

Professor Craig Butts (University of Bristol) and Dr Markus Schade (AstraZeneca)

IMPRESSION+GAMES - A Generative Machine Learning System to Elucidate Chemical Structures

Are you interested in machine learning, solving scientific problems and working in partnership with a major international company? If so, then you are invited to apply for a PhD starting, ideally, on 16 September 2024, to be supervised by Professor Craig Butts (University of Bristol).

This PhD offers the chance to work on world-leading machine learning tools for chemical structure elucidation, co-sponsored by AstraZeneca, and hosted in the School of Chemistry at the University of Bristol – ranked #1 for Research in the 7 yearly REF2021 assessment (Times Higher Education)

Summary

Working out the structure of molecules (2D/connectivity or 3D/stereochemistry/conformation) is core to drug development, and NMR spectroscopy is a key tool that enables these structures to be elucidated. Understanding and interpreting NMR data has historically required expert analysis, more recently supplemented by computational approaches.

In this project you will change this paradigm – creating a generative machine learning system (think ‘ChatGPT for molecules’) which is trained to create chemical structures that fit the experimentally measured NMR data.

Objective 1: *Optimize our existing IMPRESSION machine learning system to be the most accurate available for predicting NMR properties (3D or 2D).*

Objective 2: *Create a new generative machine learning system, GAMES (Generative Adaptive Molecule Evolving System), which will create the correct chemical structure of a sample in response to being fed the NMR properties of that chemical sample.*

Research Training: You will gain a combination of in-depth understanding of the scientific basis for the problem being solved and the coding/machine learning expertise. This program of work will provide excellent training for a PhD student in the field of scientific machine learning/data science – key skills that are in significant demand across the pharmaceutical and wider scientific sector. You will be embedded in the Butts group, recognised as the UK’s leading research group developing such NMR-based machine learning and computational tools for solution-state molecular structure determination. The research group works within a wider School which is the leading Chemistry research department in the UK, and in a University with a strong focus on high performance computing and data science through its Advanced Computing Research Centre. The School itself hosts leading doctoral training centres across a range of molecular sciences ensuring a strong culture and infrastructure to support and enable the PhD student.

Candidate Requirements

Applicants must have obtained, or be about to obtain, a First or Upper Second Class UK first undergraduate degree or postgraduate Masters degree, in a physical science or computing/data science/machine learning subject area.

How to Apply - Please make an online application for this project by searching for 'Chemistry (PhD)' and indicating within your application the title of this project. Please also name Craig Butts as your proposed supervisor: [How to apply | Study at Bristol | University of Bristol](#).

Funding - A full studentship will cover UK tuition fees, a training support fee and a stipend (£18,622p.a. in 2023/24) , updated each year) for 4 years.

Getting in Contact – If you wish to discuss this project in advance of making an application, please contact Professor Craig Butts (Craig.Butts@bris.ac.uk)