

## **ABSTRACT**

### ***Accuracy Comparison of the K-Nearest Neighbors and Random Forest Algorithms in Predicting the Risk of Diabetes Mellitus***

Dzaky Athallah Luqman <sup>1)</sup>, Nur Uddin, Ph.D. <sup>2)</sup>

<sup>1)</sup> *Student of Informatic Department, Universitas Pembangunan Jaya*

<sup>2)</sup> *Lecturer of Informatic Department, Universitas Pembangunan Jaya*

*The International Diabetes Federation (IDF) states that by 2021, diabetes mellitus has been the cause of death for as many as 6.1 million people in the world. Artificial intelligence technology or artificial intelligence can be used to predict whether a person is indicated for diabetes mellitus or not. Based on this, a necessary way to predict diabetes is to build a machine learning model and then compare the accuracy value to the machine learning model used. This study aims to build a machine learning model and to determine the accuracy of KNN and Random Forest for prediction of diabetes mellitus. In this case, the machine learning algorithm K-Nearest Neighbors (KNN) and Random Forest can be used to predict diabetes mellitus using features including cholesterol, HBA1C, age, BMI, gender, and other features based on the classification of training data provided. . This research uses qualitative research methods with literacy studies. Collecting data in the form of literacy studies, dataset retrieval, and testing using programming code based on the Confusion Matrix method in developing machine learning models. Based on what has been done, KNN and Random Forest can be used to predict diabetes based on datasets with similar accuracy results, KNN accuracy reaches 98% as well as Random Forest and evaluation of the class using the Confusion Matrix is deeper. Identifying that Random Forest is slightly more accurate than the KNN.*

**Keywords** : *Diabetes mellitus, KNN, Random Forest, Dataset, machine learning*

**Libraries** : 47

**Publication Years** : 2012 – 2021