

## **Abstract**

**The development of information technology that is so fast is able to make data increase every day because many people use the ease of technology as a place or means to find information, one of which is film. The large amount of data available on film review sites makes it difficult to obtain the information contained in it. Sentiment analysis can be a solution in processing these data because it can classify audience reviews into positive and negative categories, and make the form of an initially unstructured document into a more structured form so that it can obtain important information from the quality of a film. In this Final Project, several tests were involved to determine the effect of performance by comparing the preprocessing stage carried out between preprocessing using stemming and preprocessing using lemmatization. After that, make a comparison between using the Information Gain feature selection and not using the Information Gain feature selection. The results of this study found that preprocessing using lemmatization and Information Gain feature selection got results with an accuracy of 94.00% compared to without using the Information Gain feature selection that is equal to 92.00%.**

**Keywords: sentiment analysis, movie review, preprocessing, classification, support vector machine, information gain.**