PREDICTION OF THROMBOCYTOSIS AND INFECTIONS THROUGH BLOOD REPORTS

Hemaka Gajendra Seneviratne

A dissertation submitted in partial fulfilment of the requirement for
Bachelor of Engineering (Honours) degree in Software Engineering

Department of Computing

Informatics Institute of Technology, Sri Lanka
in collaboration with
University of Westminster, UK

2020

II.   Abstract

With the advancement of technology the revolution has begun in the field of diagnostic medicine as more and more computer based diagnostic solutions are being introduced. Such systems have a great demand as it will provide ease when taking diagnostic decisions.

Hematological diseases which can be diagnosed with the help of routine blood tests have been a very interesting topic to explore. Blood tests like complete blood count with 5 part differentials can be used to diagnose many conditions. Platelets are a types of cells in the blood that are responsible for clotting. A higher platelet count can mean a patient has a condition called thrombocytosis. This can be seen as a dangerous condition, due to the increased platelet levels it can lead to clotting tendencies. There are two forms of thrombocytosis primary and secondary, Secondary thrombocytosis can be considered very common. Infections are usually incorporated with thrombocytosis and there is greater chance that infections can result in thrombocytosis therefore this tendency is also considered in this research. The most common underlying causes of reactive thrombocytosis can be identified as Iron deficiency anemia and infections. A reasonable gap exists where there is no existing system that provide diagnosis of thrombocytosis and infections that can cause secondary thrombocytosis.

Data mining is making use of data to process it and find patterns in the datasets to incorporate and predict on the possible outcomes of the said data. Decision trees are supervised algorithms which can be used to create predictive models. This research aims at providing a predictive model that will   predict thrombocytosis.

Keywords- Data mining, Disease prediction, thrombocytosis, Decision Trees.