

ABSTRACT

The scarcity of structured data on Indonesia's investment landscape poses a significant challenge to trend analysis and decision-making. While Temporal Knowledge Graphs (TKGs) can capture market dynamics, existing benchmark datasets are heavily biased towards English-language sources and mature markets, thereby overlooking the unique context of emerging economies like Indonesia. This research addresses this gap by constructing IndovestDKG, the first TKG dataset focused on the Indonesian investment domain, sourced from the Kompas news portal. Using a Large Language Model (LLM), 7,761 news articles were extracted into 68,207 temporal facts (quadruplets). Subsequently, five state-of-the-art Temporal Knowledge Graph Completion (TKGC) methods were evaluated on IndovestDKG against two other benchmark datasets (FinDKG and ICEWS14). The results reveal a significantly lower performance of all models on IndovestDKG, highlighting the inadequacy of current model architectures for emerging markets. These markets are characterized by unique properties such as extreme sparsity, with 70.8% of entities appearing only once (singleton entities), and policy-driven dynamics. The Re-Temp model achieves the best performance on financial datasets, with an MRR of 8.39% on IndovestDKG and 28.12% on FinDKG. The primary contribution of this research is the provision of a new, small-scale benchmark dataset (IndovestDKG) and a comparative analysis that exposes the limitations of existing TKGC methods within an emerging market context.

Keywords: *temporal knowledge graph completion, indonesian investment news, large language model*