1. INTRODUCTION

In today's modern era, social media is growing rapidly along with the development of the internet and technology. People often share opinions or responses about an event, product, activity, or someone through social media, which results in interaction between users. Those who share common interests or simply like to be in a community often come together and create groups online, especially on Facebook. The reason is that Facebook social media provides convenience for its users to create a group with features that can be customized according to their needs.

Facebook is a popular social network and is widely used by Indonesian and global people. According to various online sources, one of them is statista [1], statistically, as of January 2022, Facebook is still ranked first as the most popular social network in the world and also in Indonesia with the highest number of active users. Related to that, people often share information, criticism, and opinions about something or someone to the public through social media. As a result, Facebook has a large and diverse dataset that can be used to overcome the limitations of lab-based studies by providing access to records of user behavior expressed in a natural environment [2] so that researchers no longer need to rely on traditional survey methods.

Referring to the Ministry of Education and Culture policy regarding the Tracer Study program, every university is required to conduct an alumni survey that aims to measure the performance of the college in shaping its students to be ready to enter the working world as well as input for evaluation. For this reason, sentiment analysis can be done as an attempt to conduct the survey without having to interact directly with alumni. Sentiment analysis is one of the important areas of research in social media analysis because it concentrates on detecting the polarity of opinions or emotions from texts on social media [3].

Research on sentiment analysis and topic modeling has been carried out by several researchers. In a study conducted by Handayani et al in 2020, The Support Vector Machine algorithm is used to classify positive and negative sentiments towards comments from BNI Mobile Banking application users. By applying the K-Fold Cross-Validation method, the highest accuracy value was originally 78.19% increased to 78.45%. 10-Fold Cross-Validation is also used because the method has become the standard validation method of previous research [4].

In another study conducted by Jaman et al in 2019, sentiment analysis was performed on tweets discussing online motorcycle taxi services collected from Twitter using the SVM method with the TF-IDF feature selection. The dataset is divided into three classes: positive, neutral, and negative sentiment. The classification process is carried out using several scenarios of comparing train and test data, which are 50:50, 60:40, 70:30, 80:20, and 90:10. The four kernels used in the classification process are linear, RBF, sigmoid and polynomial. The highest accuracy results were obtained in the scenario of 90% of the train data and 10% of the test data using linear and sigmoid kernels of more than 80% [5].

Moreover, in the research conducted by Kumari et al in 2017, SVM is applied to a dataset containing smartphone product reviews to determine the polarity of the sentiment, whether it is positive or negative. The highest accuracy value of 90.99% was obtained, and it was also mentioned that SVM is a better and robust method [6].

Another study, performed by Najadat et al in 2018, sentiment analysis of customer status on the official Facebook pages of 3 Jordanian telecommunications companies using and comparing several supervised learning methods, which are K Nearest Neighbors, Support Vector Machine, Naïve Bayes, and Decision Tree. The results obtained are that SVM is the most superior compared to the other three methods in terms of accuracy and F-measure in each experimental scenario [7].

Furthermore, in the research performed by Rahmadan et al in 2020, sentiment analysis and topic modeling were carried out on a collection of Indonesian-language tweets discussing the flood disaster in Jakarta. Researchers implemented a lexicon-based sentiment analysis that resulted in 10% positive sentiment, 11% neutral sentiment, and 79% negative sentiment. Then, topic modeling is carried out using the LDA method. Nine topics were obtained consisting of the distribution of words that generally contained information about flood areas, the impact caused by floods, conditions when the disaster occurred, and input from the people to the government regarding flood disaster management [8].

In this study, the sentiment analysis model was built to classify the polarity of discussions from alumni on the Telkom University Alumni Forum (FAST) Facebook group using the Support Vector Machine (SVM) method because the method proved efficient in overcoming the problem of imbalanced data in sentiment analysis [3]. In addition, based on previous studies, the SVM algorithm can produce the highest confusion matrix value compared to other classification algorithms. Proven by the result of research by Jaman et al [5] that using SVM can produce a high accuracy of more than 80%. The topics discussed can be modeled with the Latent Dirichlet Allocation (LDA) method. With this, it is hoped that the model built will be able to classify alumni sentiments correctly and accurately so that it can be input as an evaluation for the university to develop the quality of its services and maintain its title as the best private university.