

# **AN ANALYSIS OF HEPATITIS C VIRUS PREDICTION USING DIFFERENT DATA MINING TECHNIQUES**

**AHMED A. A. RADWAN, TAREK ABD-EL-HAFEEZ & HEBA MAMDOUH**

Faculty of Science, Computer Science Department, Minia University, El Minia, Egypt

## **ABSTRACT**

The prediction of hepatitis C virus (HCV) is a significant and tedious task in medicine. The healthcare environment is generally perceived as being 'information rich' yet 'knowledge poor'. There is a wealth of data available within the healthcare system. However, there is a lack of effective analysis tools to discover hidden relationships and trends in data. Knowledge discovery and data mining have found numerous applications in business and scientific domain. Valuable knowledge can be discovered from application of data mining techniques in healthcare system. Using medical profile such as age, sex, residence and (ALT, AST) enzyme blood tests it can predict the likelihood of patients getting HCV infection. It enables significant knowledge, e.g. patterns, relationships between medical factors related to HCV, to be established. It can serve a training tool to train nurses and medical students to diagnose patients infected with HCV. This paper analyses the performance of various classification function techniques in data mining for predicting the infection with HCV from the HCV data set. These Techniques are Decision Trees, Naïve Bayes and Neural Network. Results show that each technique has its unique strength in realizing the objectives of the defined mining goals. Also the performance of three data mining techniques is compared using three data sets of different size.

**KEYWORDS:** HCV Infection, Data Mining, Decision Tree, Naive Bayes, Classification, Neural Networks