GALAGAIN Calvin
calvin.galagain@gmail.com

Lab session # 7
ALTEGRAD 2022

12/21/22

# 1 Question 1

The equation is $MLP(x) = f(b + \sum_i x_i \times w_i)$. So if we want to obtain the sum of a set of integers we need to have a vector with only 0 for the bias and we need to have 1 for each $w_i$.

# 2 Question 2

The general formula for a DeepSets model is of the form with null vector for bias:
$f(X) = W2 \times ReLU(W1 \times X_1) + W2 \times ReLU(W1 \times X_2)$.

Take the following values:

- $W1 = W2 = I_2$
- null bias

$$\begin{aligned} f(X1) &= W2 \times ReLU(W1 \times X1_1) + W2 \times ReLU(W1 \times X1_2) \\ &= W2 \times ReLU(W1 \times [1.2, -0.7]) + W2 \times ReLU(W1 \times [-0.8, 0.5]) \\ &= ReLU(W1 \times [1.2, -0.7]) + ReLU(W1 \times [-0.8, 0.5]) \\ &= [1.2, 0.5] \end{aligned}$$

$$\begin{aligned} f(X2) &= W2 \times ReLU(W1 \times X2_1) + W2 \times ReLU(W1 \times X2_2) \\ &= W2 \times ReLU(W1 \times [0.2, -0.3]) + W2 \times ReLU(W1 \times [0.2, 0.1]) \\ &= [0.4, 0.1] \end{aligned}$$

And it's different as we can see.

# 3 Question 3

To classify sets of graphs, we could use an end-of-string learning approach by combining different models of graphic data processing and natural language processing.

- Data preprocessing: before you can use graphs for learning, they must be pre-processed to be usable by the models. This may include extracting graph characteristics, such as the number of vertices and edges, graph density, etc.
- Use of graphic data processing models: to process graphs, we could use models such as graphic neural networks (GNNs) or graphic attention-based data processing models (Graph Attention Networks, GANs). These models make it possible to capture the relationships between vertices and edges in graphs and use them for classification.
- Use of natural language processing models: if graphs are accompanied by textual descriptions, we could also use natural language processing models, such as Long Short-Term Memory (LSTM) or transformer models, to process these descriptions and use them for classification.
- Use of Deep Sets Model: to process sets of graphs holistically, we could use depth set models that allow you to treat the set of graphs as a whole, rather than treating each graph separately. These models make it possible to capture the relationships between the different graphs as a whole and use them for classification.
- Merger of the results of the different models: once the different models have been trained and their predictions have been obtained, they could be merged in different ways, such as using a weighted average or using a fusion model, to obtain a final prediction for each set of graphs.

# 4 Question 4

The equivariant permutation property of our message transmission layers and the permutation invariance of the reading function in our proposed architecture are appropriate since proteins are amino acid sequences. Indeed, the order of amino acids in a protein can influence its structure and functions, but this order is not necessarily crucial for the classification of proteins into different classes.

However, if we wanted to take into account the order of amino acids in our model, it would be possible to keep the current architecture of the model, but to add a sequence processing layer before the message transmission layers. This layer could use sequence processing techniques, such as convolutions or RNN on amino acid sequences, to capture the order of amino acids in proteins.

# 5 Figures

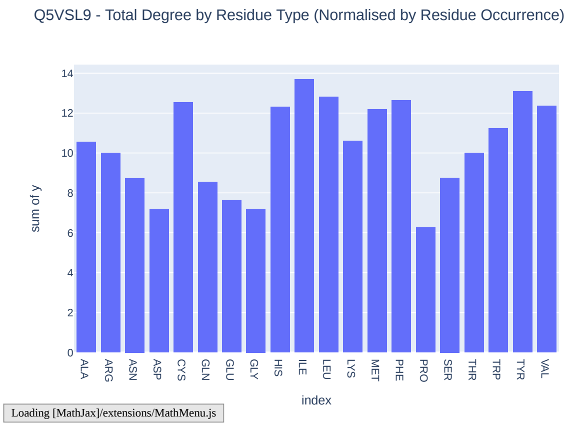

Figure 1: degree_by_residue_type

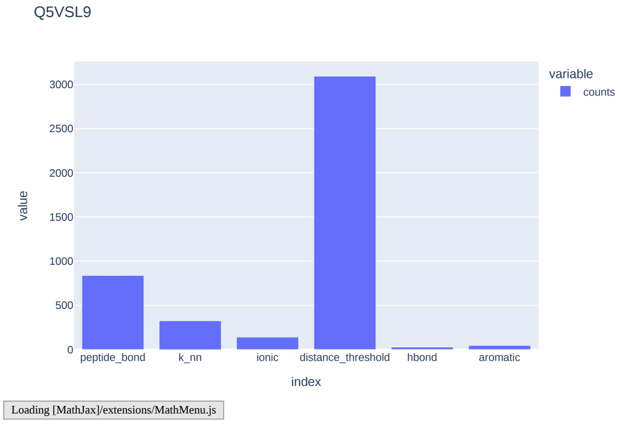

Figure 2: edge_type_distribution

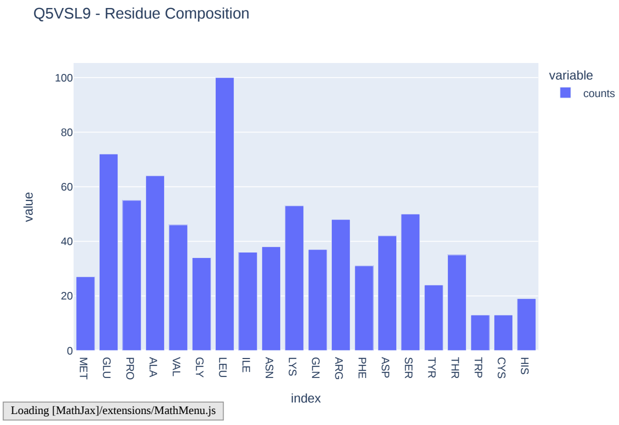

Figure 3: plot_residue_composition

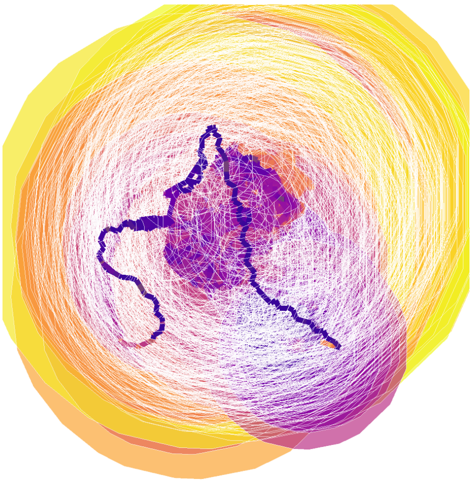

Figure 4: protein_structure_graph