

Alessandra Cabassi

Experience

- Jun – (Dec) **Intern, Google**, London, UK.
2019 Team: Advanced Measurement Technologies.
Improving existings time series-based anomaly detection algorithms (Python).
- Jun – Sep **Intern, The Alan Turing Institute**, London, UK.
2018 Supervisors: Dr Anthony Lee, Dr Ioannis Kosmidis, Dr Rajen Shah, and Dr Yi Yu.
Benchmarked machine learning libraries such as Apache Spark and TensorFlow for high-performance, large-scale linear models on the Institute's supercomputer (Python). Worked in close collaboration with software engineers from Cray EMEA Research Lab.
- Jan – May **Teaching assistant, University of Cambridge**, UK.
2018 Supervised 2nd-year Mathematics students for the Statistics module at four Cambridge Colleges.
- Oct – Dec **Visiting student, Statistical Laboratory, University of Cambridge**, UK.
2015 Supervisors: Dr Davide Pigoli and Prof Piercesare Secchi.
Developed statistical methods for hypothesis testing on covariance operators of functional data, implemented and published the corresponding software, applied the novel methodology to a study in evolutionary biology (R).
- May – Sep **Intern, INRIA**, French institute for Computer Science and Automation, Nice, France.
2013 Supervisors: Dr Paola Goatin and Prof Nicola Parolini.
Implemented mathematical traffic flow models based on partial differential equations, validated them through extensive numerical simulations, and calibrated them with data provided by industrial partners (Matlab).

Education

- Oct 2016 – (Jun 2020) **PhD Biostatistics, MRC Biostatistics Unit, University of Cambridge**, UK.
Main research interests: Bayesian Methods, Kernel Methods, Model-Based Clustering, Variational Inference.
 - Designing and implementing statistical and machine learning methods for integrative clustering of heterogeneous data, with a focus on datasets with small number of observations and large number of features (R, Matlab).
 - Collaborating with clinicians and bioinformaticians to gain insights from novel genomic datasets.
 - Selected as one of 28 early-career academics in Europe for cross-disciplinary research project.
- 2013 – 2016 **MSc (Laurea Magistrale) Engineering Mathematics, Politecnico di Milano**, Italy, 1st.
Main subjects: Statistics, Optimisation, Game Theory, Risk Analysis, Artificial Intelligence.
 - Developed Bayesian model to diagnose aortic stiffness using medical imaging data (R).
 - Used functional data analysis to infer mobility patterns in Milan via mobile phone data (R).
 - Implemented high-performance recommendation algorithm for Netflix challenge dataset (C++).
 - Won one of three prizes at *datathon* organised by the Italian Statistical Society with 20+ participating teams.
- 2011 - 2013 **MSc (Diplôme d'Ingénieur) Multidisciplinary Engineering, École Centrale Nantes**, France, 2:1.
Main subjects: Elements of Civil, Electronic, and Energy Engineering, Project Management, Communication.
 - Lead team of five students during a 4-month consulting project for a private company.
- 2009 - 2011 **BSc (Laurea Triennale) Engineering Mathematics, Politecnico di Milano**, Italy.
Main subjects: Advanced mathematics, Numerical analysis, Computer science.
 - Selected as one of five participants for double degree programme *TIME: Top Industrial Managers for Europe*.

Publications, research reports & blogs

- 2019+
 - Cabassi, A., Richardson, S., and Kirk, P. D. W. *Multiple kernel learning for integrative consensus clustering in genomic precision medicine*. In preparation.
 - Cabassi, A., Kirk, P. D. W., 2019. *Multiple kernel learning for integrative consensus clustering of genomic datasets*. arXiv preprint. arXiv:1904.07701.
- 2018
 - Cabassi, A., Casa, A., ... Farcomeni, A. *Three testing perspectives on connectome data*. In: Canale A., et al. Studies in Neural Data Science. Springer Proceedings in Mathematics & Statistics, vol 257.
 - Cabassi, A. and Wang, J. *High-performance, large-scale regression*.
Blog: turingintern2018.github.io.

- 2017 ○ Cabassi, A., Pigoli, D., Secchi, P., and Carter, P. A. *Permutation tests for the equality of covariance operators of functional data with applications to evolutionary biology*
Electronic Journal of Statistics: Vol. 11, No. 2: 3815-3840, doi:10.1214/17-EJS1347.
- 2013 ○ Cabassi, A. and Goatin, P. *Validation of traffic flow models on processed GPS data*
Research report RR-8382, INRIA, hal:00876311.

Open source software

- klic Kernel Learning Integrative Clustering, *R package*, github.com/acabassi/klic.
- coca Cluster-Of-Clusters Analysis, *R package*, github.com/acabassi/coca.
- fdcov Functional Data COVariance operators, *R package hosted on CRAN*, github.com/acabassi/fdcov.

IT skills

- | | | | |
|-----------|--|-----------|---------------------------|
| Languages | In order of proficiency: R, Python, C / C++. | Libraries | TensorFlow, Apache Spark. |
| Software | Matlab, Microsoft Office. | OS | macOS, Linux, Windows. |

Selected academic talks

- Invited ○ *Multiple kernel learning for integrative clustering in genomic precision medicine.*
11th International Conference of ERCIM on Computational & Methodological Statistics, Pisa, Italy, 12/2018.
- *Permutation tests for functional and object data.*
4th conference of the International Society for Nonparametric Statistics, Salerno, Italy, 06/2018.
- Contributed ○ *Kernel learning approaches for summarising and combining MCMC samples from the posterior distributions of Bayesian clustering models.*
Statistical Methods for Post Genomic Data workshop, Montpellier, France, 01/2018.
- *Multiple kernel learning for integrative clustering of 'omics data.*
Biometrika workshop, Cambridge, UK, 11/2017.
- Seminars ○ *Transcriptional module discovery using integrative clustering approaches*, 11/2018.
- *Integrative and semi-supervised Bayesian clustering for multi-omics data*, 09/2017.
- *Permutation tests for covariance operators of functional data*, 11/2016.
MRC Biostatistics Unit, University of Cambridge, UK.
- Other ○ *High-performance, large-scale regression*, joint talk with Junyang Wang – Newcastle University.
The Alan Turing Institute, London, UK, 09/2018.
- *Three testing perspectives on connectome data*, joint talk with Alessandro Casa – University of Padua.
Start-Up Research workshop, Palermo, Italy, 07/2018.

Scholarships

- 2017 Funding for high-cost training at University of Washington, *MRC Doctoral Training Partnership*, UK.
- 2016 3-year scholarship for PhD at University of Cambridge, *MRC Biostatistics Unit*, UK.
- 2016 Scholarship for Master's thesis (65 available for university of 42k+ students), *Politecnico di Milano*, Italy.
- 2012 1-year scholarship for outstanding international students, *Pays de la Loire Region*, France.
- 2011 1-year Erasmus scholarship for double degree, *Top Industrial Managers for Europe Association*.

Languages

- | | |
|--|--|
| Italian Native speaker. | English Fluent, <i>C2 level: IELTS 7.5.</i> |
| French Fluent, <i>C1 level: TFI 950/1000.</i> | Spanish Fluent, <i>C1 level: CULP 930/1000.</i> |

Student societies

- 2014 **President** of *AIM*, Engineering Mathematics Society, Politecnico di Milano.
Lead 9-people committee, restructured the society. It now has 1k+ members and 10+ industrial partnerships.
- 2012 **Co-founder** of *ACCENTS*, International Students Society, École Centrale Nantes.
Organised several events. Our engineering school scored 2nd in France for international openness the following year.