

Chapter 12

GA_SVM: A Classification System for Diagnosis of Diabetes

Dilip Kumar Choubey

Birla Institute of Technology Mesra, India

Sanchita Paul

Birla Institute of Technology Mesra, India

ABSTRACT

The modern society is prone to many life-threatening diseases which if diagnosis early can be easily controlled. The implementation of a disease diagnostic system has gained popularity over the years. The main aim of this research is to provide a better diagnosis of diabetes. There are already several existing methods, which have been implemented for the diagnosis of diabetes. In this manuscript, firstly, Polynomial Kernel, RBF Kernel, Sigmoid Function Kernel, Linear Kernel SVM used for the classification of PIDD. Secondly GA used as an Attribute selection method and then used Polynomial Kernel, RBF Kernel, Sigmoid Function Kernel, Linear Kernel SVM on that selected attributes of PIDD for classification. So, here compared the results with and without GA in PIDD, and Linear Kernel proved better among all of the noted above classification methods. It directly seems in the paper that GA is removing insignificant features, reducing the cost and computation time and improving the accuracy, ROC of classification. The proposed method can be also used for other kinds of medical diseases.

1. INTRODUCTION

In this era, one of the needs of this generation is healthcare and diabetes is one of the chronic diseases. Diabetes is the problem used to describe a metabolic condition of having higher than normal blood sugar levels. This is also called hyperglycemia. Diabetes occurs when insulin is not being properly produced or responded by the body, which is needed to maintain the rate of sugar. Diabetes can be controlled with the help of insulin injections, taking oral medications (pills), a controlled diet (changing eating habits), and exercise programs, but no entire cure is available.

DOI: 10.4018/978-1-5225-2128-0.ch012

Diabetes can be categorized into four types: Type 1 (Juvenile or Insulin Dependent or Brittle or Sugar) diabetes and Type 2 (Adult onset or Non Insulin Dependent) diabetes, Pre diabetes and Gestational diabetes.

Type 1 diabetes mostly occurs to children and young adults but can affect at any age, 5-10% of diabetes have Type 1 diabetes. The body does not produce or make insulin. Insulin helps the body to use sugar from food as a source of energy. People with type-1 diabetes need insulin therapy. There is a destruction of insulin secreting cells (β cells) of pancreas in our body. The cause is immune-mediated or idiopathic. It usually requires external insulin therapy hence known as Insulin Dependent Diabetes Mellitus (IDDM).

Type 2 diabetes is the most common type of diabetes, in which people are suffering at least 90-95% of all the diabetes cases. Type 2 diabetes symptoms often develop slowly. This type mostly occurs to the people more than forty years old but can also be found in younger classes. Type 2 diabetes symptoms are usually the same for men and women however, urological problems such as erectile dysfunction (ED). This is asymptomatic for many years. Body does not produce enough insulin (Insulin Deficiency) or the body cells do not respond to insulin (Insulin Resistance) or both. It can be controlled or often need lifestyle modification, taking oral medications (pills) or if required insulin therapy but no entire cure for diabetes is available.

Pre diabetes is a milder form of diabetes that is sometimes called Impaired Glucose Tolerance. It can be diagnosed with a simple blood test.

Gestational diabetes occurs during pregnancy. It raises mother's risk of getting diabetes for the rest of her life. It also raises the child's risk of being overweight and getting diabetes. It displays a high blood sugar level during pregnancy, usually occurs at around 28 weeks or later, and affects about 4% of all pregnant women. This type of diabetes usually goes away after pregnancy and causes are unknown.

In this chapter, for the analysis to diagnose of diabetes, the proposed method is implemented and evaluated by GA as an Attributes selection and used Polynomial Kernel, RBF Kernel, Sigmoid Function, Linear Kernel SVM for classification. By using GA method, 4 attributes obtained among 8 attributes. The mentioned methods have been implemented on Pima Indian Diabetes Dataset (PIDD). Here, used dataset is precise, no missing value have found, noisy free dataset.

The rest of the paper is organized as follows: Brief description of Attribute Selection, GA, Classification, and SVM are in section 2, Problem Specification is introduced in section 3, Related work is presented in section 4, Proposed methodology is discussed in section 5, Results and Discussion of proposed methodology are present in section 6, Conclusion and Work are devoted to section 7.

2. BRIEF DESCRIPTION OF ATTRIBUTE SELECTION, GA, CLASSIFICATION, AND SVM

2.1. Attribute Selection

It is a process of selecting important attributes or best subset of attributes by removing the redundant and irrelevant attributes, which will save resources while improving the classification accuracy.

In particular, the utmost aim of attribute selection is to deduce the number of attributes used in classification, while sustaining the considerable ROC and classification accuracy. However, reduction in the number of attributes or dimensionality helps us is critical in statistical learning. Notably, the attribute selection process helps us to preserve the storage capacity or memory, computation time (shorter training time and test time), computation cost, Processor requirements, and increases classification rate. As we

37 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/gasvm/179399

Related Content

Upper GI Bleed, Etiology, Role of Endoscopy in Rural Population of Punjab

Ravinder Singh Malhotra, K. S. Ded, Arun Gupta, Darpan Bansal and Harneet Singh (2010). *International Journal of Computational Models and Algorithms in Medicine* (pp. 55-68).

www.irma-international.org/article/upper-bleed-etiology-role-endoscopy/47317

Artificial Insect Algorithms for Routing in Wireless Sensor Systems

Li-Minn Ang, Adamu Murtala Zungeru, Kah Phooi Seng and Daryoush Habibi (2015). *Emerging Research on Swarm Intelligence and Algorithm Optimization* (pp. 191-213).

www.irma-international.org/chapter/artificial-insect-algorithms-for-routing-in-wireless-sensor-systems/115305

Quantum Fourier Transforms

(2021). *Examining Quantum Algorithms for Quantum Image Processing* (pp. 157-192).

www.irma-international.org/chapter/quantum-fourier-transforms/261476

Metaheuristic Algorithms for Supply Chain Management Problems

Ata Allah Taleizadeh and Leopoldo Eduardo Cárdenas-Barrón (2013). *Meta-Heuristics Optimization Algorithms in Engineering, Business, Economics, and Finance* (pp. 110-135).

www.irma-international.org/chapter/metaheuristic-algorithms-supply-chain-management/69883

Analysing the Returns-Earnings Relationship: Dempster-Shafer Theory and Evolutionary Computation Based Analyses Using the Classification and Ranking Belief Simplex

Malcolm J. Beynon and Mark Clatworthy (2013). *Meta-Heuristics Optimization Algorithms in Engineering, Business, Economics, and Finance* (pp. 198-222).

www.irma-international.org/chapter/analysing-returns-earnings-relationship/69886