

# **A Novel Feature Selection Method for Classification of Medical Data Using Filters, Wrappers, and Embedded Approaches**

## **ABSTRACT**

Feature selection is the process of identifying the most relevant features from the given data having a large feature space. Microarray datasets are comprised of high-quality features and very few samples of data. Feature selection is performed on such datasets to identify the optimal feature subset. The major goal of feature selection is to improve the accuracy by identifying a minimal feature subset. For this purpose, the proposed research focused on analyzing and identifying effective feature selection algorithms. A novel framework is proposed which utilizes different feature selection methods from filters, wrappers, and embedded algorithms. Furthermore, classification is then performed on selected features to classify the data using a support vector machine (SVM) classifier. Two publically available benchmark datasets are used, i.e., the Microarray dataset and the Cleveland Heart Disease dataset, for experimentation and analysis, and they are archived from the UCI data repository. The performance of SVM is analyzed using accuracy, sensitivity, specificity, and f-measure. The accuracy of 94.45% and 91% is achieved on each dataset, respectively.