

19- Deep Learning in Genomics

April 9, 2023

1 Application of DNNs in genomics

1.1 Gene expression prediction

DNNs, with their ability to map the input-output data, are suited for gene expression-based disease classification. For example, CNNs were successfully used to classify Alzheimer's disease.

1.2 SNP prediction

SNPs are commonly used to detect disease-causing genes in humans, predict a person's response to drugs or their susceptibility to developing the disease, and classifying complex diseases using Genomics SNP data.

1.3 Protein structure predictions

It involves modeling the relationship between the amino acids of a protein and its corresponding 3D structure. By 2020, AlphaFold's performance is impressive, and it is now considered the go-to model for predicting protein structure.

1.4 Regulatory genomics

Regulatory genomics is the study of gene regulatory elements such as promoters, enhancers, silencers, insulators, and so on. They play an important role in gene regulation and hence functionally characterizing them is very important. In addition to these gene regulatory elements, identifying sequence motifs in DNA and RNA regulatory regions is key since they represent target sites of a particular regulatory protein, such as the transcription factor (TF).

1.5 Gene regulatory networks (GRNs)

GRNs are defined as networks that are inferred by gene expression data. GRNs are an exciting area of functional genomics and represent causal relationships between the regulators and the target genes. GRNs are important to understand the causal map of network interactions, molecular marker detection, hub gene detection, and so on.

1.6 Single-cell RNA sequencing (scRNA-Seq)

scRNA-Seq enables gene expression measurements in individual cells, thereby enabling cell-type clustering. Despite its huge success, biological inference remains the major limitation because of the sparse nature of the generated data. In addition, there is a large volume of dropout events in the data.