

PROPOSAL SKRIPSI

**CLASSIFYING THE FIDELITY OF A CRYPTOCURRENCY USING
MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING**



Disusun oleh :

Gilang Islamay Putra Djuharis

11190940000055

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SYARIF HIDAYATULLAH

JAKARTA

2022 M / 1443 H

PROPOSAL SKRIPSI

CLASSIFYING THE FIDELITY OF A CRYPTOCURRENCY USING MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING

Diajukan oleh :

Gilang Islamay Putra Djuharis

11190940000055

Usulan Dosen Pembimbing

Dosen Pembimbing I

Taufik Edy Sutanto, M.Sc.Tech.,Ph.D

197905302006041002

Memiliki pendidikan terakhir S2 di *University of New South Wales Australia* serta beliau memiliki keahlian mengajar dalam bidang komputasi dan suka membimbing perwakilan UIN Syarif Hidayatullah Jakarta saat lomba.

Dosen Pembimbing II

Dr. Nur Inayah, S.Pd., M.Si.

197401252003122001

Memiliki pendidikan terakhir S3 di Institut Teknologi Bandung serta beliau memiliki keahlian mengajar dalam bidang Aljabar Graf dan sering melakukan penelitian seperti: *Strategi Pengamanan Data pada Server Komputer sebagai Implementasi Aplikatif Pelabelan Anti Ajaib Graf*, dll

Latar belakang

A. Perumusan Masalah

Perkembangan teknologi tidak hanya terlimitas pada dunia perindustrian, tetapi juga berdampingan juga dengan perkembangan teknologi pada dunia perekonomian dengan berkembangnya Cryptocurrency dimana dihilangkannya peran bank sehingga suatu sistem dapat bekerja secara mandiri tanpa harus campur tangan pihak tengah (perantara).

Cryptocurrency mengalami puncak tren pada bulan Mei 2021 (berdasarkan google trends) oleh karena itu banyak pula koin koin yang bermunculan bersamaan dengan kenaikan tren tersebut. Dengan banyaknya koin koin baru yang bermunculan, maka banyak juga koin koin yang dibangun bukan untuk investasi ataupun menjadi suatu alat tukar. Tetapi dibangun untuk keuntungan semata.

Rugpull adalah salah satu dari banyak penipuan yang terjadi pada dunia cryptocurrency, dimana investor dapat membeli koin tersebut akan tetapi tidak dapat menjualnya. Sehingga harga dari koin itu sendiri meningkat dan developer menjual seluruh kapitalisasi pasar dan mendapatkan keuntungan yang sangat amat besar. Contoh nyatanya adalah squid game koin, dimana developer menggunakan teknik rugpull dan mendapatkan keuntungan sebesar 3,38 juta dollar.

Dalam penelitian kali ini, penulis bertujuan untuk mendeteksi ketika suatu koin memiliki tendensi penipuan atau memiliki tendensi kejujuran dengan mengklasifikasikannya menggunakan peringkat dari A sampai F menggunakan analisis

terhadap whitepaper dari koin itu sendiri dan analisis terhadap faktor pendukung ataupun penentang dari koin itu sendiri.

B. Manfaat dan Tujuan Penelitian

Untuk mengetahui ketika suatu koin memiliki tendensi penipuan sehingga angka penipuan dalam dunia *cryptocurrency* menurun yang menyebabkan lebih banyak masyarakat yang tertarik untuk investasi dalam dunia *cryptocurrency*.

BAB II. TINJAUAN PUSTAKA DAN LANDASAN TEORI

A. *Cryptocurrency*

Konsep mata uang elektronik sudah ada sejak tahun 1990-an seperti DigiCash[1], B-Money, Hashcash, dll. Akan tetapi, pada tahun 2009 muncul mata uang elektronik dengan nama Bitcoin yang diluncurkan oleh pengembang anonim dengan nama Satoshi Nakamoto. Bitcoin adalah pelopor dari sistem *cryptocurrency*, dimana sistem mata uang digital yang didesain berkerja secara mandiri tanpa harus campur tangan pihak tengah (perantara) dengan sistem peer-to-peer yang menggunakan kriptografi dalam keamanannya. *Cryptocurrency* memiliki banyak sekali manfaatnya, antara lain:

1. Menghilangkan kerumitan dalam pemindahan mata uang dari satu negara ke negara lainnya.
2. Lebih aman, karena krypto menggunakan sistem peer-to-peer melalui beberapa peer sehingga ketika ada seseorang yang mencoba untuk mengubah value pada suatu transaksi maka akan ditolak karena valuenya berbeda dengan peer lainnya
3. Market krypto yang buka 24x7, berbeda dengan saham yang harus menunggu sekitar pukul 09:00 AM dan tutup pukul 04:00 PM
4. Terjamin transparan dan akuntabilitas, karena saat orang mengirimkan sesuatu melalui *blockchain* itu dapat terlihat akan tetapi tidak terlihat identitas aslinya. Tapi hanya adressnya yang berlaku sebagai nickname di dalam *blockchain* itu sendiri

B. *Naïve Bayes Classifier* [2]

Pengklasifikasi Naive Bayes adalah sekumpulan algoritma klasifikasi berdasarkan teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan decision tree and neural network. Ini bukan algoritma tunggal tetapi sekelompok

algoritma yang semuanya memiliki prinsip yang sama, yaitu setiap pasangan fitur diklasifikasikan sebagai independen satu sama lain. Pengklasifikasian ini dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Jika “d” adalah vektor masukan fitur dan “c” adalah label kelas, naïve Bayes menuliskan dengan $P(d|c)$. Bentuk umum dari Naïve Bayes Classifier antara lain:

$$P(d|c) = \frac{T_{ct} + \lambda}{N_c + \lambda V}$$

Dimana:

T_{ct} : Berapa kali kata “d” muncul dalam kelas “c”

λ : Nilai konstanta positif. Biasanya 1 untuk menghindari probabilitas nol

N_c : Jumlah kata dalam kelas “c”

V : Jumlah seluruh kata

C. Text Mining

Era informasi telah mempermudah penyimpanan data dalam jumlah besar. Dokumen yang tersedia di Web, di internet, di kabel berita, dan di tempat lain luar biasa. Namun, meskipun jumlah data yang tersedia terus meningkat, kemampuan untuk menyerap dan memproses informasi ini tetap konstan. *Search engine* hanya memperburuk masalah dengan membuat semakin banyak dokumen tersedia dalam beberapa penekanan tombol.

Text Mining adalah proses pengambilan teks yang tidak terstruktur menjadi yang terstruktur untuk dapat dianalisis dan diambil kesimpulannya. *Text Mining* adalah area penelitian baru dan menarik yang mencoba memecahkan masalah kelebihan informasi dengan menggunakan teknik dari Data Mining, Machine Learning, *Natural Language Processing* (NLP), *Information Retrival* (IR), dan manajemen pengetahuan.

Text Mining melibatkan pemrosesan awal yaitu pengumpulan dokumen (kategorisasi teks, ekstraksi informasi, serta ekstraksi istilah), penyimpanan

representasi perantara, teknik untuk menganalisis representasi perantara ini (seperti analisis distribusi, pengelompokan, analisis tren, dan asosiasi aturan), dan pada akhirnya visualisasi hasil. [3]

BAB III. METODOLOGI

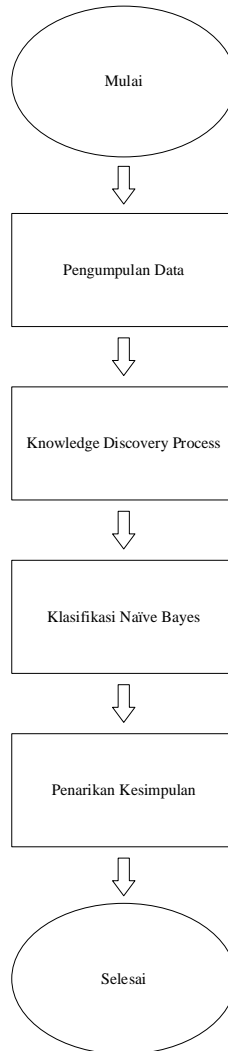
A. Sumber dan Tipe Data

Data yang digunakan adalah data yang bersifat sekunder dari whitepaper masing masing kripto.

B. Usulan Metode

Penelitian ini akan menggunakan metode Naïve Bayes Classifier

C. Diagram Alir Penelitian



BAB IV. JADWAL PENELITIAN

Tabel 1. Jadwal penelitian.

| No. | Keterangan | Bulan | | | | | |
|-----|-------------------------|-------|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| 1. | Studi Literatur | | | | | | |
| 2. | Menyusun Proposal | | | | | | |
| 3. | Pengumpulan Proposal | | | | | | |
| 4. | Pengumpulan Data | | | | | | |
| 5. | Analisis Data | | | | | | |
| 6. | Evaluasi dan Perbaikan | | | | | | |
| 7. | Pengajuan Hasil Laporan | | | | | | |

DAFTAR PUSTAKA

- [1] I. Amsyar, E. Christopher, A. Dithi, A. N. Khan, and S. Maulana, “The Challenge of Cryptocurrency in the Era of the Digital Revolution: A Review of Systematic Literature,” *Aptisi Trans. Technopreneursh.*, vol. 2, no. 2, pp. 153–159, 2020, doi: 10.34306/att.v2i2.96.
- [2] F. Rahutomo, I. Y. R. Pratiwi, and D. M. Ramadhani, “Eksperimen Naïve Bayes Pada Deteksi Berita Hoax Berbahasa Indonesia,” *J. Penelit. Komun. Dan Opini Publik*, vol. 23, no. 1, 2019, doi: 10.33299/jpkop.23.1.1805.
- [3] R. Feldman and J. Sanger, *The Text Mining Handbook*. 2006. doi: 10.1017/cbo9780511546914.
- [4] Y. Wang, J. Hodges, and B. Tang, “Classification of Web Documents Using a Naive Bayes Method,” *Proc. Int. Conf. Tools with Artif. Intell.*, no. Lcc, pp. 560–564, 2003, doi: 10.1109/tai.2003.1250241.
- [5] E. Jahani, P. M. Krafft, Y. Suhara, E. Moro, and A. S. Pentland, “ScamCoins, S*** posters, and the search for the next BitcoinTM: Collective sensemaking in cryptocurrency discussions,” *Proc. ACM Human-Computer Interact.*, vol. 2, no. CSCW, 2018, doi: 10.1145/3274348.