

1. Introduction

1.1 Background

Product reviews on e-commerce platforms applications really help us in finding good products to buy, but this can be a target for fraudsters in carrying out fake reviews because e-commerce applications users produce content that contains rich information [1]. There are many kinds of fake reviews, but there are three big categories of fake reviews, namely, fake reviews, fake reviewers, and fake reviews analysis [2]. Millions of people make transactions on e-commerce applications, making reviews on each product to consider whether this product is worth buying or not. This research significantly identifies words or sentences that frequently appear in fake reviews, which can be used to detect such reviews. The goal is to educate e-commerce platform users to be more careful in reading reviews. Deep learning methods that can be used include Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Graph Neural Network (GNN). Among the various types of GNNs, convolutional GNNs such as Graph Convolutional Networks (GCNs) are the most recognized due to their effective capabilities and ease of integration with other neural networks. GCNs, which are a variant of CNN, are optimized specifically for graph data. The model constructs layers using a learned first-order spectral filter and then applies a nonlinear activation function to form a deep representation of the graph [3]. This research uses Graph Convolutional Network (GCN) because the focus is on analyzing the relationship between words in fake reviews and identifying emerging patterns. CNN is the most effective algorithm for analyzing and understanding images [4]. Recurrent Neural Networks (RNN) are effective in approximating dynamic systems that process time-based and sequence data, such as video and audio. Long Short-Term Memory (LSTM), as a variant of RNN, is designed to handle the same type of data [5]. Research [6] employs deep learning techniques, with a focus on the LSTM algorithm, an advanced form of RNN. The findings demonstrate that LSTM is highly effective at detecting fake reviews due to its capacity to capture long-term dependencies in textual data. The use of Bi-LSTM in the study achieved an impressive accuracy of 97.21%, alongside high precision, recall, and F1-Score. This makes LSTM particularly well-suited for analyzing lengthy reviews, thanks to its strength in processing sequential information. However, challenges persist in identifying sarcasm in reviews, which requires a more sophisticated approach. In conclusion, LSTM outperforms other methods in fake review classification. Meanwhile, GCN, as a deep learning algorithm designed for graph data structures, not only classifies reviews as genuine or fake but also represents them in graph form. GCN finds patterns of relationships between graphs, which helps us understand the characteristics of fake reviews. Graphs support unstructured data such as text data, and their graphical representations provide structural relationships in the data, offering more insight compared to analyzing the data separately [7]. In this study, the GCN process in detecting fake reviews resulted in an F1-score of 0.8919, which indicates that the model works well overall.

1.2 Problem Identification

This research was conducted to find significant words or sentences in fake reviews and identify patterns or relationships between words that play a role in these reviews. The research was conducted on the Shopee e- Jurnal Tugas Akhir Fakultas Informatika September-2024 2 commerce platform with a focus on uncovering important elements that can detect patterns of fraud in user reviews. Testing was conducted in two stages: first, testing on balanced data with equal numbers of fake and genuine reviews, and second, testing on unbalanced data, where genuine reviews outnumber fake reviews.

1.3 Research Objectives

This research aims to help e-commerce users identify fake reviews by analyzing the words or sentences that often appear in the reviews. To ensure the accuracy of identification, two testing methods are conducted: one on balanced data and one on unbalanced data, in order to optimize the results of pattern and word relationship analysis.

1.4 Organization of Writing

In Section 2, we will discuss previous research related to fake review and Graph Convolutional Network (GCN). Section 3 will discuss the system design that includes classification, graph generation, and analysis

using the GCN algorithm. Section 4 will present the testing and analysis of the test results. Finally, Section 5 will present the conclusion of this research.