europass Curriculum vitae

PERSONAL INFORMATION

Simone Di Gregorio

WORK EXPERIENCE

August 2022 – September 2022

Data Science Intern

KNIME GmbH, Konstanz, Germany

KNIME GmbH is the German branch of KNIME, the data science software company developing the homonymous open source analytics platform, which enables data manipulation/handling and machine learning applications with directed graphs/workflows built through a graphical interface.

Main responsibilities were:

- Development of a KNIME component for Word2Vec and of a KNIME workflow showing a possible use case for the resulting embeddings.
- Development of a KNIME node (Python based) implementing a similar pipeline, but far more focused on performance and low-level Tensorflow.
- Writing blog articles describing theoretical and practical aspects of the above implementations.
- Collaboration with the internal Python development team in testing and perfecting a software extension for user development. The Word2Vec node was a way to test the extension to its fullest.

September 2023 - March 2024

Teaching Assistant

Sapienza University of Rome

Statistical Methods for Data Science & Laboratory I

Statistical Methods for Data Science and Laboratory is one of the main courses, spanning two semesters, in the Data Science Master's Degree, dealing with an introduction to Probability Theory before delving into Frequentist and Bayesian Inference. I won a grant (Bando n.5 2023 - I3S faculty) for tutoring in the first part of the course (9 CFUs out of 12).

This was a part time job where the main activities were:

- Holding lectures/office hours once a week to explain specific topics/revise already covered ones and tackle exercises; topics spanned from basic probability theory (non-analytic) to frequentist inference with practical examples and simulations using the R language.
- Grading homeworks (there were three in the first part of the course): each group of students was required to submit a markdown report with R code, exploring a specific (possibly advanced) topic or procedure related to what they covered in class (e.g. confidence sequences, correlation graphs with appropriate testing procedures...).
- Helping out with exams.

April 2025 - June 2025

Teaching Assistant

Sapienza University of Rome

Stochastic Processes for Applied Sciences

Stochastic Processes for Applied Sciences is one of the most relevant courses in the Master's Degree in Statistical Sciences, introducing students to probability from an analytical viewpoint, using measure-theoretic tools to tackle Markov processes, martingales, Brownian motion and Itô calculus. I won a grant (Bando n.12 2024 - I3S faculty) for assisting in teaching the course.

The main activities were:

- Holding exercise sessions (both for theoretical and practical exercises) once a week to prepare students for the final exams.
- Helping out with exams.

April 2025 - June 2025

Teaching Assistant

Sapienza University of Rome

Logic and Probabilistic Methods in Computer Science

Logic and Probabilistic Methods in Computer Science is an elective course in the last year of the Bachelor's Degree in Mathematical Sciences for Artificial Intelligence, in the Mathematics Department. The involvement is related to the second part of the course, where probabilistic tools are introduced in order to analyze randomized solutions to computationally hard algorithms. Probabilistic correctness guarantees are provided with ergodic results and measure concentration arguments, which are analyzed in the course.

The main activities were:

– Holding exercise and theory sessions once a week to go in depth with some topics and prepare students for the final exams.
– Helping out with exams.

EDUCATION AND TRAINING

2019–2022

Bachelor's Degree in Management and Computer Science

LUISS Guido Carli University

Mathematical Analysis, Multivariate Calculus, Linear Algebra, Graph Theory, Statistics, Statistical/Machine Learning, Micro/Macroeconomics, Python/R programming, Theory of Algorithms and Relational Database Design/SQL.

Thesis (statistical learning topics):

Movie box office analysis and inference with time series models and tree-based methods;

Supervisor: Dr. Francesco Iafrate

Grade: 110/110

Date of graduation: 02/11/2022

2022–2024

Master's Degree in Data Science

Sapienza Università di Roma

Frequentist and Bayesian Inference, Probability Theory, Stochastic Processes, Convex Optimization for machine learning applications, Signal Processing, Statistical Learning, Theory of Algorithms, Deep Learning, Computer Vision, Telecommunication Networks.

Thesis (mathematical statistics topics):

Estimation of the coefficients of a vector stochastic differential equation: from classical parametric methods to neural network estimators;

Supervisors: Prof. Alessandro De Gregorio, Dr. Francesco Iafrate and Prof. Pierpaolo Brutti

Honours Programme: Collaboration with Prof. Piccialli and the French Institute for Research in Computer Science and Automation on data augmentation and machine learning frameworks for a follow-up publication to *Data augmentation driven by optimization for membrane separation process synthesis* [*Computers & Chemical Engineering* 177, 2023].

Grade: 110/110 with honors

Date of graduation: 28/10/2024

2024–Currently

PhD in Data Science

Sapienza Università di Roma

PhD in Data Science under the supervision of Professor Stefano Leonardi, working on statistical learning theory in the context of algorithmic game theory, specifically in the settings of auctions and bilateral trade. I am also interested in and work on frontier results on uniform convergence theory (i.e. tight PAC-like bounds), online convex optimization/multi-armed bandits problems and multicalibration (sample complexity and applications to attaining algorithmic guarantees in linear problems).

Publications:

- *Neural Drift Estimation for Ergodic Diffusions: Non-parametric Analysis and Numerical Exploration*, with Francesco Iafrate (University of Hamburg) [*International Workshop on Functional and Operatorial Statistics*, 2025]. Main topics: neural network theory, non-asymptotic error bounds and stochastic calculus. [ArXiv, PDF]
- *Nearly Tight Regret Bounds for Profit Maximization in Bilateral Trade*, with Federico Fusco (Sapienza University of Rome), Chris Schwiegelshohn (Aarhus University) and Paul Duetting (Google Research) [*IEEE Symposium On Foundations Of Computer Science (FOCS)*, 2025]. Main topics: online learning, algorithmic game theory and computational learning theory. [ArXiv, PDF]
- *Multicalibration yields better matchings*, with Federico Fusco, Simone Fioravanti, Stefano Leonardi (Sapienza University of Rome), Matteo Russo (EPFL); joint work with Meta Central Applied Science. Main topics: multicalibration, optimization with predictions.

PERSONAL SKILLS

Language Italian (mother tongue) and English.

Language Certification English C2 CEFR certificated by C1 Advanced Cambridge Exam in June 2018/Score: 201/210

Computer skills

- **Python and R coding for data science** tasks and projects, specifically for what concerns data manipulation, data visualization and training and performance evaluation of all traditional machine learning models and techniques. For what concerns Python, strong confidence in using scikit-learn and the whole data science stack (NumPy, Pandas, SciPy, etc.), relatively in-depth knowledge of Tensorflow (and of course of Keras) and general knowledge of vanilla PyTorch. For what concerns R, fluency is both with native syntax and tidyverse-based syntax.
- **KNIME proficiency** across all the main levels of certifications (L1, L2, L3), both for local use cases and production, with additional knowledge concerning the software due to employment at KNIME GmbH in Konstanz, Germany for two months.
- General knowledge of SQL and relational database management systems.

Additional expertise, experience and/or skills

- In depth knowledge of **advanced and rigorous probability theory and real analysis**, specifically for stochastic processes. This is due to self-studying carried out during the Master's Degree, in parallel to exams and courses. This kind of knowledge gave the foundations for the Master's Degree thesis in mathematical statistics topics.
- Years of experience in entertainment journalism, having worked for important Italian web magazines, handling reviews, previews and videos/streams related to important releases. Care for **writing style, content exposition and content layout** stems from this professional experience.

Curriculum ai fini della pubblicazione

F.to Simone Di Gregorio

Data: 09/12/2025