Abstract

Cancer is one of the leading causes of death worldwide. Early detection and prevention of cancer plays a very important role in reducing deaths caused by cancer. Identification of genetic and environmental factors is very important in developing novel methods to detect and prevent cancer. Therefore a novel multi layered method combining clustering and decision tree techniques to build a cancer risk prediction system is proposed here which predicts lung, breast, oral, cervix, stomach and blood cancers and is also user friendly, time and cost saving. This research uses data mining technology such as classification, clustering and prediction to identify potential cancer patients. The gathered data is preprocessed, fed into the database and classified to yield significant patterns using decision tree algorithm. Then the data is clustered using K-means clustering algorithm to separate cancer and non cancer patient data. Further the cancer cluster is subdivided into six clusters. Finally a prediction system is developed to analyze risk levels which help in prognosis. This research helps in detection of a person’s predisposition for cancer before going for clinical and lab tests which is cost and time consuming.
Early Detection and Prevention of Cancer using Data Mining Techniques

---

**References**

- Zakaria Suliman zubi "Improves Treatment Programs of Lung Cancer using Data Mining Techniques"; Journal of Software Engineering and Applications, February 2014, 7, 69-77
- S. Santosh Kumar "Development of an Efficient Clustering Technique for Colon Dataset"; International Journal of Engineering and Innovative Technology; Volume 1, Issue 5, May 2012 ISSN: 2277-3754
Early Detection and Prevention of Cancer using Data Mining Techniques

- K. Kalaivani "Childhood Cancer-a Hospital based study using Decision Tree Techniques" Journal of Computer Science 7(12): 1819-1823, 2011 ISSN: 1549-3636
- Shomona Gracia Jacob "Data Mining in Clinical Data Sets: A. Review" International Journals of Applied Information System (IJAIS) - ISSN: 2249-0868 Foundation of Computer Science FCS, New York, USA, Volume 4-No. 6, December 2012-www. ijais. org
- K. Balachandran "Classifiers based Approach for Pre-Diagnosis of Lung Cancer Disease" International Journal of Computer Applications® (IJCA) (0975 – 8887)

Index Terms

Computer Science Data Mining

Keywords

Decision Tree  k-means  Prediction  Prognosis  Risk Levels