

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Blockchain pertama kali diperkenalkan melalui whitepaper “Bitcoin: *A Peer-to-Peer Electronic Cash System*” oleh Satoshi Nakamoto pada tahun 2008, dengan tujuan untuk memberikan solusi yang memungkinkan transaksi digital tanpa memerlukan pihak ketiga. Teknologi ini menawarkan konsep buku besar terdistribusi (*distributed ledger*) yang bersifat transparan, aman, dan tidak dapat diubah, menciptakan dasar yang kuat bagi lahirnya *cryptocurrency* pertama, yaitu Bitcoin (Larrier, 2021). Seiring dengan berjalannya waktu, *blockchain* berkembang melampaui Bitcoin, menginspirasi munculnya berbagai proyek *blockchain* lainnya, seperti Ethereum, yang memperkenalkan smart contract dan memperluas pengembangan blockchain ke sektor-sektor baru di luar keuangan (Lukianchuk, 2022).

Dalam perkembangannya, ekosistem di dalam *cryptocurrency* kini telah meluas pesat, meliputi ribuan aset digital yang mencakup berbagai fungsi dan tujuan mulai dari alat pembayaran, aset investasi, hingga utilitas dalam platform terdesentralisasi. Perkembangan ini didukung oleh inovasi dalam mekanisme konsensus, pengembangan aplikasi terdesentralisasi (DApp), serta penerapan blockchain di bidang-bidang lainnya seperti kesehatan, logistik, dan keamanan digital (Abrar & Sheikh, 2024). Meski menghadapi tantangan seperti volatilitas pasar dan ketidakpastian regulasi, kolaborasi antara pengembang, institusi keuangan, dan pemerintah menunjukkan bahwa *blockchain* dan *cryptocurrency* memiliki potensi besar untuk merevolusi sistem keuangan global (Grasic & Vidnjevic, 2024).

Whitepaper merupakan dokumen utama yang dirancang untuk memperkenalkan proyek *cryptocurrency* kepada calon investor. Biasanya, whitepaper memuat berbagai informasi seperti visi dan misi proyek, aspek teknis dari teknologi blockchain yang dikembangkan, tokenomics, serta strategi implementasi dan pengembangan masa depan. Dalam Konteks *Initial Coin Offering*

(ICO), *whitepaper* berperan penting dalam mengurangi asimetri informasi antara pengembang proyek dan calon investor. *Whitepaper* berfungsi layaknya prospectus sekuritas, bertujuan untuk membangun kepercayaan, menjamin transparansi, dan menarik investasi (Moser, 2025).

Karena peran vitalnya dalam memberikan gambaran kredibilitas sebuah proyek, calon investor sangat bergantung pada *whitepaper* untuk menilai kelayakan investasi. Penelitian menunjukkan bahwa panjang *whitepaper* berhubungan positif dengan jumlah dana yang berhasil dikumpulkan (*funding*) dan keberhasilan ICO (Samieifar & Baur, 2021). Namun, penting untuk mencermati bahwa kualitas bahasa dan struktur dokumen turut mempengaruhi keputusan investor, kesalahan linguistik seperti salah eja dan kesalahan tata bahasa dapat secara signifikan menurunkan minat investor (Thewissen et al., 2023). Oleh karena itu, kualitas penyajian informasi di dalam *whitepaper* menjadi indikator penting dalam proses pengambilan keputusan investasi.

Meskipun *whitepaper* berfungsi sebagai sumber utama informasi, banyak kasus di mana isinya bersifat menyesatkan atau mengandung elemen plagiarisme. Sebuah studi menemukan bahwa lebih dari separuh *whitepaper* proyek *crypto* menunjukkan tingkat penyalinan teks yang moderat dari proyek *cryptocurrency* lain, dengan *legal disclaimer* menjadi bagian yang paling sering disalin (Morin et al., 2021). Selain itu, walaupun beberapa studi melaporkan bahwa publikasi *whitepaper* meningkatkan jumlah investor, tidak selalu ada hubungan langsung dengan jumlah dana yang terkumpul (Belitski & Boreiko, 2022). Situasi ini menimbulkan kebutuhan mendesak untuk melakukan *information gathering* dan ekstraksi kata kunci pada *whitepaper* secara sistematis guna membantu calon investor dalam menilai kredibilitas proyek dengan lebih akurat.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka rumusan masalah dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana *whitepaper* proyek *crypto* dapat diekstraksi, disegmentasi, dan diklasifikasikan untuk mengenali struktur informasi penting yang relevan bagi calon investor?
2. Bagaimana pendekatan *information gathering*, ekstraksi kata kunci, deteksi kualitas linguistik, dan deteksi kemiripan teks dapat digunakan untuk penilaian awal *whitepaper* secara lebih sistematis?
3. Bagaimana hasil analisis *whitepaper* dapat disajikan dalam bentuk penilaian yang lebih terstruktur agar membantu investor melakukan penyaringan awal proyek secara lebih cepat dan konsisten?

1.3 Tujuan Penelitian

1. Mengembangkan *pipeline* NLP untuk mengekstraksi, menyegmentasi, dan mengklasifikasikan informasi penting yang terdapat dalam *whitepaper* proyek *cryptocurrency*.
2. Membangun mekanisme penilaian *whitepaper* berbasis beberapa sinyal, yaitu struktur dokumen, kualitas linguistik, kemiripan teks (plagiarisme), dan hasil ekstraksi kata kunci.
3. Menghasilkan sistem bantu yang dapat menyajikan penilaian awal *whitepaper* secara lebih terstruktur untuk mendukung proses *screening* atau penyaringan awal proyek *crypto*.

1.4 Manfaat Penelitian

1. Memberikan kerangka analisis yang lebih terstruktur untuk memahami isi *whitepaper* proyek *crypto* melalui identifikasi *section* penting dan sinyal-sinyal penilaian.
2. Menyediakan dasar pengembangan sistem analitik otomatis yang dapat menilai *whitepaper* secara lebih cepat, konsisten, dan terstruktur.
3. Membantu calon investor melakukan penyaringan awal terhadap proyek *crypto* secara lebih objektif sehingga resiko akibat informasi yang menyesatkan dapat ditekan.

1.5 Batasan Penelitian

Untuk menjaga fokus dan kedalaman analisis, penelitian ini dibatasi pada hal-hal berikut:

1. Penelitian hanya berfokus pada whitepaper proyek cryptocurrency yang tersedia dalam bahasa Inggris dan telah dipublikasikan secara terbuka di internet.
2. Analisis terbatas pada aspek informasi yang bersifat deskriptif dalam whitepaper, seperti visi proyek, tokenomics, teknologi yang digunakan, dan struktur tim, tanpa melakukan verifikasi eksternal terhadap kebenaran informasi tersebut.
3. Teknik *information gathering* dan *keyword extraction* dilakukan secara tekstual menggunakan metode pemrosesan bahasa alami (Natural Language Processing/NLP), tanpa mempertimbangkan elemen visual seperti diagram, tabel, atau grafik dalam whitepaper.
4. Penelitian tidak membahas aspek teknis mendalam dari blockchain yang digunakan dalam proyek, melainkan hanya sebatas bagaimana informasi tersebut disajikan dalam whitepaper.
5. Penelitian tidak menganalisis performa harga token atau indikator pasar lainnya secara langsung, melainkan berfokus pada analisis isi whitepaper sebagai dasar awal penilaian kredibilitas proyek.