# The 2nd Annual Student Research Project Conference

Stiftung Universitat Hildesheim

December 6, 2018

# Contents

**Contents** **2**

**1 Reinforcement Learning** **3**

**2 Time Series Analysis and Forecasting** **8**

**3 Large Scale Machine Learning Problems** **13**

**4 Supervised Classification Approaches** **16**

**5 External Pitches** **20**

**6 Internal Pitches** **24**

# Session 1

# Reinforcement Learning

Session Chair: Rafael Rego Drumond

## Multi-Agent Exploration with Deep Reinforcement Learning

Benjamin Hogstad, Michael Ruchte, Krithika Murugesan, Ramya Raghuraman, Mohana Ventekatesha

### Abstract

In the cooperative reinforcement learning environment, teams of agents struggle to learn the non-stationary behavior of the other agents. There are three main approaches to this multi-agent problem: decentralized, semi-centralized, and centralized. Decentralized approaches learn by applying restrictions on the behavior of the other agents simulating stationarity. Semi-centralized approaches utilize the wealth of information given in multi-agent systems to apply a global training on individual agents. Centralized approaches combine all agents into a single controller which also utilizes the information for all agents but decides actions for all agents jointly. The project implemented models from each of these approaches and evaluates the differences in a simple exploration multi-agent reinforcement learning setting.

**Keywords** multi-agent reinforcement learning

### References

- Mnih, V., (2013). Playing Atari with Deep Reinforcement Learning. NIPS Deep Learning Workshop
- Tan, M. (1993). Multi-agent reinforcement learning: Independent vs. cooperative agents. In Proceedings of the tenth international conference on machine learning (pp. 330-337)
- Sutton, R. S. & Barto, A. G. (2017). Reinforcement learning: An introduction. MIT press.

# Reinforcement Learning for Autonomous Driving

Migena Mana, Abdul Malik, Nikhil Cherian, Syed Khalid Ahmed

## Abstract

In recent years, the advancement in deep learning has enabled researchers to tackle the problem of self driving cars. Car companies use huge datasets to train their deep learning models to make autonomous cars a reality. However, this approach has certain drawbacks in that the state space of possible actions for a car is so huge that there cannot be a dataset for every possible road scenario. In order to overcome this problem, the concept of reinforcement learning is being investigated in this research. In reinforcement learning, the agent interacts with the environment and performs episodic learning by trial and error. It gets reward based on the action it performs which allows the model to find an optimal strategy. Since the problem of autonomous driving can be thought of as a simulation, therefore it lends itself naturally to the domain of reinforcement learning. Our work focuses on two RL algorithms called DDPG and PPO. We have used TORCS as the simulator of choice to train the self-driving car agent. The end result of this research project is a self-driving car agent which is able to navigate around a road track with optimal braking mechanisms, which was the core focus of our work.

**Keywords** DDPG (Deep deterministic policy gradient), PPO (Proximal policy optimization), Autonomous Driving

## References

- Sallab, Ahmad EL, et al. "Deep reinforcement learning framework for autonomous driving." Electronic Imaging 2017.19 (2017) pp. 70-76
- Kehrle, F. Optimal control of vehicles in driving simulators. Diss. Diploma thesis, Uni Heidelberg, 2010
- Kehrle, F. Optimal control of vehicles in driving simulators. Diss. Diploma thesis, Uni Heidelberg, 2010

## Deep Reinforcement Learning for Games

Benison Sam, Ankit Singh, Harish Jadhav, Fatai Oyadeyi

### Abstract

In the recent years, the Reinforcement learning (RL) area of Machine learning has taken on many different long standing challenges in the field of Artifical Intelligence (AI). Especially for the complex games in which humans are very good at but were thought to be really tough problem to be solved using AI. Nevertheless, we have seen AI advancements in computer games like Atari and classical board games like Chess, Go, etc using Deep RL methodologies. We in this Student Research Projectf researched specifically on the approaches taken by Silver, D. et al, for the AlphaGo Master and Zero game engines for the game of Go. We implemented the basic method described in the paper by implementing "selfplay" for learning the game tabula rasa. We then modified the training process by implementing Asynchronous Advantage Actor-Critic (A3C) algorithm. Then on top of it, we experimented with Model averaging as well as Supervised Critic for Momentum in Selfplay. We intend to compare the methods and share the results that we obtained for a 5x5 Go board in terms of measured Elo rating.

### Keywords

### References

- Silver, D., Schrittwieser, J., Simonyan, K. Mastering the game of Go without human knowledge. Nature 550, 354–359 (19 October 2017)
- Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." nature 529.7587 (2016): 484.
- Asynchronous Methods for Deep Reinforcement Learning arXiv:1602.01783$v2$ [cs.LG] 16 Jun 2016

## Application of autonomous driving using FPGA to accelerate the Inference.

Andre Jatmiko Wijaya, Pedram Babakhani, Naveenkumar Joshi

### Abstract

Autonomous Driving Application can be divided into three phases, Recognition, Prediction and Planning respectively. At Recognition phase, autonomous driver must be able to recognize objects around it. Prediction aims to predict object movements, such as pedestrian motion prediction. Finally, autonomous driver must make a decision accordingly. In this project we are going to focus on pedestrians motion prediction using LSTM architecture. The main idea is using FPGA (Field Programmable Gate Array) as a hardware platrform to test the model in less amount of time signifcantly. Actually we are facing with a real time application where testing time is a critical issue. We introduce FPGA as HW platform to reduce time and power exponentioally in autonomous driving application. The results demonstrate capability of FPGAs in comparison to other hardware platforms like GPUs or CPUs.

### Keywords

### References

- Nurvitadhi, Eriko, et al. "Can fpgas beat gpus in accelerating next-generation deep neural networks?." Proceedings of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays. ACM, 2017.
- Ovtcharov, Kalin, et al. "Accelerating deep convolutional neural networks using specialized hardware." Microsoft Research Whitepaper 2.11 (2015).
- Cong, Jason, et al. "Understanding Performance Differences of FPGAs and GPUs." 2018 IEEE 26th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM). IEEE, 2018.

# Session 2

# Time Series Analysis and Forecasting

Session Chair: Hadi Samer Jomaa

## Algorithmic Trading

Riccardo Lucato, Yang Qi, Eduardo Rocha, Marina Gavrilyuk, Edgar Jimenez

### Abstract

Stock trading has always been more of an art than a science. Investors typically rely on a mix of experience, intuition, knowledge of economical fundamentals and real-time information to make informed choices and try to get as high a rate of return as possible. On the other hand, the existing numerical models are typically either scarcely reliable or make too restrictive assumptions which hamper their applicability in the real world. In this work, we took on the challenge to build a realistic trading algorithm made up of a forecasting component, an optimization component and a trading component. The forecasting component makes use of an LSTM trained on historical prices to predict a daily stock price for three well-known stocks: Apple, Google and Facebook. The output of the forecasting component can be either used directly as a signal to separately trade the three stocks on the trading component or be fed into the optimization component. The later performs a portfolio optimization following Markowitz's approach and allows for the optimal trading of multiple stocks at the same time, overcoming a shortage of the baseline paper.

### Keywords

### References

- N. Nguyen, "An Analysis and implementation of the Hidden Markov Model to Technology Stock Prediction", Risks, 2017.
- Gupta, Aditya, and Bhuwan Dhingra. "Stock market prediction using hidden markov models." Engineering and Systems (SCES), 2012 Students Conference on. IEEE, 2012.
- E. Chong, C. Han, F. C. Park, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case study", Elsevier, 2017.

## Emotion Classification

Luis Alberto Barradas Chacon, Artem Fedoskin, Sutthida Neamsup, Ekaterina Shcheglakova, Arooba Khokhar

### Abstract

This project describes the attempt to classify human emotions without including any sensors measuring biosignals, using the DEAP dataset. Basing our research on the original paper, we claim that emotions (valence and arousal scores) can be classified with only one pulse detecting sensor with a comparable result to the classification based on the EEG signal. Moreover, we propose the method to extract the pulse of the person from the video based on the head movements, avoiding any sensor. Using points of interest on the person's face and a principle component analysis, signals and the average heart rate are extracted and then used in the emotion classification. First part of the project was conducted on the dataset containing 32 participants' HR values (1280 activation cases), while the second part was based on the frontal videos of the participants. Results support the idea of one-sensor emotion classification and show the possibility of zero-sensor classification.

### Keywords

### References

- Balakrishnan, Guha, Fredo Durand, and John Guttag. "Detecting pulse from head motions in video." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013.
- DEAP dataset: http://www.eecs.qmul.ac.uk/mmv/datasets/deap/
- Koelstra, S. et al (2012). Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, 3(1), 18-31

# A Dual Stage Attention Using Deep Learning For Sales Forecasting

Syed Ahmed Husaini, Sahabzada Faaizuddin Qadri , Abdullah Al Foysal, Parsa Mohammadian, Vusal Isayev

## Abstract

This paper proposes sales forecasting based on time series (dynamic) data along with static data. For performing time series predictions, one of the most widely used models is Recurrent Neural Networks (RNNs) with LSTM units. In this paper however, we use NARX neural networks. NARX, along with Echo State Networks (ESN) have been largely used in Short Term Load Forecasting (STFL). Therefore, we work to see how NARX performs in sales forecasting. We propose a model which gives output based on both dynamic and static features. We test the results of the baseline with NASDAQ 100 Stock dataset and the improvised model with Rossmann dataset with other state-of-the-art models.

**Keywords**

## References

- Yao Qin, Dongjin Song, Haifeng Cheng, Wei Cheng, Guofei Jiang, GarrisonW. Cottrell : Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction
- J.P.F. Sum, W.-K. Kan and G.H. Young : A Note on the Equivalence of NARX and RNN
- Anna Leontjeva, Ilya Kuzovkin: Combining Static and Dynamic Features for Multivariate Sequence Classification

# An Efficient Model for The Prediction And Optimization of Energy Consumption using Deep Learning

Aaqib Ali, Alda Cypi, Khushboo Kumari, Muhammad Ahsan Ali, Nazeer Basha Shaik

## Abstract

Energy management is the key to energy savings. The current level of energy production and consumption is not proportional which is one of the biggest concerns of the energy sector. The best approach to resolve this issue is to consume energy efficiently. There are many efforts made to achieve this using automated houses with the help of Machine Learning. The main idea of our project is to optimize the energy consumption of automated houses by predicting the future consumption and then optimizing the usage of energy. For this purpose, we are using forecasting models and minimization models using Deep Learning and Deep Reinforcement Learning respectively. There are many methods available in forecasting domain and for our project, we used several methods. To optimize the usage of the electric energy and to reduce the cost, we used Deep Reinforcement Learning with Deep Policy Gradient as a learning agent.

**Keywords** Forecasting, Reinforcement Learning, Policy Gradient, ARIMA, SARIMAX, Neural Networks, LSTM

## References

- Mocanu, Elena, et al. "On-line building energy optimization using deep reinforcement learning." IEEE Transactions on Smart Grid (2018).
- Li, Chengdong, et al. "Building energy consumption prediction: An extreme deep learning approach." Energies 10.10 (2017): 1525.
- Vagropoulos, Stylianos I., et al. "Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting." Energy Conference (ENERGYCON), 2016 IEEE International. IEEE, 2016.

# Session 3

# Large Scale Machine Learning Problems

Session Chair: Ahmed Rashed

# Large-Scale Dynamic Machine Learning Workflow

Danyal Javeed, Saikiran Goud Sara, Busari Eedris Adeyemi, Vijit Mehrotra, Rohith Kumar

## Abstract

Big data has become a huge part of machine learning nowadays and it has been adopted in many application domains. We are working on three different architectural designs Spark, DTF, and MPI. We use these Distributed Frameworks and analyze which framework performs the best on a given problem [1]. In Spark setup, all the data and its parameters (x, y) are stored on an RDD. Here all the target values are always sparse in case of Matrix Factorization machines. We need to use those dense representations for factor matrices which are stored as RDD for vectors. In DTF we follow a parameter server approach where each task is associated with a master which initializes the session. Workers are used to executing operations on a particular job. In MPI we have used collective communication to solve the chosen machine learning problem. The master scatters the data to the different workers and updates the parameters. Each of the workers updates their own parameters and returns the results back to the master. We are considering fault tolerance and we want to optimize these platforms to achieve a better result in faster time. We evaluate the performance on four different problems, using these architectures on both local and cluster. We are working on several machine learning problem, and the results which we had obtained we could say that Spark has performed much better than DTF and MPI because of its faster computation speed and fault tolerance mechanism.

### Keywords

## References

- YYu, Yuan, et al. "Dynamic control flow in large-scale machine learning." Proceedings of the Thirteenth EuroSys Conference. ACM, 2018.

# 3D model reconstruction from a single 2D input image

Kiran Madhusudhanan, Mohi Uddin Faruq, Saurabh Pradhan, Leandro Muñoz Giraldo

## Abstract

Humans have the ability to understand the 3D shape of an object from its 2D image based on background knowledge of the object. From the computer vision field, the task is complex and is currently considered an open research agenda. Optimal reconstruction of a 3D shape would be useful for the future of highly immersive augmented reality (AR), virtual reality (VR) and general 3D design tasks. This study addresses the research problem using the concept of Free-Form Deformations (FFD) in meshes. The proposed approach uses shape priors (templates) along with the input 2D image to generate deformations on the shape prior, in order to create a 3D mesh for the corresponding input 2D image, via FFD. Our solution is an end-to-end inexpensive learning network that is able to create a 3D mesh model of a given 2D image. Additionally, improvement of the network is intended by learning a 2.5D (Depth map) of the given 2D input image, and then using such 2.5D image to better deform the shape prior.

### Keywords

## References

- Jack, Dominic, et al. "Learning Free-Form Deformations for 3D Object Reconstruction." arXiv preprint arXiv:1803.10932 (2018).
- Eigen, David, Christian Puhrsch, and Rob Fergus. "Depth map prediction from a single image using a multi-scale deep network." Advances in neural information processing systems. 2014.
- Chang, Angel X., et al. "Shapenet: An information-rich 3d model repository." arXiv preprint arXiv:1512.03012 (2015).

# Session 4

# Supervised Classification Approaches

Session Chair: Randolf Scholz

## Prediction and Classification with Multi-Task Setups Using Satellite Image Data

Daniel Pototzky, John Rothman, Florian Pal, Nabeel Janjua, Miika Kolu

### Abstract

Satellite imagery has become more accessible for research, yet the use of satellite images is still a challenge for machine learning. This is mainly due to the lack of auxiliary data, which makes it difficult to create labels for the images. However, satellite images present a great potential for various prediction tasks of societal indicators such as population and poverty, amongst others. Our project is motivated by recent research papers that focus on predicting population, poverty, and urban patterns respectively. We created three different satellite image datasets by labelling the images based on auxiliary data for the aforementioned tasks, which we treat as classification problems. Using the images and corresponding multiple labels we explore the potential of multi- task learning in this domain by using neural network models to predict the labels associated with a particular satellite image. prevent the model from overfitting. We have already obtained results which support this thesis, and have compared the results to benchmarks achieved with single-task learning models. The concept of data distillation was applied to extract information from unlabelled images and led to a further increase in accuracy.

### Keywords

### References

- A. Albert, J. Kaur, M. Gonzales (2017): Using convolutional networks and satellite imagery to identify patterns in urban environments at large scale.
- N. Jean, M. Burke, et al. (2016): Combining satellite imagery and machine learning to predict poverty.
- C. Robinson, F. Hohman, B. Dilkina (2017): A Deep Learning Approach for Population Estimation from Satellite Imagery.

## End-to-End Motion Classification with Smartwatch Sensor Data

Tamara Belostotskaya, Valerie Chikukwa, Junaid Ahmed Ghauri, Muhammad Usman Syed, Torben Windler

### Abstract

The analysis of sensor data from Smart Devices for the classification of activities is becoming more and more interesting in industry and consumption. With the growing spread of smart- watches, this data is available to everyone. In order to obtain reliable information about what a user is doing at a certain moment, the data from accelerometers and gyroscopes are analyzed as multivariate time series. Due to specific sampling rate instabilities depending on the certain devices, previous approaches mainly work with feature extraction methods, which require a lot of time and expertise and also provide only limited helpful accuracies. To overcome this, an end-to-end model for activity classification based on convolutional neural networks without feature extraction is presented in this work. The data preprocessing is not very computationally intensive and the model is able to deal with the irregularities of the data. In addition,quasi-online prediction (with a delay of the selected window-size) of test data is possible and the accuracy increased significantly compared to feature extraction approaches.

### Keywords

### References

- Stisen, Allan, et al. "Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition." Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems. ACM, 2015.
- Yao, Shuochao, et al. "Deepsense: A unified deep learning framework for time-series mobile sensing data processing." Proceedings of the 26th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2017.

## Tumor Classification Based on MR Images of the Brain

Lucky Okehigbemen, Nazmus Sakib

### Abstract

Glioblastoma Multiforme (Gliomas) is the most common and aggressive type of primary brain tumors. Malignant tumor is usually called High Grade Gliomas (HGG) which is sub-classified into 3 and 4, while the benign is called the Low Grade Gliomas (LGG) and it is sub-classified into 1 and 2. Benign tumor grading is important for adequate treatment planning and monitoring. However, this procedure is invasive, time consuming, and prone to sampling error. We propose to use Convolutional Neural Network for predicting and classifying between the low grades gliomas directly from Magnetic Resonance Imaging(MRI) data. To avoid the misclassification rate of the tumor sub-class, we propose to use the AUC evaluation measures. We consider this part to be a simple task learning (STL). From then on we will extend it to Multi-Task Feature Learning (MTL), since our model has a hierarchical structure and thus compare the result between the single task and multi-task performance.

**Keywords** MRI, Convolution neural network, tumor classification

### References

- Mohsen, Heba, et al. "Intelligent Methodology for Brain Tumors Classification in Magnetic Resonance Images." (2017).
- Abiwinanda, Nyoman, et al. "Brain Tumor Classification Using Convolutional Neural Network." World Congress on Medical Physics and Biomedical Engineering 2018. Springer, Singapore, 2019.
- Murthy, Veda, et al. "Center-focusing multi-task cnn with injected features for classification of glioma nuclear images." Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on. IEEE, 2017.

# Session 5

# External Pitches

## Taking Reinforcement Learning a Step Further

Artus Krohn-Grimberghe

### Topics

- Deep RL for adversarial games
  - Deep RL for adversarial games: Go was yesterday. Deep Reinforcement Learning has shown first results in partially observable adversarial settings such as Dota as demonstrated by the "OpenAI 5". How about starting from there and going to more advanced settings?
- AI for Healthcare
  - We have some nice healthcare datasets and want you to focus on AI-based productivity tools for doctors and researchers. Interested in object detection for cancer or reinforcement learning and active learning for labeling assistance? The focus is on efficient tools and results.
- AI for Industry 4.0
  - If you like manufacturing and want to leave your dent there, support us in coming up with faulty part detection algorithms, condition monitoring and other ideas on the real-world datasets that we got from our partners

## Mileage Forecast

Volkswagen Financial Services
Contact Person : Anne-Christin Schwarz
(vwvg.anne-christin.schwarz@volkswagen.de)

### Abstract

The department Analytics Services & Business Development acts in the name of the dealership to offer analytics insights for e.g. campaigns. The aim is to identify individual customer needs and to be able to identify them in a personalized and timely manner. The goal is to est imate the annual mileage (km / a) of a customer last year and to proje ct a new mileage into the future.

## Analysis of (video-recorded and transcribed) classroom interaction in phases of acquisition and consolidation of knowledge and skills

Prof. Dr. Christof Wecker (christof.wecker@uni-hildesheim.de)

### Abstract

Learners' cognitive activities and verbal contributions to classroom discourse are regarded as core factors that influence learning outcomes. In research on teaching, these variables are measured based on video-recordings and trancripts of lessons. Applying data science methods to the analysis of these mechanisms can contribute both to the growth of scientific knowledge on the role of classroom interaction for learning and to the development of adaptive support for individual learners based on automated assessments of their current understanding, misconceptions, difficulties or level of engagement. In this project, students from data science and teacher education programs will collaborate on data collected by prospective teachers in fulfilment of course requirements in a project seminar runni

# Session 6

# Internal Pitches

## Time Series Analysis

Uni Hildesheim

### Modelling user behaviour using deep learning algorithms

Supervisor : Hadi S. Jomaa (hsjomaa@ismll.de)
Description : This project revolves around the users of parking applications. Particularly, this project is about better understanding the behaviour of users, i.e. the parking patterns, preferred locations, etc. through deep learning.
References


- Vieira, Armando. "Predicting online user behaviour using deep learning algorithms." arXiv preprint arXiv:1511.06247 (2015).


### Exploring the influence of data aggregation in parking prediction

Supervisor : Hadi S. Jomaa (hsjomaa@ismll.de)
Description : Parking availablility is influenced by multiple factors, which can include climatological changes, i.e. temperature and pressure, holidays, social events, number of available parking spots, etc. In this project, you will explore the impact of these external factors on parking availabilty models, by increasing the feature space or the model architecture to accommodate this new information.
References


- Yao, Huaxiu, et al. "Deep multi-view spatial-temporal network for taxi demand prediction." arXiv preprint arXiv:1802.08714 (2018).

### How long is he staying ?

Supervisor : Hadi S. Jomaa (hsjomaa@ismll.de)
Description : Parking applications offer the user the ability to park their car at a specified location for a fixed duration. Depending on the time of day, or the location, or even the task which pushed the user to drive somewhere, and ultimately park, the duration of stay can change. In this project, you will learn the expected duration in which a location is would still be occupied, given the historical data of said location.
References


- Shumway, Robert H., and David S. Stoffer. "Time series regression and exploratory data analysis." Time series analysis and its applications. Springer New York, 2011. 47-82.


### Motion Reconstruction from Scarce/Sparse Data

Supervisor : Rafael Rego Drumond (radrumond@ismll.de)
Description : In the field of entertainment such as movies and game development, there are several techniques to bring a user body motion to a digitalized environment. Several research papers have been working in mapping users limbs to a digital model. This process is not always possible when data representing the actor's body is not "complete" (i.e., not all body joints or limbs are mapped). This research task aims to recover or estimate missing data given the visible part as well as to provide an analysis of the most representative body joints using accelerometers and Motion Capture data-sets.


- May, Tobias. "Influence of binary mask estimation errors on robust speaker identification." Speech Communication 87 (2017): 40-48.

## Reinforcement-Learning

Uni Hildesheim

### Uncertainty-driven Exploration in Q-learning

Supervisor : Hadi S. Jomaa (hsjomaa@ismll.de)
Description : Several approaches have been proposed to boost exploration in reinforecement learning, which include action noise injection, parameter space noise injection, curiosity drive exploration, etc. In this project, you will focus on uncertainty-driven exploration in discrete action-spaces, where uncertainty is introduced as an additional output associated with every action.

- Kendall, Alex, and Yarin Gal. "What uncertainties do we need in bayesian deep learning for computer vision?." Advances in neural information processing systems. 2017.

### A direct RL Approach to Multi-Attribute Vehicle Routing Problems (MAVRP)

Supervisor : Jonas Falkner (falkner@ismll.de)
Description : VRP are complex extensions of the traveling salesman problem, they are a recent new areafor RL methods, since VRP normally are NP-hard, heuristic or NDP approaches like RL have to be employed

- Nazari, MohammadReza, et al. "Reinforcement Learning for Solving the Vehicle Routing Problem." Advances in Neural Information Processing Systems. 2018.

## Recommender Systems

Uni Hildesheim

### Academic Preformance Prediction via Collaborative Filtering

Supervisor : Ahmed Rashed (ahmedrashed@ismll.de)
Description : Many cognitive, behavioral, and environmental factors impact student learning curve. -Predicting the academic performance of students will give us the ability to anticipate any degrade in the learning curve which means we have an opportunity to take critical decisions to avoid such degradation before their actual occurrence. -Academic performance prediction models can also be used in application portals to rank student applications based on their predicted performance

- He, Xiangnan, et al. "Neural collaborative filtering." Proceedings of the 26th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2017.

### Multi-Relational Classification for Image-based Recommendations

Supervisor : Ahmed Rashed (rahmedrashed@ismll.de)
Description : The task of classifying multi-relational data spans a wide range of domains such as document classification in citation networks, recommended systems and protein labeling in proteins interaction graphs. Current state-of-art classification models rely on learning per-entity latent representations by mining the whole structure of the relations graph, however, they still face two major problems. Firstly, it is very challenging to generate expressive latent representations in sparse multi-relational settings with implicit feedback relations as there is very little information per-entity. Secondly, for entities with structured properties such as text in documents and images in online products, models have to be modified ad- hoc. In this project we aim to enhance the item recommendation task by utilizing different purchase relations between items which are currently not considered such as 1) users who viewed X also viewed Y 2) users who viewed X eventually bought Y 3) users who bought X also bought Y 4) users bought X and Y simultaneously'

- McAuley, J., Targett, C., Shi, Q., & Van Den Hengel, A. (2015, August). Image-based recommendations on styles and substitutes. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 43-52). ACM.

## Supervised Classification/Regression

Uni Hildesheim

### Speaker identification with Noise Filtering

Supervisor : Rafael Rego Drumond (radrumond@ismll.de)
Description : Speaker identification is an essential task in the field of security and speech processing. In the literature, we find many works that are able to identify a speaker in clear audio speech conditions fully. However, it is not feasible to consider clear audio data in a real-life situation where audio is usually distorted by background noise or device interference. In this research, your goal is to be able to filter the audio, or it's extracted features to facilitate speaker identification. Such techniques might include masking, de-noising auto-encoders, and generative adversarial networks. You must understand the most recent techniques and theory behind it and be able to improve the performance of the state-of-art SID models.

- May, Tobias. "Influence of binary mask estimation errors on robust speaker identification." Speech Communication 87 (2017): 40-48.

### Estimating Box Office Return

Supervisor : Mofassir Al-Islam Arif (mofassir@ismll.de)
Description : Before a movie's release, the hype and media attention increases, and peaks when official trailers drop. This helps producers get an estimate of how much the profit would be. However, several would be blockbuster hits have turned out to be flops. In this project, you will investigate the movie trailers and build deep learning models that are able to predict the box-office return in the opening weekend.

- Varol, Gül, Ivan Laptev, and Cordelia Schmid. "Long-term temporal convolutions for action recognition." IEEE transactions on pattern analysis and machine intelligence 40.6 (2018): 1510-1517.

## Learning data augmentation with adversarial networks

Supervisor : Lukas Brinkmeyer (lukas.brinkmeyer@ismll.de)
Description : Automate the process of data augmentation with the use of an adversarial network and create hard examples based on known operators (e.g. rotating for image data) but also learn novel augmentations with a generative approach.


- Peng, Xi, et al. "Jointly optimize data augmentation and network training: Adversarial data augmentation in human pose estimation." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.


## Deep Neural Networks as Feature Selectors for Kernel Regression

Supervisor : Randolf Scholz (rscholz@ismll.de)
Description : Artificial Neural Networks (ANNs) are the modern day workhorse of most Machine Learning applications, and older techniques from non-parametric statistic seem forgotten. However ANNs typically suffer from data inefficiency, that is, to train a Neural Network typically a very large amount of data is required. By contrast, classical Kernel Regressors and K-Nearest-Neighbor models require far less data to make useful predictions. The reasons they have fallen out of favor is that overall they do not exhibit the same generalization power and performance as ANNs when the training data is not "nice" (equally spaced points without gaps). To get best of both worlds an idea has emerged to use combine both models and use an ANN as a "feature selector" - a preprocessor which transforms the original data into a form which is exceptionally good for kernel regression. In their 2017 paper "Neural Episodic Control" Pritzel et. al. showed that this approach yielded superior data efficency in typical Reinforcement Learning problems. The goal of this student research project is to improve upon this idea and apply it to supervised classification tasks.


- Pritzel, Alexander, et al. "Neural episodic control." arXiv preprint arXiv:1703.01988 (2017).