

# **BAB I**

## **PENDAHULUAN**

Bab I adalah pendahuluan yang menceritakan keseluruhan isi dari tugas akhir. Bab ini diawali dengan penjelasan mengenai topik dan latar belakang tugas akhir, dilanjuti dengan rumusan masalah, tujuan tugas akhir, serta batasan masalah yang dibahas di tugas akhir. Selain itu, Bab I akan menjelaskan tentang pendekatan dan metode yang digunakan di metodologi.

### **I.1 Latar Belakang**

Revolusi digital telah memicu ledakan produksi data; volume data global diproyeksikan akan mencapai 181 zettabyte pada tahun 2025, meningkat drastis dari hanya 2 zettabyte pada tahun 2010 (International Data Corporation March 2021). Pertumbuhan ini melahirkan tantangan *information overload*, di mana volume informasi melampaui kapasitas kognitif individu untuk memprosesnya (Eppler dan Mengis 2004). Menurut *Cognitive Load Theory*, kapasitas memori kerja manusia yang terbatas membuat kualitas pengambilan keputusan dapat menurun hingga 50% saat dihadapkan pada beban informasi berlebih (Sweller 1988; Arnold, Goldschmitt, dan Rigotti June 2023).

Masalah ini menjadi akut dalam komunitas online di media sosial, di mana laju pesan yang tinggi dapat menenggelamkan informasi krusial (Nematzadeh dkk. 2019). Konsekuensinya adalah atensi selektif hingga kelumpuhan dalam pengambilan keputusan (Edmunds dan Morris 2000). Fenomena ini relevan di Indonesia, di mana platform seperti Telegram telah menjadi ekosistem dominan untuk komunitas finansial berskala besar. Didukung oleh fitur seperti kapasitas grup hingga 200.000 anggota, Telegram menjadi medium pilihan dengan lebih dari 27 juta pengguna di Indonesia untuk diskusi investasi *real-time* (AppMagic 2024; Perlo dkk. 2025).

Dominasi Telegram ini berjalan paralel dengan lonjakan jumlah investor ritel di In-

donesia, yang mencapai lebih dari 6 juta investor pada September 2024, dengan pertumbuhan 744.000 investor baru di tahun tersebut (Bursa Efek Indonesia September 2024). Mayoritas investor baru ini adalah generasi muda yang sangat bergantung pada komunitas media sosial untuk pengambilan keputusan investasi (Otoritas Jasa Keuangan 2022; Junaidi dan Nurhidayah 2023). Ketergantungan ini menciptakan masalah dimana akses informasi justru berujung pada *information overload* yang parah, sehingga investor kesulitan memilah sinyal berkualitas dari kebisingan informasi. Tantangan ini terefleksi secara nyata dalam komunitas finansial Michael Yeoh, yang memiliki lebih dari 22.000 anggota dalam struktur berjenjang (*multi-tiered*), di mana fragmentasi informasi antar grup semakin memperburuk masalah.

Penelitian akademis sebelumnya cenderung berfokus pada analisis sentimen dari forum publik seperti X (Twitter) dan Reddit untuk tujuan prediksi pergerakan harga saham (misalnya, Bollen, Mao, dan Zeng (2011); Pawar dkk. (2022); Cheng, Ge, dan Li (2023)). Sementara itu, solusi praktis seperti bot Telegram yang ada saat ini bersifat generik, tidak mampu melakukan sintesis lintas-grup, dan tidak dioptimalkan untuk domain finansial Indonesia yang spesifik (misalnya, Xu dkk. (2021); Beebom (2023)). Kedua pendekatan tersebut tidak dirancang untuk menyelesaikan masalah operasional internal sebuah komunitas.

Kondisi *information overload* parah dan fragmentasi informasi akibat struktur komunitas berjenjang ini menyoroti permasalahan signifikan yang perlu diteliti. Fenomena ini tidak hanya berisiko mengurangi efektivitas anggota dalam mengambil keputusan investasi, tetapi juga mengikis nilai dari komunitas itu sendiri.

## I.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan pada subbab I.1, rumusan masalah yang akan dibahas pada tugas akhir ini adalah:

1. Bagaimana merancang dan mengembangkan sistem rangkuman otomatis berbasis *AI* yang dapat mereduksi *information overload* melalui sintesis intelijen lintas-grup dalam komunitas finansial Telegram berjenjang?
2. Bagaimana meningkatkan efektivitas sistem rangkuman otomatis tersebut berdasarkan *feedback* dan evaluasi dari pengguna komunitas?

## I.3 Tujuan

Berdasarkan latar belakang dan rumusan masalah yang telah dijelaskan, tujuan dari tugas akhir ini adalah:

1. Merancang, membangun, dan mengevaluasi sistem *prototype* rangkuman otomatis berbasis *AI* yang mampu melakukan sintesis intelijen lintas-grup untuk mereduksi *information overload* dalam komunitas finansial Telegram.
2. Melakukan iterasi perbaikan sistem berdasarkan *feedback* pengguna untuk meningkatkan efektivitas dan *usability* sistem dalam konteks komunitas yang diteliti.

#### I.4 Batasan Masalah

Batasan masalah digunakan untuk menjaga pelebaran dari pokok masalah dan mengarahkan tugas akhir menjadi lebih spesifik. Batasan masalah dari tugas akhir ini adalah:

1. Konteks penelitian dan implementasi sistem terbatas secara eksklusif pada ekosistem komunitas investasi Michael Yeoh di Telegram, yang terdiri dari tiga grup utama: MY Cuap Cuap, MY Swing Plan, dan MY Advanced Group.
2. Arsitektur sistem menggunakan tumpukan teknologi (*technology stack*) yang telah ditentukan, meliputi model NLP *indonesian-roberta-base-emotion-classifier* dan *bert-base-indonesian-NER*, serta pemanfaatan Groq *API* dengan model bahasa Llama 3.1 70B untuk proses rangkuman.
3. Keluaran (*output*) utama dari sistem adalah rangkuman terstruktur yang dihasilkan secara periodik setiap jam (*hourly summary*). Penelitian ini tidak mencakup pengembangan sistem peringatan (*alert*) *real-time* atau fitur perakapan interaktif.
4. Sumber data untuk analisis terbatas pada pesan berbasis teks yang ada di dalam grup-grup Telegram yang ditentukan. Sistem tidak mengintegrasikan sumber data eksternal seperti *API* harga saham *real-time* atau umpan berita finansial dari portal media.

#### I.5 Metodologi

Pengerjaan tugas akhir ini menggunakan dua metodologi, yaitu metodologi untuk penelitian dan metodologi pengembangan.

##### I.5.1 Metodologi Penelitian

Metodologi penelitian yang digunakan adalah *Design Science Research Methodology* (DSRM). Pendekatan DSRM dipilih karena relevansinya yang tinggi dengan penelitian di bidang sistem informasi yang bertujuan untuk merancang dan menge-

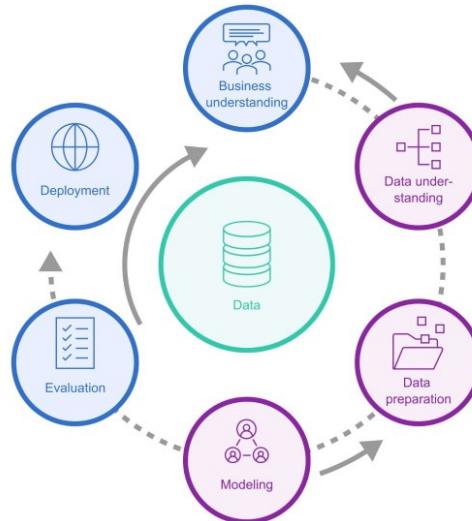
valuasi artefak teknologi inovatif guna memecahkan masalah di dunia nyata. Proses DSRM terdiri dari enam tahapan yang akan diikuti dalam penelitian ini:

1. **Identifikasi Masalah & Motivasi:** Tahap ini telah diuraikan dalam latar belakang, di mana masalah spesifik yang diidentifikasi adalah *information overload* dan fragmentasi informasi dalam komunitas finansial Telegram yang besar dan berjenjang. Motivasi utamanya adalah kebutuhan untuk meningkatkan efisiensi akses informasi bagi anggota komunitas.
2. **Perumusan Tujuan Solusi:** Berdasarkan masalah yang ada, tujuan solusi definisikan secara kuantitatif dan kualitatif. Tujuan utamanya adalah menghasilkan sebuah artefak (sistem) yang terbukti mampu mereduksi beban informasi, yang akan diukur melalui akurasi identifikasi topik dan efisiensi akses informasi yang dirasakan pengguna.
3. **Desain & Pengembangan:** Tahap ini merupakan aktivitas inti dalam menciptakan artefak penelitian. Pada tugas akhir ini, artefak yang dirancang dan dikembangkan adalah sebuah sistem purwarupa untuk rangkuman otomatis. Proses ini mencakup perancangan arsitektur, implementasi *pipeline* pemrosesan data, dan integrasi model *NLP*.
4. **Demonstrasi:** Pada tahap ini, artefak yang telah dibangun akan didemonstrasikan untuk menyelesaikan masalah yang telah diidentifikasi. Sistem purwarupa akan dioperasikan menggunakan data percakapan nyata dari komunitas Michael Yeoh untuk menunjukkan kemampuannya dalam mengubah data mentah menjadi rangkuman yang terstruktur.
5. **Evaluasi:** Artefak akan dievaluasi secara sistematis untuk mengukur sejauh mana tujuan solusi tercapai. Evaluasi akan membandingkan kondisi sebelum dan sesudah penggunaan sistem, menggunakan metrik yang telah ditetapkan seperti akurasi, tingkat kompresi pesan, dan umpan balik kualitatif dari pengguna melalui survei.
6. **Komunikasi:** Tahap terakhir adalah mengkomunikasikan masalah, artefak yang dibangun, serta hasil evaluasinya kepada audiens akademis. Komunikasi ini diwujudkan melalui penulisan laporan tugas akhir ini serta presentasi dalam forum seminar dan sidang.

### 1.5.2 Metodologi Pengembangan

Metodologi pengembangan yang digunakan dalam penelitian ini adalah *Cross Industry Standard Process for Data Mining* (CRISP-DM). Metodologi CRISP-DM dipilih karena relevansinya yang tinggi untuk proyek berbasis *machine learning* dan *natural language processing*, serta kemampuannya dalam memberikan kerang-

ka kerja terstruktur untuk proyek *data mining* di berbagai industri (Chapman dkk. 2000). CRISP-DM terdiri dari enam fase utama yang bersifat iteratif dan siklik, memungkinkan pengulangan dan penyempurnaan di setiap tahap pengembangan. Metodologi CRISP-DM dapat dilihat pada Gambar I.1.



Gambar I.1 Siklus metodologi CRISP-DM (Chapman dkk. 2000)

Metodologi CRISP-DM yang diadaptasi untuk pengembangan sistem rangkuman otomatis ini terdiri dari enam fase berikut:

1. **Business Understanding:** Memahami kebutuhan komunitas finansial Telegram dan mendefinisikan tujuan proyek dari perspektif bisnis, yaitu mereduksi *information overload* melalui rangkuman otomatis.
2. **Data Understanding:** Mengumpulkan dan menganalisis data pesan dari tiga grup Telegram target untuk memahami karakteristik, volume, dan pola komunikasi dalam komunitas.
3. **Data Preparation:** Membersihkan, mentransformasi, dan mempersiapkan data pesan untuk pemrosesan NLP, termasuk normalisasi teks dan ekstraksi fitur linguistik.
4. **Modeling:** Mengintegrasikan model NLP untuk analisis sentimen, ekstraksi entitas, identifikasi topik, dan generasi rangkuman menggunakan *large language model*.
5. **Evaluation:** Mengevaluasi kualitas rangkuman yang dihasilkan sistem melalui metrik objektif dan *feedback* subjektif dari pengguna komunitas.
6. **Deployment:** Mengimplementasikan sistem secara operasional dalam komunitas, termasuk penjadwalan otomatis dan penyampaian rangkuman periodik

ke kanal Telegram.

Penjelasan detail mengenai aktivitas dan keluaran dari setiap fase metodologi CRISP-DM dalam konteks penelitian ini akan diuraikan lebih lanjut pada Bab III (Analisis dan Perancangan Sistem).

## **BAB II**

### **STUDI LITERATUR**

Bab ini menyajikan tinjauan komprehensif terhadap literatur yang relevan dengan pengembangan sistem rangkuman otomatis berbasis *AI* untuk mereduksi *information overload* dalam komunitas finansial Telegram. Studi literatur ini mencakup tujuh domain utama: (1) *Information Overload* dan *Cognitive Load Theory*, (2) *Text Summarization* dengan *Large Language Models*, (3) Analisis Sentimen untuk Media Sosial Finansial, (4) *Named Entity Recognition* untuk Teks Finansial Indonesia, (5) *Topic Modeling* dengan BERTopic, (6) Telegram sebagai Platform Komunitas, dan (7) Arsitektur Pemrosesan Data *Real-Time*. Setiap domain memberikan landasan teoretis dan metodologis yang mendukung desain, implementasi, dan evaluasi sistem yang dikembangkan dalam penelitian ini.

#### **II.1 *Information Overload* dan *Cognitive Load Theory***

*Information overload* merupakan fenomena di mana individu dihadapkan pada volume informasi yang melampaui kapasitas kognitif mereka untuk memproses secara efektif (Eppler dan Mengis 2004). Dalam konteks komunitas online, khususnya platform media sosial seperti Telegram, masalah ini menjadi semakin akut karena tingginya laju produksi dan distribusi informasi. Bagian ini mengulas teori-teori fundamental yang menjelaskan bagaimana *information overload* memengaruhi pengambilan keputusan, serta mengidentifikasi strategi mitigasi yang dapat diterapkan melalui sistem otomatis berbasis *AI*.

##### **II.1.1 Konsep dan Dampak *Information Overload***

Penelitian seminal oleh Eppler dan Mengis (2004) menyediakan landasan teoretis yang komprehensif tentang konsep *information overload*. Berdasarkan tinjauan literatur dari berbagai disiplin ilmu termasuk *Organization Science*, Akuntansi, Pemasaran, dan Sistem Informasi Manajemen, mereka mendefinisikan *information*

*overload* sebagai kondisi di mana kualitas pengambilan keputusan menurun ketika kuantitas informasi yang tersedia melampaui kapasitas pemrosesan individu. Studi tersebut mengidentifikasi hubungan inverted U-curve antara kuantitas informasi dan kualitas keputusan, yang menunjukkan bahwa setelah mencapai titik optimal, penambahan informasi justru menurunkan performa kognitif. Temuan ini sangat relevan dengan konteks komunitas Telegram finansial yang menghasilkan ratusan hingga ribuan pesan per hari, yang berpotensi menempatkan anggota komunitas pada sisi kanan kurva tersebut di mana *overload* terjadi.

Shahrzadi dkk. (2024) memperkuat pemahaman ini melalui *scoping review* terbaru yang menganalisis penyebab, konsekuensi, dan strategi penanganan *information overload*. Penelitian mereka mengidentifikasi tiga kategori penyebab utama: (1) faktor personal seperti keterbatasan kapasitas kognitif dan kurangnya keterampilan manajemen informasi, (2) karakteristik informasi seperti ambiguitas, kompleksitas, dan ketidakpastian, dan (3) parameter teknologi informasi termasuk intensitas notifikasi dan desain antarmuka. Dampak yang diidentifikasi mencakup penurunan produktivitas, kualitas keputusan yang buruk, stres, dan kelelahan kognitif. Yang paling signifikan bagi penelitian ini adalah identifikasi mereka terhadap strategi mitigasi berbasis teknologi, khususnya penggunaan *filtering tools* dan sistem otomatis untuk mereduksi volume informasi, yang memberikan justifikasi langsung untuk pengembangan bot rangkuman otomatis.

### **II.1.2 Landasan Teori: *Cognitive Load Theory***

*Cognitive Load Theory* (CLT), yang dikembangkan oleh Sweller (1988), memberikan kerangka teoretis untuk memahami keterbatasan kapasitas memori kerja manusia dalam pemrosesan informasi. Teori ini mengidentifikasi tiga jenis beban kognitif: (1) *intrinsic load*, yang inheren dengan kompleksitas materi yang dipelajari, (2) *extraneous load*, yang dihasilkan oleh cara informasi disajikan, dan (3) *germane load*, yang berkontribusi pada pembentukan skema dan otomatisasi. Dalam konteks komunitas online, anggota menghadapi *intrinsic load* yang tinggi karena kompleksitas informasi finansial, sementara desain antarmuka platform media sosial sering kali meningkatkan *extraneous load* melalui notifikasi, pesan yang tidak terstruktur, dan kurangnya organisasi informasi.

Skulmowski dan Rey (2022) mengembangkan pemahaman CLT dengan mengadaptasinya untuk lingkungan digital dan pembelajaran daring. Penelitian mereka mengusulkan perspektif baru tentang *extraneous cognitive load* yang mempertimbangkan karakteristik unik media digital, seperti konten yang kaya secara perceptual, distrak-

si visual, dan desain antarmuka yang kompleks. Mereka berargumen bahwa media digital menciptakan bentuk *extraneous load* yang berbeda dari media tradisional, terutama melalui elemen-elemen yang tidak relevan dengan tujuan pembelajaran atau pemrosesan informasi inti. Temuan ini sangat relevan untuk memahami bagaimana antarmuka Telegram yang penuh dengan pesan, emoji, dan media visual dapat meningkatkan beban kognitif ekstrinsik bagi anggota yang mencoba meng-ekstrak informasi investasi yang relevan.

Penelitian empiris oleh Sidnam-Mauch dan Monge (2024) menerapkan CLT secara khusus pada konteks distraksi media sosial melalui studi dengan 1.026 partisipan. Mereka mengintegrasikan *load theory of attention* dengan kerangka *uses and gratifications* untuk menjelaskan bagaimana media sosial mobile menciptakan distraksi melalui sumber daya kognitif yang terbatas. Studi mereka mendemonstrasikan bahwa individu dengan kapasitas memori kerja yang lebih rendah lebih rentan terhadap distraksi media sosial, dan bahwa *affordances* teknologi seperti notifikasi *push* dan aksesibilitas konstan memperburuk masalah ini. Temuan ini memperkuat argumen bahwa komunitas Telegram finansial dengan lebih dari 22.000 anggota dan aliran pesan yang konstan menciptakan lingkungan yang sangat rentan terhadap *cognitive overload* dan distraksi, yang pada akhirnya dapat menurunkan kualitas pengambilan keputusan investasi.

### **II.1.3 Ekonomi Perhatian di Media Sosial**

Heitmayer (2025) memberikan perspektif teoretis terbaru dengan mengusulkan bahwa perhatian (*attention*) berfungsi sebagai mata uang simbolis universal di media sosial. Penelitian ini memperkenalkan model *dual-stream* yang membedakan antara perhatian yang "mengalir" (*flow attention*) dan perhatian yang "mengeras" (*calcified attention*). *Flow attention* mengacu pada alokasi perhatian yang dinamis dan sementara terhadap konten yang terus berubah, sementara *calcified attention* merujuk pada investasi perhatian jangka panjang yang termanifestasi dalam bentuk *followers*, *likes*, atau keanggotaan komunitas. Dalam konteks komunitas finansial Telegram, kedua bentuk perhatian ini bersaing: anggota harus mengalokasikan *flow attention* mereka untuk memproses aliran pesan yang konstan, sementara *calcified attention* mereka termanifestasi dalam komitmen berkelanjutan terhadap komunitas dan kepercayaan terhadap sumber informasi tertentu seperti Michael Yeoh. Kelangkaan perhatian (*attention scarcity*) yang diidentifikasi dalam penelitian ini menjadi masalah sentral yang harus diatasi oleh sistem rangkuman otomatis dengan memaksimalkan efisiensi alokasi perhatian anggota.

Analisis empiris oleh Nematzadeh dkk. (2019) tentang *information overload* dalam komunikasi grup menggunakan data dari platform Twitch menunjukkan bagaimana peningkatan volume pesan dalam grup dapat menyebabkan transisi dari percakapan yang koheren menjadi "kakofoni" di mana informasi bermakna tenggelam dalam kebisingan. Mereka mengidentifikasi bahwa ketika laju pesan melebihi ambang batas tertentu, kualitas percakapan menurun drastis, partisipasi menjadi tidak merata, dan informasi penting hilang. Fenomena ini sangat relevan dengan komunitas Telegram finansial yang diteliti, di mana pesan dapat muncul dengan laju ratusan per jam selama jam perdagangan aktif. Sistem rangkuman otomatis yang dirancang dalam penelitian ini bertujuan untuk mengatasi transisi menuju kakofoni ini dengan menyaring dan mensintesis informasi kunci secara periodik.

## **II.2 *Text Summarization* dengan Large Language Models**

Peringkasan teks otomatis telah mengalami evolusi signifikan dengan munculnya model bahasa berbasis *transformer* dan *large language models* (LLMs). Bagian ini mengulas perkembangan metodologi peringkasan teks dari pendekatan berbasis ekstraksi hingga generasi abstraktif dengan LLMs, dengan fokus khusus pada tantangan peringkasan multi-dokumen dan percakapan yang relevan dengan konteks komunitas Telegram.

### **II.2.1 Peringkasan Berbasis *Transformer***

Liu dan Lapata (2019) memperkenalkan pendekatan *BERTSum* yang memanfaatkan representasi BERT untuk peringkasan ekstraktif dan abstraktif. Penelitian mereka mendemonstrasikan bahwa model *encoder* tingkat dokumen yang dibangun di atas BERT dengan lapisan *Transformer* antar-kalimat dapat menghasilkan rangkuman berkualitas tinggi. Pendekatan dua tahap *fine-tuning* yang mereka usulkan, di mana model pertama-tama di-*fine-tune* untuk tugas ekstraksi kalimat sebelum dilatih untuk generasi abstraktif, menetapkan metodologi dasar yang diadopsi oleh banyak penelitian berikutnya. Kontribusi kunci dari penelitian ini adalah demonstrasi bahwa model bahasa pra-latih dapat secara efektif menangkap struktur dokumen dan hubungan semantik antar-kalimat, yang esensial untuk menghasilkan rangkuman yang koheren dan informatif.

Pemahaman tentang *trade-off* antara peringkasan ekstraktif dan abstraktif diperluas oleh penelitian Pilault dkk. (2020), yang membandingkan kedua pendekatan menggunakan model *transformer* pada dokumen panjang. Evaluasi manusia mereka mengungkapkan temuan penting: sementara pendekatan abstraktif berbasis *tran-*

*sfomer* unggul dalam koherensi dan kelancaran, metode ekstraktif cenderung mencetak lebih tinggi dalam hal informativeness. Temuan ini menginformasikan desain sistem dalam penelitian ini, di mana pendekatan hibrida digunakan: BERTopic untuk ekstraksi topik dan entitas kunci (pendekatan ekstraktif), dikombinasikan dengan Llama 3.1 70B untuk generasi rangkuman naratif yang koheren (pendekatan abstraktif).

### **II.2.2 Peringkasan Multi-Dokumen dan Sintesis Informasi**

Tantangan peringkasan multi-dokumen, yang sangat relevan dengan konteks sintesis lintas tiga grup Telegram, dibahas secara komprehensif oleh DeYoung dkk. (2024). Penelitian mereka mengajukan pertanyaan fundamental: apakah model peringkasan multi-dokumen modern benar-benar dapat mensintesis informasi, atau hanya melakukan agregasi sederhana? Melalui evaluasi terhadap berbagai model dari *fine-tuned transformers* hingga GPT-4, mereka menemukan bahwa sebagian besar model mengalami kesulitan dalam mensintesis informasi yang bertentangan atau berbeda perspektif, dan sangat sensitif terhadap urutan input dokumen. Temuan kritis ini memiliki implikasi langsung bagi desain sistem dalam penelitian ini: sistem harus dirancang dengan mekanisme eksplisit untuk menangani informasi yang mungkin bertentangan antara grup "Cuap Cuap", "Swing Plan", dan "Advanced Group", dan tidak boleh bergantung semata-mata pada kemampuan sintesis implisit LLM.

Ravaut dkk. (2024) menyediakan wawasan penting tentang bagaimana LLMs memanfaatkan konteks panjang dalam peringkasan. Penelitian mereka mengungkapkan masalah *position bias*, di mana LLMs cenderung memberikan bobot yang tidak merata terhadap informasi berdasarkan posisinya dalam konteks. Mereka meng-evaluasi enam LLMs berbeda pada sepuluh dataset dengan lima metrik evaluasi, menemukan bahwa model cenderung memberikan perhatian berlebihan pada informasi di awal dan akhir konteks, sementara mengabaikan informasi di tengah. Untuk mengatasi masalah ini, mereka memperkenalkan benchmark *MiddleSum* dan meng-evaluasi metode hierarkis dan inkremental untuk peringkasan. Temuan ini sangat relevan untuk implementasi Llama 3.1 70B dalam penelitian ini, yang harus memproses konteks panjang (potensial ratusan pesan dalam satu jam). Desain sistem harus mempertimbangkan strategi mitigasi *position bias*, seperti pengelompokan pesan berdasarkan topik sebelum peringkasan atau penggunaan rangkuman inkremental.

### **II.2.3 Peringkasan Percakapan dan Dialog**

Gliwa dkk. (2019) memperkenalkan *SAMSum Corpus*, dataset dialog yang dianotasi secara manual untuk peringkasan abstraktif, yang menetapkan bahwa peringkasan percakapan memiliki tantangan unik dibandingkan dengan peringkasan artikel berita. Dialog dan percakapan grup memiliki struktur yang lebih kompleks dengan pergantian pembicara, konteks implisit, referensi anaforik, dan topik yang bercabang. Dataset mereka mendemonstrasikan pentingnya memahami struktur dialog dan dinamika percakapan untuk menghasilkan rangkuman yang akurat dan informatif. Karakteristik ini sangat relevan dengan konteks Telegram, di mana percakapan dapat melibatkan puluhan pembicara dengan topik yang berubah cepat dan sering kali tumpang tindih.

Tian, Xia, dan Song (2024) mengembangkan metodologi peringkasan dialog lebih lanjut dengan mengusulkan pendekatan *Mixture of Experts* (MoE) berbasis LLM. Penelitian mereka memperkenalkan mekanisme *role-oriented routing* yang memilih model atau strategi peringkasan yang sesuai berdasarkan karakteristik informasi percakapan yang berbeda. Pendekatan ini sangat relevan dengan konteks komunitas finansial Telegram yang berjenjang, di mana pesan dari berbagai peran (misalnya, Michael Yeoh sebagai pemimpin komunitas versus anggota biasa) mungkin memerlukan perlakuan berbeda dalam proses peringkasan. Sistem yang dirancang dalam penelitian ini mengadopsi konsep serupa melalui *weighted sentiment analysis* yang memberikan bobot berbeda berdasarkan peran pengirim pesan.

### **II.2.4 Memori Jangka Panjang dalam Sistem Dialog**

Konsep memori jangka panjang dalam sistem percakapan dieksplorasi oleh Wang dkk. (2023), yang mengusulkan metode peringkasan rekursif untuk memungkinkan LLMs mempertahankan konteks percakapan yang panjang. Pendekatan mereka melibatkan pembuatan rangkuman hierarkis dari konteks dialog yang kemudian digunakan untuk menjaga konsistensi dalam percakapan yang berkepanjangan. Konsep ini paralel dengan desain sistem rangkuman per jam dalam penelitian ini, di mana setiap rangkuman per jam dapat dipandang sebagai pembentukan "memori jangka panjang" bagi komunitas. Akumulasi rangkuman per jam ini dapat membentuk representasi terkompresi dari diskusi komunitas yang memungkinkan anggota untuk memahami tren dan perkembangan informasi tanpa harus membaca setiap pesan individual.

### **II.2.5 Metrik Evaluasi untuk Peringkasan**

Evaluasi kualitas rangkuman otomatis merupakan aspek kritis yang dibahas oleh Zhang dkk. (2020). Mereka memperkenalkan *BERTScore*, metrik evaluasi yang menggunakan *contextual embeddings* dari BERT untuk menghitung kesamaan antara rangkuman yang dihasilkan dan rangkuman referensi, mengatasi keterbatasan metrik berbasis *n-gram* seperti ROUGE yang hanya menghitung kecocokan kata eksak. *BERTScore* mendemonstrasikan korelasi yang lebih baik dengan penilaian manusia karena dapat mengenali parafrase dan kesamaan semantik yang tidak tertangkap oleh metrik tradisional. Penelitian ini mengadopsi *BERTScore* bersama dengan metrik lain untuk evaluasi komprehensif kualitas rangkuman yang dihasilkan sistem.

Namun, Shen dkk. (2023) memberikan peringatan penting tentang penggunaan LLMs sebagai evaluator otomatis untuk peringkasan abstraktif. Melalui analisis ekstensif terhadap ChatGPT dan GPT-4 sebagai evaluator, mereka menemukan masalah signifikan dalam stabilitas, reliabilitas, dan ketergantungan pada dimensi evaluasi tertentu. Temuan mereka menunjukkan bahwa evaluasi otomatis berbasis LLM tidak dapat sepenuhnya menggantikan evaluasi manusia, terutama untuk aspek kualitatif seperti koherensi naratif dan relevansi kontekstual. Implikasi bagi penelitian ini adalah perlunya kombinasi antara metrik otomatis (*BERTScore*, ROUGE) dan evaluasi subjektif melalui *feedback* pengguna komunitas untuk mendapatkan penilaian holistik terhadap kualitas rangkuman.

## **II.3 Analisis Sentimen untuk Media Sosial Finansial**

Analisis sentimen merupakan komponen esensial dalam memahami dinamika diskusi finansial di media sosial. Bagian ini mengulas perkembangan metodologi analisis sentimen, dengan fokus khusus pada domain finansial dan pemrosesan bahasa Indonesia, serta mengeksplorasi bagaimana sinyal sentimen dari komunitas dapat dimanfaatkan untuk meningkatkan kualitas rangkuman.

### **II.3.1 Model Bahasa untuk Analisis Sentimen Indonesia**

Wilie dkk. (2020) memperkenalkan *IndoNLU*, benchmark komprehensif untuk evaluasi pemahaman bahasa Indonesia yang mencakup dua belas tugas termasuk analisis sentimen. Dataset mereka dilatih pada *Indo4B*, korpus bahasa Indonesia dengan empat miliar kata, dan menyediakan model IndoBERT yang mengungguli *multilingual BERT* (mBERT) pada berbagai tugas. Penelitian ini mendemonstra-

sikan bahwa model bahasa monolingual yang dilatih khusus pada korpus bahasa Indonesia menghasilkan performa superior dibandingkan model multilingual untuk tugas-tugas NLP Indonesia. Kontribusi mereka menetapkan IndoBERT sebagai baseline untuk berbagai aplikasi NLP Indonesia, termasuk implementasi analisis sentimen dalam penelitian ini yang menggunakan *indonesian-roberta-base-emotion-classifier*.

Penelitian oleh Koto dkk. (2020) memperkuat temuan ini dengan memperkenalkan *IndoLEM*, benchmark dataset dan model IndoBERT untuk NLP Indonesia. Mereka mendemonstrasikan bahwa *pre-training* pada korpus bahasa Indonesia yang besar (220 juta kata) menghasilkan representasi bahasa yang lebih baik untuk tugas-tugas *downstream* termasuk analisis sentimen, yang mencapai F1-score 84,13%. Penelitian mereka juga mengeksplorasi *transfer learning* dari model multibahasa, menunjukkan bahwa meskipun model multibahasa dapat memberikan titik awal yang baik, *fine-tuning* pada data Indonesia sangat meningkatkan performa.

Nugroho dkk. (2021) menerapkan *fine-tuning* BERT untuk analisis sentimen pada ulasan aplikasi mobile berbahasa Indonesia, konteks yang memiliki kesamaan dengan teks media sosial karena sifatnya yang informal dan sering mengandung bahasa gaul atau singkatan. Penelitian mereka membandingkan model multibahasa dengan model khusus Indonesia, mendemonstrasikan keunggulan *transfer learning* untuk bahasa Indonesia dengan data berlabel terbatas. Temuan mereka sangat relevan dengan penelitian ini, karena pesan Telegram finansial sering kali menggunakan bahasa informal, singkatan, dan istilah khusus yang memerlukan model yang robust terhadap variasi linguistik.

### **II.3.2 Analisis Sentimen Domain Finansial**

Araci (2019) memperkenalkan *FinBERT*, model BERT pertama yang dilatih khusus pada korpus finansial untuk analisis sentimen. Model ini mencapai peningkatan akurasi 15% dibandingkan BERT umum pada dataset *Financial PhraseBank* dan FiQA. Penelitian ini mendemonstrasikan pentingnya *domain-specific pre-training* untuk teks finansial, yang memiliki karakteristik linguistik unik termasuk terminologi khusus, struktur kalimat formal, dan nuansa sentimen yang berbeda dari domain umum. Konsep sentimen dalam konteks finansial lebih kompleks daripada klasifikasi positif-negatif sederhana; istilah seperti "volatilitas tinggi" dapat memiliki konotasi positif atau negatif tergantung konteks strategi investasi. Temuan ini memotivasi penggunaan model analisis sentimen yang disesuaikan untuk konteks finansial dalam penelitian ini, meskipun disesuaikan untuk bahasa Indonesia.

Si dkk. (2014) mengeksplorasi pemanfaatan sentimen media sosial untuk prediksi pergerakan saham dengan mengintegrasikan analisis sentimen Twitter dengan struktur relasi sosial. Penelitian mereka mendemonstrasikan bahwa menggabungkan sentimen dengan informasi jaringan sosial (siapa yang mempengaruhi siapa) secara signifikan meningkatkan akurasi prediksi pergerakan harga saham. Temuan kunci mereka adalah bahwa tidak semua opini memiliki bobot yang sama; pendapat dari individu yang lebih berpengaruh atau kredibel dalam jaringan sosial harus diberi bobot lebih tinggi. Konsep ini memiliki paralel langsung dengan desain sistem dalam penelitian ini, di mana implementasi *weighted sentiment analysis* memberikan bobot berbeda berdasarkan peran pengirim (Michael Yeoh versus anggota biasa), mengakui bahwa tidak semua kontributor dalam komunitas memiliki tingkat pengaruh atau keahlian yang sama.

### **II.3.3 Pemanfaatan Sinyal Komunitas dalam Peringkasan**

Kano dkk. (2018) mengusulkan metodologi inovatif untuk memanfaatkan sinyal popularitas dalam media sosial (seperti *votes*, *shares*, *bookmarks*) sebagai label jarak jauh (*distant labels*) untuk peringkasan ekstraktif percakapan online. Penelitian mereka mendemonstrasikan bahwa metrik keterlibatan komunitas dapat berfungsi sebagai indikator kualitas dan relevansi konten, memisahkan kontribusi konten dari faktor kontekstual. Pendekatan ini sangat relevan dengan konteks komunitas Telegram finansial berjenjang dalam penelitian ini, di mana struktur hierarkis (grup "Cuap Cuap", "Swing Plan", "Advanced Group") dan peran pengguna (pemimpin komunitas versus anggota biasa) dapat dimanfaatkan sebagai sinyal implisit tentang pentingnya informasi. Prinsip pemanfaatan sinyal komunitas ini diadopsi secara langsung dalam penelitian ini, di mana "peran pengguna" (misalnya, admin Michael Yeoh versus anggota biasa) digunakan sebagai proksi untuk "popularitas" atau "kredibilitas" informasi, yang kemudian menjadi dasar bagi fitur *weighted sentiment*. Sistem yang dikembangkan mengimplementasikan *weighted sentiment* yang memberikan bobot lebih tinggi pada pesan dari sumber yang lebih kredibel atau berpengaruh dalam komunitas.

### **II.3.4 Klasifikasi Emosi versus Sentimen untuk Investor Ritel**

Pemilihan *indonesian-roberta-base-emotion-classifier* dalam penelitian ini memerlukan justifikasi teoretis, karena berbeda dari pendekatan analisis sentimen finansial konvensional yang fokus pada klasifikasi bullish/bearish atau positif/negatif. Perbedaan fundamental ini didasarkan pada karakteristik unik investor ritel dibandingkan dengan investor institusional atau analis profesional.

Penelitian perilaku keuangan menunjukkan bahwa keputusan investasi ritel sangat dipengaruhi oleh emosi, terutama ketakutan (*fear*) dan keserakahan (*greed*). Berbeda dengan analis profesional yang dilatih untuk membuat keputusan berdasarkan analisis rasional, investor ritel sering kali membuat keputusan yang didorong oleh reaksi emosional terhadap pergerakan pasar atau diskusi komunitas. Dalam konteks komunitas Telegram dengan lebih dari 22.000 anggota yang sebagian besar adalah investor ritel muda, emosi kolektif yang terekspresikan dalam diskusi dapat berfungsi sebagai indikator sentimen pasar yang lebih akurat daripada sekadar klasifikasi positif/negatif.

Model klasifikasi emosi yang digunakan dalam penelitian ini dapat mendeteksi spektrum emosi yang lebih nuansa, termasuk ketakutan, kegembiraan, kemarahan, dan kesedihan. Emosi-emosi ini memberikan sinyal yang lebih kaya tentang psikologi pasar dibandingkan dengan sentimen biner. Misalnya, ekspresi ketakutan tinggi dalam diskusi komunitas dapat mengindikasikan potensi *panic selling*, sementara kegembiraan berlebihan dapat menandakan *euphoria* yang sering kali mendahului koreksi pasar. Pendekatan berbasis emosi ini selaras dengan teori *behavioral finance* yang mengidentifikasi bias psikologis sebagai faktor kunci dalam pengambilan keputusan investasi ritel.

Lebih lanjut, konteks diskusi informal di Telegram lebih cenderung mengekspresikan emosi autentik dibandingkan dengan laporan finansial formal atau analisis profesional. Anggota komunitas sering kali berbagi reaksi emosional mereka terhadap pergerakan harga, keputusan investasi, atau berita pasar dengan cara yang langsung dan tidak tersaring. Klasifikasi emosi dapat menangkap sinyal-sinyal psikologis ini yang hilang dalam analisis sentimen tradisional, memberikan wawasan yang lebih mendalam tentang suasana hati (*mood*) komunitas dan potensi implikasinya terhadap perilaku investasi kolektif.

#### **II.4 *Named Entity Recognition* untuk Teks Finansial Indonesia**

*Named Entity Recognition* (NER) merupakan tugas fundamental dalam NLP yang bertujuan mengidentifikasi dan mengklasifikasikan entitas bernama dalam teks, seperti nama orang, organisasi, lokasi, dan dalam konteks finansial, nama perusahaan, kode saham (ticker), dan istilah finansial spesifik. Bagian ini mengulas perkembangan metodologi NER untuk bahasa Indonesia dan domain finansial, yang menjadi dasar implementasi ekstraksi entitas dalam sistem yang dikembangkan.

#### **II.4.1 NER untuk Bahasa Indonesia**

Koto dkk. (2020) menetapkan benchmark untuk NER bahasa Indonesia melalui dataset *IndoLEM* dan model IndoBERT. Penelitian mereka mendemonstrasikan bahwa IndoBERT mencapai performa *state-of-the-art* pada tugas NER Indonesia, mengungguli mBERT dengan margin yang signifikan. Keberhasilan ini dikaitkan dengan kemampuan model untuk menangkap nuansa linguistik bahasa Indonesia, termasuk struktur morfologi, pola afiksasi, dan konteks penggunaan yang spesifik pada korpus Indonesia. Temuan ini memberikan justifikasi kuat untuk penggunaan *bert-base-indonesian-NER* dalam sistem yang dikembangkan, yang dibangun di atas arsitektur IndoBERT.

Penelitian Khairunnisa, Imankulova, dan Komachi (2020) mengidentifikasi masalah kritis dalam dataset NER Indonesia yang ada: inkonsistensi anotasi yang dapat menurunkan performa model. Mereka menyediakan dataset yang dianotasi ulang dengan standar yang lebih konsisten dan mengevaluasi berbagai pendekatan *deep learning* termasuk BiLSTM-CRF dengan berbagai jenis *embeddings*. Penelitian mereka menekankan pentingnya kualitas data pelatihan dan konsistensi skema anotasi untuk mencapai performa NER yang optimal. Wawasan ini relevan dengan penelitian ini, terutama dalam konteks *fine-tuning* model NER untuk mengenali entitas finansial spesifik yang mungkin tidak terwakili dengan baik dalam dataset NER umum Indonesia.

#### **II.4.2 NER Domain Finansial**

Zhang dan Zhang (2022) mengembangkan *FinBERT-MRC*, pendekatan NER finansial yang merumuskan tugas ekstraksi entitas sebagai masalah *machine reading comprehension*. Model mereka mencapai F1-score 92,78% dan 96,80% pada dataset finansial Tiongkok, mengungguli pendekatan BiLSTM-CRF dan BERT-CRF tradisional. Penelitian ini mendemonstrasikan bahwa NER domain finansial memiliki tantangan unik karena variasi semantik dan leksikal yang spesifik, seperti perbedaan antara nama perusahaan dalam konteks berita versus dokumen formal, atau ambiguitas antara nama organisasi dan produk finansial. Meskipun penelitian mereka fokus pada bahasa Tiongkok, prinsip-prinsip metodologis dapat diadaptasi untuk konteks Indonesia, terutama dalam menangani ambiguitas nama perusahaan versus kode saham.

Shah dkk. (2023) menyediakan sumber daya penting untuk NER finansial melalui *FiNER-ORD*, dataset NER finansial *open research* pertama dengan kualitas tinggi.

Dataset ini mengatasi variasi semantik dan leksikal yang unik dalam domain finansial dan menyediakan benchmark untuk berbagai model bahasa pra-latih termasuk varian BERT dan LLMs. Penelitian mereka menekankan bahwa entitas finansial sering kali memiliki representasi multipel (misalnya, "Apple Inc.", "Apple", "AAPL") yang harus dikenali sebagai entitas yang sama. Wawasan ini sangat relevan dengan konteks pasar saham Indonesia, di mana perusahaan mungkin dirujuk dengan nama lengkap, nama singkat, atau kode ticker (misalnya, "Bank Central Asia", "BCA", "BBCA").

## II.5 *Topic Modeling* dengan BERTopic

*Topic modeling* merupakan teknik untuk mengidentifikasi tema atau topik abstrak dalam kumpulan dokumen. Bagian ini mengulas evolusi metodologi *topic modeling* dari pendekatan statistik tradisional ke pendekatan berbasis *neural embeddings*, dengan fokus khusus pada BERTopic yang digunakan dalam sistem yang dikembangkan.

### II.5.1 BERTopic: Metodologi dan Keunggulan

Grootendorst (2022) memperkenalkan BERTopic, metodologi *topic modeling* neural yang menggabungkan model bahasa berbasis *transformer* (BERT), *document embeddings*, *clustering* dengan HDBSCAN, dan prosedur *class-based TF-IDF* (c-TF-IDF) untuk mengekstrak representasi topik yang koheren. Arsitektur BERTopic mengatasi keterbatasan pendekatan tradisional seperti *Latent Dirichlet Allocation* (LDA) dengan memanfaatkan representasi semantik kontekstual dari *transformer models*. Prosedur c-TF-IDF yang diusulkan menghasilkan representasi topik yang lebih interpretatif dibandingkan dengan metode konvensional dengan memperlakukan semua dokumen dalam satu kluster sebagai satu dokumen besar dan membandingkannya dengan dokumen dari kluster lain. Metodologi ini sangat sesuai untuk teks media sosial yang pendek dan informal, karena tidak memerlukan asumsi distribusional yang ketat seperti LDA.

Keunggulan BERTopic untuk teks media sosial divalidasi secara empiris oleh Egger dan Yu (2022), yang membandingkan BERTopic dengan LDA, NMF, dan Top2Vec pada dataset 31.800 *tweet* terkait COVID-19 dan perjalanan. Evaluasi mereka mendemonstrasikan superioritas BERTopic untuk teks pendek dan tidak terstruktur, khususnya dalam hal koherensi topik dan interpretabilitas. Metrik evaluasi menunjukkan bahwa BERTopic menghasilkan topik yang lebih koheren dan mudah dipahami dibandingkan dengan metode tradisional, terutama untuk data yang memiliki karakteristik yang beragam.

teristik serupa dengan media sosial: panjang dokumen pendek, penggunaan bahasa informal, dan struktur gramatikal yang longgar. Temuan ini memberikan justifikasi kuat untuk pemilihan BERTopic dalam penelitian ini, mengingat pesan Telegram finansial memiliki karakteristik yang sangat mirip dengan *tweet*: pendek, informal, dan sering menggunakan singkatan atau jargon.

### **II.5.2 Perkembangan Terkini dalam *Topic Modeling* Neural**

Angelov dan Inkpen (2024) mengembangkan metodologi *topic modeling* lebih lanjut dengan memperkenalkan *Contextual-Top2Vec*, yang menggunakan *contextual token embeddings* dari dokumen untuk menghasilkan topik hierarkis, *topic spans* dalam dokumen, dan label topik berbasis frasa. Penelitian mereka mengusulkan penggunaan *BERTScore* untuk mengevaluasi koherensi dan informativeness topik, mengatasi keterbatasan metrik evaluasi tradisional seperti koherensi berbasis PMI. Model mereka mengungguli pendekatan *state-of-the-art* pada metrik evaluasi komprehensif, mendemonstrasikan bahwa representasi kontekstual dapat meningkatkan kualitas topik yang dihasilkan. Meskipun penelitian ini menggunakan BERTopic sebagai metodologi dasar, wawasan dari *Contextual-Top2Vec* menginformasikan pertimbangan desain, terutama dalam hal evaluasi kualitas topik dan interpretasi hasil *clustering*.

## **II.6 Telegram sebagai Platform Komunitas**

Telegram telah berkembang menjadi platform komunikasi yang dominan untuk komunitas online, terutama di Indonesia. Bagian ini mengulas karakteristik unik Telegram sebagai platform untuk komunitas berskala besar, pola komunikasi yang terjadi di dalamnya, dan implikasi untuk desain sistem analitik dan rangkuman otomatis.

### **II.6.1 Karakteristik dan Dinamika Platform Telegram**

Nobari, Reshadatmand, dan Neshati (2017) menyediakan analisis pionir tentang struktur dan aspek topikal Telegram menggunakan data yang di-*crawl* dari kanal publik. Penelitian mereka mengeksplorasi pola jaringan, klasifikasi pesan (spam/ham), dan arsitektur Telegram sebagai platform *messaging* hibrida yang menggabungkan fitur komunikasi pribadi dan publik. Mereka mengidentifikasi bahwa Telegram memiliki karakteristik unik dibandingkan platform media sosial lain: kombinasi antara privasi komunikasi satu-satu dengan kemampuan penyebaran informasi massal melalui kanal dan grup besar. Struktur ini menciptakan dinamika komunikasi yang berbeda dari platform seperti Twitter atau Facebook, di mana informasi dapat menyebar

dengan cepat dalam komunitas tertutup yang besar tanpa eksposur publik yang luas. Pemahaman tentang arsitektur ini penting untuk desain sistem rangkuman yang harus mempertimbangkan privasi komunitas sambil memfasilitasi akses informasi yang efisien.

Hashemi dan Chahooki (2019) mengembangkan metodologi untuk mengukur kualitas grup Telegram melalui analisis perilaku pengguna, mengeksplorasi lebih dari 900.000 kanal dan 300.000 *supergroup* berbahasa Persia. Mereka mengusulkan fitur-fitur pengukuran kualitas yang membedakan grup berkualitas tinggi dari grup berkualitas rendah berdasarkan pola aktivitas pengguna, konsistensi partisipasi, dan struktur komunikasi. Penelitian mereka menunjukkan bahwa grup berkualitas tinggi cenderung memiliki partisipasi yang lebih seimbang, topik diskusi yang lebih fokus, dan tingkat spam yang lebih rendah. Temuan ini relevan dengan konteks komunitas Michael Yeoh yang memiliki struktur berjenjang, di mana grup "Advanced" diasumsikan memiliki kualitas diskusi yang lebih tinggi dibandingkan grup "Cuap Cuap" umum. Sistem rangkuman harus dirancang dengan mempertimbangkan perbedaan kualitas dan karakteristik diskusi antar grup.

### **II.6.2 Analisis Skala Besar Komunitas Telegram**

Morgia dkk. (2024) menyediakan dataset Telegram terbesar hingga saat ini dengan 120.979 kanal dan lebih dari 400 juta pesan, bersama dengan infrastruktur untuk pengumpulan dan analisis data skala besar. Penelitian mereka menyediakan alat dan metodologi untuk penelitian komunitas Telegram yang mencakup analisis dinamika komunitas, propagasi konten, dan pola komunikasi tingkat platform. Kontribusi mereka mendemonstrasikan bahwa Telegram telah berkembang menjadi ekosistem informasi yang kompleks dengan dinamika komunitas yang beragam. Skala dataset ini memberikan konteks untuk memahami bahwa komunitas finansial dengan 22.000+ anggota seperti yang diteliti dalam penelitian ini merupakan bagian dari fenomena komunikasi online yang lebih luas, di mana platform seperti Telegram telah menjadi infrastruktur kritis untuk pertukaran informasi dalam komunitas tematik berskala besar.

Penelitian terbaru oleh Perlo dkk. (2025) menyediakan analisis per-topik pertama dari grup Telegram, mengeksplorasi 51 juta pesan dari 669 grup di berbagai topik termasuk Pendidikan, Politik, dan Cryptocurrency. Penelitian mereka menganalisis pola aktivitas pengguna, kehadiran bot, dan berbagi media, serta menyediakan alat *open-source* untuk pengumpulan pesan otomatis. Temuan kunci mereka adalah bahwa grup dengan topik berbeda menunjukkan pola komunikasi yang sangat berbeda

dalam hal volume pesan, distribusi waktu aktivitas, dan tipe konten yang dibagikan. Khususnya untuk kategori Cryptocurrency, yang memiliki kesamaan dengan komunitas finansial yang diteliti, mereka menemukan tingkat aktivitas yang sangat tinggi dengan puncak aktivitas yang berkorelasi dengan jam perdagangan pasar. Temuan ini memvalidasi desain sistem rangkuman per jam dalam penelitian ini, karena pola aktivitas temporal yang kuat mengindikasikan bahwa rangkuman periodik sesuai dengan ritme alami diskusi komunitas.

## II.7 Arsitektur Pemrosesan Data *Real-Time*

Implementasi sistem rangkuman otomatis untuk komunitas Telegram yang aktif memerlukan arsitektur pemrosesan data yang dapat menangani aliran pesan secara *real-time* atau *near-real-time*. Bagian ini mengulas pendekatan arsitektural untuk pemrosesan *stream* teks dengan NLP, pola *event-driven architecture*, dan pertimbangan desain untuk sistem yang responsif dan *scalable*.

### II.7.1 *Pipeline* NLP untuk Pemrosesan *Stream* Teks

Hamidi dkk. (2021) mengusulkan arsitektur *pipeline* NLP yang dinamis dan terdistribusi untuk pemrosesan *stream* teks *real-time*. Penelitian mereka menggunakan Apache Storm dan Apache Kafka untuk menerapkan tugas-tugas NLP pada aliran data tekstual, memungkinkan pengembang untuk menginjeksi modul NLP melalui berbagai bahasa pemrograman dan mendukung sumber data multipel. Arsitektur mereka mendemonstrasikan bagaimana komponen NLP yang berbeda (tokenisasi, *part-of-speech tagging*, *named entity recognition*, analisis sentimen) dapat dikomposisikan dalam *pipeline* yang dapat di-*scale* secara horizontal. Prinsip-prinsip design ini relevan dengan penelitian ini, di mana sistem harus mengintegrasikan beberapa komponen NLP (analisis sentimen, NER, *topic modeling*, generasi rangkuman) dalam arsitektur yang koheren dan dapat diperluas.

Isah dkk. (2019) menyediakan survei komprehensif tentang *framework* pemrosesan *stream* data terdistribusi, termasuk Apache Storm, Spark Streaming, Apache Flink, dan Kafka Streams. Penelitian mereka membandingkan model pemrosesan, karakteristik operasional, dan pola arsitektural dari berbagai *framework*, memberikan taksonomi yang membantu dalam pemilihan teknologi yang sesuai untuk kasus penggunaan spesifik. Untuk konteks komunitas Telegram dengan volume pesan yang tinggi namun tidak ekstrem (ratusan hingga ribuan pesan per jam), mereka mengidentifikasi bahwa pendekatan *micro-batch processing* dengan Kafka atau arsitektur berbasis *event* yang lebih sederhana dapat memberikan keseimbangan optimal anta-

ra latensi, throughput, dan kompleksitas implementasi. Wawasan ini menginformasikan desain sistem dalam penelitian ini, yang mengadopsi arsitektur berbasis *event* dengan penjadwalan periodik daripada pemrosesan *stream* kontinu, sesuai dengan kebutuhan rangkuman per jam.

### **II.7.2 Arsitektur Berbasis *Event* dan Performa**

Rubert dan Farias (2023) menyediakan bukti empiris tentang dampak *event-driven architecture* (EDA) terhadap performa melalui studi eksploratif yang membandingkan arsitektur berbasis *event* dengan arsitektur monolitik. Penelitian mereka mengukur penggunaan CPU, memori, waktu respons, *throughput*, dan paket jaringan, mendemonstrasikan bahwa EDA dapat memberikan keunggulan signifikan dalam *scalability* dan responsivitas, terutama untuk sistem yang menangani beban kerja yang bervariasi dan tidak dapat diprediksi. Temuan mereka menunjukkan bahwa EDA sangat cocok untuk skenario di mana komponen sistem perlu berkomunikasi secara asinkron dan dapat *di-scale* secara independen. Konteks komunitas Telegram finansial, di mana volume pesan bervariasi drastis sepanjang hari (tinggi selama jam perdagangan, rendah di malam hari), sangat sesuai dengan karakteristik ini. Sistem yang dikembangkan dalam penelitian ini mengadopsi prinsip EDA dengan pemisahan antara komponen pengumpulan pesan, pemrosesan NLP, dan generasi rangkuman, memungkinkan setiap komponen untuk beroperasi secara independen dan *di-scale* sesuai kebutuhan.

### **II.7.3 Pertimbangan *Micro-batching* dan Frekuensi Data**

Meskipun pemrosesan *stream* murni memberikan latensi terendah, penelitian menunjukkan bahwa pendekatan *micro-batching* sering kali lebih praktis untuk banyak aplikasi NLP. Konsep rangkuman per jam dalam penelitian ini dapat dipandang sebagai bentuk *macro-batching*, di mana pesan dikumpulkan selama satu jam sebelum diproses sebagai satu *batch*. Pendekatan ini memberikan beberapa keuntungan: (1) memungkinkan agregasi konteks yang lebih kaya untuk tugas-tugas seperti *topic modeling* yang memerlukan pandangan holistik terhadap data, (2) mengurangi beban komputasi dengan menghindari pemrosesan setiap pesan individual secara *real-time*, dan (3) sesuai dengan ritme natural konsumsi informasi pengguna, yang jarang memerlukan pembaruan lebih sering dari setiap jam untuk analisis tren komunitas.

#### **II.7.4 Tantangan Inferensi LLM dalam Sistem *Real-Time***

Bottleneck arsitektural yang paling signifikan dalam sistem rangkuman otomatis berbasis LLM adalah latensi inferensi model bahasa besar. Tidak seperti komponen NLP tradisional seperti tokenisasi atau *named entity recognition* yang dapat dieksekusi dalam milidetik, generasi teks dengan LLM skala besar (puluhan hingga ratusan miliar parameter) dapat memerlukan waktu detik hingga menit untuk satu panggilan inferensi, tergantung pada panjang konteks input dan panjang output yang dihasilkan. Tantangan ini menjadi kritis dalam konteks sistem rangkuman per jam, di mana seluruh *pipeline* pemrosesan harus selesai dalam jendela waktu yang terbatas (ide-alnya kurang dari 5-10 menit dari akhir periode pengumpulan data) untuk menjaga relevansi temporal rangkuman.

Penelitian terkini menunjukkan bahwa kecepatan inferensi LLM sangat dipengaruhi oleh tiga faktor utama: (1) ukuran model (jumlah parameter), (2) panjang konteks input, dan (3) infrastruktur komputasi yang digunakan (CPU, GPU, atau akselerator khusus seperti TPU). Untuk konteks rangkuman komunitas Telegram dengan potensi ratusan pesan per jam, panjang konteks input dapat dengan mudah mencapai puluhan ribu *token*, yang secara signifikan meningkatkan waktu inferensi pada infrastruktur standar. Lebih lanjut, model yang lebih besar seperti Llama 3.1 70B, meskipun memberikan kualitas output superior dibandingkan model yang lebih kecil, memerlukan sumber daya komputasi yang substansial dan dapat mengalami latensi yang tidak dapat diterima pada infrastruktur konvensional.

Pemilihan Groq API dengan model Llama 3.1 70B dalam penelitian ini merupakan solusi arsitektural yang dirancang khusus untuk mengatasi bottleneck inferensi LLM ini. Groq menggunakan *Language Processing Unit* (LPU), arsitektur akselerator khusus yang dirancang untuk inferensi LLM dengan kecepatan ekstrem. Platform Groq telah mendemonstrasikan kemampuan untuk menghasilkan ratusan hingga ribuan *tokens per second*, secara signifikan melampaui solusi berbasis GPU tradisional. Kecepatan inferensi yang tinggi ini adalah faktor pendukung (*enabler*) teknis yang krusial yang memungkinkan implementasi rangkuman per jam yang praktis; tanpa throughput inferensi yang tinggi, sistem akan mengalami akumulasi *backlog* dan keterlambatan yang tidak dapat diterima, merusak nilai temporal dari rangkuman yang dihasilkan.

Selain itu, penggunaan *API* cloud seperti Groq menghilangkan kompleksitas operasional dalam mengelola infrastruktur inferensi LLM lokal, termasuk manajemen GPU, optimasi *batch size*, dan penjadwalan sumber daya. Pendekatan ini memung-

kinkan sistem untuk fokus pada logika bisnis dan integrasi komponen NLP, sementara kebutuhan komputasi intensif untuk inferensi LLM di-*offload* ke platform yang dioptimalkan. Arsitektur ini mencerminkan prinsip *separation of concerns* dalam desain sistem terdistribusi, di mana setiap komponen (pengumpulan data, pemrosesan NLP, inferensi LLM) diimplementasikan menggunakan teknologi yang paling sesuai untuk karakteristik komputasinya.

## II.8 Peringkasan untuk Media Sosial dan Percakapan Online

Peringkasan konten media sosial memiliki tantangan unik dibandingkan dengan peringkasan dokumen formal atau artikel berita. Bagian ini mengulas penelitian tentang peringkasan media sosial, khususnya Reddit dan forum diskusi online, yang memiliki karakteristik serupa dengan komunitas Telegram.

### II.8.1 Pemanfaatan Sinyal Komunitas dalam Peringkasan

Kano dkk. (2018) mengusulkan pendekatan inovatif untuk peringkasan ekstraktif percakapan online dengan memanfaatkan ukuran popularitas dalam media sosial (seperti *votes*, *shares*, *bookmarks*) sebagai *distant labels*. Penelitian mereka mendemonstrasikan bahwa metrik keterlibatan komunitas dapat berfungsi sebagai indikator kualitas dan relevansi konten, dengan memisahkan kontribusi konten dari faktor kontekstual. Mereka menemukan bahwa pesan dengan *upvotes* tinggi dalam forum diskusi cenderung mengandung informasi yang lebih substantif dan relevan dengan topik diskusi utama. Pendekatan ini sangat relevan dengan penelitian ini, di mana struktur hierarkis komunitas Telegram (pimpinan komunitas versus anggota biasa) dapat dimanfaatkan sebagai sinyal implisit tentang kredibilitas dan pentingnya informasi. Implementasi *weighted sentiment analysis* dalam sistem yang dikembangkan mengadopsi prinsip ini dengan memberikan bobot yang berbeda berdasarkan peran pengirim pesan.

Fabbri dkk. (2021) memperkenalkan *ConvoSumm*, benchmark peringkasan percakapan yang mencakup empat dataset dari berbagai bentuk percakapan online: komentar berita, forum diskusi, forum Q&A komunitas, dan *thread email*. Penelitian mereka menggunakan *framework issues-viewpoints-assertions* yang mengintegrasikan *argument mining* melalui konstruksi graf untuk memodelkan struktur percakapan. Pendekatan ini mengakui bahwa percakapan online sering kali melibatkan pertukaran pandangan yang kompleks dengan argumen pendukung, yang memerlukan pemahaman struktural lebih mendalam daripada sekadar ekstraksi kalimat penting. Kontribusi metodologis mereka relevan dengan penelitian ini karena

na diskusi finansial di Telegram sering kali mengandung perdebatan antara strategi investasi berbeda, analisis yang bertentangan, atau perspektif pasar yang beragam. Sistem rangkuman yang efektif harus dapat mensintesis pandangan-pandangan yang berbeda ini daripada hanya mengekstrak opini dominan.

### **II.8.2 Peringkasan Percakapan *Multi-speaker***

Kompleksitas percakapan *multi-speaker* dibahas secara komprehensif oleh Chen dan Yang (2021), yang mengusulkan pendekatan yang sadar struktur (*structure-aware*) untuk peringkasan percakapan abstraktif. Penelitian mereka memodelkan relasi wacana dan *action triples* ("siapa-melakukan-apa") menggunakan graf tersusunan untuk menangkap karakteristik kompleks interaksi manusia-manusia di media sosial. Mereka mengimplementasikan mekanisme *decoding* multi-granularitas yang dapat menghasilkan rangkuman pada tingkat abstraksi berbeda. Pendekatan ini sangat relevan dengan konteks komunitas Telegram berjenjang, di mana percakapan melibatkan puluhan hingga ratusan partisipan dengan peran dan tingkat keahlian yang berbeda. Sistem harus dapat mengenali struktur interaksi (misalnya, pertanyaan-jawaban, klaim-dukungan, analisis-counter-analisis) untuk menghasilkan rangkuman yang koheren dan informatif.

### **II.8.3 Peringkasan Ekstrim untuk Konten Media Sosial**

Sotudeh dkk. (2021) memperkenalkan dataset TLDR9+ dengan lebih dari 9 juta *instance* pelatihan dari Reddit untuk peringkasan ekstrim (rangkuman satu kalimat). Penelitian mereka mengatasi agregasi konten media sosial skala besar dan mencakup subset berkualitas tinggi (TLDRHQ) melalui anotasi manusia. Mereka mendemonstrasikan bahwa peringkasan ekstrim, yang menghasilkan rangkuman sangat singkat dengan tingkat kompresi dan abstraksi tinggi, dapat efektif untuk konten media sosial yang cenderung verbose dan mengandung banyak informasi redundant. Konsep peringkasan ekstrim ini relevan dengan kebutuhan komunitas Telegram finansial, di mana anggota mungkin memerlukan "snapshot" cepat dari diskusi komunitas tanpa membaca rangkuman panjang. Sistem yang dikembangkan dalam penelitian ini dapat mengadaptasi prinsip ini dengan menyediakan "headline" atau "key takeaway" di awal setiap rangkuman per jam sebelum menyajikan detail yang lebih mendalam.

## **II.9 Kesenjangan Penelitian dan Kontribusi**

Tinjauan literatur komprehensif di atas mengidentifikasi beberapa kesenjangan penelitian kritis yang ditangani oleh penelitian ini. Pertama, tidak ada penelitian yang

mengintegrasikan NLP bahasa Indonesia (khususnya model berbasis RoBERTa dan BERT) dengan analisis domain finansial untuk konteks media sosial. Sebagian besar penelitian analisis sentimen finansial fokus pada bahasa Inggris atau Tiongkok, sementara penelitian NLP Indonesia cenderung menggunakan domain umum atau ulasan produk. Penelitian ini menjembatani kesenjangan ini dengan mengembangkan sistem yang mengintegrasikan *indonesian-roberta-base-emotion-classifier* dan *bert-base-indonesian-NER* untuk analisis teks finansial Indonesia dalam konteks media sosial.

Kedua, meskipun terdapat penelitian ekstensif tentang komunitas Telegram dan peringkasan media sosial secara terpisah, sangat sedikit penelitian yang menangani sistem peringkasan otomatis khusus untuk aplikasi bot Telegram. Penelitian tentang bot Telegram cenderung fokus pada interaksi percakapan sederhana atau otomasi tugas administratif, bukan pada sintesis intelijen informasi yang kompleks. Penelitian ini memberikan kontribusi dengan mendemonstrasikan implementasi praktis sistem peringkasan otomatis yang diintegrasikan dengan Telegram Bot API untuk penyampaian rangkuman periodik.

Ketiga, penelitian tentang *pipeline* NLP *real-time* yang mengintegrasikan multiple model (analisis sentimen, NER, *topic modeling*, generasi rangkuman dengan LLM) dalam arsitektur *event-driven* untuk platform *messaging* masih terbatas. Sebagian besar penelitian mengevaluasi model-model ini secara terpisah atau dalam konteks dataset benchmark offline. Penelitian ini memberikan kontribusi metodologis dengan mendemonstrasikan integrasi empat komponen NLP berbeda dalam sistem yang koheren dan operasional.

Keempat, tidak ada penelitian yang menangani aliran informasi hierarkis dan sintesis lintas grup dalam komunitas berjenjang. Sebagian besar penelitian peringkasan media sosial mengasumsikan satu sumber data atau agregasi horizontal dari multiple sumber yang setara. Penelitian ini menangani skenario unik di mana informasi mengalir melalui struktur hierarkis (grup Cuap Cuap, Swing Plan, Advanced) dengan karakteristik diskusi yang berbeda, dan sistem harus mensintesis informasi lintas tingkat hierarki ini.

Kelima, pendekatan rangkuman per jam untuk diskusi finansial dengan *weighted sentiment* berbasis peran pengguna merupakan kontribusi novel. Sementara penelitian sebelumnya mengeksplorasi pemanfaatan sinyal popularitas atau metrik keterlibatan, penelitian ini mengimplementasikan *weighting* eksplisit berdasarkan struktur peran komunitas yang mencerminkan kredibilitas dan pengaruh relatif kontributor

dalam ekosistem informasi finansial.

Secara keseluruhan, penelitian ini mengintegrasikan teori *information overload* dan *cognitive load*, metodologi NLP terkini untuk bahasa Indonesia dan domain finansial, arsitektur pemrosesan *event-driven*, dan pemahaman tentang dinamika komunitas Telegram untuk mengembangkan solusi holistik yang mengatasi tantangan praktis dalam komunitas finansial Indonesia. Kontribusi utama penelitian ini terletak pada demonstrasi bahwa sistem otomatis berbasis *AI* dapat secara efektif mereduksi *information overload* dalam komunitas online berskala besar melalui sintesis intelejen multi-sumber, lintas-grup (*cross-group*), dan hierarkis yang disesuaikan dengan konteks spesifik ekosistem informasi komunitas.