

ABSTRACT

Student complaints at Telkom University submitted through the Telyufess account on the social media platform X have not been optimally utilized as input for evaluating campus services. This study aims to classify tweets from the Telyufess account—a complaint platform for Telkom University students—into two categories: domain-related (related to rectorate units such as academics, finance, and campus services) and non-domain-related (general complaints not associated with official units). The main issue addressed is the need for automatic mapping of student complaints to support campus service evaluation. The classification method used is Naïve Bayes, with stages including manual labeling by the researcher and assistant annotators (validated using inter-rater consistency), text preprocessing (normalization using a standard dictionary and the Sastrawi library, removal of special characters, stopword filtering based on the standard Indonesian list with additional specific words such as “telyu!”), tokenization, stemming, TF-IDF weighting, and data splitting with ratios of 65:35, 70:30, 80:20, and 90:10. A total of 1,090 tweets were collected (January 1, 2023 – January 1, 2025) using the Tweet Harvest API, based on criteria of complaints, opinions, and suggestions (retweets were excluded). The highest accuracy achieved was 87.27% with a 90:10 split, followed by 84.93% (80:20), 81.10% (70:30), and 79.32% (65:35). The results show that Naïve Bayes is effective, although the model experienced overfitting (training data evaluation >99%). This study contributes to the use of social media as a data source for campus service evaluation through an automatic classification system. Visualization was also carried out on the evaluation results to help the Telkom University rectorate identify frequently reported issues more easily.

Keywords: *field related, classification, Naïve Bayes, non field related, telyufess*