Projects for UCL Computer Science M.Sc. Programmes and 4th Year Undergraduate Students (2025/26)

Ωmega Research Group
Department of Computer Science
University College London

N.B. If you are interested, please reach out via an email at v.lampos@ucl.ac.uk as soon as possible (do not wait for the formal announcement of the projects). I am also open to discuss different research directions that fit within my group's research interests (see omega.cs.ucl.ac.uk) and can be supported by data sets that are already at our disposal or do not require an additional ethics approval.

Project 1 — Time series forecasting

Time series models play an important role in many domains, from public or personalised health to physics, but also climate science and finance. Most traditional machine learning forecasting models have various limitations such as stationarity assumptions, need for extensive feature engineering, and ineffective handling of longer input sequences. Recent research efforts have focused on the deployment of modern neural network architectures (esp. Transformer) to improve time series forecasting models. However, these models have limited capacity in providing statistically accurate uncertainty bounds for their predictions, and oftentimes their development introduces irrational modelling discounts (e.g. removing focus from inter- and intra-variable dependencies) that seem to improve forecasting performance in a set of established benchmarks although in practical terms they might not (caveats explained in the Appendix of Shu and Lampos (2025)). Our research (see also Shu and Lampos (2026)) seeks to improve forecasting performance under scenarios with tangible outcomes (meteorology, epidemiology, finance), generate foundational (task-agnostic) models that can learn from a diverse set of predictors, and rigorously compare to almost forgotten (but still quite competitive) traditional forecasting approaches. We are also interested in investigating improved ways for deriving uncertainty estimates going beyond Bayesian neural networks (Morris et al., 2023), and starting from conformal prediction approaches, with the aim to retain (or improve) mean forecasting accuracy while maintaining a similar level of computational complexity. Please feel free to specify your own project within this context.

Project 2 — Machine-generated text detection

With the emergence of large language models (LLMs), online text (and text content in general) is increasingly becoming machine-generated. In this project, you can attempt to build your own machine-generated text detector and assess its accuracy under different scenarios (e.g. different base language models, different domains of text and so on). Please have a look at the paper by Li et al. (2024) to understand the task better. Our recent work on model collapse in LLMs (Drayson et al., 2025) highlights aspects of the importance of machine-generated text detection.

Data and computational resources

Data for the aforementioned projects is either publicly available or will be provided by us. You will be able to leverage computational resources provided by the Computer Science department. In exceptional circumstances (e.g. highly research-oriented work) and upon agreement with our group, you may also be able to leverage our dedicated computational resources.

Ethics

These projects will use aggregate information. Data for individual users is not available to us. Given this, the projects described here have obtained an ethics exemption by UCL Computer Science.

References

George Drayson, Emine Yilmaz, and Vasileios Lampos. 2025. Machine-generated text detection prevents language model collapse. In *Proc. of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 29645–29661.

Yafu Li, Qintong Li, Leyang Cui, Wei Bi, Zhilin Wang, Longyue Wang, Linyi Yang, Shuming Shi, and Yue Zhang. 2024. MAGE: Machine-generated text detection in the wild. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 36–53.

Michael Morris, Peter Hayes, Ingemar J. Cox, and Vasileios Lampos. 2023. Neural network models for influenza forecasting with associated uncertainty using Web search activity trends. *PLOS Computational Biology*, 19(8).

Yuxuan Shu and Vasileios Lampos. 2025. DeformTime: capturing variable dependencies with deformable attention for time series forecasting. *Transactions on Machine Learning Research*.

Yuxuan Shu and Vasileios Lampos. 2026. Sonnet: Spectral Operator Neural Network for Multivariable Time Series Forecasting. In *Proc. of the 40th Annual Conference on Artificial Intelligence*.