Pré-traitement et analyse de données génomiques à l'aide d'outils de fouille de données

Théo FIGINI
M2 Computer Science
Academic year 2024-2025

Organisme d'accueil : *Université des Antilles*

Enseignant :
XX XX

June, 20th 2025

Contents

Introduction

1 Related Work

Genomics is a field that has been growing rapidly in the past few years. The advent of high-throughput sequencing technologies has made it possible to sequence the entire genome of an organism in a matter of days. This has led to an explosion of data, with the number of sequenced genomes increasing exponentially. This has created a need for new tools and algorithms to analyze this data. In this chapter, we review some of the existing tools and algorithms for analyzing genomic data.

1.1 Genomic data analysis

1.1.1 Random forests

Another method that has been used for genomic data analysis is random forests (RF) [1]. This method is based on the idea of ensemble learning, where multiple decision trees are trained on different subsets of the data and then combined to make a final prediction.

1.2 AI applications in genomic analysis

Travail réalisé

Réalisation

Conclusion

Conclusion

Bibliography

[1] Xi Chen and Hemant Ishwaran. "Random forests for genomic data analysis". In: *Genomics* 99.6 (2012), pp. 323–329. ISSN: 0888-7543. DOI: https://doi.org/10.1016/j.ygeno.2012.04.003. URL: https://www.sciencedirect.com/science/article/pii/S0888754312000626.