

## Devoir 3: Revue de la Littérature

Guillermo Martinez (matricule x)  
Dereck Piché (matricule 20177385)  
Jonas Gabirot (matricule 20185863)

February 24, 2023

### **Aptamères**

#### **Description des aptamères**

\*insert\*

#### **Motivation de l'étude des aptamères**

\*insert\*

### **Algorithmes d'apprentissage automatique**

#### **Multilayer Perceptrons**

\*insert\*

#### **Reccurent Neural Networks**

\*insert\*

#### **Transformers**

Selon nos hypothèses, l'architecture \*transformer\* est de loin la plus appropriée pour notre tâche. les transformers utilisent un mecanisme d'attention multi-tête et l'auto-attention. Soit un séquence de tokens d'entrée. Alors le mécanisme d'attention créera pour chaque token d'entrée un autre token qui est une combinaison linéaire des autres vecteurs d'entrée donnés. Ainsi, on [1]. Comme on suppose une forte corrélation entre les éléments de l'entrée, (...) Positional encodings ()

#### **G-flow nets**

\*insert\*

## Algorithmes utilisés présentement (State of The Art)

Il existe présentement peu de recherches et d'articles sur l'apprentissage des aptamères avec des algorithmes d'apprentissage profond. Des algorithmes propres à la biologie sont plutôt favorisés, ainsi que des algorithmes de regroupement (clustering). Par exemple, cet article de janvier 2023 utilise un algorithme original qui combine des méthodes de regroupement pour trouver un aptamère optimal parmi une sélection. <https://pubs.acs.org/doi/pdf/10.1021/acssynbio.2c00462>. Cependant, quelques articles récents utilisent l'apprentissage profond. "Machine learning guided aptamer refinement and discovery" (<https://www.nature.com/articles/s41467-021-22555-9>) fait appel à un réseau neuronal standard MLP pour trouver les aptamères les plus compatibles (high affinity) avec des molécules cibles. L'estimation de l'énergie libre est une sous-étape du calcul de l'affinité. Il performe une étape de troncage pour minimiser la longueur de l'aptamère sans altérer ses propriétés. Un autre modèle d'apprentissage profond avec des aptamères est AptaNet (<https://www.nature.com/articles/s41598-021-85629-0>). Ce modèle utilise un MLP et un CNN pour apprendre la relation entre des aptamères et des protéines cibles (Aptamere-protein relations ou API). Le MLP fonctionne le mieux, avec une précision de test de 91.38. Aucun modèle sur les aptamères publié n'utilise des transformers ou des RNNs pour prédire l'énergie libre, donc notre méthode serait originale dans ce domaine.

## References

- [1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *CoRR*, abs/1706.03762, 2017.