ABSTRACT

The American Cancer Society estimates that one in every nine men sixty and above will be diagnosed with this disease. A major challenge with the diagnosis of this disease is that it can go a long time without detection and when detected, tests are difficult and usually time consuming. This project tends to proffer a solution to this challenge of testing and diagnosis. This project outlines the development of a decision support system for people who are at the risk of being diagnosed with prostate cancer. Approaching the problem as a classification problem, the major aim of the project is to use the Support Vector Machine model in the design and development of an Application Programming Interface to assist the diagnosis of the disease. I performed a comparative analysis of the Support Vector Machine to determine the effectiveness of the model in the analysis of the test cases. This is to effectively predict prostate cancer in test samples. I compared the model with the Naïve Bayes model and the Logistic regression. It was found to have an approximate accuracy of 90% above the 83% and the 80% of the other two models, respectively. The model was used in the creation of an Application Programming Interface and was found to be an effective solution given its high accuracy and its effective management of large datasets. In conclusion, the use of the Support Vector Machine was found to be an effective model in the classification, analysis and prediction of prostate cancer samples for effective analysis and diagnosis. However, recommendations suggest that the model be tested with wider range data to improve system optimization. This would further ensure accurate diagnosis of the disease and establish optimal accuracy levels for the decision support system.

**Key words: prostate cancer, diagnosis, tests, Machine learning techniques, Support Vector Machine, Logistic Regression, Naive Bayes.**