Clinical-Genomic Analysis using Machine Learning Techniques to Predict Risk of Disease

Yesha, Yelena[1]; Rishe, Naphtali[2]; Lucic, Tajana[3]
[1]University of Maryland Baltimore County, USA; [2]Florida International University, USA; [3]NOA Inc / Polymedicine.com, USA

Recent advances in genomic research have generated vast amounts of information that can help identify individuals who differ in their susceptibility to a particular disease or in responsiveness to a specific treatment. This may offer solutions for the treatment of complex chronic diseases that are influenced by a wide array of factors. Creation of this knowledge brings critical challenges in applying advanced technology to synthesize clinical and genomic patient data. Synthesis is necessary to derive knowledge that would empower physicians to provide personalized care with the best customized therapeutic intervention.

We have employed statistical methods and data mining approaches to understand clinical and genomic risk factors that differentiate Type II Diabetes, Cardio-vascular diseases, Breast Cancer and Prostate Cancer cases from healthy controls. We have mined the data from DbGAP (Databases of genotype and Phenotype) for these diseases. We have investigated whether inclusion of genomic risk factors in conjunction with clinical information improves classification accuracy. We have also demonstrated how a biased and an unbiased method for selection of risk associated single nucleotide polymorphisms (SNPs) affects Clustering (unsupervised learning) along with clinical information. We have determined an optimal method for clustering based on its accuracy and performance. We have also used Classification techniques (supervised learning) to classify patients based on high risk, various socio-demographic parameters and pedigree data. Our goal in this project is to come up with a decision model based on genomic criteria as well as clinical factors the patient. Healthcare professionals will be able to utilize this decision support system in personalizing treatment to patients.