

**PERBANDINGAN ALGORITMA SVM, MNB DAN ENSEMBLE
LEARNING DALAM SENTIMEN ANALISIS TERHADAP
CRYPTOCURRENCY CRASH LUNA & UST**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer



disusun oleh

ACHMAD NAFIK

18.83.0304

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**PERBANDINGAN ALGORITMA SVM, MNB DAN ENSEMBLE
LEARNING DALAM SENTIMEN ANALISIS TERHADAP
CRYPTOCURRENCY CRASH LUNA & UST**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer



disusun oleh
ACHMAD NAFIK
18.83.0304

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023

HALAMAN PERSETUJUAN

SKRIPSI

**PERBANDINGAN ALGORITMA SVM, MNB DAN ENSEMBLE LEARNING
DALAM SENTIMEN ANALISIS TERHADAP CRYPTOCURRENCY CRASH
LUNA & UST**

yang disusun dan diajukan oleh

Achmad Nafik

18.83.0304

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 25 Januari 2023

Dosen Pembimbing,

Dony Ariyus, M.Kom

NIK. 190302128

HALAMAN PENGESAHAN

SKRIPSI

**PERBANDINGAN ALGORITMA SVM, MNB DAN ENSEMBLE LEARNING
DALAM SENTIMEN ANALISIS TERHADAP CRYPTOCURRENCY CRASH
LUNA & UST**

yang disusun dan diajukan oleh

Achmad Nafik

18.83.0304

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 Januari 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Jeki Kuswanto, M.Kom
NIK. 190302456

Senie Destya, M.Kom
NIK. 190302312

Wahid Miftahul Ashari, S.Kom., M.T
NIK. 190302452



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 25 Januari 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Achmad Nafik
NIM : 18.83.0304

Menyatakan bahwa Skripsi dengan judul berikut:

Perbandingan Algoritma SVM, MNB dan Ensemble Learning dalam Sentimen Analisis Terhadap Cryptocurrency Crash Luna & UST

Dosen Pembimbing : Dony Ariyus, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 25, Januari 2023

Yang Menyatakan,



Achmad Nafik

HALAMAN PERSEMBAHAN

Puji syukur kehadirat Allah SWT atas berkat dan rahmat yang telah diberikan kepada saya sehingga dapat menyelesaikan skripsi ini. Oleh karena itu dengan rasa syukur dan bangga saya ucapkan terimakasih kepada :

1. Allah SWT, berkat rahmat kasih sayang dan pertolongan-Nya saya bisa sampai sejauh ini.
2. Kedua orang tua saya, Bapak dan Ibu, yang selalu mendukung dan mendidik saya sejak kecil hingga saat ini. Mereka selalu memberikan dukungan moril dan doa yang tulus sehingga saya dapat menyelesaikan masa studi S1 dengan baik. Tanpa dukungan dan doa mereka, saya tidak akan dapat mencapai titik ini.
3. Dosen pembimbing saya, Bapak Dony Ariyus, M.Kom yang selalu memacu dan membimbing saya dalam proses menyelesaikan skripsi.
4. Bapak/Ibu dosen pembimbing, pengajar, serta penguji di Universitas AMIKOM Yogyakarta yang selalu memberikan pengalaman ilmu dengan tulus dan mengarahkan saya sehingga saya dapat menyelesaikan masa studi S1 dengan baik. Saya berharap segala ilmu yang telah diberikan oleh Bapak/Ibu akan menjadi ladang amal yang memberikan pahala jariyah dan selalu diberikan keberkahan ilmu dan rezeki oleh Tuhan yang Maha Esa.
5. Seluruh teman, pengurus, bapak kyai serta seluruh elemen Pondok Pesantren Sunan Pandanaran yang telah memberikan dukungan dan bantuan selama ini.
6. Serta semua pihak yang tidak bisa penulis sebutkan satu persatu.

KATA PENGANTAR

Dengan rasa syukur yang mendalam kepada Tuhan yang Maha Esa karena dengan rahmat-Nya saya dapat menyelesaikan penyusunan skripsi ini. Judul skripsi yang diajukan adalah "Perbandingan Algoritma SVM, MNB dan Ensemble Learning dalam Sentimen Analisis Terhadap *Cryptocurrency Crash Luna & UST*". Skripsi ini diajukan untuk memenuhi persyaratan kelulusan mata kuliah skripsi di Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.

Oleh karena itu, saya ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Allah SWT, berkat rahmat kasih sayang dan pertolongan-Nya saya bisa sampai sejauh ini.
2. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
3. Bapak Hanif Al Fatta, S.Kom., M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
4. Bapak Dony Ariyus, M.Kom. Selaku Kaprodi Teknik Komputer Universitas AMIKOM Yogyakarta sekaligus dosen pembimbing saya.
5. Dosen penguji dan segenap Dosen serta Karyawan Universitas AMIKOM Yogyakarta yang telah berbagi ilmu dan pengalamannya.
6. Kedua orang tua yang telah mendoakan, mendukung dan memberikan semangat kepada saya.
7. Semua pihak yang telah membantu dan tidak dapat disebutkan satu persatu.

Pada akhir paragraf, penulis menyadari bahwa tidak ada yang sempurna dan pasti terdapat kesalahan dalam penyusunan skripsi. Oleh karena itu, penulis meminta maaf atas segala kesalahan yang mungkin terjadi. Penulis berharap agar skripsi ini dapat bermanfaat bagi pembaca dan dapat dijadikan sebagai referensi untuk pengembangan ilmu yang lebih baik.

Yogyakarta, 6 Januari 2023

Penulis

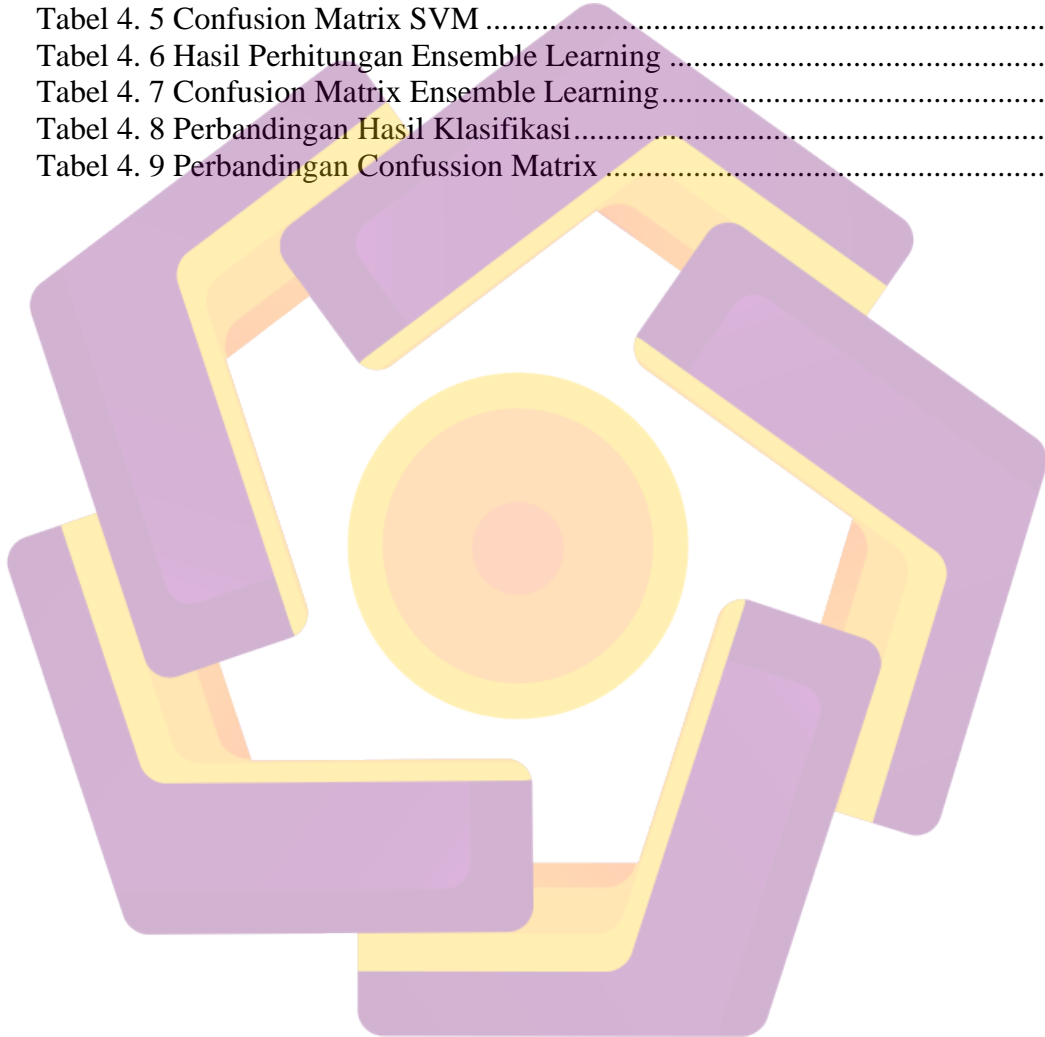
DAFTAR ISI

| | |
|---|----------|
| HALAMAN PERSETUJUAN..... | ii |
| HALAMAN PENGESAHAN | iii |
| HALAMAN PERNYATAAN KEASLIAN SKRIPSI | iv |
| HALAMAN PERSEMBAHAN | v |
| KATA PENGANTAR | vi |
| DAFTAR ISI..... | vii |
| DAFTAR TABEL..... | ix |
| DAFTAR GAMBAR | x |
| INTISARI | xi |
| ABSTRACT..... | xii |
| BAB I PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang..... | 1 |
| 1.2 Rumusan Masalah..... | 2 |
| 1.3 Batasan Masalah | 3 |
| 1.4 Tujuan Penelitian | 3 |
| 1.5 Manfaat Penelitian | 4 |
| 1.6 Sistematika Penulisan | 4 |
| BAB II TINJAUAN PUSTAKA | 6 |
| 2.1 Studi Literatur | 6 |
| 2.2 Dasar Teori | 10 |
| 2.2.1 Blockchain dan Cryptocurrency..... | 10 |
| 2.2.2 Terra (UST & Luna) | 11 |
| 2.2.3 Metode CRISP-DM | 11 |
| 2.2.4 Sentimen Analisis | 12 |
| 2.2.5 Sosial Media Twitter..... | 13 |
| 2.2.6 AI, ML dan DL | 13 |
| 2.2.7 Supervised Learning | 14 |
| 2.2.8 Unsupervised Learning | 14 |
| 2.2.9 Data Mining | 15 |
| 2.2.10 Text Mining | 15 |
| 2.2.11 Scraping Data..... | 15 |
| 2.2.12 Preprocessing Data..... | 16 |
| 2.2.13 Lexicon Based..... | 17 |
| 2.2.14 Naïve Bayes Classifier | 18 |
| 2.2.15 Support Vector Machine | 18 |

| | | |
|------------------|---|-----------|
| 2.2.16 | Ensemble Learning | 19 |
| 2.2.17 | Confusion Matrix | 20 |
| 2.2.18 | Accuracy | 20 |
| 2.2.19 | Precision..... | 20 |
| 2.2.20 | Recall | 21 |
| 2.2.21 | F1 Score | 21 |
| 2.2.22 | Word cloud..... | 21 |
| 2.2.23 | Receiver Operating Characteristic Curve | 22 |
| 2.2.24 | Area Under Curve | 23 |
| BAB III | METODE PENELITIAN | 24 |
| 3.1 | Alur Penelitian | 24 |
| 3.1.1 | Identifikasi masalah | 25 |
| 3.1.2 | Collecting Data | 25 |
| 3.1.3 | Data Understanding | 25 |
| 3.1.4 | Preprocessing Data..... | 26 |
| 3.1.5 | Splitting Data | 28 |
| 3.1.6 | Modelling Klasifikasi..... | 28 |
| 3.1.7 | Testing dan Validasi | 29 |
| 3.1.8 | Hasil | 29 |
| 3.2 | Alat dan Bahan..... | 30 |
| BAB IV | HASIL DAN PEMBAHASAN | 31 |
| 4.1 | Collecting Data | 31 |
| 4.2 | Preprocessing Data | 32 |
| 4.2.1 | Data Cleaning dan Case folding | 32 |
| 4.2.2 | POS tagging, Stopword removal dan Lemmatisasi | 34 |
| 4.2.3 | Labelling Data..... | 34 |
| 4.2.4 | Splitting Data dan Text Transformation | 35 |
| 4.3 | Model Klasifikasi MNB..... | 37 |
| 4.4 | Model Klasifikasi SVM..... | 40 |
| 4.5 | Model Klasifikasi Ensemble Learning | 43 |
| 4.6 | Perbandingan Klasifikasi | 46 |
| 4.7 | Visualisasi Kata Barplot dan WordCloud..... | 47 |
| BAB V | PENUTUP | 50 |
| 5.1 | Kesimpulan | 50 |
| 5.2 | Saran | 50 |
| REFERENSI | | 51 |

DAFTAR TABEL

| | |
|--|----|
| Tabel 2. 1 Referensi Penelitian | 6 |
| Tabel 3. 1 Tabel Alat | 30 |
| Tabel 3. 2 Tabel Bahan | 30 |
| Tabel 4. 1 Splitting Data | 35 |
| Tabel 4. 2 Hasil Perhitungan MNB..... | 38 |
| Tabel 4. 3 Tabel Confusion Matrix MNB..... | 38 |
| Tabel 4. 4 Hasil Perhitungan SVM | 41 |
| Tabel 4. 5 Confusion Matrix SVM | 41 |
| Tabel 4. 6 Hasil Perhitungan Ensemble Learning | 44 |
| Tabel 4. 7 Confusion Matrix Ensemble Learning..... | 44 |
| Tabel 4. 8 Perbandingan Hasil Klasifikasi..... | 46 |
| Tabel 4. 9 Perbandingan Confussion Matrix | 47 |



DAFTAR GAMBAR

| | |
|--|----|
| Gambar 2. 1 Tahapan CRISP-DM[19] | 12 |
| Gambar 2. 2 AI, ML dan DL[22]..... | 13 |
| Gambar 2. 3 Contoh Word Cloud | 22 |
| Gambar 2. 4 Contoh bentuk ROC curve | 23 |
| Gambar 3. 1 Alur Penelitian | 24 |
| Gambar 3. 2 Model Stacking Ensemble Learning | 28 |
| Gambar 4. 1 Informasi dataset yang berhasil di dapat..... | 31 |
| Gambar 4. 2 Dataset yang di dapat | 32 |
| Gambar 4. 3 Jumlah data yang memiliki duplikat | 33 |
| Gambar 4. 4 Data setelah dibersihkan dan case folding | 33 |
| Gambar 4. 5 Data yang sudah dilakukan POS tagging, Stopword removal dan Lemmatisasi..... | 34 |
| Gambar 4. 6 Jumlah data sentimen yang sudah dilabeli | 35 |
| Gambar 4. 7 Data yang Sudah Dilabeli | 35 |
| Gambar 4. 8 Splitting data dan pengubahan dalam bentuk matriks vektor | 36 |
| Gambar 4. 9 Label Sentimen yang Sudah Dikonversi | 36 |
| Gambar 4. 11 Confusion Matrix MNB | 38 |
| Gambar 4. 12 Kurva ROC pada model MNB | 39 |
| Gambar 4. 14 Confusion Matrix SVM | 41 |
| Gambar 4. 15 Kurva ROC SVM..... | 42 |
| Gambar 4. 17 Confusion Matrix Ensemble Learning | 45 |
| Gambar 4. 18 Kurva ROC Stacking Ensemble Learning | 46 |
| Gambar 4. 19 Kata yang Sering Muncul (Barplot)..... | 48 |
| Gambar 4. 20 Kata yang Sering Muncul (Wordcloud)..... | 48 |
| Gambar 4. 21 Wordcloud Kata Negatif | 49 |
| Gambar 4. 22 Wordcloud Kata Positif..... | 49 |

INTISARI

Penelitian ini bertujuan untuk melakukan perbandingan pada algoritma klasifikasi *Machine Learning* yaitu *Multinomial Naïve Bayes Classifier* (MNB), *Support Vector Machine* (SVM) dan *Stacking Ensemble Learning* dalam melakukan sentimen analisis pada media sosial twitter terkait insiden *cryptocurrency crash* yaitu UST dan Luna.

Data yang digunakan dalam penelitian ini adalah dataset sentimen yang diambil dengan teknik *scraping* pada sosial media Twitter sebanyak 30806 data tweet. Hasil pengujian menunjukkan bahwa model SVM memiliki performa terbaik dengan akurasi 90,3%, *precision* 90,6%, *recall* 88%, *F1 score* 95%, dan AUC 0,952. Di sisi lain, model MNB memiliki akurasi 81,7%, *precision* 79,7%, *recall* 80,8%, *F1 score* 88,7%, dan AUC 0,882. Sedangkan model Ensemble Learning memiliki akurasi 86,9%, *precision* 85%, *recall* 86,2%, *F1 score* 91,8%, dan AUC 0,904.

Dari hasil tersebut dapat disimpulkan bahwa model SVM adalah metode klasifikasi yang paling baik dalam melakukan klasifikasi pada data sentimen terkait insiden *cryptocurrency crash* UST dan Luna.

Kata kunci: Sentimen analisis, *Machine Learning*, *Support Vector Machine*, *Multinomial Naïve Bayes*, *Ensemble Learning*.

ABSTRACT

The study aims to compare machine learning classification algorithms such as Multinomial Naïve Bayes Classifier (MNB), Support Vector Machine (SVM) and Stacking Ensemble Learning in conducting sentimental analysis on social media twitter related to cryptocurrency crash incidents UST and Luna.

The data used in the study was a sentiment datasets taken with scraping techniques on social media Twitter of 30806 tweets data. Test results showed that the SVM model had the best performance with accuracy of 90.3%, precision of 90.6%, recall of 88%, F1 score of 95%, and AUC of 0.952. On the other hand, the MNB model has an accuracy of 81.7%, precision of 79.7%, recall of 80.8%, F1 score of 88.7%, and AUC of 0.882. The Ensemble Learning model has accuracy of 86.9%, precision of 85%, recall 86.2%, F1 score of 91.8%, and AUC of 0.904.

From the results, it can be concluded that the SVM model is the best classification method in conducting classification on sentimental data related to cryptocurrency crash incidents UST and Luna

Keyword: *Sentiment analysis, Machine Learning, Support Vector Machine, Multinomial Naïve Bayes, Ensemble Learning.*