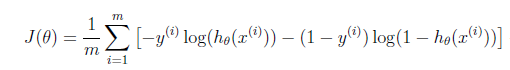
Machine Learning in Genomic Medicine: A Review of Computational Problems

And Data Sets

In the Earlier stages there is no clear info about the genomes but now due to advancement of medical field genome data are available in abundant. A genome is the book which has instructions for building an organism there are two types of genomes they are protein- coding genes and non-coding genes, the human genes contains 20,000 and 25,000 protein-coding and non-coding genes respectively in which some are crucial for life meanwhile some can be removed entirely without any apparent harm. One of the most important structures within a typical gene is the presence of introns and exons. In some genetic diseases normal mechanism is more complex as the disease with multiple causal pathways all leading to similar symptoms but requiring different treatments E.g. Cancer in which genomic data are becoming essential for providing more detailed diagnoses and targeted treatments.

Logistic Regression

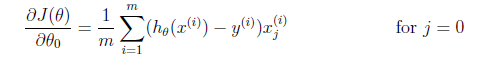
It is used to classify the data into two or more parts. Initially the cost function J is calculated using

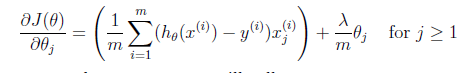


Since our application has more features to avoid overfitting the regularization factor is added

Regularization factor =

Further the cost can be reduced using gradient decent





The multiple classification can be done using one vs all classifier