



WT2.9: University of Bradford

Report

Activities performed during the visit

in Wrocław University of Technology

period: 24/10/2015 -29/10/2015

author: Anna Palczewska, Urszula Markowska-Kaczmar



Personal Information

Mr./Ms. Anna Palczewska, faculty member of University of Bradford, UK visited

(name of sending institution, country)

Wrocław University of Technology in the period from 24/10/2015 to 29/10/2015 in

(name of the visited institution, country)

order to carry out research and training activities in the field of machine learning and data mining.

(give the area)

Information about Seminars

The seminar presentation was organized on 26/10/2015

the date

It was entitled:

Interpreting Random Forest Models Using a Feature Contributions Method.



Description of scientific activities

(Please describe value added to the ENGINE project i.e. new knowledge, new skills with respect to the objectives of the project, the assigned common area of future cooperation with the partner, plans for common research, projects, publications and how it could be used in the scope of ENGINE)

During the visit in Wrocław University of Technology Anna Palczewska presented her former research related to interpretation of black-box predictive model on a seminar held by the Computer Science Departments of Computer Science and Management Faculty. The presentation was titled: “Interpreting Random Forest Models Using a Feature Contributions Method”. Dr Anna Palczewska presented a novel method called feature contributions and its application to interpreting random forest models.

Anna Palczewska (AP) and Urszula Markowska-Kaczma (UMK) started collaboration on “Interpretation of Artificial Neural Networks (ANN) Models” in September during the Urszula’s visit in University of Bradford in UK. The visit in Wrocław University of Technology was a continuation of this collaboration.

In this study UMK and AP decided to use a pedagogical approach of rule extractions from ANN. This was done by applying a technique called decision tree. To compare rules extracted by various methods, three types of decision trees were used: recursive partitioning, evolutionary and conditional trees. These were implemented in the R programming language for the following dataset: IRIS, Breast Cancer Wisconsin, German Credit Score and Cox2.

During the AP’s visit in Wrocław, AP and UMK had a several meeting to discuss obtained results: sampling dataset, analysing different rules extracted by each technique for the same model and their accuracies. To compare rules the cross-frequency table was proposed with pattern coverage.

AP and UMK decided to use Random Forest model to extract rules and apply the Feature Contributions method for assessment of the ANN model for a new observation.

As results of the meeting, UMK and AP posed a number of research questions that will be investigated in the following months:

- 1) Generalisation of rules for a forest: applying the cross-frequency tables (pattern coverage), rule pruning by frequency threshold and Feature Contribution
- 2) Analysis of Feature Contributions for a new observation
- 3) Location of a feature in tree vs its feature contribution
- 4) Mixing and sampled non labelled patterns
- 5) Concept drift for ANN

The results of this study will be published in the conference or journal paper.

Actions: AP - software implementation and gathering results, test on a larger set of benchmark datasets, UMK and AP - results analysis, weekly meeting on Skype.

Information referring to the intellectual property

(the generally binding law in this area in the visited country and procedures of patenting);

.....

Description of the cooperation between universities and industry

(how it is organized in partner's organization, the sources of funding, the opinions about drawbacks and strengths of existing solution).

.....

Other activities

.....

REMARK: Apart from this information also a program of the visit and the presentation in electronic version should be given to the project office (please send all of them to Urszula.Markowska-Kaczmar@pwr.wroc.pl). Please respond to the points 1-5 for outgoing visit and points 1-3 for incoming visit. Point 6 is for extra activities that are not put in points 1-5.

