

Paper Title: Pattern recognition in bioinformatics

Link: <https://www.nature.com/articles/s41467-021-23458-5>

1 Summary

1.1 Motivation

There is huge increase in the volumes of sequencing data produced lately as the cost of next-generation sequencing for genomics and transcriptomics dropped significantly. As a result, there is a rapid growth of bioinformatics over the past decade leading to the requirement of computational tools to manipulate and handle and making sense out of sequencing data. Pattern recognition has been highly successful in helping to explore and exploit high-throughput measurement data.

1.2 Contribution

This paper is course on pattern recognition (or machine learning) at the core of any bioinformatics education program. In this review, we discuss the main elements of a pattern recognition course, based on material developed for courses taught at the BSc, MSc and PhD levels to an audience of bioinformaticians, computer scientists and life scientists. We pay attention to common problems and pitfalls encountered in applications and in interpretation of the results obtained.

1.3 Methodology

The methodology of the paper mainly explained some of the important steps in pattern recognition pipeline, i.e, a number of computational steps to achieve the goal. This steps includes data representation, preprocessing of data, performance evaluation some pattern recognition algorithms, new developments in pattern recognition, etc.

1.4 Conclusion

In this review, the core elements of a pattern recognition course: data representation; the problems of clustering, dimensionality reduction and classification; performance evaluation; and model selection is discussed.

2 Limitations

2.1 First Limitation

It will be increasingly challenging to tie together various heterogeneous data sources in a single application.

2.2 Second Limitation

As tools for measuring and particularly manipulating the cell become more widely available, pattern recognition should help close the systems biology loop by supporting researchers in setting up experiments.