

**The implementation of K-means clustering to determine the interests of
social media users based on text postings**

Rajih Kharissuha

**Study Program of Informatics Engineering
Majoring in Information Technology**

ABSTRACT

Instagram is a social media platform that provides communication services, including image, video, and text communication, offering various personal and public information. As a communication medium, the abundance of posts from each social media user creates a new problem where other users struggle to determine with whom to collaborate to market their products because they are unaware of the specific preferences of each profile. Therefore, this research aims to develop a system that can assist in determining compatibility among social media users. Hence, the author conducted a study on the implementation of k-means *clustering* for user profile interests on Instagram. In this study, the researcher used 1000 data from 4 Instagram user profiles as test data. Based on the Instagram data, *clustering* was performed using the K-Means method. The research results, obtained through accuracy testing by comparing the system's results with those from RapidMiner software, showed an accuracy of 82.68%. This accuracy depends on the amount of data, indicating that the accuracy level in reading Instagram data has a significant impact. For further research, additional Instagram data and updated preprocessing are needed to enable language translation for documents containing both English and Indonesian words