

Abstracts of invited talks and posters

Editor: Alex Gammerman, Vladimir Vovk, Zhiyuan Luo, and Evgeni Smirnov

1. Invited talks

“Mathematics of Deep Learning” by Professor Alexander Balinsky (Cardiff School of Mathematics, UK) Deep Learning is another name for a set of algorithms that use a neural network as an architecture. In the past few years, Deep Learning has generated much excitement due to many breakthrough results in speech recognition, computer vision and text processing. This recent success has been due to new mathematical techniques, the availability of inexpensive, parallel hardware (GPUs, computer clusters) and massive amounts of data. This powerful way of processing data can be used to address an ever-growing number of problems, and its impact on science and society is increasing exponentially. In this talk we present mathematical foundations of Deep Learning, relations with physics, features extraction and interpretability. We also explain mathematics behind adversarial attacks and how to protect against them. Several new mathematical problems will be presented.

“Prediction and exchangeable processes” by Professor Peter McCullagh, FRS (Department of Statistics, University of Chicago, USA) This talk will present an overview of recent and not-so-recent work on a variety of exchangeable processes, including partition processes where the observation space is the set of partitions of $[n]$ or equivalence relations on $[n]$, and ordered partition processes where the observation space is the set of partial rankings of $[n]$. The connection with fiducial processes, Dirichlet processes and survival processes will be discussed.

2. Posters

“Conformal clustering for functional variables, with an application to electricity consumption curves” by Ilia Nouretdinov, Matteo Fontana, James Gammerman, Laura Shelmit and Daljit Rehal Conformal Clustering (CC) is a clustering technique which also allows for anomaly detection. It involves finding a so-called “region of conformity” in the feature space: at a pre-selected significance level ϵ , the data points within this region are grouped into clusters, while all data points outside it are considered anomalies. As part of the conformal prediction (CP) framework, CC inherits the property of CP that allows for a variety of underlying algorithms.

In the existing literature on CC—and CP in general—the data has typically been considered as finite-dimensional feature vectors. However, much existing data is found in functional form—for example, data presented in the form of a time series. In this work we generalise the CC technique to the domain of functional data. More specifically, we use CC

to clean data from energy consumption curves, with the aim of subsequently disaggregating these energy curves into their components.

Our experiments play two roles. Firstly, to validate our expectation that conformal clustering is an effective technique for data cleaning. And secondly, to confirm that a functional approach to data analysis can provide new insights that are lacking in a vector approach.

“Conformal Anomaly Detection based on Association Rules” by Ilia Nouretdinov, James Gammerman and Daljit Rehal We propose a novel technique based on a combination of association rule learning and conformal prediction in its label-dependent form. The non-conformity score is not based directly on a classification algorithm, but aggregates information of many association rules that can be extracted from the data, and exceptions from them.

As an application, we use customer data from UKB. There are multiple fields indicating if a customer is an Industrial Corporation (INC) or Small/Medium-sized Enterprise (SME). We will consider these as labels. Often these labels are incorrect or inconsistent across the SAP system. The aim is to use machine learning to identify inconsistencies and potential errors.

“Sorting email with natural language processing and conformal prediction” by Patrizio Giovannotti and Daljit Rehal Large companies receive thousands of email messages every day. Most of these messages come from their customers and must be forwarded to the right support agents, depending on the customer needs. This sorting procedure can be done manually, which is cost- and time-inefficient, or automatically using natural language processing (NLP) techniques. However, each wrong classification may result in a message being forwarded to the wrong agent, creating further costs. For this reason a reliable estimation of the classification confidence is required. This poster describes how Centrica combined Conformal Prediction and NLP to reliably classify the topic and sentiment of incoming email.

“A note on posterior probability estimation for classifiers” by Georgi Nalbantov and Svetoslav Ivanov One of the central themes in the classification task is the estimation of class posterior probability at a new point \mathbf{x} . The vast majority of classifiers output a score for \mathbf{x} , which is monotonically related to the posterior probability via an unknown relationship. There are many attempts in the literature to estimate this latter relationship. Here, we provide a direct way to estimate the posterior probability without resorting to using classification scores.