COPA 2018: The 7th Symposium on Conformal and Probabilistic Prediction with Applications

Aula, School of Business and Economics, Maastricht University, Tongersestraat 53, Maastricht 6211 LM, The Netherlands

June 11-13, 2018

Final Program and Abstracts

Program: The 7th Symposium on Conformal and Probabilistic Prediction with Applications

Monday, June 11, 2018	
09:30 – 10:00	**Registration**
10:00 – 10:10	**Symposium Opening**
10:10 – 11:25	**Tutorial 1:** *Confidence Prediction: Introduction to Conformal Prediction*, Lars Carlsson and Henrik Linusson
11:25 – 11:45	Coffee Break
11:45 – 13:00	**Tutorial 2:** *Probabilistic Prediction: Venn-ABERS Prediction*, Paolo Toccaceli
13:00 – 14:00	Lunch
14:00 – 15:15	**Tutorial 3:** *Conformal Instance Transfer*, Shuang Zhou
15:15 – 15:45	Coffee Break
15:45 – 16:00	Walk to Aula Minderbroedersberg of Maastricht University, *(reserved for the Kolmogorov lecture) Minderbroedersberg 4-6, Maastricht*
16:00 – 16:15	**Introduction to the Kolmogorov Lecture,** prof. Alex Gammerman
16:15 – 17:15	**The Kolmogorov Lecture:** *Rethinking Statistical Learning Theory: Learning Using Statistical Invariants*, prof. Vladimir Vapnik
17:15 – 17:25	**Award Ceremony,** prof. Vladimir Vovk
17:25 – 17:30	Walk to Grand Café Soiron, *Vrijthof 18, Maastricht*
17:30 – 18:30	Reception Grand Café Soiron
18:30 – 20:00	City-guided walk (that starts from Grand Café Soiron)

Tuesday, June 12, 2018	
09:00 – 10:30	**Plenary Session 1**, **Predictive Distribution,** Chair: Ralf Peeters *Cross conformal predictive distributions*, Vladimir Vovk, Ilia Nouretdinov, Valery Manokhin and Alexander Gammerman *Conformal predictive decision making*, Vladimir Vovk and Claus Bendtsen *Inductive Venn-Abers Predictive Distribution*, Ilia Nouretdinov, Denis Volkhonskiy, Pitt Lim, Paolo Toccaceli and Alexander Gammerman
10:30 – 11:00	Coffee Break
11:00 – 12:30	**Plenary Session 2, Drug Discovery**, Chair: Lars Carlsson *Using Venn-ABERS Predictors to assess Cardio-Vascular Risk*, Ernst Ahlberg, Ruben Buendia and Lars Carlsson *Conformal Prediction in Learning Under Privileged Information Paradigm with Applications in Drug Discovery*, Niharika Gauraha, Lars Carlsson and Ola Spjuth *Venn-Abers* *Venn-ABERS Predictors for Improved Compound Iterative Screening in Drug Discovery*, Ruben Buendia, Ola Engkvist, Lars Carlsson, Thierry Kogej and Ernst Ahlberg
12:30 – 14:00	Lunch
14:00 – 15:00	**Invited Lecture:** *Safe Testing*, prof. Peter Grünwald
15:00 – 15:30	Coffee Break
15:30 – 17:00	**Plenary Session 3: Posters** *Conformal prediction for predictive maintenance: A case study at a commercial gas terminal,* Ilia Nouretdinov, James Gammerman, and Daljit Rehal *Seemingly Unrelated Regression Tree,* Jaeoh Kim and Hyungjun Cho *Parameters Estimation in Mixed Membership Stochastic Block Model and its Applications in Community Detection*, Evgeny Marshakov, Roman Ushakov, Nikita Mokrov and Maxim Panov *A comparison of machine learning methods for epilepsy and depression classification based on structural MRI data,* Ekaterina Kondrateva, Marina Pominova and Nikolay Skuratov *A comparison of preprocessing and graph-based classification methods for functional MRI data analysis,* Svetlana Sushchinskaya, Anastasia Demidova and Anna Kuzina *Dropout-based Active Learning for Regression,* Evgenii Tsymbalov, Maxim Panov, and Alexander Shapeev *fMRI: preprocessing, classification and pattern recognition,* Maxim Sharaev, Alexander Andreev, Alexey Artemov, Alexander Bernstein, Evgeny Burnaev, Ekaterina Kondratyeva, Svetlana Sushchinskaya, and Renat Akzhigitov *Machine Learning pipeline for discovering neuroimaging-based biomarkers in neurology and psychiatry*, Alexander Bernstein, Evgeny Burnaev, Ekaterina Kondratyeva, Svetlana Sushchinskaya, Svetlana Sushchinskaya, Maxim Sharaev, Alexander Andreev, Alexey Artemov , and Renat Akzhigitov
18:45 – 23:00	Dinner: Ship “Jekervallei”, *Maaspromenade 58, Maastricht*

Wednesday, June 13, 2018	
09:00 – 10:30	**Plenary Session 4**, **Other Applications of Conformal Prediction,** Chair: Harris Papadopoulos *Detecting seizures in EEG recordings using Conformal Prediction*, Charalambos Eliades and Harris Papadopoulos *Cover Your Cough: Detection of Respiratory Events with Confidence Using a Smartwatch,* Khuong An Nguyen and Zhiyuan Luo *Conformal Stacked Weather Forecasting*, Jelmer Neeven and Evgueni Smirnov
10:30 – 11:00	Coffee Break
11:00 – 12:30	**Plenary Session 5, Plenary Session 5: Transfer Learning, Feature Selection,** Chair: Evgueni Smirnov *Conformal Feature-Selection Wrappers for Instance Transfer*, Shuang Zhou, Evgueni Smirnov, Gijs Schoenmakers, Ralf Peeters and Tao Jiang *Exchangeability Martingales for Selecting Features in Anomaly Detection*, Giovanni Cherubin, Adrian Baldwin and Jonathan Griffin *Transfer learning for the Probabilistic Classification Vector Machine*, Christoph Raab and Frank-Michael Schleif
12:30 – 14:00	Lunch
14:00 – 16:00	**Plenary Session 6: Conformal Prediction and Other Methods**, Chair: Vladimir Vovk *Conformal Prediction in Manifold Learning*, Alexander Kuleshov, Alexander Berstein and Evgeny Burnaev *Venn Predictors for Well-Calibrated Probability Estimation Trees*, Ulf Johansson, Henrik Bostrom, Henrik Linusson, Tuwe Lofstrom, Haken Sundell, and Anders Gidenstam *Aggregating Strategies for Long-term Forecasting,* Alexander Korotin, Vladimir V'Yugin and Evgeny Burnaev *Interpolation error of Gaussian process regression for misspecified case*, Alexey Zaytsev, Evgenya Romanenkova and Dmitry Ermilov
16:00 – 16:15	Coffee Break
16:15 – 17:00	**Discussions & Closing Ceremony**

Abstracts: The 7th Symposium on Conformal and Probabilistic Prediction with Applications

The 2018 University of London Kolmogorov Lecture

Rethinking Statistical Learning Theory: Learning Using Statistical Invariants (LUSI)

Prof. Vladimir Vapnik (Facebook, Royal Holloway University of London, Columbia University)

Abstract

The talk introduces a new learning paradigm, called Learning Using Statistical Invariants (LUSI), which is different from the classical one.

In the classical paradigm, learning machine constructs, using data, a classification rule that minimizes frequency of training error; it is data-driven model of learning.

In the LUSI paradigm, in order to construct the desired classification rule using data, learning machine introduces statistical invariants that are specific for the problem, and then constructs the rule that then minimizes the frequency of training error and preserves these invariants; it is thus both data- and invariant-driven learning. New paradigm allow learning machine to extract more information from available training data.

From a mathematical point of view, methods of the classical paradigm employ mechanisms of strong convergence of approximations to the desired function, whereas methods of the new paradigm employ both strong and weak convergence mechanisms. This can significantly increase the rate of convergence.

Sponsor: Yandex

Invited Talk

Safe Testing

Prof. Peter Grünwald (CWI, Amsterdam, The Netherlands)

Abstract

In recent years, standard p-value based hypothesis testing has come under intense scrutiny. One of its many problems is the following: if our test result is promising but nonconclusive (say, p = 0.07) we cannot simply decide to gather a new batch of data. While this practice is ubiquitous in science, it invalidates p-values and error guarantees.

Here we propose an alternative hypothesis testing methodology that allows us to do so after all. For simple null hypotheses, our proposal coincides with earlier work by Vovk and collaborators, and amounts to using Bayes factors and/or general test martingales. Here, we work out the composite null case, which allows us to formulate safe, nonasymptotic versions of the most popular tests such as the t-test and the chi square tests. Safe tests for composite H0 are not always Bayesian or test martingale-based but rather based on the Barron-Li 'reverse information projection'. A central innovation is to distinguish between 'optional stopping' and 'optional continuation' which may have repercussions in the conformal prediction world as well.

Posters

Conformal prediction for predictive maintenance: A case study at a commercial gas terminal

Ilia Nouretdinov (I.R.Nouretdinov@cs.rhul.ac.uk), James Gammerman (james.gammerman@centrica.com) and Daljit Rehal (daljit.rehal@centrica.com)

Abstract

This poster presents an approach for predictive maintenance comprising various machine learning methods such as conformal clustering and probabilistic prediction using K-Nearest Neighbours.

The Morecambe Bay gas terminal (owned by Centrica plc) consists of many components, one of which is the field gas compressor (FGC). This component compresses gas which has been extracted offshore to the required pressure for the UK gas system. However, it breaks very frequently so the ability to predict its failure would be commercially valuable. The FGC contains sensors which produce data on 183 attributes of the system which we take as the dataset for this problem.

We show that using a combination of t-SNE dimensionality reduction and the conformal clustering algorithm it is possible to produce prediction regions and identify anomalies which lie outside these regions. We also suggest an approach for identifying which attributes are the most likely cause of failure in a given cluster.

Furthermore we predict likelihood of failure of the FGC within a 24-hour period using a K-Nearest Neighbours classifier on the most statistically significant attributes. We find that when making predictions with high confidence the classifier is highly accurate.

Seemingly Unrelated Regression Tree

Jaeoh Kim (c14180@naver.com) and Hyungjun Cho (hj4cho@korea.ac.kr)

Abstract

Nonparametric seemingly unrelated regression provides a powerful alternative to parametric seemingly unrelated regression for relaxing the linearity assumption. The existing methods are limited, particularly with sharp changes in the relationship between the predictor variables and the corresponding response variable. We propose a new nonparametric method for seemingly unrelated regression, which adopts a tree-structured regression framework, has satisfiable prediction accuracy and interpretability, no restriction on the inclusion of categorical variables, and is less vulnerable to the curse of dimensionality. Moreover, an important feature is constructing a unified tree-structured model for multivariate data, even though the predictor variables corresponding to the response variable are entirely different.

This unified model can offer revelatory insights such as underlying economic meaning. We propose the key factors of tree-structured regression, which are an impurity function detecting complex nonlinear relationships between the predictor variables and the response variable, split rule selection with negligible selection bias, and tree size determination solving underfitting and overfitting problems. We demonstrate our proposed method using simulated data and illustrate it using data from the Korea stock exchange sector indices.

Parameters Estimation in Mixed Membership Stochastic Block Model and its Applications in Community Detection

Evgeny Marshakov (emarshakov@gmail.com), Roman Ushakov (ushakov.ra@phystech.edu), Nikita Mokrov (mokrov@frtk.ru) and Maxim Panov (m.panov@skoltech.ru)

Abstract

Community detection is an important problem in modern network analysis. It has wide applications in analysis of social and biological networks (Girvan and Newman, 2002; Backstrom et al., 2006), designing network protocols (Lu et al., 2015) and many other areas. Recently, much attention has been paid to detection of overlapping communities, where each node in a network may belong to multiple communities. Such situation is quite common, and most prominent examples include overlapping communities in social networks (Leskovec and Mcauley, 2012), where each user may belong to several social circles, and protein-protein interaction networks (Palla et al., 2005), where a protein may belong to multiple protein complexes.

In this paper, we consider the parameter estimation problem in Mixed Membership Stochastic Block Model (MMSB), which is a quite general instance of random graph model allowing for overlapping community structure. We present the new algorithm successive projection overlapping clustering (SPOC) which combines the ideas of spectral clustering and geometric approach for separable non-negative matrix factorization. The proposed algorithm is provably consistent and sub-optimal under MMSB with general conditions on the parameters of the model. SPOC is also shown to perform well experimentally in comparison to other algorithms.

A comparison of machine learning methods for epilepsy and depression classification based on structural MRI data

Ekaterina Kondrateva (ekaterina.kondratyeva@skolkovotech.ru), Marina Pominova (mspominova@edu.hse.ru) and Nikolay Skuratov (nikolay.skuratov@phystech.edu)

Abstract

Background. Depression is the most frequent psychiatric disorder concomitant in epilepsy patients. These diseases have been extensively studied yet separately. However, the diagnosis of depression in patients with epilepsy can be complicated as it potentially shares the common pathological pathways with epilepsy. Thus, the distinctive biomarkers of depression, as well as their relation to epilepsy, remain to be revealed.

Aims. We sought to examine the diagnostic potential of the structural neuroanatomy patterns of epilepsy and depression under the comparison of the most common methods of MRI data analysis for further investigation of depression in patients with epilepsy.

Methods. Machine Learning (ML) and Deep Learning (DL) approaches with the most-used architectures and dimensionality reduction methods were tested solving the following classification problems: epilepsy, depression, and recognition of the depression in patients with epilepsy. SPM and FreeSurfer toolboxes were applied for the data dimensionality reduction and preprocessing.

Evaluation. Structural MRI data was provided by Russian Scientific and practical psychoneurological center named after Z.P. Solovyov under the Skoltech Biomedical Initiative program. The explored dataset includes 25 healthy volunteers and 25 patients with the major depressive disorder in an acute depressive episode, as well as 25 epilepsy patients and 25 epilepsy patients with the major depressive disorder (mean age 33 ± 9 years). In addition, the preprocessed neuroimaging data from the Autism Brain Imaging Data Exchange (ABIDE) were exploited for the neural network pretraining.

Results. The Epilepsy classification with 3D convolutional network VoxCNN shows the accuracy with the most stable pattern being at 0.69 STD (0.09) on ROC/AUC scoring; the accuracy is comparable to the one on DL network pre-trained with ABIDE dataset. The Epilepsy classification on morphometrical features yielded 0.75 STD (0.16) accuracy, then the temporal lobe epilepsy identification exhibits an accuracy of 0.81 STD (0.12) with high sensitivity and specificity; Depression classification shows the accuracy reaching 0.72 STD (0.15).

Conclusions. The ML methods applied indicated common recognition structural patterns of epilepsy associated with Cerebellum, Precuneus, Amygdala and Temporal Lobes; they partially coincide with the depression patterns additionally associated with Temporal Pole, Insula and Superior Occipital Gyrus. These findings provide an important step in the development of potential neuroimaging-based diagnostic and prognostic tools for epilepsy concomitant with depression.

A comparison of preprocessing and graph-based classification methods for functional MRI data analysis

Svetlana Sushchinskaya (Kuzinasvetlana.sushchinskaya@skolkovotech.ru), Anastasia Demidova (aademidova_2@hse.edu.ru) and Anna Kuzina (anna.kuzzina@skoltech.ru)

Abstract

Background. Nowadays, there are more than one third of people with epilepsy suffering from depression. Meanwhile, there are no biological tests which can contribute to it's diagnosis. Neuroimaging of patients with major depression has revealed abnormal functional connectivity measured during the resting state, showing a good potential for diagnostics of different mental disorders. However, the connectivity patterns of major depression as the concomitant disease are still unclear as well as their correlation with epilepsy.

Evaluation. The proposed comparison was carried out on the dataset provided by Russian Scientific and practical psycho-neurological center named after Z.P. Solovyov under the Skoltech Biomedical Initiative program. The dataset consists of one hundred 1.5T resting state functional MRI (fMRI) divided into four groups: 25 patients with depression, 25 patients with epilepsy, 25 patients with both depression and epilepsy and 25 healthy controls.

Methods. fMRI sequences were preprocessed in SPM toolbox, then the denoising stage was performed through three methods to be consequently compared: the default noise spectral filtering in Nilearn toolbox, automated Independent Component Analysis (ICA) noise cancelling with the use of FSL Fix toolbox and finally manual ICA denoising in FSL Melodic toolbox. Further, the functional connectivity matrices for 117 brain regions were calculated in Nilearn using the previously denoised data; the matrices were sparced under different correlation thresholds to construct the functional brain network graph. Then, the common graph features (clustering coefficient, betweenness centrality, etc.) were extracted, resulting in a vector of 587 features for

each person. The informativeness of the fMRI functional connectivity graph features was compared with the Machine Learning classifiers pipeline, which arranges dimensionality reduction methods and the most common *classifiers (SVM, LR, RFC, KNN) in a hyper parameter grid search. Finally, the best models were tested with leave-one-out validation.*

Results. The accuracy 0.77 STD (0.15) on ROC/AUC scoring for depression classification on the fMRI data with default spectrum denoising; then ICA preprocessing improves the classification performance up to 0.88 STD (0.09) accuracy with high sensitivity and specificity. The improvement also was obtained for temporal lobe epilepsy classification: from accuracy of 0.61 STD (0.20) to 0.78 STD (0.15). Noticeable, that denoising FSL Fix toolbox does not yield any improvement in the classification results.

Conclusion. fMRI data denoising was revealed to significantly affect the ML model classification accuracy. The considered graph-based functional connectivity analysis provides an important step in the development of potential neuroimaging-*based tools for diagnosis of epilepsy and depression with high sensitivity and specificity.* Such task-independent functional brain measures could be potentially used to promote an early intervention to reduce the likelihood of developing depression.

Estimating Confidence for Top-N Recommendations in Neighbourhood based Algorithms

Tadiparthi V R Himabindu (himaworld_06@yahoo.com), Vineet Padmanabhan (vineetcs@uohyd.ernet.in), Arun K Pujari, Abdul Sattar (A.Sattar@griffith.edu.au) and Abhaya Nayak (Abhaya.Nayak@mq.edu.au)

Abstract

Confidence is an important concept in recommender systems which tells us about a system's trust in its predictions/recommendations. Attaching confidence values to a set of recommended items not only gives the relevance of those items but also can help a user in his/her decision making process based on the recommendation's confidence.Most of the confidence estimation approaches proposed in the recommender systems' literature are designed for prediction task. Some of these approaches are non-personalized or limited to finding the confidence of the entire algorithm and others are applicable to only specific algorithms and are not generalized. Furthermore, none of these approaches provide guarantees on the error rate of these predictions. Conformal prediction is a generalized framework to associate confidence values to individual predictions made by machine learning algorithms with an upper bound on the error rate. Though conformal prediction is originally designed for prediction task, in this work, we show how it can be adopted to provide confidence values to the set of recommendations produced by neighbourhood-based collaborative filtering algorithms. Experimental results demonstrate that the recommendation accuracy of our proposed algorithms is very close to their underlying algorithms and the recommendation regions produced by our algorithms at different confidence levels are efficient while satisfying the validity property. We also show the precision and recall values at different confidence levels.

CP-MCNN : Reliable Convolutional Neural Networks based on confidence prediction for Multi-Label Chest X-ray diagnosis

Huazhen Wang (wanghuazhen@hqu.edu.cn), Nengguang Wu (444736647@qq.com), Xiaozheng Li (2463840089@qq.com), Feng He (835175514@qq.com), Ting He (xuantinghe@hit.edu.cn), Jinzhun Wu (1923731201@qq.com) and Fan Yang (yang@xmu.edu.cn)

Abstract

The absence of calibrated confidence evaluation for Multi-label Convolutional Neural Network (MCNN) hinders its effective practice. In this paper, a new framework that plugs MCNN into the framework of conformal predictor (CP) is proposed as CP-MCNN, which not only provides multi-label prediction but appends a calibrated confidence to the prediction. Specifically, the whole given dataset is divided into two parts, the former is applied to construct the MCNN model, and the latter is used for confidence prediction. A label-wise p-value computation method is developed to address different label with distinct subset of conformity scores. And the output of MCNN has been regulated for nonconformity measure, which makes the nonconformity score of each example agree very well with the i.i.d characteristic. The application of CP-MCNN focuses on to-date the largest open source of chest x-ray (Chest X-ray14 dataset). The experimental results show that CP-MCNN performs comparative effeciency with state-of-art MCNN by traditional MLL evaluation metrics, such as set-accuracy, Hamming-Loss, ranking-loss and so on. More importantly, CP-MCNN provides well-calibrated confidence prediction, which enhances the reliability and interpretability for computer-assisted chest x-ray diagnosis.

Dropout-based Active Learning for Regression

Evgenii Tsymbalov (etsymbalov@gmail.com), Maxim Panov (panov.maxim@gmail.com)and Alexander Shapeev (a.shapeev@skoltech.ru)

Abstract

Active learning is relevant and challenging for high-dimensional regression models when the annotation of the samples is expensive. Yet most of designed sampling methods cannot be applied to large-scale problems, consuming too much time for data processing. In this paper, we propose a fast active learning algorithm for regression, tailored for neural network models. It is based on uncertainty estimation from stochastic dropout output of the network. Experiments on both synthetic and real-world data sets show comparable or better performance (depending on the accuracy metric) as compared to baselines. This approach can be generalized to various deep learning architectures. It improve the existing model due to offering a computationally cheap and efficient way of sampling more data.

fMRI: preprocessing, classification and pattern recognition

Maxim Sharaev (M.Sharaev@skoltech.ru), Alexander Andreev (aandreev@deloitte.ru), Alexey Artemov (A.Artemov@skoltech.ru), Alexander Bernstein (A.Bernstein@skoltech.ru), Evgeny Burnaev (e.burnaev@skoltech.ru), Ekaterina Kondratyeva (Ekaterina.Kondratyeva@skoltech.ru), Svetlana Sushchinskaya (Svetlana.Sushchinskaya@skolkovotech.ru) and Renat Akzhigitov (Barms@yandex.ru)

Abstract

As machine learning continues to gain momentum in the neuroscience community, we witness the emergence of novel applications such as diagnostics, characterization, and treatment outcome prediction for psychiatric and neurological disorders, for instance, epilepsy and depression. Systematic research into these mental disorders increasingly involves drawing clinical conclusions on the basis of data-driven approaches; to this end, structural and functional neuroimaging serve as key source modalities. Identification of informative neuroimaging markers requires establishing a comprehensive preparation pipeline for data which may be severely corrupted by artifactual signal fluctuations. In this work, we review a large body of literature to provide ample evidence for the advantages of pattern recognition approaches in clinical applications, overview advanced graph-based pattern recognition approaches, and propose a noise-aware neuroimaging data processing pipeline. To demonstrate the effectiveness of our approach, we provide results from a pilot study, which show a significant improvement in classification accuracy, indicating a promising research direction.

Machine Learning pipeline for discovering neuroimaging-based biomarkers in neurology and psychiatry

Alexander Bernstein (A.BERNSTEIN@SKOLTECH.RU), Evgeny Burnaev (e.burnaev@skoltech.ru), Ekaterina Kondratyeva (EKATERINA.KONDRATYEVA@SKOLKOVOTECH.RU), Svetlana Sushchinskaya (SVETLANA.SUSHCHINSKAYA@SKOLKOVOTECH.RU), Maxim Sharaev (M.SHARAEV@SKOLTECH.RU), Alexander Andreev (AANDREEV@DELOITTE.RU), Alexey Artemov (A.ARTEMOV@SKOLTECH.RU) and Renat Akzhigitov (BARMS@YANDEX.RU)

Abstract

We consider the problems of diagnostic pattern recognition/classification from neuroimaging data. We propose a common data analysis pipeline for neuroimaging-based diagnostic classification problems using various ML algorithms and brain imagery processing toolboxes. We illustrate the pipeline application by discovering new biomarkers for diagnostic of epilepsy and depression based on clinical and MRI/fMRI data for patients and healthy volunteers.