

**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP
PERISTIWA BITCOIN HALVING PADA DATA TEKS TWITTER
MENGUNAKAN METODE NAÏVE BAYES DAN PEMBOBOTAN
FITUR TF-IDF**

SKRIPSI

Diajukan oleh:

Andi Nur Halim

2011102441038



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
JULI 2024**

**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP
PERISTIWA BITCOIN HALVING PADA DATA TEKS TWITTER
MENGUNAKAN METODE NAÏVE BAYES DAN PEMBOBOTAN
FITUR TF-IDF**

SKRIPSI

Diajukan Sebagai Salah Satu Persyaratan Untuk Memperoleh Gelar Sarjana Teknik
Informatika Fakultas Sains Dan Teknologi Universitas Muhammadiyah Kalimantan
Timur

Diajukan oleh:

Andi Nur Halim

2011102441038



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
JULI 2024**

LEMBAR PERSETUJUAN

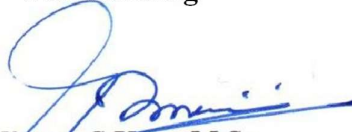
ANALISIS SENTIMEN OPINI PUBLIK TERHADAP PERISTIWA BITCOIN HALVING PADA DATA TEKS TWITTER MENGGUNAKAN METODE NAIVE BAYES DAN PEMBOBOTAN FITUR TF-IDF

SKRIPSI

**Diajukan Oleh :
Andi Nur Halim
2011102441038**

**Disetujui untuk diujikan
Pada tanggal 29 Juni 2024**

Pembimbing



**Rudiman, S.Kom, M.Sc
NIDN. 1105068202**

Mengetahui,

Koordinator Skripsi



**Abdul Rahim S.Kom M.cs
NIDN. 0009047901**

LEMBAR PENGESAHAN

ANALISIS SENTIMEN OPINI PUBLIK TERHADAP PERISTIWA BITCOIN HALVING PADA DATA TEKS TWITTER MENGGUNAKAN METODE NAIVE BAYES DAN PEMBOBOTAN FITUR TF-IDF

SKRIPSI

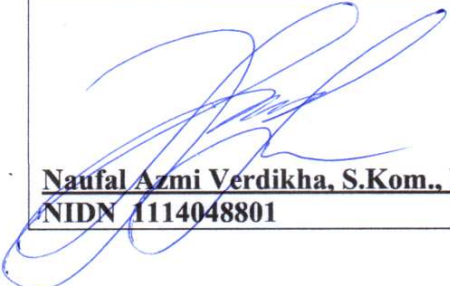
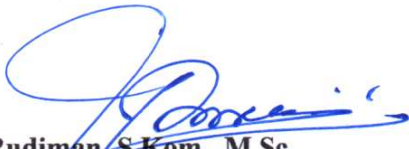
Diajukan oleh:

Andi Nur Halim

2011102441038

Diseminarkan dan Diujikan

Pada Tanggal 8 Juli 2024

Penguji I	Penguji II
 <u>Naufal Azmi Verdikha, S.Kom., M.Eng</u> NIDN 1114048801	 <u>Rudiman, S.Kom., M.Sc</u> NIDN 1105068202

Mengetahui,

Ketua

Program Studi Teknik Informatika



PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan di bawah ini:

Nama : Andi Nur Halim

NIM : 2011102441038

Program Studi : S1 Teknik Informatika

Judul Penelitian : Analisis Sentimen Opini Publik terhadap Peristiwa Bitcoin Halving pada Data Teks Twitter Menggunakan Metode Naïve Bayes dan Pembobotan Fitur TF-IDF

Menyatakan bahwa Skripsi yang saya tulis ini benar-benar hasil karya saya sendiri, dan bukan merupakan hasil plagiasi/falsifikasi/fabrikasi baik sebagian atau seluruhnya.

Atas pernyataan ini, saya siap menanggung resiko atau sanksi yang dijatuhkan kepada saya apa bila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam skripsi saya ini, atau klaim dari pihak lain terhadap keaslian karya saya ini.

Samarinda 30 Juni 2024
Yang membuat pernyataan



Andi Nur Halim
NIM: 2011102441038

ABSTRAK

Perkembangan teknologi telah mengubah cara orang berinteraksi dan melakukan transaksi. Salah satu inovasi penting adalah mata uang digital, yang biasa dikenal sebagai cryptocurrency. Baru-baru ini, topik Bitcoin Halving telah menarik perhatian besar di Twitter, bahkan menjadi trending topic di seluruh dunia. Peristiwa ini memicu banyak opini dan komentar dari pengguna Twitter. Mengingat banyaknya tweet yang terkait dengan Bitcoin Halving, sangat sulit untuk menentukan secara manual apakah sentimennya positif atau negatif. Oleh karena itu, diperlukan text mining untuk mengklasifikasikan sentimen tersebut, baik positif maupun negatif. Penelitian ini bertujuan memanfaatkan algoritma Naïve Bayes Classifier dan pembobotan fitur menggunakan TF-IDF (Term Frequency – Inverse Document Frequency). Dari total 538 data tweet yang diperoleh dari proses crawling di media sosial Twitter, dilakukan preprocessing dan pembobotan kata menggunakan TF-IDF, serta pembagian data untuk pelatihan dan pengujian model. Dengan beberapa rasio data latih dan data uji yaitu 90:10, 80:20, dan 80:20, hasil penelitian ini menunjukkan bahwa model Naïve Bayes dengan rasio 70:30 mendapatkan hasil akurasi terbaik yaitu 74%.

Kata kunci: Bitcoin Halving, Twitter, Analisis Sentimen, Naïve Bayes, TF-IDF

ABSTRACT

Technological advancements have revolutionized how people interact and conduct transactions. One significant innovation is digital currency, commonly known as cryptocurrency. Recently, the topic of Bitcoin Halving has garnered substantial attention on Twitter, even trending worldwide. This event has sparked numerous opinions and comments from Twitter users. Given the large volume of tweets related to Bitcoin Halving, manually determining whether the sentiment is positive or negative is extremely challenging. Therefore, text mining is necessary to classify these sentiments as either positive or negative. This study aims to utilize the Naïve Bayes Classifier algorithm and feature weighting using TF-IDF (Term Frequency – Inverse Document Frequency). From a total of 538 tweet data collected through crawling on social media Twitter, preprocessing and word weighting using TF-IDF were conducted, followed by data splitting for model training and testing. With various training and testing data ratios of 90:10, 80:20, and 80:20, the results of this study indicate that the Naïve Bayes model with a 90:30 ratio achieved the best accuracy of 74%.

Keywords: *Bitcoin Halving, Twitter, Sentiment Analysis, Naïve Bayes, TF-IDF*

PRAKATA

Assalamu'alaikum warahmatullahi wabarkatuh.

Alhamdulillah, dengan mengucapkan puji syukur kehadiran Allah SWT, atas segala nikmat yang di karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan tepat waktu, dengan judul “ Analisis Sentimen Opini Publik Terhadap Peristiwa Bitcoin Halving Pada Data Teks Twitter Menggunakan Metode Naïve Bayes Dan Pembobotan Fitur TF-IDF ” yang merupakan tugas akhir selama menempuh pendidikan di Universitas Muhammadiyah Kalimantan Timur dan merupakan salah satu syarat untuk kelulusan dan mendapatkan gelar sarjana yang harus di penuhi.

Dalam menyelesaikan Tugas Akhir ini banyak pihak yang telah memberikan bantuan dan dukungan kepada penulis dengan secara langsung maupun tidak secara langsung, penulis ingin menyampaikan rasa terima kasih kepada :

1. Puji syukur saya ingin ucapkan kepada Allah SWT atas segala limpahan rahmat dan berkah-Nya yang telah membimbing langkah-langkah saya dalam menyelesaikan skripsi ini.
2. Bapak Rudiman, S.Kom., M.Sc selaku dosen pembimbing yang selalu memberikan masukan, dukungan, semangat dan arahan kepada penulis untuk menyelesaikan skripsi ini.
3. Bapak Naufal Azmi Verdikha, S.Kom., M.Eng., sebagai dosen penguji pertama fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur.
4. Bapak Arbansyah, S.Kom., M.TI sebagai ketua Program Studi Teknik Informatika yang telah memberikan dukungan selama masa perkuliahan di Teknik Informatika.
5. Bapak Abdul Rahim S.Kom., M.Cs selaku Koordinator Skripsi yang telah memberikan dukungan dan arahan dalam proses penyelesaian skripsi.
6. Bapak Taghfirul Yoga Azhima Yoga Siswa selaku dosen pembimbing Akademik yang telah membimbing dari masa perkuliahan.
7. Dr. Muhammad Musiyam, M.T, selaku Rektor Universitas Muhammadiyah Kalimantan Timur yang telah memberikan dukungan, inspirasi, dan motivasi yang sangat berarti selama penulisan skripsi ini.
8. Seluruh dosen Teknik Informatika yang telah memberikan banyak ilmu dan pengetahuan kepada penulis, agar penulis bisa menyelesaikan skripsi ini dengan baik.
9. Kepada pahlawan dan panutanku, Ayahanda Abrani. Terima kasih selalu berjuang untuk kehidupan penulis, beliau memang tidak sempat menyelesaikan pendidikan sampai bangku perkuliahan, namun beliau mampu mendidik penulis, memotivasi dan memberikan dukungan hingga mampu menyelesaikan studinya sampai sarjana.
10. Kepada ibunda tercinta, Fitriah yang tidak henti hentinya memberikan kasih sayang dengan penuh cinta dan selalu memberikan motivasi serta do'a hingga penulis mampu menyelesaikan studinya sampai sarjana.

11. Kepada rekan-rekan sahabat dan teman atas dukungan dan kerjasamanya selama menempuh pendidikan hingga menyelesaikan penyusunan skripsi ini.

Akhir kata, penulis berharap agar skripsi ini dapat memberikan kontribusi yang bermanfaat bagi perkembangan ilmu pengetahuan, khususnya dalam bidang analisis sentimen dan text mining. Semoga penelitian ini dapat menjadi referensi bagi penelitian selanjutnya. Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, oleh karena itu penulis sangat mengharapkan saran dan kritik yang membangun dari berbagai pihak. Terima kasih.

Samarinda 30 Juni

2024

Penyusun

A handwritten signature in black ink, appearing to read 'Andi Nur Halim', with a stylized flourish at the end.

Andi Nur Halim

DAFTAR ISI

Halaman

LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN.....	iii
PERNYATAAN KEASLIAN PENELITIAN.....	iv
ABSTRAK	v
ABSTRACT	vi
PRAKATA.....	vii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
BAB I	1
PENDAHULUAN.....	1
1. Latar belakang Masalah	1
1.1. Rumusan masalah	3
1.2. Tujuan penelitian	4
1.3. Manfaat Penelitian.....	4
BAB II.....	5
METODE PENELITIAN	5
2.1 Objek penelitian.....	5
2.2 Alat dan bahan	5
2.3 Prosedur Penelitian	6
2.3.1 Pengumpulan data.....	7
2.3.2 Labeling Data.....	8
2.3.3 <i>Pre-Processing</i>	9
2.3.4 TF-IDF (<i>Term Frequency - Inverse Document Frequency</i>).....	11
2.3.5 Split Data	13
2.3.6 <i>Naïve Bayes Classifier</i>	14
2.3.7 Evaluasi.....	16
BAB III.....	17
HASIL ANALISIS DAN PEMBAHASAN.....	17
3.1 Dataset	17
3.2 <i>Labeling</i>	18
3.3 <i>Pre-processing</i>	19

a. <i>Case folding</i>	20
b. <i>Cleansing</i>	21
c. <i>Tokenizing</i>	21
d. <i>Stopword Removal</i>	22
e. <i>Stemming</i>	23
f. <i>Delete Duplicates</i>	23
3.4 <i>Wordcloud</i>	24
3.5 <i>Pembobotan Kata TF-IDF</i>	25
3.6 <i>Split Data</i>	26
3.7 <i>Klasifikasi</i>	27
3.8 <i>Evaluasi</i>	30
BAB 4.....	33
PENUTUP.....	33
4.1 <i>Kesimpulan</i>	33
4.2 <i>Saran</i>	33
DAFTAR PUSTAKA.....	35
RIWAYAT HIDUP PENULIS	38
LAMPIRAN	39

DAFTAR TABEL

Halaman

Tabel 2. 1 Confusion Matrix	16
Tabel 3. 1 Dataset	17
Tabel 3. 2 Labeling Dataset.....	18
Tabel 3. 3 Hasil case folding	20
Tabel 3. 4 Hasil Cleansing.....	21
Tabel 3. 5 Hasil Tokenizing.....	21
Tabel 3. 6 Hasil Stopword Removal.....	22
Tabel 3. 7 Hasil Stemming	23
Tabel 3. 8 Hasil TF-IDF	25

DAFTAR GAMBAR

Halaman

Gambar 2. 1 Kerangka Penelitian	7
Gambar 2. 2 Tahapan Preprocessing	10
Gambar 2. 3 Split Data.....	13
Gambar 3. 1 Hasil Crawling.....	17
Gambar 3. 2 Visualisasi Persentase Sentimen.....	19
Gambar 3. 3 Delete Duplicates	24
Gambar 3. 4 Visualisasi Wordcloud	25
Gambar 3. 6 Split Dataset	27
<i>Gambar 3. 7 Confusion Matrix Rasio 90:10.....</i>	<i>28</i>
Gambar 3. 8 Confusion Matrix 80:20	29
Gambar 3. 9 Confusion Matrix Rasio 70:30	30
Gambar 3. 10 Perbandingan Accuracy Naïve Bayyes.....	31

DAFTAR LAMPIRAN

Halaman

Lampiran 1. 1 CV Expert Labelling	39
Lampiran 2. 1 Code Crawling Twitter.....	40
Lampiran 2. 2 Import Library & Pip Install	41
Lampiran 2. 3 Read Dataset & Count Sentiment	42
Lampiran 2. 4 Preprocessing	42
Lampiran 2. 5 Code Untuk Menyimpan Hasil Teks Preprocessing	43
Lampiran 2. 6 Code Delete Duplicate	43
Lampiran 2. 7 Code Visualisasi Persentase Sentimen.....	44
Lampiran 2. 8 Code Wordcloud Sebelum Teks Preprocessing	44
Lampiran 2. 9 Code Wordcloud Setelah Teks Preprocessing.....	44
Lampiran 2. 10 TF-IDF (Term Frequency – Inverse Document Frequency).....	45
Lampiran 2. 11 Naive Bayes Classification	45
Lampiran 2. 12 Code Confusion Matrix	46
Lampiran 3. 1 Kartu Kendali Bimbingan.....	47