

***Classification Of Social Media Influencers Category Based On Text-Post Using
K-NN Method
Supervisor (1 people)***

**Faiz Iqbal Darmawan
Study Program of Informatics Engineering
Majoring of Information Technology**

ABSTRACT

The popularity of Instagram in Indonesia, particularly among millennials (Generation Y), has resulted in a significant rise in influencers. However, both the audience and companies face challenges in identifying relevant influencers for their desired themes. Often, the audience selects influencers randomly without considering relevancy. In light of this, the idea of applying Machine Learning concepts to the influencer selection process emerged, specifically using the K-Nearest Neighbor algorithm, aiming to facilitate the audience and vendors in choosing influencers based on relevant categories. This research utilized a dataset of 1500 labeled instances, divided into 500 instances labeled as Beauty, 500 instances labeled as Food, and 500 instances labeled as Fashion. The evaluation was conducted using a confusion matrix to compare three different data ratios with $k=3$. The first ratio, 9:1, achieved an accuracy of 99%. The second ratio, 8:2, yielded an accuracy of 97%, and the third ratio, 7:3, also resulted in an accuracy of 97%. From the testing results, it can be concluded that the number of training data significantly affects the model's performance and accuracy.

Keywords : instagram, influencer, machine learning, classification, k-nearest neighbor