Timo Hromádka

Master's Student in Computer Science

+421 918 871 381
timo.hromadka@gmail.com
th716@cam.ac.uk
GitHub Profile
LinkedIn Profile
Google Scholar

Education

- **University of Cambridge** *2023-24*
  MPhil Advanced Computer Science
- **King's College London** *2019-22*
  BSc Computer Science First Class w/ Honours: 82%

Experience

- **Data Scientist/Bioinformatician** *2023*
  CircaGene London, UK
  – Developed ML models for oral disease prediction and experimentation, including data cleaning, pipeline development, data analysis, and literature reviewing.
  – Largely contributed to InnovateUK Grant deliverables, including patent writing and presentation preparation.
  – Designed end-to-end Variant Calling (VC) bioinformatics pipeline (Perl, GATK, LaTex) to create PDFs displaying customers' susceptible variants.
- **NLP Researcher/Engineer** *2022-23*
  KInIT Bratislava, Slovakia
  – Worked on two publications (Found under *Publications* section 3). Tasks included developing ML experimentation pipeline, data analysis (on noisy social media data primarily), literature reviewing, and paper writing.
- **Teaching Assistant** *2020-2022*
  King's College London London, UK
  – Taught undergraduate courses in Logic and Computer Systems, recognized for excellent teaching performance.
  – Developed customized teaching materials and engaged in active communication with students.
- **Undergraduate Research Fellow** *2022*
  King's College London London, UK
  – Researched in image segmentation and computer vision for converting handwritten UML diagrams into digital format.
- **Computer Vision Intern** *2020*
  Slovak Academy of Sciences (Institute of Measurement) Bratislava, Slovakia
  – Programmed a CNC machine for high-resolution photography stitching of gallery paintings using Python and OpenCV.

Publications

- **Multilingual Previously Fact-Checked Claim Retrieval** [arxiv] [pdf] [dataset] *2022-2023*
  KInIT
  – Information Retrieval to multilingually match claims from social media posts to our custom-collected dataset of fact-checked claims.
- **Multilingual Persuasion Techniques Detection** [arxiv][pdf] *2022-2023*
  KInIT
  – Investigating techniques to develop language-agnostic solutions for detecting persuasion techniques in news articles and social media posts.
  – Led my team to be overall winner of Semeval2023 Task 3 Subtask 3

Projects

- **Blindfolded Rubik's Cube Solving Algorithm Finder** *2023 - present*
  Open-source tool to help people learn blindfold solving.
  - Dynamic Programming to efficiently retrieve the algorithms in an NP-hard space for blindfolded Rubik's Cube Solving
  - With the use of dynamic programming I have been able to discover a brand new set of commutator algorithms to more efficiently solve a Rubik's cube blindfolded.

- **Diffusion Models for Personalized Insomnia Music-Treatment** *2023 - present*
  Master's Thesis (in progress)
  - Investigating the efficacy of diffusion models in customizing sound/audio/music samples through text-guided editing/impainting.

- **Benchmarking Misinformation Detection Approaches with Paraphrasing**
  Bachelor's Thesis
  - Implemented and combined various techniques in feature extraction, classification, paraphrasing, model architectures, algorithm development, dataset cleansing/noise reduction to detect Covid fake news. Skills included model training, dataset management/filtering, machine learning, and research/implementation in NLP topics (NN architectures, embeddings, seq2seq).

- **Bias Detection in NLP Using Adversarial Data and Compression Techniques** *2023*
  L101: Machine Learning for Language Processing (Research Project) (in progress)
  - The research project investigates if large models, unlike compressed ones, inherently develop biases through shortcut learning, building on the concept of training dynamics outlined in this paper.

- **Knowledge Distillation with Training Dynamics Data Selection** *2023*
  L46: Principles of Machine Learning Systems (Research Project) (in progress)
  - The research project investigates whether certain data points, according to their training dynamics, are more beneficial for knowledge distillation transfer between a teacher and a student model.

- **Transformer From Scratch** *2023*
  L90: Overview of Natural Language Processing (Technical Project) (in progress)
  - Building a transformer architecture from scratch and training an abstractive summarizer.

Technical Skills and Languages

Languages: Python, Java, LaTex, C++, Perl
Developer Tools: Weights and Biases, Docker, MLFlow, Linux (WSL)
Frameworks: PyTorch, PyTorch Lightning, Tensorflow
Spoken Languages: English (bilingual), Slovak (bilingual), German (B2)

Leadership and Extracurriculars

- **Lucy Cavendish College Social Secretary** University of Cambridge *2023-present*
- **Informatics SSLC Representative** King's College London *2019-2022*
- **President and Founder of University Calisthenics Club** King's College London *2022-2023*
- **Blindfold Rubik's Cube Solving National Champion** WCA Profile *2016-present*
- **First Team Player - Chess Club** King's College London *2019-2022*
- **Charity Volunteer** Charity Begins at Home, London, UK *2022-2023*