DeepSF: Deep CNN for mapping protein sequences to folds

Zhiyang's presentation about [1]

Zhiyang Ong

Department of Electrical and Computer Engineering
Dwight Look College of Engineering,
Texas A&M University
College Station, TX

March 25, 2020

Table of Contents

Context of the Problem
Background Information

Context of the problem.

1. The structures of most ($> 99\%$) proteins are unknown
2. Protein fold recognition enables us to associate a protein sequence to a protein fold
   - With this protein fold, determine its protein structure
   - With this protein structure, determine its protein function
     # Insert References!!!
   - Knowing about protein function facilitates Insert References!!!:
     - drug/medication design
     - biotechnology
     - synthetic biology
     - personalized/precision medicine

Problem Definition and Importance
What is the problem that [1] is solving?

Problem description.

1 Sequence-alignment methods for mapping protein sequences to protein folds are indirect

2 Protein fold recognition enables us to associate a protein sequence to a protein fold

- With this protein fold, we can determine its protein structure
- With this protein structure, we can determine its protein function
  - DOI:10.1073/pnas.1102727108
  - https://www.nature.com/scitable/ebooks/essentials-of-cell-biology-14749010/122996920/
  - https://www.nature.com/scitable/ebooks/cell-biology-for-seminars-14760004/122995569/
  - DOI:
  - https://www.ncbi.nlm.nih.gov/books/NBK21177/
  - DOI:10.1006/jmbi.1999.2661

Problem Importance
Why is it important?

Why is this problem important?

- This facilitates protein structure prediction, which is important for drug/medication design and biotechnology.

What are the Knowledge Gaps in [1]?
How would [1] address these knowledge gaps?

List of knowledge gaps:

-

Table of Contents

Proposed Solution(s) - 1
What do they proposed to address the knowledge gap?

Proposed Solution(s).

-

Proposed Solution(s) - 2
What do they proposed to address the knowledge gap?

Proposed Solution(s).

•

Table of Contents

Experimental Results - 1
How do their proposed solutions compare with existing solutions?

Experimental Results - 1.

-

Experimental Results - 2
How do their proposed solutions compare with existing solutions?

Experimental Results - 2.

-

Table of Contents

Discussion of Experimental Results
What do the experimental results tell us?

Discuss the experimental results.

-

Muddy Points
What do I not understand about [1]?

What do I not understand about [1]?

-

References

Jie Hou, Badri Adhikari, and Jianlin Cheng.
DeepSF: deep convolutional neural network for mapping
protein sequences to folds.
Bioinformatics, 34(8):1295–1303, April 15 2018.