

Informatics Institute of Technology

In Collaboration with

University of Westminster, UK.

Infectious disease outbreak pattern recognition & prediction system

A dissertation by

Mr. I. D. D. Madushanka

Supervised by

Mr. Saman Hettiarachchi

Submitted in partial fulfillment of the requirements for the

BSc (Hons) Software Engineering degree

Department of Computing

May 2017

© The copyright for this project and all its associated products resides with Informatics
Institute of Technology

Abstract

It is a certainty that there is a huge effect on morbidity and mortality rate of Sri Lanka due to infectious disease occurrences in every year. The best way to reduce the burden of infectious diseases is, identifying disease occurrences in early stages and take actions to prevent.

According to the World Health Organization publications, there is a considerable variation in the nature of each infectious disease and also a variety of causes and factors involving for each infectious disease from country to country, reflecting economic, social, cultural and epidemiological contrasts. Because of that each country cannot rely on a general solution.

Also, the most of the existing solutions are mainly focused on generating predictive results for fewest numbers of diseases. So the purpose of the project is to research and develop an infectious disease occurrences prediction system for Sri Lanka based on its economic, social, cultural and epidemiological contrasts.

Sri Lanka mortality and morbidity data records of the past four years will be used as the main data source of the system to train the machine learning program in order to create predictive data models. And also data records of causes and factors such as climate changes, festivals & events will be collected in order to find matching patterns against disease occurrences.

Subject Descriptors:

• **Computing methodologies~Ranking** • **Computing methodologies~Supervised learning by classification** • *Computing methodologies~Classification and regression trees* • *Computing methodologies~Perceptron algorithm* • *Computing methodologies~Feature selection* • **Information systems~Relational database model**

Keywords:

Infectious diseases, epidemiology, machine learning, supervised learning, data mining, classification, feature ranking.