



025-BM

Cancer Classification using Data Mining Applications

Ahmed Ramadan Youssef, Ahmed Adel Ragab, Aya Mohamed Sadek, Shima Moustafa Korany,
Abdelrahman Gamal Youssef, Abdelrahman Ibrahim
Fayoum University, Egypt, ar1521@fayoum.edu.eg, ahmedhoyhy@gmail.com, ayasadeksaleh@gmail.com,,
shimaamoustafafci95@gmail.com, abdelrahman.gamal137@gmail.com, abdelrhman.ibrahim101@gmail.com

Supervisor: Rania Ahmed Abdel Azeem Abul Seoud

Department of Electrical Engineering, Communication and Electronics Section; Faculty of Engineering, Fayoum University;
Fayoum, Egypt,
r-abulseoud@k-space.org, raa00@fayoum.edu

Supervisor: Dina Ahmed Salem

Dept. of Computer Eng.-Faculty of Eng. Misr University for Science and Technology, 6th of October, Giza, Egypt
dina.almahdy@must.edu.eg

The correct interpretation of the biological data is the main goal of Bioinformatics. One emerging and reliable source of data is the microarray technology which is considered a breakthrough in Bioinformatics. Cancer classification using microarray data is a challenge due to the enormous number of features compared to the samples. In the current work, an algorithm was developed in order to classify cancer samples. The developed algorithm was conducted on two steps. In the first step, the feature selection technique was applied on the data to eliminate any undesired features of little or no predictive information. The feature selection technique was based on Entropy and F-score measurements. Then, the classification process was performed using linear support vector machine (SVM), K-Nearest Neighbor (KNN) and Naive Bayes (NB) algorithms, the results achieved were 100% using Naive Bayes, 97% using Linear SVM and 94% using KNN on leukemia dataset. The ability of the developed algorithm for classifying the samples was practically examined using leukemia microarray dataset. The results showed that the developed algorithm could detect and classify all the samples. Then we generalized the algorithm to be applied on different microarray datasets such as Prostate and Colon.