

Meta-heuristic Swarm Intelligence based algorithm for feature selection and prediction of Arrhythmia

Nandini Nayar, Sachin Ahuja, Shaily Jain

Abstract—The healthcare industry is developing extensively, thus more high-dimensional datasets are emerging. Thus it is imperative for health maintenance industries to mine healthcare data effectively, that can be used for accurate prognosis of diseases. Since, all the available features in these datasets are not valuable and more number of irrelevant features may negatively affect the performance of classifier model. Therefore, accuracy of models can be augmented by using a prudently selected subset of features, instead of using each available feature. Feature subsets yield superior result as compared to entire set of features. However, choosing effectual and key features is a tedious task in healthcare datasets. Therefore, in this paper, an algorithm based on Ant Colony Optimization and Simulated Annealing concept is proposed for selecting valuable feature subset in high-dimensional arrhythmia dataset for detection of disease using decision tree. As arrhythmia is considered to be a grievous disease, so timely detection and prevention would be valuable for patients. Experimental results indicate that the proposed algorithm outperforms other approaches in terms of accuracy.

Index Terms— Swarm Intelligence, Ant Colony Optimization, Simulated Annealing, Temperature, Feature Selection

1 INTRODUCTION

Numerous data mining techniques are proposed to extract useful patterns from medical data that are valuable for healthcare experts in the process of decision-making. However, in the real world, the domain of features is ever-expanding, that leads to difficulty in pattern detection. Researchers have proved that reduced subset of features yield more accuracy as compared to whole set of features. Thus, it is imperative to select informative and useful subset of features. Numerous techniques are developed to perform feature selection in healthcare datasets. As a part of our research, we will focus on Ant Colony Optimization (ACO) that is used for feature selection for arrhythmia dataset. Arrhythmia is a problem associated with irregular heart-rate. The heart may beat at a slow pace or very fast pace or with some irregular pattern. The dataset is fetched from UCI repository.

The key idea of this approach is to utilize the concept of Simulated Annealing and Decision tree with Ant Colony Optimization to elevate the accuracy. Section 1 outlines the basic concepts and overview of Feature selection, Swarm Intelligence, Ant Colony Optimization and Simulated Annealing. The proposed algorithm and experimental results

are discussed in section 3 and section 4 describes the conclusion and future scope.

1.1 FEATURE SELECTION

The utmost issue involved in analyzing healthcare datasets is huge volume and heterogeneity. As this data is accumulated from numerous sources, thus the data may be complex and may exist in numerous formats. Moreover, the datasets may have enormous features, often termed as 'High Dimensional Datasets'. In order to derive a model that is consistent and more accurate, the number of features must be minimized.

Feature Selection is the procedure of selecting pertinent features that's the key factor to be used in model construction i.e. to derive an optimal feature subset comprising of only relevant features that are valuable to proliferate the performance of classifier, which further facilitates the reduction of prediction errors [1]. Feature selection methods are categorized as: Wrapper, Filter and Embedded models [2]. The benefits of feature selection are summarized in Figure 1.

- Nandini Nayar, Chitkara University School of Engineering and Technology, Chitkara University Himachal Pradesh, India. nandini.nayar@chitkarauniversity.edu.in
- Sachin Ahuja, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India. sachin.ahuja@chitkara.edu.in
- Shaily Jain, Chitkara University School of Engineering and Technology, Chitkara University, Himachal Pradesh, India. shaily.jain@chitkarauniversity.edu.in