

Plan de travail prévisionnel

Thème : Prédiction de la structure secondaire des protéines à partir de leurs séquences d'acides aminés à l'aide de réseaux de neurones.

Mon github associé :

https://github.com/MathieuxTony/projet_7

Dataset : Il s'agit d'une partie des données se trouvant dans le dossier « data/humans » du github ci-dessous.

<https://github.com/malhotra-sidharth/protein-structure-prediction>

Nature de la prédiction : Il ne s'agit pas de donner une réponse binaire sur le fait qu'il y ait ou non une hélice à un endroit donné de la séquence d'acides aminés, mais d'une probabilité qu'il y ait une hélice à un endroit donné de la séquence d'acides aminés.

Baseline : La probabilité correspond aux fréquences observées sur un jeu d'entraînement.

Méthode mise en œuvre : Divers réseaux de neurones seront testés, en commençant par un simple MLP combiné avec une fenêtre glissante, et en complexifiant suivant les résultats obtenus et le temps disponible.

Sources :

<https://arxiv.org/pdf/2208.11248.pdf>

<https://medium.com/analytics-vidhya/stock-price-prediction-single-neural-network-with-tensorflow-75b64af74ed6>