PREDICTION OF HEMATOLOGICAL DISEASES USING ARTIFICIAL NEURAL NETWORK ENSEMBLES

S. Sandharuwan Abeykoon

A dissertation submitted in partial fulfillment for the requirement for
Master of Science degree in Advanced Software Engineering

Department of Computing

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

2020

Abstract

Anemia and Thrombocytopenia are hematological diseases that affect blood cells. Primarily, doctors use a complete blood count report (CBC) to analyze various parameters that correlate with the diagnosis of hematological diseases. In this research, the author has applied deep learning techniques with artificial neural network ensembles to discover anemia and thrombocytopenia from the parameters of the complete blood count (CBC) test.

Anemia is a disorder of the blood that does not have enough healthy red blood cells to carry enough oxygen to the tissues of the body. Anemia causes a person to become tired and unhealthy. There are several classifications of anemia and they differ from each other. Anemia can be interpreted as the most frequent hematological disease in the world, anemia hits 1.62 billion people, in proportion to 24.8% of the population. According to the latest stats, globally highest prevalence is in preschool children, and the lowest prevalence is in men. Thrombocytopenia, or fewer platelets, is way less normal platelets in the blood. Platelets are produced in the bone marrow with other blood cells. This research focuses on the classification of immune thrombocytopenia (ITP) that the autoimmune disorder calls abnormal low blood cell platelets. Globally, ITP is supposed to influence more than 200,000 people.

This procedure belongs to the post-analytical stage of diagnosis. Due to having many varieties of anemia and thrombocytopenia, when the variant of the disease analyzing, trained and unspecialized doctors can produce failures. To overcome this, a system will be developed followed by a deep learning approach using artificial neural network ensemble techniques. To generate the first three training models, different activation and optimization techniques are used under artificial neural networks. Finally, the author created an integrated stacking model by concatenating the first three models to improve the final accuracy and best-fitting. In conclusion, this prediction system can predict four anemia classifications and one thrombocytopenia classification using seven CBC parameters.

Keywords: Anemia, Thrombocytopenia, Deep Learning, Artificial Neural Network, Complete Blood Count.