**Review on State of Art Data Mining and Pattern Recognizing Models for Identifies Inherited Diseases.**

# Abstract

Data mining and pattern recognizing methods are reveal interesting findings in genetic studies, specifically on the inherited diseases. Various researchers have proposed data mining models for biomedical approaches to find pathogenesis of inherited disease. The main challenges in genetic inherited disease predict is correctly classifies the genetic factor.

The main aim of the paper is review state-of-art data mining and pattern recognizing models for identifies inherited diseases. Binary classification methods and scoring based prioritization methods are explained in order to provide the scientific background of the problem. There are different approaches are reviewed in this paper such as SNP based binary classification, scoring based genes prioritization, in order to identifies inherited diseases.

# Introduction

The diseases are main threat of the human life. Inherit factor acts main role of the diseases paradigm. Advancement of technology and science contributes, to identify inherited diseases based on DNA. There are various methods introduces in many researches to identify inherited disease based on DNA, genetic variations in DNA sequences ([1](#_ENREF_1)), prioritize candidate genes and identify inherited diseases([2](#_ENREF_2)) and Single nucleotide based genetic mapping method for complex diseases([3](#_ENREF_3)).

It is a challenging task to identify inherited diseases without applying bioinformatics techniques and algorithms([4](#_ENREF_4)).Finding the inherited disease gene from human genome is one of the most major task in bioinformatics research ([5](#_ENREF_5)).Data mining techniques in genetic studies are helpful to identify inherited diseases ([6](#_ENREF_6)).

We examine following areas, in order to better understanding of the DNA contribution of inherited diseases.

1. Single Nucleotide Polymorphism (SNP).
2. Non-synonymous single nucleotide polymorphisms (nsSNPs).
3. Gene ontology for inherited genes.
4. Protein-protein interaction network.

Evaluation will be discussed through advantages and disadvantages of previous work to find appropriate solution for identify inherited diseases based on DNA system.

# Single Nucleotide Polymorphism based approaches for identify inherited diseases

SNPs are genetic variants in single bases of DNA sequence [[1](file:///D:\IIT\l\Final_Year_Project\DNA_Inherit_Disease_Indentiyer\IIDDNA_Review_Paper\IIDDNA_Review_Paper_ijcsit_V1.0-edited.docx#_ENREF_1)]. It has been suggested that different traditional classification methods and novel data mining techniques are applied to classification problem. Also explains effective of identifying the disease associate with SNPs and performance of the different algorithms.

1. Jiaxin W, Wangshu Z, Rui J, editors. Comparative study of ensemble learning approaches in the identification of disease mutations. Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on; 2010 16-18 Oct. 2010.

2. Xie B, Agam G, Sulakhe D, Maltsev N, Chitturi B, Gilliam TC. Prediction of candidate genes for neuropsychiatric disorders using feature-based enrichment. Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine; Orlando, Florida. 2383022: ACM; 2012. p. 564-6.

3. Dawy Z, Sarkis M, Hagenauer J, Mueller JC. Fine-Scale Genetic Mapping Using Independent Component Analysis. Computational Biology and Bioinformatics, IEEE/ACM Transactions on. 2008;5(3):448-60.

4. Tranchevent L-C, Capdevila FB, Nitsch D, De Moor B, De Causmaecker P, Moreau Y. A guide to web tools to prioritize candidate genes. Briefings in Bioinformatics. 2010.

5. Xu J, Li Y. Discovering disease-genes by topological features in human protein–protein interaction network. Bioinformatics. 2006;22(22):2800-5.

6. Fiaschi L, Garibaldi JM, Krasnogor N, editors. A framework for the application of decision trees to the analysis of SNPs data. Computational Intelligence in Bioinformatics and Computational Biology, 2009 CIBCB '09 IEEE Symposium on; 2009 March 30 2009-April 2 2009.