Big Bio-Data Analysis (Artificial Intelligence and Machine Learning)

22 July 2022

Machine Learning & BioInformatics Case Studies

Individual Assignment

By

Richard Sserunjogi
Department of Computer Science,
Makerere University, Uganda
sserurich@gmail.com

AFRICAN CENTERS OF EXCELLENCE IN BIOINFORMATICS & DATA INTENSIVE SCIENCE

House Keeping

Resources

- Introduction to Machine Learning and Bioinformatics

https://github.com/aceuganda/BigDataClass-ACE-Sem-II-2021-2022/tree/main/Resources

CLASS ASSIGNMENT : Expected Outcome

- Get a topic/title from the book [summaries in the slide]
- Presentation [slides for class]
- Technical/scientific write up
- Sample practical solution
- Expected deadline [18 August 2022] : submission of assignment]

Groups

Name	Topic
James	Bayesian Machine-Learning Methods for Tumor Classification Using Gene Expression Data
Farida	Connections between Machine Learning and Bioinformatics
Frank	Statistical Methods for Classifying Mass Spectrometry Database Search Results
Robert	Machine Learning in Structural Biology: Interpreting 3D Protein Images
Baker	Modeling and Analysis of Quantitative Proteomics Data Obtained from iTRAQ Experiments

Probabilistic and Model-Based Learning

"Why are probabilistic and model-based learning relevant in the context of biological systems?"

This is a pertinent question because, after all, probability theory deals with uncertainty and probabilistic models are a way of quantifying uncertainty

Classification Techniques

- **Applications of Classification Techniques to Bioinformatics Problems**
  - Most active area has been the class prediction problem (e.g.different stage of cancer of patients) using primarily gene but more recently protein microarray data
  - Peptide and protein identification in mass spectrometry

Connections between Machine Learning and Bioinformatics

Direct applications of standard machine learning algorithms or specializations of them to particular contexts.
Focus on three problem areas:

- DNA and amino acid sequence analysis,
- gene expression analysis
- network inference.

Machine Learning in Structural Biology: Interpreting 3D Protein Images

Bayesian Machine-Learning Methods for Tumor Classification Using Gene Expression Data

Modeling and Analysis of Quantitative Proteomics Data Obtained from iTRAQ Experiments

Statistical Methods for Classifying Mass Spectrometry Database Search Results

References

Introduction to Machine Learning and Bioinformatics

Thank you!

If you have any questions feel free to email me:

sserurich@gmail.com