1. Introduction

Social media has become an integral part of the daily lives of numerous internet users. These platforms serve as a primary channel for searching and sharing information, as well as communicating with people worldwide[1]. Social media platforms such as Facebook and Twitter have the ability to make a topic go viral quickly [2]. Social media also facilitates rapid information retrieval and enables the exchange of opinions due to the real-time dissemination of information on these platforms. Consequently, there is a continuous flow of information available at any given time[3].

Twitter is one of the most widely used social media platforms in Indonesia, On Twitter, users can search for information, stay updated with the latest news, and communicate with other users. The information is conveyed through a tweet. One of the features on Twitter is retweet, which indicates a social connection between users. When a user retweets a tweet, that tweet will be shared again to their followers [4]. As a result, the tweet will reach a wider audience as it gets shared through retweet. Modeling information diffusion is important to understand the spread of information that occurs on Twitter [5].

Fans of K-pop are a group of people who appreciate South Korean culture, including music, dramas, and films. K-pop music has gained immense popularity internationally due to effective marketing strategies, and its music has been embraced in various parts of the world[6]. With the advancement of social media, K-pop fans now have numerous platforms to discuss and interact with fellow fans. These platforms provide spaces for fans to engage in conversations about their favorite idols, preferred groups, genres, dramas, and more. It has created a vibrant community where fans can connect and share their passion for K-pop. One of the social media platforms widely used by K-pop fans is Twitter. In this research, "KPOP" will be used as a keyword to collect tweet data that will be utilized for retweet prediction.

In the previous study, machine learning was applied using a multi-class classification method with three groups of features utilized in the modeling: user-based, time-based, and content-based features [5]. In the user-based feature, it includes features related to the user, such as the number of followers. The time-based feature includes information about when the tweet was created. The content-based feature includes attributes related to the content of the tweet, such as whether the tweet contains videos or photos. In the study conducted by [7], the prediction of whether a tweet would be retweeted or not was performed using classification methods such as Decision Tree J48, Support Vector Machine, and Logistic Regression.

In this research, the author aims to predict retweets by utilizing user-based and content-based features using Artificial Neural Network (ANN) method, which will be optimized with the Firefly Algorithm.