

SICSA DVF Prof. Juho Rousu - final report

Prof. Rousu (Aalto University, Finland) visited the University of Glasgow from April 8th until June 20th 2019. Prof. Rousu is an international expert in the development and application of kernel methods (an area of machine learning) for solving problems in the life sciences, particularly those involving small molecules such as metabolites and drugs.

His main contact during the visit was Dr. Simon Rogers (University of Glasgow) with whom he worked throughout his visit along with shorter visits to other SICSA institutions: The University of Edinburgh, The University of St. Andrews, and The University of Aberdeen.

His work in Glasgow with SR (with Grimur Hjorleifsson) was focused on investigating ways in which state-of-the-art metabolite identification methods (based around input output kernel regression) could be made faster without significant drops in performance through the removal of the costly fragmentation tree generation step. We anticipate a publication in this area. This work supports SR's BBSRC TRDF funded project on the development of software to accelerate the discovery of new antimicrobials. A collaboration was also initiated with Dr. John Williamson on the visualisation of uncertainty in metabolite identification across large data sets, which will likely result in a publication and become one of the exemplars in Dr. Williamson's work package in the EPSRC closed-loop data science project. Finally, based upon discussions during the visit Dr. Rogers and Prof. Rousu are in the early stages of planning a European funding application. A further research collaboration link arose between Prof. Guido Sanguinetti's group (Edinburgh) on linking epigenomic models into drug-response models developed by Prof. Rousu's group.

In Glasgow, Prof. Rousu have a research seminar "Small Molecule Identification through Machine Learning: CSI:FingerID and beyond" (details here: <https://samoa.dcs.gla.ac.uk/events/viewtalk.jsp?id=16053>) that was attended by researchers from CS, other Schools in the College of Science and Engineering, and MVLS. In addition, he gave a masterclass "Predicting multi-view and structured data with kernel methods" (details here: <https://samoa.dcs.gla.ac.uk/events/viewtalk.jsp?id=16097>) that was attended by ~30 researchers (from PhD students to professors) from across the University.

A three day trip to The University of St. Andrews (hosted by Dr. Juliana Bowles) and The University of Aberdeen (hosted by Dr. Wei Pang) was organized in April 30 - May 2. The research seminar ("Prediction of drug interactions") at the University of St. Andrews attracted an active audience of ca. 15 people and some interesting research discussions followed. At the University of Aberdeen the programme involved a repeat of the aforementioned masterclass and a seminar ("Sparse non-linear canonical correlation analysis"). Both the Masterclass presentation and the research seminar attracted an audience around 20 people, and most of the audience was composed of PhD students. The students were very active both during and after the presentations, and the presentation materials were distributed to the interested audience after the talk. Finally, a two day trip to Edinburgh (hosted by Prof. Guido Sanguinetti) involved a research seminar ("Predicting drug interactions with kernel

methods”) that was also broadcasted to researchers at the MRC Institute of Genetics and Molecular Medicine (IGMM). With the live broadcast, the seminar attendance was ca. 40 people. The Edinburgh schedule was otherwise filled with a full schedule one-on-one discussions with local researchers, and several interesting links in research were uncovered, with the above-mentioned link to Prof. Sanguinetti’s work being the most concrete.

The materials of all the presentations given during the fellowship stay are publicly available in ResearchGate:

- Masterclass on “Predicting Multi-view and Structured Data with Kernel Methods”: a Crash Course (Aberdeen and Glasgow):
https://www.researchgate.net/publication/333220416_Predicting_Multi-view_and_Structured_Data_with_Kernel_Methods_a_Crash_Course DOI: 10.13140/RG.2.2.36652.51846
- Research seminar on “Predicting drug interactions with kernel methods” (Edinburgh and St. Andrews): DOI: 10.13140/RG.2.2.24203.13605
https://www.researchgate.net/publication/333613878_Predicting_drug_interactions_with_kernel_methods
- Research seminar on “Small Molecule Identification through Machine Learning: CSI:FingerID and Beyond” (Glasgow): DOI: 10.13140/RG.2.2.21057.40809
https://www.researchgate.net/publication/333610098_Small_Molecule_Identification_through_Machine_Learning_CSI:FingerID_and_Beyond