.5 Tujuan Penelitian

Tujuan dari penelitian ini untuk menganalisa sentiemen pengguna twitter terhadap kasus gugatan sengketa perselisihan hasil pemilihan umum (PHPU) 2019 ke Mahkamah Konstitusi (MK) berdasarkan tagar media sosial *twitter* menggunakan metode *multi layer perceptron.*

1.6 Manfaat Penelitian

Adapun manfaat penelitian ini "memberikan manfaat baik kepada penulis maupun pembaca tentang gambaran sentimen para pengguna Twitter terhadap kasus gugatan sengketa perselisihan hasil pemilihan umum (PHPU) 2019 ke Mahkamah Konstitusi (MK) yang trending dibicarakan oleh masyarakat".

BAB II
LANDASAN TEORI

2.1 Studi Kepustakaan

Studi pustaka ini digunakan sebagai pembanding antara penelitian yang sudah dilakukan dan yang akan dilakukan peneliti. Penelitian pertama yang dilakukan oleh (Petrix Nomleni 2015) dalam judul "Sentiment Analysis menggunakan Support Vector Machine (SVM) "pada penelitian ini dibahas klasifikasi keluhan masyarakat terhadap pemerintah pada media sosial *Facebook* dan *twitter* sapa warga dan berbahasa Indonesia menggunakan metode *Support Vector Machine (SVM)* yang di jalankan dalam komputasi terdistribusi dengan menggunakan *hadoop.* Pengujian dilakukan dengan perhitungan *precision, recall, F1-Measure* dan accuracy. akurasi yang dihasilkan rata-rata diatas 80% dengan akurasi tertinggi 84.4086% precision 81% serta F-Measure 80 %.

Selanjutnya penelitian kedua yang dilakukan oleh (Grace Tika, dkk 2019) yang berjudul "Klasifikasi Topik Berita Berbahasa Indonesia Menggunakan Multi layer Perceptron", pada penelitian ini dikhususkan untuk melakukan analisis sentimen berita ekonimi. Pada penelitian ini membahas klasifikasi terhadap berita agar berita, dan di kelompokkan berdasarkan kategori seperti teknologi, budaya, pendidikan dan lain-lainnya. Implementasi metode jaringan syaraf tiruan ini dapat menciptakan suatu pola pengetahuan dengan kemampuan belajar (*self organizing*) secara optimasi, sehingga metode jaringan syaraf tiruan dinyatakan dalam *F1-Measure Micro-average* dengan nilai performansi mencapai 77.44 % .

Selanjut penelitian ketiga yang dilakukan oleh (Anita Novantirani,dkk 2015) tentang "Analisis Sentimen Pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota Dengan Metode *Support Vector Machine*". Pada penelitian ini membahas pendapat dan opini masyarakat mengenai penggunaan transportasi umum dalam kota melalui *twitter*. Analisis dimanfaatkan untuk mengetahui penilaian pelayanan transportasi umum darat dalam kota apakah positif atau negatif. Hasil analisis tersebut dapat membantu dalam penilaian dan evaluasi terhadap penggunaan transportasi umum darat dalam kota. Dengan dilakukannya peningkatan fasilitas dan pelayanan berdasarkan hasil analisis sentimen. Analisis sentimen dari hasil pengujian pada penelitian ini didapatkan bahwa SVM dapat diimplementasikan dengan nilai akurasi mencapai 78,12%. Variabel berpengaruh terhadap akurasi adalah jumlah data, perbandingan jumlah data latih dan uji, serta perbandingan jumlah data positif dan negatif yang digunakan.

Berdasarkan hasil ketiga penelitian diatas dapat dikemukakan sistem monitoring yang dibuat memiliki perbedaan studi kasus, maupun metode yang digunakan. Hasil penelitian diatas dapat dijadikan rujukan atau referensi dalam penelitian tentang analisis anemo masyarakat terhadap isu kecurangan paska pilpres 2019 berdasarkan sosial media twitter.

2.2 Dasar Teori

2.2.1 Analisi Sentimen

Analisi Sentimen secara umum dibagi menjadi informasi opini dan fakta. Fakta adalah ekspresi obyektif terhadap suatu benda, kejadian atau kepunyaan benda tersebut. Sedangkan opini biasanya berupa ekspresi subyektif yang menggambarkan atau menjelaskan penilaian atau perasaan seseorang terhadap suatu benda, atau kepunyaan dari benda tersebut. Analisis Sentimen merupakan bagian dari pekerjaan yang meninjau sesuatu yang berhubungan dengan pendapat komputasi, sentimen dan subjektivitas teks.

Analisis Sentimen merupakan alat untuk memproses koleksi hasil pencarian yang tujuanya untuk mencari atribut suatu produk (kualitas, fitur) dan prosesnya memperoleh hasil pendapat. Tugas dasar dalam analisis sentimen yaitu mengelompokkan polaritas dari teks yang ada pada dokumen yang bersifat positif maupun negatif. Sejak tahun 2003 penelitian analisis sentimen telah berkembang dan *text mining* merupakan bagian penelitian komputasi berdasarkan sentimen, emoticon, pendapat, komentar dan setiap ekspresi yang diungkap oleh teks.

Analisis sentimen lebih difokuskan untuk review klasifikasi berdasarkan polaritas. Berdasarkan klasifikasi, analisis sentimen dibagi jadi dua kelompok. Dokumen klasifikasi ke pendapat atau fakta, atau dikenal sebagai klasifikasi subjektivitas (*subjectivity classification*) dan dokumen klasifikasi ke dalam positif atau negatif, dikenal sebagai analisis sentimen. Hal ini adalah proses yang penting untuk menetukan dokumen yang memiliki opini atau fakta dan dokumen yang menghasilkan nilai positif atau negatif (Nurhuda & Shiwi , 2014).

2.2.2 Twitter

Twitter didirikan oleh Jack Dorsey pada 2006 dan diluncurkan didunia sosial pada juli 2006. *Twitter* mempunyai *application programming interface* dan *developer* bisa mengluaskan aplikasi sesuai dengan kebutuhannyan tiap-tiap pengguna. Di twitter setiap orang yang terdaftar dapat mengirim dan membaca teks pendek yang disebut tweet dan banyak orang mengomentari hal-hal.

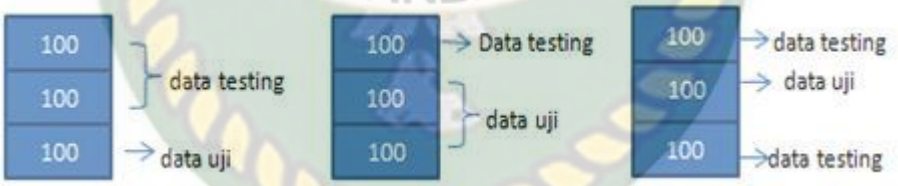
Tweet memiliki karakteristik yang unik berdasarkan laporan analisis GO baik pada panjang teks, ketersediaan data, model bahasa, dan domain. Panjang teks pada tweet maksimum 140 karakter dengan rata-rata tweet terdiri dari 14 kata atau 78 karakter. *Twitter* API dapat dikumpulkan begitu banyak tweet sebagai data set. Model bahasa pada *tweet* tidak terstruktur bahkan terkadang ada kesalahan pengetikan. Serta domain pembahasan pada twitter sangat luas terhadap banyak sekali topik apapun, berbeda dengan penelitian lain pada domain terbatas seperti review movie. (Assuja & Saniati, 2016).

2.2.3 K-fold Cross Validation

Cross-validation juga dapat disebut dengan estimasi rotasi yaitu sebuah teknik validasi model untuk menilai hasil statistik analisis akan menggeneralisasi data independen. Teknik ini biasa digunakan untuk melakukan prediksi dan memperkirakan seberapa akuratnya prediktif ketika dijalankan dalam praktek. Salah satu teknik validasi yaitu *k-fold cross validation*, yaitu dengan memecah data dengan ukuran yang sama. Penggunaan *k-fold* untuk menghilangkan bias pada data. Pelatihan dan pengujian dilakukan sebanyak k kali. Pada percobaan pertama, subset S1 diperlakukan sebagai data pengujian dan subset lainnya

sebagai data pelatihan, pada percobaan kedua subset S1,S3,…Sk menjadi data pelatihan dan S2 menjadi data pengujian, dan setererusnya (Tempola et al., 2018).

Validation merupakan salah satu teknik menilai atau memvalidasikan keakuratan sebuah model yang dibangun berdasarkan dataset tertentu. Pembuatan model biasanya bertujuan untuk melakukan prediksi maupun klasifikasi terhadap suatu data. Data yang di gunakan dalam proses pembangunan model disebut data training, sedangkan data yang akan digunakan untuk memvalidasi model disebut sebagai data uji. Salah satu metode *Cross Validation* yang populer adalah *K-fold validation*, dalam teknik ini data dibagi menjadi sejumlah K-buah secara acak. Kemudian dilakukan sejumlah K-kali eksperimen, dimana masing-masing eksperimen menggunakan data partisi ke-K sebagai data testing dan sisa data lainnya sebagai data training. Sebagai gambaran jika kita melakukan 3-fold maka desain data eksperimenya sebagai berikut: (Hariri & Pamungkas, 2016)



Gambar 2.1 3-fold validation

Untuk mendapatkan total dari keseluruhan Perhitungan akurasi, menggunakan persamaan,*k-fold cross validation* sebagai berikut :(Assiva et al., 2019):

a. Total data dibagi menjadi beberapa bagian.
b. *Fold* ke-1 adalah ketika bagian ke-1 jadi data uji dan sisanya jadi data training. Lanjut menghitung akurasi berdasarkan data tersebut. Perhitungan total akurasi menggunakan rumus sebagai berikut :

$$akurasi = \frac{\sum data\ uji\ benar\ klasifikasi}{\sum \text{total data uji}} \times 100\% \quad (21)$$

c. *Fold* ke-2 ketika bagian ke-2 jadi data uji dan sisanya jadi data training. Lanjut hitung akurasi berdasarkan data tersebut.

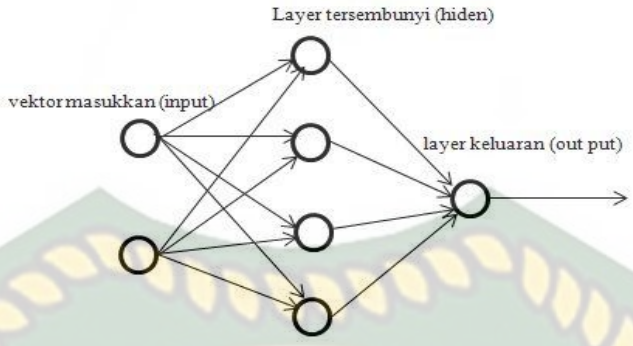
d. Dan seterusnya hingga selesai di fold K. Lalu hitung rata-rata akurasi dari K buah akurasi diatas. Rata-rata akurasi akhir akan menjadi akurasi final. Validasi menggunakan *3 Fold cross Validation* dimana data di bagi menjadi 3 bagian dengan jumlah data yang sama.

2.2.4 Klasifikasi

Klasifikasi disini yaitu untuk menentukan sebuah kalimat sebagai kelompok positif atau negatif berdasarkan nilai perhitunganya. Teknik *klasifikasi* yang akan digunakan untuk penelitian ini adalah *multilayer perceptron*. Hasil klasifikasi teks dalam bentuk kelas positif atau kelas negatif. Pengukuran akurasi berdasarkan *multilayer perceptron* sebelum atau sesudah pembobotan (Assiva et al., 2019).

2.2.5 Multilayer Perceptron

Multilayer perceptron merupakan ANN atau turunan perceptron, umpan balik (*feedforward*) terdiri dari sejumlah neuron yang dihubungkan oleh bobot-bobot penghubung dengan satu layer tersembunyi (*hiddenlayer*) atau lebih. Biasanya, jaringan terdapat satu layer masukan, satu layer neuron tersembunyi ditengah, dan sebuah layer neuron komputasi keluaran. Sinyal masukkan ditambah dengan arah maju pada layer perlayer. Contoh gambar MLP sebagai berikut ini (Syadid, 2019).



Gambar 2.2 Multi Layer Perceptron

gambar di atas terlihat ada satu layer masukkan (input), satu layer tersembunyi (hiden layer), satu layer keluar (output). Sebenarnya ada satu layer lagi didalam MLP, yaitu layer masukkan atau vektor masukkan (bias). Tapi dalam layer masukkan ini tidak perhitungan, yang dilakukan hanya menuju ke sinyal atau vektor masukkan yang diterima ke layer di depannya di sebut bias.

Tiap-tiap layer MLP mempunyai fungsi khusus. Layer masukkan berfungsi menerima sinyal masukkan dari luar dan mendistribusikan kesemua neuron dalam layer tersembunyi. Layer keluaran menerima sinyal keluaran atau dengan kata lain, stimulus pola dari layer tersembunyi dan memunculkan sinyal nilai atau kelas keluaran dari keseluruhan jaringan.

Neuron dalam layer tersembunyi mendeteksi fitur-fitur yang tersembunyi. Bobot dari neuron dalam layer tersembunyi untuk merepresentasikan fitur-fitur tersembunyi dalam vektor masukkan. Fitur tersembunyi ini kemudian digunakan oleh layer keluaran untuk menentukan pola atau kelas keluaran. Dan satu layer tersembunyi dapat merepresentasikan fungsi kontiniu dari sinyal masukkan, dengan dua layer tersembunyi, fungsi diskontiniu pun dapat dihasilkan. Hidden layer “menyembunyikan” keluaran yang diinginkan.

Neuron dalam layer tersembunyi tidak dapat diamati dari prilaku masukkan atau keluaran jaringan secara seluruhnya. Dan juga tidak ada cara yang jelas untuk mengetahui keluara yang diinginkan oleh layer tersebut. Dengan kata lain, keluaran yang diinginkan oleh layer ditentukan oleh layer itu sendiri. Dibawah ini algoritma Multilayer Perceptron (Muliantara & widiartha, 2011):

1. Inisialisasikan nilai bobot dengan nilai acak kecil.
2. Jika kondisi belum dipenuhi, lakukan langkah 2-8.
3. Untuk setiap pasangan pelatihan, lakukan langkah 3-8.
4. Tiap unit masukkan menerima sinyal dan meneruskan ke unit tersembunyi.
5. Hitung semua keluaran di unit tersembunyi z_j (j= 1,2...,p).

$$z\_net_j = v_{j0} + \sum_{i=1}^{n} x_i \quad \text{.....(2.2)}$$

$$z_i = (z\_net_j) = \frac{1}{1 + e^{-z\_net_j}} \quad \text{.....(2.3)}$$

6. Hitung keluaran jaringan di unit keluaran y_k (k=1,2...,m).

$$y\_net_k = w_{k0} + \sum_{j=1}^{p} z_j w_{kj} \quad \text{.....(2.4)}$$

$$y_k = (y_{net_k}) = \frac{1}{1 + e^{-z\_net_j}} \quad \text{.....(2.5)}$$

7. Hitung faktor δ unit keluaran berdasarkan kesalahan di setiap unit keluaran y_k (k = 1, 2,..., m).

$$\delta = (t_k - y_k)'(y\_net_k) = (t_k - y_k)y_k(1 - y_k) \quad \text{.....(2.6)}$$

$$t_k = target$$

δk merupakan unit kesalahan yang akan dipakai dalam perubahan bobot layer dibawahnya. Hitung perubahan bobot w_{kj} dengan laju pemahaman α.

$\Delta w_{kj} = \alpha\delta_k Z_j,$ k= 1,2,.., m; j=0,1,... p .....(2.7)

8. Hitung faktor δ unit tersembunyi berdasarkan kesalahan di setiap unit tersembunyi z_j (j = 1).

$$\delta\_net_j = \sum_{k=1}^{m} \delta_k w_{kj} \quad \text{.....(2.8)}$$

Faktor δ unit tersembunyi.

$$\delta_j = \delta\_net_j f'(Z\_net_j) = \delta\_net_j z_j(1 - z_j) \quad \text{.....(2.9)}$$

Hitung suku perubahan bobot v_{ji}

$$\Delta v_{ji} = \alpha\delta_j, \quad j = 1,2,\dots,p\,; 1 = 1,2,\dots,n \quad \text{.....(3.0)}$$

9. Hitung semua perubahan bobot. Perubahan bobot garis yang menuju ke unit keluaran, yaitu :

$w_{kj}\,(baru) = w_{kj}\,(lama) + \Delta w_{kj}$, ($k$ = 12,..., m ; j = 0,1....,p).....(3.1)

Perubahan bobot garis yang menuju ke unit tersembunyi, yaitu :

$V_{ji}(baru) = (lama) + \Delta v$, $(j = 1,2,\dots,p; i = 0,1,\dots,n)$ .....(3.2)

2.2.6 Trem Frequancy-Inverse Document Frequency (TF – IDF)

Metode *Term frequency-iverse document frequency* adalah cara pemberian bobot hubungan suatu kata (term) terhadap dokumen. Untuk dokumen tunggal tiap kalimat dianggap sebagai dokumen, metode ini menggabungkan dua konsep untuk perhitungan bobot *Trem frequency* (DF) adalah banyak kalimat diamana suatu kata (t) muncul.

Frekuensi kemunculan kata didalam dokumen yang dimasukkan menunjukkan seberapa penting kata itu didalam dokumen tersebut. Frekuensi dokumen yang mengandung kata tersebut menunjukkan berapa umum kata tersebut. Bobot kata

semakin sering muncul di dalam satu dokumen semakin besar, dan semakin kecil jika muncul dalam banyak dokumen. pada algoritma tf-idf digunakan untuk menghitung bobot (w) masing-masing dokumen terhadap kata kunci dengan persamaan (Nurjannah & Astuti, 2013).

$$tf = 0.5 + 0.5 * \frac{ft,d}{\max(ft,d)} \quad \text{.....(3.3)}$$

$$idf = \log\left(\frac{N}{dft}\right) \quad \text{.....(3.4)}$$

$$W = tf * idf \quad \text{.....(3.5)}$$

Keterangan :

d : dokumen

t : kata pada dokumen

ft.d : frekuensi kata pada d

tf : banyaknya kata i pada sebuah dokumen

N : total jumlah dokumen

dft : banyak dokumen yang mengandung kata i

idf : Inversed Dokumen Frequency

W : bobot dokumen ke-d terhadap kata ke-t

2.2.7 Evalution Measure

Evaluation Measure Pengukuran performa dari model yang telah dibangun diukur menggunakan nilai yang diperoleh dari hasil prediksi yang sudah didapat. Evaluasi yang digunakan adalah akurasi. Akurasi untuk memprediksi postif dan negatif. (Syah et al, 2017).

Rumus perhitungan evaluasinya adalah sebagai berikut (Syah et al., 2017):

$$Accuracy = \frac{TP+TN}{(TP+FP+FN+TN)} \quad \text{.....(3.6)}$$

Keterangan :

True Positif (TP) = suatu kondisi dimana sistem mendektesi kelas positif dan fakta pun positif

True Negative (TN) = suatu kondisi dimana sistem mendekteksi kelas negatif dan faktanya pun negatif

False Positif (FP) = suatu kondisi dimana sistem mendekteksi kelas positif namun faktanya negatif

False Negative (FN) = suatu kondisi dimana sistem mendekteksi kelas negatif namun faktanya positif

2.2.8 Text Mining

Text mining merupakan istilah yang mengacu pada teknik penambangan data untuk menganalisa dan memproses data yang tidak terstruktur dan semu terstruktur. *Text mining* memiliki proses yang sama dengan *data mining* tetapi memiliki inputan yang berbeda. *Text Mining* memiliki proses yang sama dengan *data mining* tetapi memiliki inputan yang berbeda. Pada *text mining* pertama diperlukan pengambilan data kemudian data tersebut di *pre-processing* sebelum proses klasifikas dengan penambahan beberapa fitur dan penghilangan beberapa diantaranya,dan penyisipan *subsequent* kedalam database, menentukan pola dalam data terstruktur, mengevaluasi dan menginterpretasi out put, (Lidya et al.,2015)

2.2.9 Pre-processing Text

Preprocessing merupakan proses awal sebelum melakukan klasifikasi dokumen. Ada beberapa tahap menyiapkan dokumen teks yang tidak disusun menjadi data yang disusun yang siap digunakan untuk proses pembersian noise dan siap digunakan pada tahap *text processing* berikut ini (Syadid, 2019)

a. Cleansing proses pembersihan kalimat dari hashtag, mention, dan juga tanda baca.
b. Tokenisasi proses perubahan kalimat menjadi kata.
c. Case Folding proses penggantian huruf dari huruf yang bercampur (*lowercase* dan *uppercase*) menjadi semua huruf kecil.
d. Stopword penghapus kata yang tidak penting.
e. Stemming proses mencari kata dasar dari kata hasil stopword.

2.2.10 Repid Miner

Rapid Miner adalah platform perangkat lunak ilmu data yang dikembangkan oleh perusahaan bernama sama dengan yang menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam, penambangan teks, seperti *data mining*, *text mining* dan analisis prediktif. Repidminer menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik.

Repidminer memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. Repidminer merupakan perangkat lunakyang berdiri sendiri untuk analisis data dan menggunakan bahasa java sehingga bisa berjalan disemua sistem. (Köksal & Penez,2015).

2.2.11 Python

Python adalah bahasa pemograman interpretatif yang dianggap mudah dipelajari serta berfokus pada keterbacaan kode. Bahasa python ini muncul pertama kali pada tahun 1991, dan dirancang oleh Guido Van Rossum. Sampai saat ini python masih di kembangkan oleh *python software foundation*. Bahasa *python* mendukung hampir semua sistem operasi, bahkan untuk sistem operasi *linux*. Selain itu python juga produk yang *opensource* juga multiplatform Dengan kata lain, *python* adalah bahasa pemograman yang memiliki kode pemograman yang jelas, lengkap dan mudah dipahami.

Secara umum python berbentuk pemograman berorientasi objek, pemograman imperatif dan pemograman fungsional. Python juga digunakan untuk berbagai keperluan pembangunan perangkat lunak dan dapat berjalan semua platfrorm sistem operasi (Syadid, 2019):

2.2.12 Jupyter Notebook

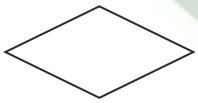
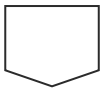
Jupyter Notebook adalah (file yang berekstensi *ipynb*) adalah dokumen yang dihasilkan oleh *Jupyter Notebook App* yang berisikan kode komputer dan *rich text element* seperti paragraf, persamaan matematik, gambar dan tautan (*links*). *Jupyter Notebook* sebelumnya dikenal sebagai *IPython Notebook* yang berbasis bahasa pemograman Python. *IPython Notebook* di ciptakan oleh perez dan granger, *ipython notebook* berbasis web browser yang telah banyak digunakan dibidang pendidikan dan penelitia. Aplikasi ini belum tersosialisasikan dengan baik di indonesia. perkembangan teknologi informasi dan komputer yang sangat cepat,

Dukungan yang masif dari komunitas, maka secara alami dalam waktu dekat *Jupyter Notebook* akan berevolusi menjadi *JupyterLab* (Setiabudidaya, 2018).

2.2.13 Flowchart

Flowchart adalah penyajian yang sistematis tentang proses dan logika dari penggambaran langkah-langkah dan urut-urutan prosedur suatu program, dapat di lihat pada tabel 2.1 berikut ini: (Anharku, 2009).

Tabel 2.1 Simbol *Flowchart*

No.	Simbol	Nama	Fungsi
1		Terminator	Permulaan atau akhir program.
2		Proses	Proses pengolahan data.
3		Garis Alir	Arah aliran program.
4		Preparation	Proses inisialisasi.
5		Input atau output	Proses data input atau output,parameter, informasi
6		Sub program	Proses menjalankan sub program atau permulaan
7		decision	Penyeleksian data, perbandingan pertanyaan selanjutnya
8		Off page connector	Penghubung bagian flowchart pada halaman berbeda.

BAB III
METODOLOGI PENELITAN

3.1 Metodologi Penelitian

3.1.1 Metode Penelitian

Metode penelitian merupakan tahapan-tahapan yang dilalui oleh peneliti untuk memperoleh gambaran yang jelas mengenai penelitiaan. Tahapan yang dilalui dalam penelitian ini adalah sebagai berikut ini:

1. Pengumpulan Data

   Data yang dikumpul yaitu data dari *twitter*, *tweet* yang diangkat dengan topik #mahkamahkonstitusi. Data tersebut didapat dengan cara melakukan pencarian data di twitter melalui aplikasi *repidminer*, aplikasi yang bisa terhubung dengan *twitter* untuk mengambil data sesuai dengan tagar atau topik yang kita angkat.

2. Studi literatur

   Studi literatur yang dilakukan yaitu mengumpulkan dan mempelajari segala informasi yang berhubungan dengan Analisis. Analisis Sentimen pada twitter menggunakan metode *multilayer perceptron.*

3. Perancangan

   dilakukan perancangan terhadap perangkat lunak yang akan dibuat berdasarkan study literatur yang ada. Perancangan ini meliputi desain struktur data, desain aliran informasi, desain alogirtma dan pemograman.

4. Tahap Implementasi

Tahapan implementasi dilakukan secara bertahap dengan studi literatur. Perancangan tersebut akan di implementasikan pada bahasa pemograman yang telah disepakati.

5. Pengujian dan evaluasi

Tahap ini dilakukan uji coba untuk mencari permasalahan yang mungkin terjadi, mengevaluasi jalan sistem dan melakukan perbaikan apabila dibutuhkan.

6. Penyususan laporan penelitian

Penyusanan laporan dilakukan pada tahap akhir untuk dokumentasi.

3.1.2 Spesifikasi Perangkat Keras (Hardware)

spesifikasi komputer digunakan untuk perancangan dengan *hardware* sebagai berikut ini:

1. Computer name : compaq-PC
2. Processor : Intel(R) core(TM) i3 CPU M 350 @2.27GHz 2.27GHz
3. RAM : 2,00 GB
4. System type : 32-bit operating System

3.1.3 Spesifikasi Perangkat Lunak (software)

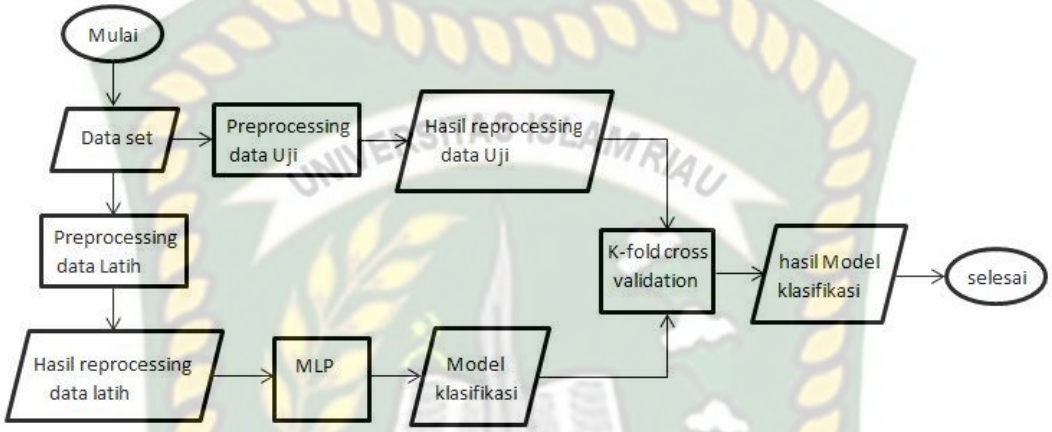
Perangkat lunak yang digunakan dalam pembuatan analisis sentimen pada *twitter* adalah sebagai berikut ini :

1. Sistem Operasi : Windows 7 Home
2. Bahasa Pemograman : Python dan Jupyter Notebook
3. Desain Logika Program : visio 2013

3.2 Perancangan Sistem

3.2.1 Gambaran Umum

Analisis sentimen yang akan dibangun dapat digambarkan secara detail melalui perancangan sistem yang bisa dilihat pada gambar 3.1 dibawah ini :



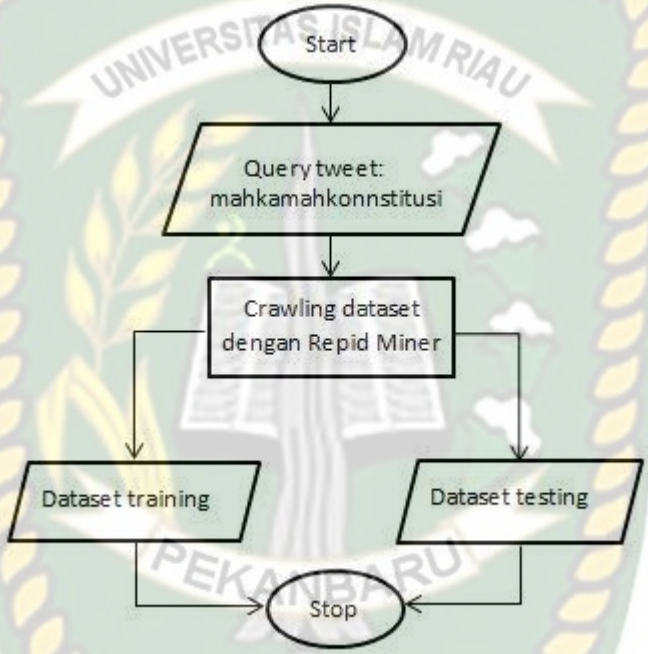
Gambar 3.1 Gambaran Umum Analisis Sentimen

Berikut penjelasan dari gambar 3.1 ini:

1. Mulai mengambil data dari twitter menggunakan rapid miner dengan tagar #mahkamahkonstitusi. Data yang diambil berupa data latih dan data uji
2. Data set dipecah menjadi *Preprocessing* data latih dan *Preprocessing* data uji.
3. Setelah preprocessing didapatkan hasil data latih dan data uji, yang kemudian akan diterapkan metode *multi layer perceptron.*
4. Setelah menerapkan metode *multi layer perceptron*, didalam metode tersebut akan melakukan model klasifikasi dan melakukan perbandingan dengan *k-fold cross validation*
5. Begitu juga hasil data preprocessing uji diproses dalam model klasifikasi dan melakukan perbandingan *k-fold cross validation..*
6. Tahap selanjutnya adalah hasil evaluasi klasifikasi

3.2.2 Proses Dataset

Dataset berupa teks berbahasa indonesia yang diambil dari twitter. Data yang diambil untuk penelitian ini menggunakan query ‘mahkamahkonstitusi’. *Query* tersebut merupakan tagar yang sedang *trending topik* di twitter. Diagram alur proses pengambilan *dataset* seperti pada Gambar 3.2 dibawah ini .



Gambar 3.2 Diagram aliran dataset

Dataset dari hasil *crawling* ini akan dibagi menjadi dua bagian yaitu data *training* (data latih) dan data *testing* (data uji) yang dipresentasikan pada gambar 3.2. Data latih diklasifikan menggunakan *multilayer perceptron* dengan label negatif dan positif. Sedangkan data uji dikasifikasikan secara manual dengan label negatif dan positif. Data uji nantinya akan digunakan pada saat evaluasi untuk menentukan keakuratan data pada sistem menggunakan *k-fold corss validation*. pengambilan data *training* dan data *testing* dilakukan dengan waktu pengambilan yang berbeda.

3.2.3 Preprocessing

Proses *preprocessing* merupakan hal yang penting untuk mengurangi atribut yang tidak penting terhadap proses klasifikasi. Data yang dimasukkan pada tahap ini masih berupa data mentah, sehingga hasil dari proses ini adalah dokumen berkualitas yang harapannya mempermudah dalam proses klasifikasi. Berikut proses *preprocessing* yang tejadi:

a. Cleansing

Cleansing atau pembersihan untuk menghapus koma, titik, seluruh tanda baca, angka dalam tweet dan beberapa bagian khas yang ada pada tweet, yaknik *@username, URL,* simbol html dan hastag (#). Karna tidak memiliki pengaruh apapun dalam proses analisis sentimen, maka bagian tersebut akan dihilangkan dengan tujuan mengurangi noise. Proses cleansing di dilihat pada tabel 3.1 berikut ini.

Tebal 3.1 Contoh Cleansing

No	Sebelum	Sesudah
1	@shidayat_alqus Alhamdulillah Putusan 9 hakim MK untuk melegalkan kemenangan pasangan 01, semakin membuktikan bahwa KECURANGAN yang terjadi pada Pemilu 2019 ini Terstruktur,Sistematis dan Masih karena Mahkamah Konstitusi ikut berkontribusi dalam Melegalkan	Alhamdulillah Putusan hakim MK untuk melegalkan kemenangan pasangan semakin membuktikan bahwa KECURANGAN yang terjadi pada Pemilu ini Terstruktur Sistematis dan Masih karena Mahkamah Konstitusi ikut berkontribusi dalam Melegalkan

b. Tokenisasi

Tokenisasi yaitu memperoses bagian untuk memisahkan deretan kata dalam bentuk kalimat, paragraf dan halaman menjadi potongan kata tunggal *term* med

word. Bersamaan tokenisasi juga membuang kata yang tidak penting atau tanda baca. Proses tokenisasi dapat dilihat pada tabel 3.2 berikut ini :

Tebal 3.2 Contoh Tokenisasi

No	Sebelum	Sesudah
1	Alhamdulillah Putusan hakim MK untuk melegalkan kemenangan pasangan semakin membuktikan bahwa KECURANGAN yang terjadi pada Pemilu ini Terstruktur ,Sistematis dan Masih karena Mahkamah Konstitusi ikut berkontribusi dalam Melegalkan	‘Alhamdulillah’,‘Putusan’,‘ hakim’, ‘MK’,‘untuk’,‘melegalkan’, ‘kemenangan’,‘pasangan’,‘semakin’, ‘membuktikan’,‘bahwa’, ‘KECURANGAN’,‘yang’,‘terjadi’, ‘pada’,‘Pemilu’,‘ini’,‘Terstruktur’, ‘Sistematis’,‘dan’,‘Masih’,‘karena’, ‘Mahkamah’,‘Konstitusi’,‘ikut’, ‘berkontribusi’,‘dalam’,‘Melegalkan’,

c. Case folding

Case folding yaitu proses ubah seluruh ukuran huruf pada kata menjadi ukuran huruf yang sama, karena tidak semua tweet konsisten dalam penggunaan ukuran huruf. *Case folding* juga mengubah kata menjadi lower case atau huruf kecil. Proses case folding dapat dilihat pada tabel 3.3 berikut ini :

Tebal 3.3 Contoh Case Folding

No	Sebelum	Sesudah
2	Alhamdulillah Putusan hakim MK untuk melegalkan kemenangan pasangan semakin membuktikan bahwa KECURANGAN yang terjadi pada Pemilu iniTerstruktur, Sistematis dan Masih karena Mahkamah Konstitusi ikut	alhamdulillah putusan hakim mk untuk melegalkan kemenangan pasangan semakin membuktikan bahwa kecurangan yang terjadi pada pemilu ini terstruktur sistematis dan masih karena mahkamah konstitusi ikut berkontribusi dalam melegalkan

	berkontribusi dalam Melegalkan	

d. Stopword

Stopword yaitu proses penghapusan kata yang dianggap tidak penting dan tidak berpengaruh terhadap proses kategori. Contohnya kata penghubung seperti ’atau’,‘untuk’,’dan’,’ke’,’di’, dan seterusnya dan seterusnya. contoh penghilangan stopword dapat dilihat pada tabel 3.4 berikut ini:

Tebal 3.4 Contoh stopword

No	Sesudah	Sebelum
3	alhamdulillah putusan hakim mk untuk melegalkan kemenangan pasangan semakin membuktikan bahwa kecurangan yang terjadi pada pemilu ini terstruktur sistematis dan masih karena mahkamah konstitusi ikut berkontribusi dalam melegalkan	|alhamdulillah||putusan||hakim| |mk ||melegalkan|kemenangan| |pasangan||semakin||membuktikan| |bahwa||kecurangan||pemilu||terstruktur| |sistematis||mahkamah||konstitusi||ikut| |berkontribusi| |melegalkan|

3.2.4 Term Frequency – Inverse Document Frequency (TF-IDF)

Setelah melakukan *preprocessing* langkah selanjutnya adalah pembobotan kata menggunakan perhitungan tf-idf. Tf-idf yaitu cara beri bobot hubungan suatu kata *term* terhadap kata. kata tunggal setiap kalimat dianggap sebagai dokumen. Berikut contoh perhitungan tf-idf dengan dokumen A positif dan dokumen B dan C negatif dapat dilihat pada tabel 3.5 dibawah ini:

Tabel 3.5 Contoh Dokumen

Doc A	**Mahkama konstitusi Adil, Semua Dapat Penolakan**
Doc B	**Ending drama, ditolak mahkamah konstitusi**
Doc C	**Gagal dalam Inovasi Hukum Cacat**

1. Dari contoh dokumen pada tabel 3.2, hitung frekuensi kemunculan kata pada tiap dokumen sebagai langkah awal.
2. Setelah didapatkan frekuensi kata pada tiap dokumen, langkah selanjutnya adalah menghitung tf dengan rumus **tf = 0.5+0.5** $\frac{ft,d}{\max(ft,d)}$, misalkan pada doc A pada kata adil maka tf = 0.5+0.5 ($\frac{1}{1}$) = 1
3. Untuk mencari nilai **IDF** adalah dengan rumus **log (** $\frac{N}{dft}$**)** ,maka log 3/1 =0.477
4. Terakhir dengan mencari nilai hasil bobot setiap dokumen berdasarkan hasil dari tf-idf yaitu dengan **tf×idf**, maka 1*0.477 = 0.477. Berikut hasil dari perhitungan tf idf .

Tabel 3.6 TF-IDF Doc A

Kata	**Frekuensi Kemunculan Kata**			**TF (0.5+0.5 (ft,d/max(ftd)**	**IDF (log (** $\frac{N}{dft}$**)**	**W (TF*IDF)**
	A	**B**	**C**	**A**	**A**	**A**
Mahkamah	1	1	0	0,5	0,176	0,088
Konstitusi	1	1	0	0,5	0,176	0,088
Adil	1	0	0	1	0.477	0.477
Semua	1	0	0	1	0.477	0.477
Dapat	1	0	0	1	0.477	0.477
Penolakan	1	0	0	1	0.477	0.477
Ending	0	1	0	0	0,477	0
Drama	0	1	0	0	0,477	0
Ditolak	0	1	0	0	0,477	0
Gagal	0	0	1	0	0,477	0
Dalam	0	0	1	0	0,477	0
Inovasi	0	0	1	0	0,477	0

Hukum	0	0	1	0	0,477	0
Cacat	0	0	1	0	0,477	0

Tabel 3.7 TF-IDF Doc B

Kata	**Frekuensi Kemunculan Kata**			**TF (0.5+0.5 (ft,d/max(ftd)**	**IDF (log ($\frac{N}{dft}$)**	**W (TF*IDF)**
	A	**B**	**C**	**B**	**B**	**B**
Mahkamah	1	1	0	0,5	0,176	0,088
Konstitusi	1	1	0	0,5	0,176	0,088
Adil	1	0	0	0	0,477	0
Semua	1	0	0	0	0,477	0
Dapat	1	0	0	0	0,477	0
Penolakan	1	0	0	0	0,477	0
Ending	0	1	0	1	0,477	0,477
Drama	0	1	0	1	0,477	0,477
Ditolak	0	1	0	1	0,477	0,477
Gagal	0	0	1	0	0,477	0
Dalam	0	0	1	0	0,477	0
Inovasi	0	0	1	0	0,477	0
Hukum	0	0	1	0	0,477	0
Cacat	0	0	1	0	0,477	0

Tabel 3. 8 TF-IDF Doc C

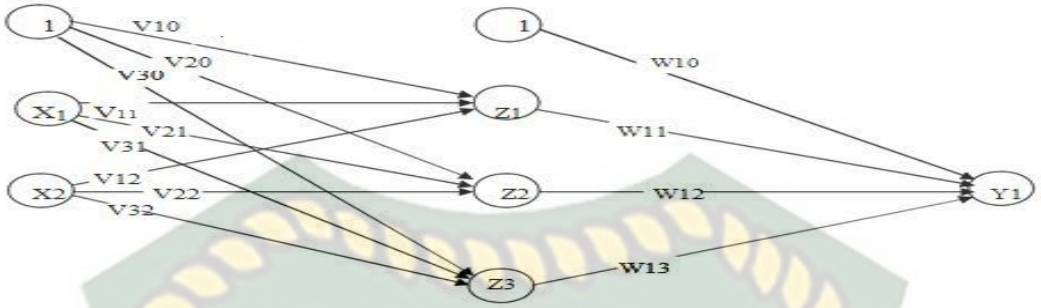
Kata	**Frekuensi Kemunculan Kata**			**TF (0.5+0.5 (ft,d/max(ftd)**	**IDF (log ($\frac{N}{dft}$)**	**W (TF*IDF)**
	A	**B**	**C**	**C**	**C**	**C**
Mahkamah	1	1	0	0	0,176	0

Konstitusi	1	1	0	0	0,176	0
Adil	1	0	0	0	0,477	0
Semua	1	0	0	0	0,477	0
Dapat	1	0	0	0	0,477	0
Penolakan	1	0	0	0	0,477	0
Ending	0	1	0	0	0,477	0
Drama	0	1	0	0	0,477	0
Ditolak	0	1	0	0	0,477	0
Gagal	0	0	1	1	0,477	0,477
Dalam	0	0	1	1	0,477	0,477
Inovasi	0	0	1	1	0,477	0,477
Hukum	0	0	1	1	0,477	0,477
Cacat	0	0	1	1	0,477	0,477

3.2.5 Perhitungan Multi-layer Perceptron

Berikut ini perhitungan manual metode *multi layer perceptron* (MLP). Jika di ketahui nilai tf-idf yang di ambil dari dua data berbeda yaitu (mahkamah 0,088) dan (Adil 0,0477), dengan menggunakan tahap atau algoritma backpropagation standar dengan sebuah layer tersembunyi (dengan 3 unit), input 1, hiden 1, output 1 dan untuk menghitung bobot jaringan dengan fungsi aktivasi sigmoid biner dengan learning-reat constant $\alpha = 0{,}1$. Bobot-bobot diberikan nilai acak, Misal bobot dari layer input ke layer tersembunyi seperti pada tabel A dan bobot-bobot dari layer tersembunyi ke layer output seperti pada tabel B.

Penyelesaian :



Gambar 3.3 MLP

1. Inisialisasi semua bobot dengan bilangan acak kecil.

Dengan Tabel A

x	Z_1	Z_2	Z_3
Mahkamah X_1 (0. 088)	0.02	0.03	-0.01
Adil X_2 (0.477)	0.03	0.01	0.03
1 (bias)/ ketetapan	-0.01	0.01	-0.01

Tabel B

	Y
Z_1	0.05
Z_2	-0.04
Z_3	0.02
1	-0.01

2. Jika kondisi penghentian belum terpenuhi, lakukan langkah 2 sampai dengan 8

3. Untuk setiap pasang data pelatihan, lakukan langkah 3 sampai dengan 8

4. Tiap unit masukkan menerima sinyal dan meneruskan ke unit tersembunyi

5. Hitung semua keluaran di unit tersembunyi z_j (j = 1, 2,..., p).

$$z\_net_j = v_{j0} + \sum_{i=1}^{n} x_i \, v_{ji}$$

z_net_1 = (-0.01) + (0.088) (0.02) + (0.477) (0.03) = 0,006

z_net_2 = (0.01) + (0.088) (0.03) + (0.477) (0.01) = 0,017

z_net_3 = (-0.01) + (0.088) (-0.01) + (0.477) (0.03) = 0,003

Hitung sigmoid :

$$z_j = f(z\_net_j)\frac{1}{1+\ e^{-z\_net_j}}$$

$z_1 = f\ (z_{net_1}) = \frac{1}{1+e^{-0,006}} = 0{,}994$

$z_2 = f\ (z_{net_2}) = \frac{1}{1+e^{-0,017}} = 0{,}983$

$z_2 = f\ (z_{net_2}) = \frac{1}{1+e^{-0,003}} = 0.997$

6. Hitung semua jaringan di unit keluaran y_k (k= 1,2.....m).

$$y\_net_k = w_{k0} + \sum_{j=1}^{p} z_j\, w_{kj}$$

$$y_{net1} = w_{10} + \sum_{j=1}^{p} z_j\, w_{kj} = w\ 10 + z_1\ w_{11} + z_2\ w_{12} + z_3\ w_{13}$$

= -0,01 + 0,994 . 0,05 + 0,983 . -0,04 + 0,997 . 0,02= 0,020

$$y_k = (y_{net_k}) = \frac{1}{1+e^{-0,020}} = 0{,}980$$

7. Hitung faktor δ unit keluaran berdasarkan kesalahan di setiap unit keluaran y_k (k=1,2,...m)

$\delta_k = (t_k - y_k)'(y\_net_k) = (t_k - y_k)\ y_k\ (1- y_k) = target$

$\delta_1 = (t_k - y_k)'(ynetk) = (t_k - y_k)y_k(1 - y_k)$

$= (0 - 0{,}980)\ 0{,}980\ (1 - 0{,}980) = -0{,}019$

δ_k merupakan unit kesalahan yang akan dipekai dalam perubahan bobot layer dibawahnya. Hitung perubahan bobot w_{kj} dengan laju pemahaman α.

$\Delta w_{kj} = \alpha\delta_k Z_j$ k= 1,2,...,m; j=0,1 ... p

$\Delta w_{kj} = \alpha\ \delta_k\ z_j$

$\Delta w_{10} = \alpha\ \delta_1\ (1) = 0{,}1$. (-0,019) . (1)= -0,001

$\Delta w_{11} = \alpha\ \delta_1\ (z_1) = 0{,}1$. (-0,019). (0,994)= -0,001

$\Delta w_{12} = \alpha\ \delta_1\ (z_2) = 0{,}1$. (-0,019) . (0,983)= -0,001

$\Delta w_{13}= \alpha\ \delta_1\ (z_3)=0{,}1\ .\ (-0{,}019)\ .\ (0{,}997)= -0{,}001$

8. Hitung faktor δ unit tersembunyi berdasarkan kesalahan disetiap unit tersembunyi z_j (j=1)

$$\delta\_net_j = \sum_{k=1}^{m} \delta_k w_{kj}$$

$\delta\_net_1 = \delta_{1.w11} = (-0{,}019).(0{,}05) = -0{,}001$

$\delta\_net_2 = \delta_{1.w12} = (-0{,}019).(-0{,}04) = 0{,}001$

$\delta\_net_3 = \delta_{1.w13} = (-0{,}019).(0{,}02) = -0{,}001$

Faktor δ unit tersembunyi

$\delta_j = \delta\_net_j f'(Z\_net_j) = \delta\_net_j z_j (1 - z_j)$

$\delta_1 = \delta\_net\ z_1\ (1\text{-}z_1) = (-0{,}001).\ 0{,}994.(1\text{-}\ (0{,}994)) = -5{,}964$

$\delta_2 = \delta\_net\ z_2\ (1\text{-}z_2) = (0{,}001).\ 0{,}983.\ (1\text{-}\ (0{,}983)) = 1{,}671$

$\delta_3 = \delta\_net\ z_3\ (1\text{-}z_3) = (-0{,}001).\ 0{,}997.\ (1\text{-}\ (0{,}997)) = -2.991$

Hitung suku perubahan bobot

$\Delta v_{ji} = \alpha\ \delta_j x_i$

$\Delta v10 = \alpha\ \delta 1 = 0{,}1.\ (-5{,}964)\ .1 = -0{,}596$

$\Delta v20 = \alpha\ \delta 2 = 0{,}1.\ (1{,}671)\ .1 = 0{,}167$

$\Delta v30 = \alpha\ \delta 3 = 0{,}1.\ (-2.991)\ .1 = -0{,}299$

$\Delta v11 = \alpha\ \delta 1 x1 = 0{,}1.\ (-5{,}964)\ .0{,}088 = -0{,}052$

$\Delta v21 = \alpha\ \delta 2 x1 = 0{,}1.\ (1{,}671)\ .0{,}088 = 0{,}014$

$\Delta v31 = \alpha\ \delta 3 x1 = 0{,}1.\ (-2.991)\ .0{,}088 = -0{,}026$

$\Delta v12 = \alpha\ \delta 1 x2 = 0{,}1.\ (-5{,}964)\ .\ 0{,}477 = -0{,}284$

$\Delta v22 = \alpha\ \delta 2 x2 = 0{,}1.\ (1{,}671)\ .0{,}477 = 0{,}079$

$\Delta v32 = \alpha\ \delta 3 x2 = 0{,}1.\ (-2.991)\ .\ 0{,}477 = -0{,}142$

9. Hitung semua perubahan bobot. Perubahan bobot garis yang menuju ke unit keluaran, yaitu :

w_{kj} (baru) = w_{kj}(lama) + Δw_{kj} (k = 1,2, ...m; j=0,1..., p)

w_{10} (baru)= w_{10} (lama) + Δw_{10}= -0,1+ -0,001= -0,101

w_{11} (baru)= w_{11} (lama) + Δw_{11}= 0,5+ -0,001= 0,499

w_{12} (baru)= w_{12} (lama) + Δw_{12}= -0,4+ -0,001= -0,401

w_{13} (baru)= w_{13} (lama) + Δw_{13}= 0,2+ -0,001= 0,199

perubahan bobot garis yang menuju ke unit tersembunyi yaitu :

V_{ji}(baru) = v_{ji} (lama) + Δv_{ji} , (j= 1,2,... , p ; i = 0,1, … . ,n)

V_{10} (baru)= v_{10} (lama) + Δv_{10} = -0,01+ -0,596 = -0,606

V_{20} (baru)= v_{20} (lama) + Δv_{20} = 0,01+ 0,167 = 0,177

V_{30} (baru)= v_{30} (lama) + Δv_{30} = -0,01+ -0,299 = -0,309

V_{11} (baru)= v_{11} (lama) + Δv_{11} = 0,02+ -0,052 = -0,032

V_{21} (baru)= v_{21} (lama) + Δv_{21} = 0,03+ 0,014 = 0,044

V_{31} (baru)= v_{31} (lama) + Δv_{31}= -0,01+ -0,026 = -0,036

V_{12} (baru)= v_{12} (lama) + Δv_{12} = 0,03+ -0,284 = -0,254

V_{22} (baru)= v_{22} (lama) + Δv_{22} = 0,01+ 0,079 = 0,169

V_{32} (baru)= v_{32} (lama) + Δv_{32}= 0,03+ -0,142 = -0,112

3.2.6 K-Fold Cross Validation

Berikut perhitngan metode Cross-Validation menggunakan 3-Fold Cross Validation dilihat pada tabel dibawah ini :

Gambar 3.4 Pembagian Data k-fold

Dengan Prosedur 3-Cross validation dari 300 data yang ada, dibagi menjadi iga bagian dengan komposisi jumlah data seperti tabel 3.9 di bawah :

Tabel 3.9 Rincian pembagian data

Data set	Jumlah Data
Data set 1	100
Data set 2	100
Data set 3	100

Setelah dilakukan uji coba didapat hasil seperti pada tabel 3.10 di bawah ini :

Tabel 3.10 hasil uji coba

3-fold	Jumlah Data		Akurasi
	Testing	Uji	
1	200	100	89%
2	100	200	76%
3	100	100	80%

untuk mendapatkan total dari keseluruhan Perhitungan akurasi tersebut dengan menggunakan persamaan, cara kerja *k-fold* adalah sebagai berikut ini :

a. Total instance dibagi menjadi N bagian.

b. Fold ke-1 adalah ketika bagian ke-1 menjadu data uji dan sisanya menjadi data latih. Selanjutnya, hitung akurasi berdasarkan porsi data tersebut. Perhitungan akurasi tersebut dengan menggunakan persamaan sebagai berikut:

c. Fold ke-2 adalah ketika bagian ke-2 menjadi data uji dan sisanya menjadi data latih. Selanjutnya, hitung akurasi berdasarkan porsi data tersebut.

d. Demikian seterusnya hingga mencapai fold ke-K. Hitung rata-rata akurasi dari K buah akurasi diatas. Rata-rata akurasi ini menjadi akurasi final. Validation menggunakan 3 fold cross validation dimana data dibagi secara acak menjadi 3bagiandata dengan jumlah yang sama. Sehingga dilakukan proses validasi sebanyak 3 kali secara berulang, dari perhitungan diatas dapat dari hasil sebagai berikut:

$$\text{Akurasi} = \frac{89 + 76 + 80}{3} \times 100\% = \frac{245}{3} \times 100\% = 8{,}1\%$$

BAB IV
HASIL DAN PEMBAHASAN

4.1 Model Analisis Sentimen

Penelitian ini menggunakan jupyter notebook sebagai tempat pembuatan analisis sentimen, karena jupyter notebook memiliki interface yang sederhana dan mudah dipahami. Jupyter ini juga memiliki kelebihan, yaitu dapat memperlihatkan alur program yang berhasil dijalankan atau error pada alur program tertentu. Berikut tahapan-tahapan analisis sentimen.

4.1.1 Penelitian

Proses pelatihan ini bertujuan untuk membentuk model prediksi dari data yang telah diketahui kelasnya. Pada kasus ini penulis menggunakan data latih sebanyak 200. Awal data yang gunakan 100 data, akan tetapi akurasi yang dihasilkan oleh sistem untuk 100 data latih hanya 0.6. Lalu data ditambahkan 100 data lagi menjadi 200 data untuk digunakan sebagai data latih dan akurasi yang dihasilkan adalah 0.706. Angka yang cukup bagus untuk digunakan sebagai data latih. Terdapat dua jenis kelas pada data latih ini yaitu “POSITIF” dan “NEGATIF”. Berikut contoh tampilan data latih pada sistem dapat di lihat pada gambar 4.1 berikut ini:

Out[3]:

	liked	txt
0	0	#BreakingNews Mahkamah Konstitusi (MK) menolak...
1	0	Gugatan @prabowo @sandiuno DITOLAK SELURUHNYA ...
2	0	RT @hermana_t: MK tdk bisa lepas tangan, MK be...
3	0	RT @helmifelis: Perjuangan tidak berhenti kare...
4	0	RT @dema_arraniry: * Desakan dari Aceh untuk P...
5	0	@hartiwansyah @Humas_MKRI Mampus jagoan lu kal...
6	0	Jimly Tegaskan Tak Ada Celah Hukum Lagi untuk ...
7	0	@AllamisKoto1 @maubelimimpi @ustadtengkuzul @H...
8	0	RT @KompasTV: Mahkamah Konstitusi Tolak Dalil ...
9	0	RT @bawaslu_RI: 2. Keterangan tersebut berpera...
10	0	@CNNIndonesia Mau kmn lagi...??? Mungkin ke BM...

Gambar 4.1 Tampilan Data Latih

Liked yang terdapat pada gambar 4.1 adalah kelas dari kalimat dimana 0 adalah negatif dan 1 adalah positif. Sedangkan txt adalah kalimat atau *tweet* yang diambil dari twitter. Data yang diinput berupa data dari excel dengan format csv. Gambar 4.1 adalah 10 contoh data yang ditampilkan dari data yang ada.

4.1.2 Pengujian Dengan Data Uji

Pengujian dengan data uji adalah proses puncak pada analisis sentimen. Pada proses ini pengguna akan melakukan penginputan data dan memperoleh hasil berupa kelas dalam bentuk binner yaitu negatif dan positif. Hasil ini didapatkan dari sistem melalui pembelajaran dari data latih yang sudah diinputkan. Untuk menginputkan data uji, dilakukan secara otomatis ditempat yang telah disediakan. bisa di lihat pada gambar 4.2 berikut ini:

```
In [13]: print (clf.predict(uji_sistem_vector))
[1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 1 1 0 0 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1]
```

Gambar 4.2 Tampilan *Form* Input Data Uji.

Dalam gambar 4.2 Hasil dari sistem ini di uji dengan hasil yang didapatkan oleh user, dengan cara manual untuk mendapatkan ketepatan prediksi. Hasil dari

uji manual yang sudah dilakukan oleh user diinputkan pada sistem. Disini penulis menggunakan 100 data baru yang diambil secara acak. Untuk menghasilkan ketepatan prediksi, penulis menggunakan pengujian ketepatan prediksi antara uji pada sistem dan uji manual. Pengujian ini salah satu metode pengujian perogram, untuk mengetahui hasil berjalan dengan baik dan menghasilkan *output* yang diinginkan. Hasil dari algoritma yang ada pada program akan menampilkan tabel ketepatan prediksi seperti pada gambar 4.3 di bawah ini:

Out[15]:

	Tweet	Sentiment	Uji Manual	Ketepatan Prediksi
0	@habibthink dan 2 lainnya Terhadap keputusaan...	1	0	Tidak Tepat
1	ampyuunnn... forum sidang Mahkamah Konstitusi....	1	0	Tidak Tepat
2	Atas dasar itu UU SBT diajukan ke Mahkamah Kon...	0	0	Tepat
3	Pembahasan tindak lanjut putusan mahkamah Kons...	0	0	Tepat
4	MOHON DOAKAN KAMI, AGAR #MENANG DI #MAHKAMAH K...	0	0	Tepat
5	Tunggalisasi wadah gerakan koperasi secara lan...	0	0	Tepat
6	Berkunjung ke museum sejarah Konstitusi, #Mahk...	0	0	Tepat
7	Bawaslu Kota Cirebon Mengucapkan Selamat Hari ...	0	0	Tepat
8	Aku sih gak tegang nunggu putusan MK, yang leb...	0	0	Tepat
9	Ending drama, ditolak mahkamah konstitusi wkwk...	1	0	Tidak Tepat
10	jokowi @Pak_JK @DPR_RI @bgiyer Akankah RUU Per...	0	0	Tepat

Gambar 4.3 Tampilan Ketepatan Prediksi

Sentimen merupakan hasil uji dari sistem sedangkan uji manual merupakan hasil uji dari user. Jika hasil uji dari sistem sama dengan hasil uji dari *user* maka prediksinya tepat, dan jika hasil uji tidak sama maka prediksinya tidak tepat.

4.2 *Evaluation Measure*

Evaluasi *Measure* atau evaluasi performansi dilakukan untuk menguji hasil klasifikasi dengan mengukur nilai kebenaran dari sistem. Ketika dataset memiliki dua kelas yaitu positif dan negatif. Dalam kasus ini ada dua baris dan kolom confusion matrix dirujuk sebagai *true and false positives* dan *true and false negatives.*

Tabel 4.1 *Confusion Matrix*

		Predicted class	
		Positif	Negatif
Actual class	Positif	TP	FN
	Negatif	FP	TN

True positive (TP) adalah jumlah record positif yang diklasifikasikan sebagai positif, *false positive* (FP) adalah jumlah record negatif yang diklasifikasikan sebagai positif, *false negative* (FN) adalah jumlah record positif yang diklasifikasikan sebagai negatif, *true negative* (TN) adalah jumlah record negatif yang diklasifikasikan sebagai negatif. Untuk klasifikiasi text, menggunakan pengukuran, yaitu *accuracy*

4.2.1 Accuracy

Accuracy yaitu persentase dokumen yang berhasil diklasifikasikan dengan tepat oleh sistem. *Accuracy* diperoleh dari hasil perhitungan yang ditunjukkan pada persamaan dibawah ini.

$$Accuracy = \frac{TP+TN}{(TP+FP+FN+TN)}$$

Dari 100 data *testing* (uji) yang sudah diproses ketepatan prediksi nya, didapatkan hasil TP : 45, TN : 44, FP : 6, FN : 5 . Perhitungan *accuracy* nya adalah sebagai berikut:

$$Accuracy = \frac{45+44}{(45+6+5+44)} = \frac{89}{(100)} = 0.89$$

Disini penulis menggunakan bilangan bulat karna pada perhitungan ini penulis menentukan *range* untuk hasilnya dengan nilai 1-100 maka hasil nya dikalikan

dengan 100, dan didapatkan hasil dari *accuracy* antara uji manual dan uji pada sistem adalah 89. Angka yang cukup tinggi karna mendekati 100. Dokumen yang diklasifikasikan dengan tepat oleh sistem memiliki persentase yang cukup tinggi. proses ini menggunakan pengujian perhitungan *accuracy*, dan hasil yang diperoleh dari fungsi accuracy dapat di lihat pada gambar 4.4 berikut ini:

```
In [21]: print('akurasi :', compute_akurasi(tp, tn, fn, fp))
         akurasi : 89.0
```

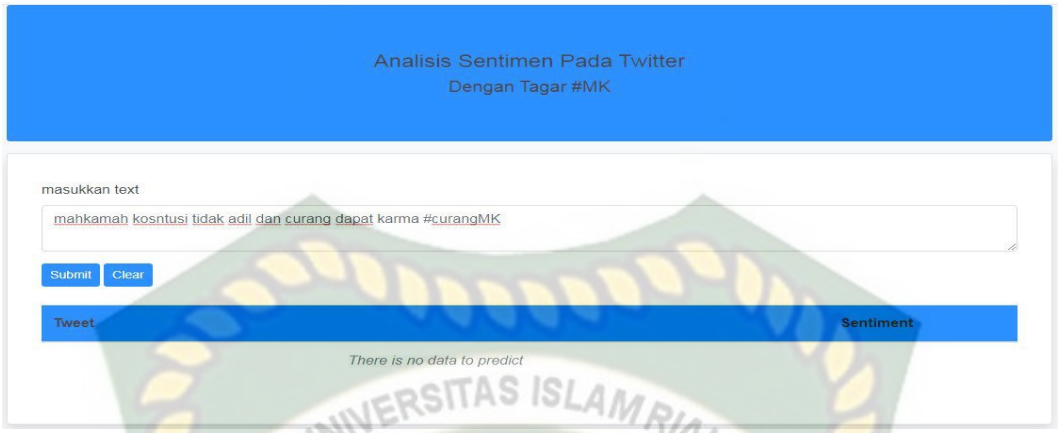
Gambar 4.4 Perhitungan *Accuracy* Pada Sistem

Hasil dari pengujian yang dilakukan ini dapat disimpulkan bahwa perhitungan accuracy sama dengan hasil yang dilakukan secara manual, pengujian ini mendapatkan hasil sesuai dengan yang diharapkan.

Evaluation measure pada sistem sesuai dengan perhitungan yang sudah dilakukan menghasilkan nilai cukup tinggi karna data *testing* yang diinputkan memiliki kosa kata yang ada pada data *training*, sehingga sistem dapat mempelajari dari data *training* dan dapat memproses data *testing* dengan baik.

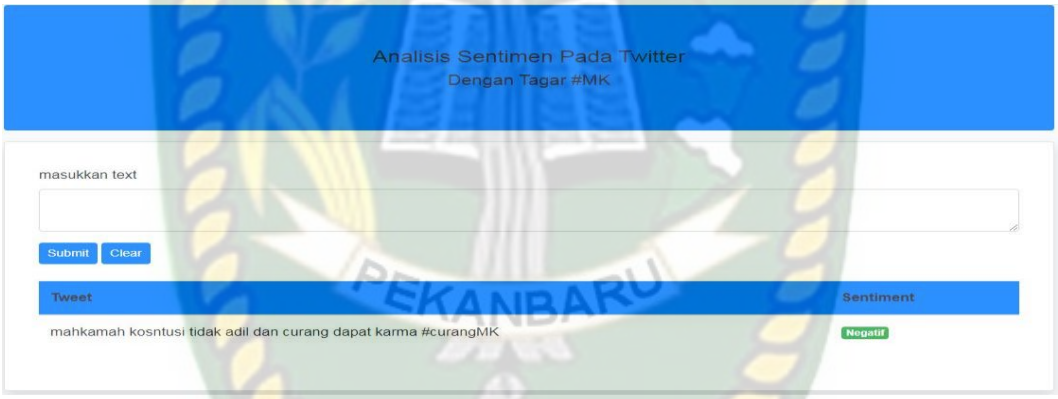
4.3 Antarmuka Pada Analisis Sentimen

Pada penelitian ini, pengguna umum juga bisa menggunakan analisis sentimen untuk melihat apakah sebuah *tweet* memiliki kelas positif atau negatif. Antarmuka yang disediakan untuk pengguna umum adalah penginputan berupa text, yang nantinya akan di proses dengan data training yang ada dan menghasilkan ouput berupa tabel yang berisi *tweet* yang diinputkan dan sentimen berupa positif atau negatif dari *tweet* tersebut. Berikut antarmuka input untuk pengguna umum bisa dilihat pada gambar 4.5 berikut ini:



Gambar 4.5 Form *Input* User

Antarmuka bagian output untuk pengguna umum bisa dilihat pada gambar 4.6 dibawah ini :



Gambar 4.6 Form *Output* User

4.4 Pengujian Kepada User

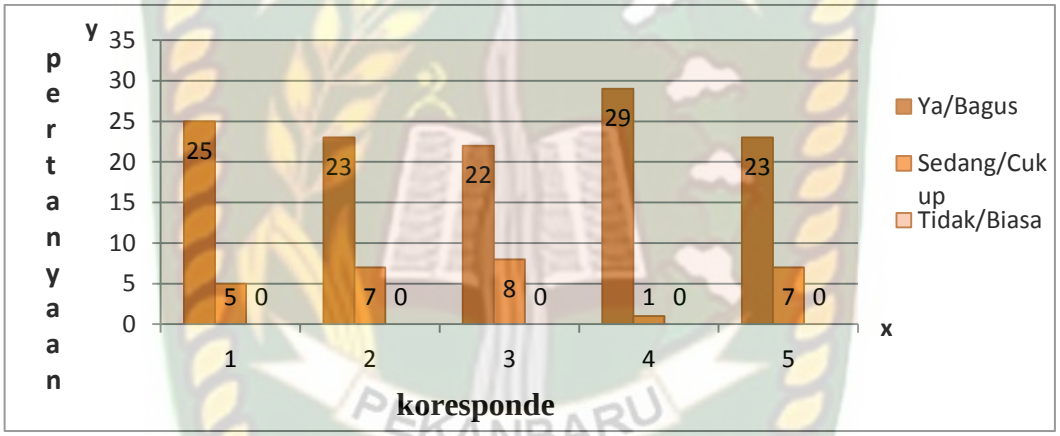
4.4.1 Implementasi User

Implementasi yang dilakukan yaitu membuat kuisioner dengan 5 pertanyaan dan 30 koresponden yang mana ditujukan kepada pengguna yang ingin memakai sistem ini. kepada 30 koresponden diajukan pertanyaan yang terkait dengan kinerja atau *performance* dari sistem. Adapun kelima pertanyaan yang akan diajukan sebagai berikut :

1. Apakah informasi yang ditampilkan mudah dimengerti oleh user?

2. Apakah bahasa yang digunakan pada analisis sentimen mudah dimengerti?
3. Bagaimana pendapat anda tentang tampilan pada analisis sentimen ini?
4. Apakah analisis sentimen ini cukup mudah digunakan (dioperasikan)?
5. Menurut anda, apakah analisis sentimen ini sudah layak dipublikasikan?

Dari pertanyaan-pertanyaan diatas, maka hasil jawaban atau tanggapan dari koresponden terhadap kinerja dan *performance* dari sistem berdasarkan pertanyaan yang diajukan sebagai berikut:



Gambar 4.7 Hasil Kuisioner

Keterangan:

1. Apakah informasi yang ditampilkan mudah dimengerti oleh user memiliki nilai BAGUS : 25 koresponden, SEDANG : 5 koresponden, TIDAK : 0 koresponden.

2. Apakah bahasa yang digunakan pada analisis sentimen mudah dimengerti memiliki nilai BAGUS : 23 koresponden, SEDANG : 7 koresponden, TIDAK : 0 koresponden.

3. Bagaimana pendapat anda mengenai tampilan pada analisis sentimen ini memiliki nilai BAGUS : 22 koresponden, CUKUP : 8 koresponden, BIASA : 0 koresponden.

4. Apakah analisis sentimen ini cukup mudah digunakan (dijalankan) memiliki nilai BAGUS : 29 koresponden, SEDANG : 1 koresponden, TIDAK : 0 koresponden.

5. Menurut anda, apakah analisis sentimen ini sudah layak dipublikasikan memiliki nilai BAGUS : 23 koresponden, CUKUP : 7 koresponden, TIDAK : 0 koresponden.

Dari keterangan diatas sumbu Y adalah hasil dari orang yang menjawab pertanyaan yang ada di sumbu X.

4.4.2 Kesimpulan Implementasi Sistem

Berdasarkan hasil dari kuisioner tersebut maka dapat disimpulkan bahwa aplikasi manajemen jurnal online ini memiliki presentase sebagai berikut:

Tabel 4.2 Hasil Presentasi Tiap Pertanyaan Kuisioner

No	Pertanyaan	Jumlah Persentase Koresponden		
		Ya / Bagus	Sedang / Cukup	Tidak / Biasa
1	Apakah informasi yang ditampilkan mudah dimengerti oleh user?	83%	17%	0%
2	Apakah bahasa yang digunakan pada analisis sentimen mudah dimengerti?	77%	23%	0%
3	Bagaimana pendapat anda mengenai tampilan pada analisis sentimen ini?	73%	27%	0%
4	Apakah analisis sentimen ini cukup mudah digunakan (dijalankan)	97%	3%	0%
5	Menurut anda, apakah analisis sentimen ini sudah layak dipublikasikan	77%	23%	0%

Dari hasil persentase tabel 4.1 di atas, nilai presentase tiap-tiap pertanyaan kuisioner, analisis sentimen pada twitter dengan tagar #Mahkamahkosntitusi memiliki performance baik dengan nilai persentase rata-ratanya sebesar 81.4%, sehingga aplikasi ini dapat diimplementasikan ke publik.

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan pembahasan yang sudah dijelaskan pada bab-bab sebelumnya, maka dapat ditarik kesimpulnya bahwa :

1. jumlah data dan keanekaragam data berpengaruh terhadap akurasi pada sistem perhitungan *evaluation measure*.
2. Metode *multilayer perceptron* telah berhasil dijalankan terhadap analisis sentimen twitter, dimana sentimen menghasilkan akurasi yang cukup tinggi dengan akurasi pada sistem 89% dengan 300 data.

5.2 Saran

Penenulis sadar bahwa aplikasi analisis sentimen ini masih banyak kelemahan dan kekurangannya. untuk penelitian berikutnya, dapat menggunakan metode algoritma klasifikasi lainnya seperti *K Nearest Neighbor (KNN) Support Vector Machine (SVM), Decision Tree* dan Naive Bayesian *(naïve Bayesian classifier)* dan menambahkan data yang lebih banyak untuk akurasi yang lebih tinggi.

DAFTAR PUSTAKA

Sarjana, P. P., Elektro, J. T., & Industri, F. T. (2015). *SENTIMENT ANALYSIS MENGGUNAKAN SUPPORT VECTOR MACHINE (SVM) SENTIMENT ANALYSIS USING SUPPORT VECTOR MACHINE (SVM)*.

Tika, G., Informatika, F., & Telkom, U. (2019). *Klasifikasi Topik Berita Berbahasa Indonesia menggunakan Multilayer Perceptron*. *6*(1), 2137–2143.

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. *E-Proceeeding of Engineering*, *2*(1), 1–7.

Nurhuda, F., & Sihwi, S. W. (2014). *Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier*. *2*(2).

Assuja, M. A., & Saniati, S. (2016). Analisis Sentimen Tweet Menggunakan Backpropagation Neural Network. *Jurnal Teknoinfo*, *10*(2), 48. https://doi.org/10.33365/jti.v10i2.20

Tempola, F., Muhammad, M., & Khairan, A. (2018). Perbandingan Klasifikasi Antara Knn Dan Naive Bayes Pada Penentuan Status Gunung Berapi Dengan K-Fold Cross Validation Comparison of Classification Between Knn and Naive Bayes At the Determination of the Volcanic Status With K-Fold Cross. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *5*(5), 577–584. https://doi.org/10.25126/jtiik20185983

Hariri, F. R., & Pamungkas, D. P. (2016). Implementasi Naïve Bayes Classifier. *Seminar Nasional Teknologi Informasi Dan Multimedia*, 6–7.

Assiva, M. A., Santoso, H. A., & Supriyanto, C. (2019). *METODE FASTICA UNTUK REDUKSI DATA DIMENSI TINGGI PADA ANALISIS SENTIMEN PARIWISATA KOTA SEMARANG*. *15*, 45–60.

Syadid, F. (2019). Analisis Sentimen Komentar Netizen Terhadap Calon Presiden Indonesia 2019 Dari Twitter Menggunakan Algoritma Term Frequency-Invers Document Frequency (Tf- Idf) Dan Metode Multi Layer Perceptron (Mlp) Neural Network. *Skripsi Universitas Islam Negeri Syarif Hidayatullah Jakarta*, 1–89

Muliantara, A., & Widiartha, I. M. (2011). Penerapan Multi Layer Perceptron Dalam Anotasi Image Secara Otomatis. *Jurnal Ilmu Komputer*, *4*(1), 9–15. http://ojs.unud.ac.id/index.php/jik/article/view/2686%5Cnfiles/586/Muliantar

a and Widiartha - 2011 - PENERAPAN MULTI LAYER PERCEPTRON DALAM ANOTASI IMA.pdf%5Cnfiles/587/2686.html

Nurjannah, M., & Astuti, I. F. (2013). *PENERAPAN ALGORITMA TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF) UNTUK TEXT MINING*. *8*(3), 110–113.

Syah, A. P., Faraby, S. Al, Informatika, F., & Telkom, U. (2017). *ANALISIS SENTIMEN PADA DATA ULASAN PRODUK TOKO ONLINE DENGAN METODE MAXIMUM ENTROPY SENTIMENT ANALYSIS ON ONLINE STORE PRODUCT REVIEWS WITH MAXIMUM*. *4*(3), 4632–4640.

Lidya, S. K., Sitompul, O. S., & Efendi, S. (2015). Sentiment Analysis Pada Teks Bahasa Indonesia Menggunakan Support Vector Machine (Svm). *Seminar Nasional Teknologi Dan Komunikasi 2015*, *2015*(Sentika), 1–8. https://doi.org/10.1016/j.eswa.2013.08.047

Nur Azizah Vidya (2015). *TWITTER SENTIMENT ANALIYSIS TERHADAP BRAND REPUTATION: STUDI KASUS PT XL AXIATI Tbk.* https://doi.org/10.1590/s1809-98232013000400007

Setiabudidaya, D. (2018). *Penggunaan Piranti Lunak Jupyter Notebook dalam Upaya Mensosialisasikan Open Science Dedi Setiabudidaya*. 2–5.

Dokumen, L., & Seluruh, C. (2009). *Flowchart*. 1–4.