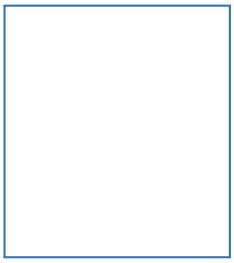

Lorenzo Di Rocco

Education

2021– now **Phd in Methodological Statistics**, *Sapienza University of Rome*, Italy.

2017–2020 **Statistical and Decision Sciences**, *Sapienza University of Rome*, Italy.
Master of Science
Final Mark:110/110 cum laude

2014–2017 **Economic Sciences**, *Roma Tre University*, Italy.
Bachelor of Science
Final Mark: 105/110

Master Thesis

title *Applicazione di Apache Spark e del framework FADE per l'analisi alignment-free di sequenze genomiche*

description My thesis project proposes a novel distributed algorithm, based on the MapReduce paradigm, for estimating a measure of similarity between genomic sequences.

Research Experience

- **Research Project Associate**, *Sapienza University of Rome*, Italy. 08/2020 – 01/2021

  I collaborated with the Department of Statistical Sciences at Sapienza - University of Rome, at the design and the development of distributed algorithms for the analysis of protein-toprotein interaction networks, under the framework of the EIT Health European project on "Big Data Software Services for Decision Support in Precision Medicine" . One of the main outcomes of this activity has been the development of a library of classes and functions specialized for the analysis of huge protein-to-protein interaction networks through Apache Spark and the MapReduce paradigm (a scientific paper describing the results of this research activity is under preparation).

Academic positions

- **Temporary Research Fellow**, *Sapienza University of Rome*, Italy.
  03/2021 – 02/2022
  Winner of a research fellowship (literally, "assegno di ricerca, categoria A, Tipologia I") granted by the Department of Statistical Sciences at Sapienza University of Rome (supervisor: Prof. Umberto Ferraro Petrillo). In this context, I am developing a research project about the development of different classes of distributed algorithms for the analysis of genomic data by means of Apache Spark. The algorithms being developed are of interest also in other

application domains such as the analysis of large-scale computer network for cybersecurity purposes.

Publications

- **Large Scale Graph Based Network Forensics Analysis**.
  Di Rocco, U. Ferraro-Petrillo, F. Palini
  International Workshop on Biometric Data Analysis and Forensics at 25th International Conference on Pattern Recognition. (ICPR), January 10–15, 2021, Proceedings, Part V, pp. 457-469
- **Scheduling K-mers Counting in a Distributed Environment**.
  Amorosi,L Di Rocco, UF Petrillo
  International Conference on Optimization in Artificial Intelligence and Data Sciences (ODS2021), September 14-17, 2021, Rome - (Paper accepted for publication)

Awards

- **Google academic research grant.** .
  Project title: "MapReduce Approach toward Reference-Based Genotyping for Genomic Sequences".
  Value: 1000 USD (to be spent on the Google Cloud Platform).

International Conference Presentations

- **Large Scale Graph Based Network Forensics Analysis**.
  International Workshop on Biometric Data Analysis and Forensics at 25th International Conference on Pattern Recognition. (ICPR), January 10–15, 2021

Reviewing Activity

- International Journal on Artificial Intelligence Tools

Springer Soft Computing

Computer Skills

Programming

- Java
- Scala
- R
- Python:
  - Tensor Flow
  - Scikit-Learn

Distributed Computing

- Apache Spark:
  - SparkSQL
  - GraphX
  - GraphFrame

DBMS

- MySql
- Neo4j
- MongoDb