Abstract

Diabetes is a metabolic disease characterized by hyperglycemia resulting from defects in insulin secretion, insulin action, or both. Diabetes mellitus is a health problem that is getting attention in the 21st century. Several studies related to diabetes mellitus prove that late and improper handling of people with diabetes mellitus causes blood glucose to be out of control for a long time. This condition causes serious changes in the heart, blood vessels of the brain and leg veins, nerves, kidneys, and eyes. Based on this, to prevent and find out someone has diabetes or not is very important. This study uses the Naïve Bayes method to predict people with diabetes based on drug review content. The N-Gram and TF-IDF (Term Frequency - Inverse Document Frequency) methods are used to extract features. The results of the study using unigram feature extraction resulted in an accuracy of 92.1% and an f-measure of 92.5%.

Keywords: Diabetes, Naïve Bayes, N-Gram, TF-IDF