Mon 2024.05.06

Neural Context Flows for Learning Generalizable Dynamical Systems

Roussel Desmond Nzoyem, David A. W. Barton, Tom Deakin

Neural Ordinary Differential Equations typically struggle to generalize to new dynamical behaviors created by parameter changes in the underlying system, even when the dynamics are close to previously seen behaviors. The issue gets worse when the changing parameters are unobserved, i.e., their value or influence is not directly measurable when collecting data. We introduce Neural Context Flow (NCF), a framework that encodes said unobserved parameters in a latent context vector as input to a vector field. NCFs leverage differentiability of the vector field with respect to the parameters, along with first-order Taylor expansion to allow any context vector to influence trajectories from other parameters. We validate our method and compare it to established Multi-Task and Meta-Learning alternatives, showing competitive performance in mean squared error for in-domain and out-of-distribution evaluation on the Lotka-Volterra, Glycolytic Oscillator, and Gray-Scott problems. This study holds practical implications for foundational models in science and related areas that benefit from conditional neural ODEs. Our code is openly available at https://github.com/ddrous/ncflow.

link: http://arxiv.org/abs/2405.02154v1

Multi-method Integration with Confidence-based Weighting for Zero-shot Image Classification

Siqi Yin, Lifan Jiang

This paper introduces a novel framework for zero-shot learning (ZSL), i.e., to recognize new categories that are unseen during training, by using a multi-model and multi-alignment integration method. Specifically, we propose three strategies to enhance the model's performance to handle ZSL: 1) Utilizing the extensive knowledge of ChatGPT and the powerful image generation capabilities of DALL-E to create reference images that can precisely describe unseen categories and classification boundaries, thereby alleviating the information bottleneck issue; 2) Integrating the results of text-image alignment and image-image alignment from CLIP, along with the image-image alignment results from DINO, to achieve more accurate predictions; 3) Introducing an adaptive weighting mechanism based on confidence levels to aggregate the outcomes from different prediction methods. Experimental results on multiple datasets, including CIFAR-10, CIFAR-100, and TinyImageNet, demonstrate that our model can significantly improve classification accuracy compared to single-model approaches, achieving AUROC scores above 96% across all test datasets, and notably surpassing 99% on the CIFAR-10 dataset.

link: http://arxiv.org/abs/2405.02155v1

Simulating the economic impact of rationality through reinforcement learning and agent-based modelling

Simone Brusatin, Tommaso Padoan, Andrea Coletta, Domenico Delli Gatti, Aldo Glielmo

Agent-based models (ABMs) are simulation models used in economics to overcome some of the limitations of traditional frameworks based on general equilibrium assumptions. However, agents within an ABM follow predetermined, not fully rational, behavioural rules which can be cumbersome to design and difficult to justify. Here we leverage multi-agent reinforcement learning (RL) to expand the capabilities of ABMs with the introduction of fully rational agents that learn their policy by interacting with the environment and maximising a reward function. Specifically, we propose a 'Rational macro ABM' (R-MABM) framework by extending a paradigmatic macro ABM from the economic literature. We show that gradually substituting ABM firms in the model with RL agents, trained to maximise profits, allows for a thorough study of the impact of rationality on the economy. We find that RL agents spontaneously learn three distinct strategies for maximising profits, with the optimal strategy depending on the level of market competition and rationality. We also find that RL agents with independent policies, and without the ability to communicate with each other,

spontaneously learn to segregate into different strategic groups, thus increasing market power and overall profits. Finally, we find that a higher degree of rationality in the economy always improves the macroeconomic environment as measured by total output, depending on the specific rational policy, this can come at the cost of higher instability. Our R-MABM framework is general, it allows for stable multi-agent learning, and represents a principled and robust direction to extend existing economic simulators.

link: http://arxiv.org/abs/2405.02161v1

Mapping the Unseen: Unified Promptable Panoptic Mapping with Dynamic Labeling using Foundation Models

Mohamad Al Mdfaa, Raghad Salameh, Sergey Zagoruyko, Gonzalo Ferrer

In the field of robotics and computer vision, efficient and accurate semantic mapping remains a significant challenge due to the growing demand for intelligent machines that can comprehend and interact with complex environments. Conventional panoptic mapping methods, however, are limited by predefined semantic classes, thus making them ineffective for handling novel or unforeseen objects. In response to this limitation, we introduce the Unified Promptable Panoptic Mapping (UPPM) method. UPPM utilizes recent advances in foundation models to enable real-time, on-demand label generation using natural language prompts. By incorporating a dynamic labeling strategy into traditional panoptic mapping techniques, UPPM provides significant improvements in adaptability and versatility while maintaining high performance levels in map reconstruction. We demonstrate our approach on real-world and simulated datasets. Results show that UPPM can accurately reconstruct scenes and segment objects while generating rich semantic labels through natural language interactions. A series of ablation experiments validated the advantages of foundation model-based labeling over fixed label sets.

link: http://arxiv.org/abs/2405.02162v1

EEG2TEXT: Open Vocabulary EEG-to-Text Decoding with EEG Pre-Training and Multi-View Transformer

Hanwen Liu, Daniel Hajialigol, Benny Antony, Aiguo Han, Xuan Wang

Deciphering the intricacies of the human brain has captivated curiosity for centuries. Recent strides in Brain-Computer Interface (BCI) technology, particularly using motor imagery, have restored motor functions such as reaching, grasping, and walking in paralyzed individuals. However, unraveling natural language from brain signals remains a formidable challenge. Electroencephalography (EEG) is a non-invasive technique used to record electrical activity in the brain by placing electrodes on the scalp. Previous studies of EEG-to-text decoding have achieved high accuracy on small closed vocabularies, but still fall short of high accuracy when dealing with large open vocabularies. We propose a novel method, EEG2TEXT, to improve the accuracy of open vocabulary EEG-to-text decoding. Specifically, EEG2TEXT leverages EEG pre-training to enhance the learning of semantics from EEG signals and proposes a multi-view transformer to model the EEG signal processing by different spatial regions of the brain. Experiments show that EEG2TEXT has superior performance, outperforming the state-of-the-art baseline methods by a large margin of up to 5% in absolute BLEU and ROUGE scores. EEG2TEXT shows great potential for a high-performance open-vocabulary brain-to-text system to facilitate communication.

link: http://arxiv.org/abs/2405.02165v1

Self-Supervised Learning for Real-World Super-Resolution from Dual and Multiple Zoomed Observations

Zhilu Zhang, Ruohao Wang, Hongzhi Zhang, Wangmeng Zuo

In this paper, we consider two challenging issues in reference-based super-resolution (RefSR) for smartphone, (i) how to choose a proper reference image, and (ii) how to learn RefSR in a self-supervised manner. Particularly, we propose a novel self-supervised learning approach for real-world RefSR from observations at dual and multiple camera zooms. Firstly, considering the popularity of multiple cameras in modern smartphones, the more zoomed (telephoto) image can be

naturally leveraged as the reference to guide the super-resolution (SR) of the lesser zoomed (ultra-wide) image, which gives us a chance to learn a deep network that performs SR from the dual zoomed observations (DZSR). Secondly, for self-supervised learning of DZSR, we take the telephoto image instead of an additional high-resolution image as the supervision information, and select a center patch from it as the reference to super-resolve the corresponding ultra-wide image patch. To mitigate the effect of the misalignment between ultra-wide low-resolution (LR) patch and telephoto ground-truth (GT) image during training, we first adopt patch-based optical flow alignment and then design an auxiliary-LR to guide the deforming of the warped LR features. To generate visually pleasing results, we present local overlapped sliced Wasserstein loss to better represent the perceptual difference between GT and output in the feature space. During testing, DZSR can be directly deployed to super-solve the whole ultra-wide image with the reference of the telephoto image. In addition, we further take multiple zoomed observations to explore self-supervised RefSR, and present a progressive fusion scheme for the effective utilization of reference images. Experiments show that our methods achieve better quantitative and qualitative performance against state-of-the-arts. Codes are available at https://github.com/cszhilu1998/SelfDZSR_PlusPlus.

link: http://arxiv.org/abs/2405.02171v1

Hoaxpedia: A Unified Wikipedia Hoax Articles Dataset

Hsuvas Borkakoty, Luis Espinosa-Anke

Hoaxes are a recognised form of disinformation created deliberately, with potential serious implications in the credibility of reference knowledge resources such as Wikipedia. What makes detecting Wikipedia hoaxes hard is that they often are written according to the official style guidelines. In this work, we first provide a systematic analysis of the similarities and discrepancies between legitimate and hoax Wikipedia articles, and introduce Hoaxpedia, a collection of 311 Hoax articles (from existing literature as well as official Wikipedia lists) alongside semantically similar real articles. We report results of binary classification experiments in the task of predicting whether a Wikipedia article is real or hoax, and analyze several settings as well as a range of language models. Our results suggest that detecting deceitful content in Wikipedia based on content alone, despite not having been explored much in the past, is a promising direction.

link: http://arxiv.org/abs/2405.02175v1

Assessing and Verifying Task Utility in LLM-Powered Applications

Negar Arabzadeh, Siging Huo, Nikhil Mehta, Qinqyun Wu, Chi Wang, Ahmed Awadallah, Charles L. A. Clarke, Julia Kiseleva

The rapid development of Large Language Models (LLMs) has led to a surge in applications that facilitate collaboration among multiple agents, assisting humans in their daily tasks. However, a significant gap remains in assessing to what extent LLM-powered applications genuinely enhance user experience and task execution efficiency. This highlights the need to verify utility of LLM-powered applications, particularly by ensuring alignment between the application's functionality and end-user needs. We introduce AgentEval, a novel framework designed to simplify the utility verification process by automatically proposing a set of criteria tailored to the unique purpose of any given application. This allows for a comprehensive assessment, quantifying the utility of an application against the suggested criteria. We present a comprehensive analysis of the effectiveness and robustness of AgentEval for two open source datasets including Math Problem solving and ALFWorld House-hold related tasks. For reproducibility purposes, we make the data, code and all the logs publicly available at https://bit.ly/3w3yKcS .

link: http://arxiv.org/abs/2405.02178v1

Training-Free Deepfake Voice Recognition by Leveraging Large-Scale Pre-Trained Models

Alessandro Pianese, Davide Cozzolino, Giovanni Poggi, Luisa Verdoliva

Generalization is a main issue for current audio deepfake detectors, which struggle to provide reliable results on out-of-distribution data. Given the speed at which more and more accurate

synthesis methods are developed, it is very important to design techniques that work well also on data they were not trained for.In this paper we study the potential of large-scale pre-trained models for audio deepfake detection, with special focus on generalization ability. To this end, the detection problem is reformulated in a speaker verification framework and fake audios are exposed by the mismatch between the voice sample under test and the voice of the claimed identity. With this paradigm, no fake speech sample is necessary in training, cutting off any link with the generation method at the root, and ensuring full generalization ability. Features are extracted by general-purpose large pre-trained models, with no need for training or fine-tuning on specific fake detection or speaker verification datasets. At detection time only a limited set of voice fragments of the identity under test is required. Experiments on several datasets widespread in the community show that detectors based on pre-trained models achieve excellent performance and show strong generalization ability, rivaling supervised methods on in-distribution data and largely overcoming them on out-of-distribution data.

link: http://arxiv.org/abs/2405.02179v1

A Flow-Based Model for Conditional and Probabilistic Electricity Consumption Profile Generation and Prediction

Weijie Xia, Chenguang Wang, Peter Palensky, Pedro P. Vergara

Residential Load Profile (RLP) generation and prediction are critical for the operation and planning of distribution networks, particularly as diverse low-carbon technologies are increasingly integrated. This paper introduces a novel flow-based generative model, termed Full Convolutional Profile Flow (FCPFlow), which is uniquely designed for both conditional and unconditional RLP generation, and for probabilistic load forecasting. By introducing two new layers--the invertible linear layer and the invertible normalization layer--the proposed FCPFlow architecture shows three main advantages compared to traditional statistical and contemporary deep generative models: 1) it is well-suited for RLP generation under continuous conditions, such as varying weather and annual electricity consumption, 2) it shows superior scalability in different datasets compared to traditional statistical, and 3) it also demonstrates better modeling capabilities in capturing the complex correlation of RLPs compared with deep generative models.

link: http://arxiv.org/abs/2405.02180v1

Imitation Learning in Discounted Linear MDPs without exploration assumptions

Luca Viano, Stratis Skoulakis, Volkan Cevher

We present a new algorithm for imitation learning in infinite horizon linear MDPs dubbed ILARL which greatly improves the bound on the number of trajectories that the learner needs to sample from the environment. In particular, we remove exploration assumptions required in previous works and we improve the dependence on the desired accuracy ϵ from $\mathcal{O}\br{\epsilon^{-5}}$ to $\mathcal{O}\br{\epsilon^{-4}}$. Our result relies on a connection between imitation learning and online learning in MDPs with adversarial losses. For the latter setting, we present the first result for infinite horizon linear MDP which may be of independent interest. Moreover, we are able to provide a strengthen result for the finite horizon case where we achieve $\mathcal{O}\br{\epsilon^{-2}}$. Numerical experiments with linear function approximation shows that ILARL outperforms other commonly used algorithms.

link: http://arxiv.org/abs/2405.02181v1

Metalearners for Ranking Treatment Effects

Toon Vanderschueren, Wouter Verbeke, Felipe Moraes, Hugo Manuel Proença

Efficiently allocating treatments with a budget constraint constitutes an important challenge across various domains. In marketing, for example, the use of promotions to target potential customers and boost conversions is limited by the available budget. While much research focuses on estimating causal effects, there is relatively limited work on learning to allocate treatments while considering the operational context. Existing methods for uplift modeling or causal inference primarily estimate treatment effects, without considering how this relates to a profit maximizing

allocation policy that respects budget constraints. The potential downside of using these methods is that the resulting predictive model is not aligned with the operational context. Therefore, prediction errors are propagated to the optimization of the budget allocation problem, subsequently leading to a suboptimal allocation policy. We propose an alternative approach based on learning to rank. Our proposed methodology directly learns an allocation policy by prioritizing instances in terms of their incremental profit. We propose an efficient sampling procedure for the optimization of the ranking model to scale our methodology to large-scale data sets. Theoretically, we show how learning to rank can maximize the area under a policy's incremental profit curve. Empirically, we validate our methodology and show its effectiveness in practice through a series of experiments on both synthetic and real-world data.

link: http://arxiv.org/abs/2405.02183v1

Optimistic Regret Bounds for Online Learning in Adversarial Markov Decision Processes

Sang Bin Moon, Abolfazl Hashemi

The Adversarial Markov Decision Process (AMDP) is a learning framework that deals with unknown and varying tasks in decision-making applications like robotics and recommendation systems. A major limitation of the AMDP formalism, however, is pessimistic regret analysis results in the sense that although the cost function can change from one episode to the next, the evolution in many settings is not adversarial. To address this, we introduce and study a new variant of AMDP, which aims to minimize regret while utilizing a set of cost predictors. For this setting, we develop a new policy search method that achieves a sublinear optimistic regret with high probability, that is a regret bound which gracefully degrades with the estimation power of the cost predictors. Establishing such optimistic regret bounds is nontrivial given that (i) as we demonstrate, the existing importance-weighted cost estimators cannot establish optimistic bounds, and (ii) the feedback model of AMDP is different (and more realistic) than the existing optimistic online learning works. Our result, in particular, hinges upon developing a novel optimistically biased cost estimator that leverages cost predictors and enables a high-probability regret analysis without imposing restrictive assumptions. We further discuss practical extensions of the proposed scheme and demonstrate its efficacy numerically.

link: http://arxiv.org/abs/2405.02188v1

Non-Destructive Peat Analysis using Hyperspectral Imaging and Machine Learning

Yijun Yan, Jinchang Ren, Barry Harrison, Oliver Lewis, Yinhe Li, Ping Ma

Peat, a crucial component in whisky production, imparts distinctive and irreplaceable flavours to the final product. However, the extraction of peat disrupts ancient ecosystems and releases significant amounts of carbon, contributing to climate change. This paper aims to address this issue by conducting a feasibility study on enhancing peat use efficiency in whisky manufacturing through non-destructive analysis using hyperspectral imaging. Results show that shot-wave infrared (SWIR) data is more effective for analyzing peat samples and predicting total phenol levels, with accuracies up to 99.81%.

link: http://arxiv.org/abs/2405.02191v1

Impact of emoji exclusion on the performance of Arabic sarcasm detection models

Ghalyah H. Aleryani, Wael Deabes, Khaled Albishre, Alaa E. Abdel-Hakim

The complex challenge of detecting sarcasm in Arabic speech on social media is increased by the language diversity and the nature of sarcastic expressions. There is a significant gap in the capability of existing models to effectively interpret sarcasm in Arabic, which mandates the necessity for more sophisticated and precise detection methods. In this paper, we investigate the impact of a fundamental preprocessing component on sarcasm speech detection. While emojis play a crucial role in mitigating the absence effect of body language and facial expressions in modern communication, their impact on automated text analysis, particularly in sarcasm detection, remains underexplored. We investigate the impact of emoji exclusion from datasets on the performance of

sarcasm detection models in social media content for Arabic as a vocabulary-super rich language. This investigation includes the adaptation and enhancement of AraBERT pre-training models, specifically by excluding emojis, to improve sarcasm detection capabilities. We use AraBERT pre-training to refine the specified models, demonstrating that the removal of emojis can significantly boost the accuracy of sarcasm detection. This approach facilitates a more refined interpretation of language, eliminating the potential confusion introduced by non-textual elements. The evaluated AraBERT models, through the focused strategy of emoji removal, adeptly navigate the complexities of Arabic sarcasm. This study establishes new benchmarks in Arabic natural language processing and presents valuable insights for social media platforms.

link: http://arxiv.org/abs/2405.02195v1