

1. Introduction

As a human being, we need guidelines or rules to follow in order to live as a decent person in this planet. According to Islamic, in addition to following the instructions of the Qur'an, we are also expected to follow the Hadith. Hadith is referred to everything pertaining to the Prophet Muhammad SAW before and after his prophethood, including his deeds, words, and personality traits[1].

There are various principles in the Hadith that we might take to heart. These principles can be divided into three groups: the first group is guidance, the second in prohibition, and the last group is information[2]. It is essential that we are able to recognize hadiths as well as their categories so that there are no inaccuracies when studying or reading hadith. So categorization is expected to make it efficient for us to read or study hadiths, since it is important to be able to distinguish hadiths along with their categories[2].

Research on the classification of Indonesian translations of hadiths has been carried out by Ghina Zain Nabiilah, Said Al Farah, and Mahendra Dwifabri Purbalaksono with a study entitled "Topic Classification of Indonesian Translated Hadith Using K-Nearest Neighbor and Chi-Square"[2]. In this study, the classification used was a multi-label classification with Chi-Square feature extraction and the K-Nearest Neighbor model with an accuracy of 88.29%. In this study, the researcher wanted to use the Chi-Square selection feature which had previously been proven effective in research[2] to select features in the Indonesian translation hadith dataset. The classification model that will be used in this study is the XGBoost model which has previously been used for sentiment analysis research on Indonesian motorcycle taxi services[3] which obtains an accuracy value of 92%, using the Chi-Square extraction feature which was previously also used in Indonesian translation of hadiths topic classification research[2] for feature extraction in his research. By looking at the results of sentiment analysis research on Indonesian gojek services[3] and research on the topic classification of hadith translations using the KNN method[2], XGBoost model is chosen as a model and Chi-Square as a feature selection because it has been proven effective and managed to get good results and optimal in previous studies.