Gergő Bartha

EIT Digital Budapest DTC

PhD topic: Applications of data mining algorithms on spatial data

PhD supervisor: Dr. János Tapolcai, BME

Industrial partner: Csaba Gáspár, Dmlab Ltd.

Contact: barta@tmit.bme.hu

'The EIT Digital initiative broadens my perspective and lets me work towards recognition beyond my immediate proximity. It also gives me inspiration and firsthand experience in converting my research efforts into business.'

Achievements & further plans

Gergő is in the final phase of his PhD studies; his initial research focused on **spatio-temporal data modelling**. Over time, his interest gradually shifted towards time series analysis, including high-performance forecasting methods, particularly in the energy industry. His results are published in a handful of conference papers and in an article in the International Journal of Forecasting. He also participated in the OR module and he is currently spending his geographical mobility in Warsaw at the Interdisciplinary Centre for Mathematical and Computational Modelling. His passion is taking part in challenging data mining competitions, which has led him to some top rankings and globally recognized achievements.

The unique selling point of the expected model ensemble framework is the capacity to deliver high accuracy forecasts to renewable power generators and transmission system operators that ultimately leads to a smarter and more sustainable energy system, and generates substantial savings.

EIT Digital DTC educational phase at Spring semester of 2016:

RA	OR	BMD	GH	Mobility	BDExp.
✓	✓	✓	✓		

Reserach topic

There is a great deal of interest in forecasting energy-related time series, specifically regarding the generation of renewable energy sources, load forecasting, electricity price forecasting or gas consumption forecasting. In the past, he worked on applying efficient forecasting methods using the state-of-the-art in machine learning industrial problems.

In the field of time series forecasting, multiple regressors often provide more accurate results than individual models. This comes from the fact that the combination of model outputs efficiently limits or even extinguishes the effect of outliers, erroneous data and even bias to some extent. But finding the optimal mixture of models is a non-trivial task; solving this problem is Gergő's number one goal.

His focus is on researching the theoretical background and gathering experimental results to lay down the foundations of a model ensemble framework. A suitable framework takes the statistical properties of individual model outputs into consideration to gain insight into how to utilize them in a more effective way. Observing the performance degradation over time and rebalancing the modelling mixture accordingly are also key issues of his ongoing research.