Early Gastric Cancer prediction from Diet and Epidemiological data using Machine Learning

Brindha Senthil Kumar[1]*, Senthil Kumar Nachimuthu[2], Arindam Maitra[3], Prem Nath[1]

[1]Department of Computer Engineering, Mizoram University, India

[2] Department of Biotechnology, Mizoram University, , India

[3] National Institute of Biomedical Genomics, Kalayani, West Bengal– 741251, India

Abstract

Gastric cancer is predominantly caused by demographic and diet factors as compared to other cancer types. The aim of the study is to predict Early Gastric Cancer (ECG) from diet and lifestyle factors using supervised machine learning algorithms. For this study, 80 patients and 160 healthy individual cases were selected. A dataset containing 10 features that are core risk factors for the gastric cancer were extracted. Supervised machine algorithms: Logistic Regression, Naive Bayes, Support Vector Machine (SVM), Multilayer perceptron, and Random Forest were used to analyze the dataset using Python Jupyter Notebook Version 3. The obtained classified results had been evaluated using metrics parameters: minimum_false_positives, brier_score, accuracy, precision, recall, F1_score, and Receiver Operating Characteristics (ROC) curve. Data analysis results showed Naive Bayes - 88, 0.11; Random Forest - 83, 0.16; SVM - 77, 0.22; Logistic Regression - 75, 0.25 and Multilayer perceptron - 72, 0.27 with respect to accuracy and brier_score in percent. Naive Bayes algorithm classification results in predicting ECG showed very satisfactory results using only diet cum lifestyle factors which will be very helpful for the physicians to educate the patients and public, thereby mortality of gastric cancer can be reduced/avoided with this knowledge mining work.

Biography:

Brindha Senthil Kumar is Postgraduate Scholar pursuing her M.Tech, final year in Mizoram University. She is well aquinted in C, C++, Java and Python. She is interested in applying Machine Learning algorithms to mine the information from Healthcare Data.

Speaker Publications:
1. "Characterization of Bacillus thuringiensis Cry1 class proteins in relation to their insecticidal action"
2. "Analysis of aminoacids pattern in receptor tyrosine kinase using Boolean Association Rule"

7th World Machine Learning and Deep Learning Congress, Webinar- June 18-19, 2020

Abstract Citation:

Brindh Senthilkumar, Early Gastric Cancer prediction from Diet and Epidemiological data using Machine Learning, Machine Learning 2020, 7th World Machine Learning and Deep Learning Congress, Webinar, June 18-19, 2020

https://machinelearning.conferenceseries.com/2020