Generative Collective Intelligence

The rapid spread of fake news in digital ecosystems demands innovative detection frameworks, especially for linguistically nuanced languages like Vietnamese. This paper presents Generative Collective Intelligence (GCI), a pioneering multiagent ensemble system that integrates ContentAnalyzer and EmotionSentimentAnalyzer to detect misinformation. Built on PhoBERT embeddings and enhanced by custom Vietnamese lexicons, GCI employs a dynamic ensemble mechanism to fuse content quality and affective analysis. Beyond detection, GCI embodies generative intelligence by synthesizing realistic fake and true articles using the Gemini API, enriching its training data. Its collective intelligence emerges from agent collaboration and human feedback loops, adapting weights dynamically. Trained on the Vietnamese Health Misinformation Dataset (VHMD), GCI achieves promising results, with ContentAnalyzer at 0.8081 accuracy and EmotionSentimentAnalyzer at 0.8005. This work highlights GCI’s scalable, adaptive architecture as a transformative approach to combating misinformation

\section{Introduction}

The rapid proliferation of misinformation, particularly in health-related contexts \cite{suarez2021prevalence}, poses significant challenges to public trust and decision-making. False claims regarding medical treatments, vaccines, and public health policies have led to widespread misinformation, exacerbating health risks and undermining institutional credibility. Detecting and mitigating such misinformation requires sophisticated AI-driven approaches that leverage multiple dimensions of content analysis.

Generative Collective Intelligence (GCI) represents an innovative approach that integrates collective intelligence mechanisms with generative AI \cite{zou2023wireless} to create an adaptive, self-improving system for misinformation detection. Unlike traditional AI models that operate in isolation, GCI enhances decision-making by aggregating insights from multiple specialized agents. In this study, we propose a GCI-based misinformation detection system designed to process and verify Vietnamese health-related information, leveraging a multi-agent architecture to enhance accuracy and robustness.

The system comprises three key components: (1) ContentAnalyzer, which applies machine learning techniques such as TF-IDF + Logistic Regression, PhoBERT-based \cite{nguyen2020phobert} semantic analysis, and gradient boosting to evaluate content reliability; (2) EmotionalPatternAnalyzer, which detects sentiment anomalies using multi-GPU pattern recognition and AdamW optimization; and (3) SourceCredibilityAgent, which assesses domain reputation through pattern matching and historical credibility tracking, without directly utilizing dataset-based information.

To evaluate the system, we utilize the Vietnamese Health Misinformation Dataset (VHMD) \cite{vhmd}, a large-scale dataset containing 140,000 real health-related news articles from verified sources and 140,000 AI-generated misinformation cases. Unlike previous datasets, VHMD offers structured metadata and diverse misinformation patterns, enabling robust training for AI models. Experimental results indicate that the ensemble approach that combines multiple agents significantly improves misinformation detection accuracy compared to traditional single-model approaches.

This paper is structured as follows: Section 2 not only reviews related work on collective intelligence and fake news detection but also details the system architecture and agent design; Section 3 presents experimental setup and evaluation results; Section 4 discusses implications and limitations; and finally, Section 5 concludes with potential future research directions.