

Deleterious SNP's Detection Using Machine Learning

Authors : Hamza Zidoum

Abstract : This paper investigates the impact of human genetic variation on the function of human proteins using machine-learning algorithms. Single-Nucleotide Polymorphism represents the most common form of human genome variation. We focus on the single amino-acid polymorphism located in the coding region as they can affect the protein function leading to pathologic phenotypic change. We use several supervised Machine Learning methods to identify structural properties correlated with increased risk of the missense mutation being damaging. SVM associated with Principal Component Analysis give the best performance.

Keywords : single-nucleotide polymorphism, machine learning, feature selection, SVM

Conference Title : ICCSEA 2016 : International Conference on Computer Science, Engineering and Applications

Conference Location : Madrid, Spain

Conference Dates : March 24-25, 2016