Kabarak University International Conference on Computing and Information Systems 2019

Contribution ID: **19** Type: **Abstract for Research Paper**

# MACHINE LEARNING SMS SPAM DETECTION MODEL

Millions of shillings are lost by mobile phone users every year in Kenya due to SMS Spam, a social engineering skill attempting to obtain sensitive information such as passwords, Personal identification numbers and other details by masquerading as a trustworthy entity in an electronic commerce. The design of efficient fraud detection algorithm and techniques is key to reducing these losses. Fraud detection using machine learning is a new approach of detecting fraud especially in Mobile commerce. The design of fraud detection techniques in a mobile platform is challenging due to the non-stationary distribution of the data. Most machine learning techniques especially in SMs Spam deal with one language. It is in this background that the study will focus on a client side SMs Spam detection in Kenya's mobile environment for both English and Swahili text messages using machine learning. Naive's Bayes algorithm was used for this purpose because it is highly scalable in text classification. The contributors of Corpus include mobile service providers in Kenya and selected mobile phone users. Machine learning and data mining experiments were conducted using WEKA (Waikato Environment for Knowledge Analysis). The results and discussions are presented in form of descriptive statistics and detection metrics. At the end an android based prototype implementation of the SMs Spam detection model is demonstrated that based on Naive Bayes machine learning algorithm .This model gave an overall classification accuracy of 96.1039% .

## Keywords

Algorithm, Classification, Detection, Machine learning, Naïve bayes, WEKA.

**Primary author:** Mr KIPKEBUT, ANDREW (KABARAK UNIVERSITY)

**Track Classification:** Artificial Intelligence Advances and Applications for Development