Review for Reports of Team 1 to Team 10 for Course "Learning from Data"

Reviewer: Zifeng Wang

Group 1: Network-based Anomaly Detection against IoT Cyberattacks using CNN-LSTM Autoencoder

Member: 王平东 李若愚 王策

Unknown attacks can easily fool the existing supervised learning based IoT security system, hence this paper presents a CNN-LSTM autoencoder for learning from normal data only, and employed for detecting anomaly data. Experiments on real-world IoT traffic data validate the effectiveness of the proposed framework.

pros.

- Experiments are performed on real-world IoT traffic data, which enhances the credibility of this work.

cons.

- Introduction is relatively weak, as the topic of this paper is anomaly detection, presenting too much about IoT and attacks are not necessary.
- This paper claims that "autoencoders were not used at all" in the literature, referring to a survey in 2014. I think it is not the truth now, the authors should survey more recent papers before claim it.
- How about analysis of the limitations of existing approach? It could be better if there is a section about future work.

Group 2: Intelligent Diagnosis of Cervical Cancer Risk

Member: 杨品慈 彭天任 张可

This work focus on Cervical cancer automatic detection via pathology board photo. Retinanet, with ResNet as its backbone, is selected for this task.

pros.

- A comprehensive introduction of used techniques, the implementation details about parameters, and results demonstration on different metrics.

- Theoretical analysis is performed trying to explain the results of different models.
- The analysis of limitations of the current work is complete, from data side to result side.
- Good visualization for understanding.

cons.

- The data set used in this work comes from Aliyun Tianchi challenge, so how about this work's result compared to the winner of this challenge? Even this work cannot reach the best, it is still reasonable to analyze the cause of the gap.
- It should be claimed the source code, from github or by authors. Or the specific modifications made by authors.

Group 3: Federated learning system targeted on MNIST

Member: 刘心怡 蒋慧玲 毛思梅

This work introduces conception of federated learning (FL), and the challenge of statistical heterogeneity (i.e. not i.i.d.) of different data source, faced by the current most popular FL algorithm FedAvg. Then comes with the proposed metric based on probability distribution distance, to reweighting parameters from different devices in aggregation phase, namely FedJS.

pros.

- An introduction to background of FL, and point a key challenge which is still an open problem in the literature.
- A good attempt to propose an original idea for solving the problem.
- Whole writing concentrates on a specific point.

cons.

- No good labeling of Equations for reference in this work.
- About Non-i.i.d.'s effects, only one point: slow convergence of training, is mentioned. I am wondering the detailed reason of it.
- Followed by above point, it seems that only testing accuracy of different methods is evaluated, how about the convergence speed mentioned in introduction?
- Furthermore, from the viewpoint of this work, why the client with more distant distribution to the union distribution needs be assigned larger weights during aggregation? I am wondering that it may cause training deviating from the optima for others closer to the union distribution.
- Figure 3 is not well-organized for demonstration.
- It could be better to validate the choice of distance metric, there is no attempt to explain why JS-Divergence is picked in this work. For finite dimensional distribution, like the histogram here, the Wasserstein distance might be better choice.

Group 4: The Application of Gradient-based Meta-learning Scheme on Few-shot Learning

Member: 谭杨 郑子严 王伟达

Few-short learning attracts concern for the dilemma in requirements of large amount of data by high-performance neural networks. A recent state-of-the-art gradient-based meta-learning pipeline MAML is proved being succeed in improving few-short learning performance as well as adaptation capacity to extensive tasks. This work casts on generalization ability of it, that is, how well does MAML perform compared with pre-training methods. The experimental results validate the MAML's superiority.

pros.

- The experimental setting, as well as the description of results, is technically sound, which is convincing for illustrating performance of the two methods.
- Detailed introduction of definition of few-short learning, and its current challenges mentioned in the literature.

cons.

- On generalization ability, the first term comes to my head is deriving generalization error bound. It is ok to empirically evaluate it, while it should be more appealing and useful to explore its statistical property, and might shed lights on better approaches.
- For reproducing the existing method, which may have available code or not. It should be noted where this work's code comes, by authors themselves, or drawn from github.

Group 5: Multi-Zone Taxi Order Demand Prediction

Member: 李鹏舜 程晨曦 田宸宇

This work focuses on the key challenge for resource allocation and travel management, the travel order demand prediction in real-time. First, the kmeans++ is employed for zone division in Shenzhen. Then, feature selection is conducted, followed by 6 methods for predicting order demand. The kNN fusion based method is the best regarding to the case study results.

pros.

- Good to claim the contributions of group members. It means the team plays synergy well.
- The abstract concludes the whole picture of this work, from motivation, method, to the final case study.

- Good to find an unique point which is hardly exploited in the literature: the multi-zone prediction analysis considering order spatial distribution, and zone prediction.
- Data visualization is great to help the readers understand this work.

cons.

- Selected methods, like SVR, RF, etc. are not the state-of-the-art approaches for predicting spatial temporal traffic. Though they may be satisfying in practice, it could be better to compare their performance with the GCN-based approaches, which have been widely used for traffic prediction recently.
- I am bit confused about the conception of multi-zone prediction throughout this paper. Does it mean we first clustering the data, then apply same method with different parameters for different clusters of data?

Group 6: Debugging Neural Networks

Member: 贝辰峰 余文才 亓熙烁

This work presents to visualize a classifier's (e.g. a neural network) robustness by synthesizing human understandable images, by a GAN as an image prior. Deep Visualization (DV) technique, such as Activation maximization, is proposed in the literature to generate images for any specific neuron in a DNN, while the generated images are not interpretable. Besides, the existing works are not easy-to-use. Considering the challenges, this work proposes to utilize a trained GAN as the generator network, and testify the GAN-based Activation maximization framework on the selected 100 image per class ImageNet.

pros.

- The article is well-formatted, with architecture, implementation demonstration, as well as the results visualization.
- Very novel and interesting topic within scope of interpretable deep learning, the introduction has a well review over the literature, and explains how does DV technique work in practice.
- The authors attempt to propose their original idea for improving method's interpretability, and the finds that appearance of a boy when maximizing the output neuron with the label of ball are interesting. The authors give their explanation about it.
- Good to claim the code drawn from github.

cons.

- This work is trying to grab an existing GAN, and incorporate it in the Activation maximization framework. From Fig.6, the trained GAN might not be good enough, which degenerates the credibility of the experiments.
- 70% for training, 20% for validation, and 10% for testing might not be a good choice. Empirically we tend to select a larger test set, and a smaller validation set.

Group 7: Stock Price Prediction via Various Models

Member: 陈梦玄 安志成 纪鸿璐

112 alpha factors are selected, with three genres of regression models: linear regression, tree models and neural networks, to do stock price prediction in the secondary market. Interesting results show that feature selection is hardly helpful for improving effectiveness of prediction. This work ends with building stock portfolio based on previously mentioned best algorithm, and doing back test to evaluate the portfolio return.

pros.

- A complete technical report about quantitative investment using machine learning model. From factor selection, model selection and building portfolio. The experiments performed are complete, and the main framework of this article is clear.

cons.

- Not titled. A title is important for readers at first glance to understand this work's theme.
- For time series prediction, a gold standard is to ensure the input series are stationary. In this work, the return is used as target, with defining a start point in one date. In long period, the return series must not be stationary, there is no mention about how to do preprocessing considering it.
- It is somewhat weird that this model can maintain very good performance in out-of-sample test within 2 years, on the prediction of daily return. In practice, this hardly happen because stock price is far from than i.i.d. over time. It could be better for clarification of which target used in the experiments, and more implementation details.

Group 8: MR-Image to CT-Image Translation

Member: 张晓筱 张文杰 黄欣欣

CT scan is harmful for human beings, but is useful for detecting extensive diseases; MRI is safer, but not so powerful as CT. This work attempts to translate MRI to CTI, such that we can obtain a safe synthetizing CTI for supporting disease diagnosis. It adopts a deep convolutional neural network framework proposed in the literature, for MRI to CTI translation, and evaluates the results.

pros.

- A good attempt to apply computer vision techniques for healthcare.

cons.

- This report replicates the paper "MR-based synthetic CT generation using a deep convolutional neural network method", but it does appear in the

reference.
- Describing the 62th, 1062th, ..., 5062th epochs of learned models as model1 ot model4 is somewhat weird. And selection of the epoch is also magic number, it might be better to select like 100, 200, ..., etc.
- When reproducing existing work, it is recommended to claim the source code, if it is achieved by authors, or adapted from a github repo.
- No explanation to selecting DCNN for experiment, even if in introduction it mentions that there exist various GAN-based approach for this multi-modal learning task.
- The Figure 3 seems to be drawn from online or other papers, it should be claimed the source.

Group 9: Solving AI safety gridworld problems using auxiliary rewards

Member: 苑乐文 邓翔天 李国栋

AI system's robustness towards adversarial attacks is a great concern recently. Eight different AI safety environments are considered in this work, for empirically test the reinforcement learning algorithm's robustness, faced with different safety challenges. This paper develops empirical solutions for specific 3 safety problems, by optimizing on cumulated reward and considering auxiliary reward functions.

pros.

- Good to claim contributions of the team members.
- Interesting topic about AI safety problems in RL, which might not be fully explored currently. The RL in real-world application is complicated, considering multiple objectives instead of only maximizing a predefined reward.
- This work tries to propose novel ideas to solve specific problems, though not further associated theoretical explanation, it is still a good attempt.

cons.

- This work concentrates on the 3 out of all 8 problems mentioned, so it might be better concentrate on the 3, to press the readers about this work's contribution.

Group 10: Graph Neural Networks with Applications to Neural Machine Translation

Member: 倪登伟 史蒂文

Before GCN becoming emerging recently, Neural machine translation (NMT) is always performed on encoder-decoder network with long short-term memory (e.g. Seq2Seq). The GCN approach, which has succeeded in extensive tasks, becomes appealing for NMT as well. This work implement both basic RNN and

attention GNN for comparison on NMT task between Portuguese and English. The experiment shows GNN provides meaningful benefits and interpretability.

pros.

- Good to claim contributions of team members.
- Comprehensive thinking for current challenges in GNN for NMT, as distance metric definition, word property identification, word placement for grammar understanding, etc. Though they are sidestepped, it is still worth exploring in future works.

cons.

- The figures (a), (b), .. in Result section is not labeled, and they are too small for read.
- Using attention for GNN is a good idea, while it could be better for exploiting deeper about interpretability brought by attention mechanism, e.g. the attention weights point out the different importance of input features.
- Analysis of superiority brought by GNN says that attention might be the major contributor, while it could be better followed by an ablation study to support or reject this idea.