BioMed AI: Protein–Protein–Interaction (PPI) Modeling Platform

Project Description

Protein–protein interactions (PPIs) are essential for almost all biological processes in the human body. Proteins rarely act alone. Instead, they work together in networks to regulate cell growth, signaling, metabolism, and disease development. Understanding how proteins interact is especially important in cancer research, drug discovery, and precision medicine.

This project aims to build a reliable and scalable artificial intelligence (AI) system to model the full human protein interaction network. This system integrates protein sequence data, AlphaFold structural information, and large-scale interaction databases into a unified computational framework.

The project has several structured phases; starting from data collection and cleaning, where the task is to collect 20,431 reviewed human proteins and integrate over 2.6 million protein interaction records from public databases such as BioGRID, STRING, IntAct, and BioPlex. Then, preprocessing the collected datasets and prepared for deep learning models utilization. After that, protein sequences will be encoded using the pretrained ESM-2 language model, and other language models for comparison purposes. AlphaFold structures will also be utilized for future structural modeling. Followed by developing deep learning models including MLP pairwise model (sequence-only), Siamese neural network, GraphSAGE, GCN (Graph Convolutional Network), and GSMFormer (Graph Transformer-based model).

Required Skills

- Programming skills in Python.
- Familiarity with deep learning models and knowledge of biological/ bioinformatics data is a plus.
- Problem solving and analytical skills.

Learning Opportunities

- Work on real projects.
- Elevate bioinformatics knowledge.
- Gain experience in using bioinformatics tools.
- Potential opportunity to contribute to research publications.
- Contribute to creation of a new Python platform/app.
- Develop expertise in deep learning modeling.

Expected Team Size: *2-4*

Mentors

Name: Dr. Mohamed Elshrif   email: melshrif@hbku.edu.qa