Fri 2024.03.08

Joint Sparsity Pattern Learning Based Channel Estimation for Massive MIMO-OTFS Systems

Kuo Meng, Shaoshi Yang, Xiao-Yang Wang, Yan Bu, Yurong Tang, Jianhua Zhang, Lajos Hanzo

We propose a channel estimation scheme based on joint sparsity pattern learning (JSPL) for massive multi-input multi-output (MIMO) orthogonal time-frequency-space (OTFS) modulation aided systems. By exploiting the potential joint sparsity of the delay-Doppler-angle (DDA) domain channel, the channel estimation problem is transformed into a sparse recovery problem. To solve it, we first apply the spike and slab prior model to iteratively estimate the support set of the channel matrix, and a higher-accuracy parameter update rule relying on the identified support set is introduced into the iteration. Then the specific values of the channel elements corresponding to the support set are estimated by the orthogonal matching pursuit (OMP) method. Both our simulation results and analysis demonstrate that the proposed JSPL channel estimation scheme achieves an improved performance over the representative state-of-the-art baseline schemes, despite its reduced pilot overhead.

link: http://arxiv.org/abs/2403.03771v1

AcceleratedLiNGAM: Learning Causal DAGs at the speed of GPUs

Victor Akinwande, J. Zico Kolter

Existing causal discovery methods based on combinatorial optimization or search are slow, prohibiting their application on large-scale datasets. In response, more recent methods attempt to address this limitation by formulating causal discovery as structure learning with continuous optimization but such approaches thus far provide no statistical guarantees. In this paper, we show that by efficiently parallelizing existing causal discovery methods, we can in fact scale them to thousands of dimensions, making them practical for substantially larger-scale problems. In particular, we parallelize the LiNGAM method, which is quadratic in the number of variables, obtaining up to a 32-fold speed-up on benchmark datasets when compared with existing sequential implementations. Specifically, we focus on the causal ordering subprocedure in DirectLiNGAM and implement GPU kernels to accelerate it. This allows us to apply DirectLiNGAM to causal inference on large-scale gene expression data with genetic interventions yielding competitive results compared with specialized continuous optimization methods, and Var-LiNGAM for causal discovery on U.S. stock data.

link: http://arxiv.org/abs/2403.03772v1

Verified Training for Counterfactual Explanation Robustness under Data Shift

Anna P. Meyer, Yuhao Zhang, Aws Albarghouthi, Loris D'Antoni

Counterfactual explanations (CEs) enhance the interpretability of machine learning models by describing what changes to an input are necessary to change its prediction to a desired class. These explanations are commonly used to guide users' actions, e.g., by describing how a user whose loan application was denied can be approved for a loan in the future. Existing approaches generate CEs by focusing on a single, fixed model, and do not provide any formal guarantees on the CEs' future validity. When models are updated periodically to account for data shift, if the generated CEs are not robust to the shifts, users' actions may no longer have the desired impacts on their predictions. This paper introduces VeriTraCER, an approach that jointly trains a classifier and an explainer to explicitly consider the robustness of the generated CEs to small model shifts. VeriTraCER optimizes over a carefully designed loss function that ensures the verifiable robustness of CEs to local model updates, thus providing deterministic guarantees to CE validity. Our empirical evaluation demonstrates that VeriTraCER generates CEs that (1) are verifiably robust to small model updates and (2) display competitive robustness to state-of-the-art approaches in handling empirical model updates including random initialization, leave-one-out, and distribution shifts.

link: http://arxiv.org/abs/2403.03773v1

ENOT: Expectile Regularization for Fast and Accurate Training of Neural Optimal Transport

Nazar Buzun, Maksim Bobrin, Dmitry V. Dylov

We present a new extension for Neural Optimal Transport (NOT) training procedure, capable of accurately and efficiently estimating optimal transportation plan via specific regularisation on conjugate potentials. The main bottleneck of existing NOT solvers is associated with the procedure of finding a near-exact approximation of the conjugate operator (i.e., the c-transform), which is done either by optimizing over maximin objectives or by the computationally-intensive fine-tuning of the initial approximated prediction. We resolve both issues by proposing a new, theoretically justified loss in the form of expectile regularization that enforces binding conditions on the learning dual potentials. Such a regularization provides the upper bound estimation over the distribution of possible conjugate potentials and makes the learning stable, eliminating the need for additional extensive finetuning. We formally justify the efficiency of our method, called Expectile-Regularised Neural Optimal Transport (ENOT). ENOT outperforms previous state-of-the-art approaches on the Wasserstein-2 benchmark tasks by a large margin (up to a 3-fold improvement in quality and up to a 10-fold improvement in runtime).

link: http://arxiv.org/abs/2403.03777v1

Neural Architecture Search using Particle Swarm and Ant Colony Optimization

Séamus Lankford, Diarmuid Grimes

Neural network models have a number of hyperparameters that must be chosen along with their architecture. This can be a heavy burden on a novice user, choosing which architecture and what values to assign to parameters. In most cases, default hyperparameters and architectures are used. Significant improvements to model accuracy can be achieved through the evaluation of multiple architectures. A process known as Neural Architecture Search (NAS) may be applied to automatically evaluate a large number of such architectures. A system integrating open source tools for Neural Architecture Search (OpenNAS), in the classification of images, has been developed as part of this research. OpenNAS takes any dataset of grayscale, or RBG images, and generates Convolutional Neural Network (CNN) architectures based on a range of metaheuristics using either an AutoKeras, a transfer learning or a Swarm Intelligence (SI) approach. Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) are used as the SI algorithms. Furthermore, models developed through such metaheuristics may be combined using stacking ensembles. In the context of this paper, we focus on training and optimizing CNNs using the Swarm Intelligence (SI) components of OpenNAS. Two major types of SI algorithms, namely PSO and ACO, are compared to see which is more effective in generating higher model accuracies. It is shown, with our experimental design, that the PSO algorithm performs better than ACO. The performance improvement of PSO is most notable with a more complex dataset. As a baseline, the performance of fine-tuned pre-trained models is also evaluated.

link: http://arxiv.org/abs/2403.03781v1

A machine learning workflow to address credit default prediction

Rambod Rahmani, Marco Parola, Mario G. C. A. Cimino

Due to the recent increase in interest in Financial Technology (FinTech), applications like credit default prediction (CDP) are gaining significant industrial and academic attention. In this regard, CDP plays a crucial role in assessing the creditworthiness of individuals and businesses, enabling lenders to make informed decisions regarding loan approvals and risk management. In this paper, we propose a workflow-based approach to improve CDP, which refers to the task of assessing the probability that a borrower will default on his or her credit obligations. The workflow consists of multiple steps, each designed to leverage the strengths of different techniques featured in machine learning pipelines and, thus best solve the CDP task. We employ a comprehensive and systematic approach starting with data preprocessing using Weight of Evidence encoding, a technique that

ensures in a single-shot data scaling by removing outliers, handling missing values, and making data uniform for models working with different data types. Next, we train several families of learning models, introducing ensemble techniques to build more robust models and hyperparameter optimization via multi-objective genetic algorithms to consider both predictive accuracy and financial aspects. Our research aims at contributing to the FinTech industry in providing a tool to move toward more accurate and reliable credit risk assessment, benefiting both lenders and borrowers.

link: http://arxiv.org/abs/2403.03785v1

PPTC-R benchmark: Towards Evaluating the Robustness of Large Language Models for PowerPoint Task Completion

Zekai Zhang, Yiduo Guo, Yaobo Liang, Dongyan Zhao, Nan Duan

The growing dependence on Large Language Models (LLMs) for finishing user instructions necessitates a comprehensive understanding of their robustness to complex task completion in real-world situations. To address this critical need, we propose the PowerPoint Task Completion Robustness benchmark (PPTC-R) to measure LLMs' robustness to the user PPT task instruction and software version. Specifically, we construct adversarial user instructions by attacking user instructions at sentence, semantic, and multi-language levels. To assess the robustness of Language Models to software versions, we vary the number of provided APIs to simulate both the newest version and earlier version settings. Subsequently, we test 3 closed-source and 4 open-source LLMs using a benchmark that incorporates these robustness settings, aiming to evaluate how deviations impact LLMs' API calls for task completion. We find that GPT-4 exhibits the highest performance and strong robustness in our benchmark, particularly in the version update and the multilingual settings. However, we find that all LLMs lose their robustness when confronted with multiple challenges (e.g., multi-turn) simultaneously, leading to significant performance drops. We further analyze the robustness behavior and error reasons of LLMs in our benchmark, which provide valuable insights for researchers to understand the LLM's robustness in task completion and develop more robust LLMs and agents. We release the code and data at \url{https://github.com/ZekaiGalaxy/PPTCR}.

link: http://arxiv.org/abs/2403.03788v1

Popeye: A Unified Visual-Language Model for Multi-Source Ship Detection from Remote Sensing Imagery

Wei Zhang, Miaoxin Cai, Tong Zhang, Guoqiang Lei, Yin Zhuang, Xuerui Mao

Ship detection needs to identify ship locations from remote sensing (RS) scenes. However, due to different imaging payloads, various appearances of ships, and complicated background interference from the bird's eye view, it is difficult to set up a unified paradigm for achieving multi-source ship detection. Therefore, in this article, considering that the large language models (LLMs) emerge the powerful generalization ability, a novel unified visual-language model called Popeye is proposed for multi-source ship detection from RS imagery. First, to bridge the interpretation gap between multi-source images for ship detection, a novel image-instruction-answer way is designed to integrate the various ship detection ways (e.g., horizontal bounding box (HBB), oriented bounding box (OBB)) into a unified labeling paradigm. Then, in view of this, a cross-modal image interpretation method is developed for the proposed Popeye to enhance interactive comprehension ability between visual and language content, which can be easily migrated into any multi-source ship detection task. Subsequently, owing to objective domain differences, a knowledge adaption mechanism is designed to adapt the pre-trained visual-language knowledge from the nature scene into the RS domain for multi-source ship detection. In addition, the segment anything model (SAM) is also seamlessly integrated into the proposed Popeye to achieve pixel-level ship segmentation without additional training costs. Finally, extensive experiments are conducted on the newly constructed instruction dataset named MMShip, and the results indicate that the proposed Popeye outperforms current specialist, open-vocabulary, and other visual-language models for zero-shot multi-source ship detection.

link: http://arxiv.org/abs/2403.03790v1

KG-TREAT: Pre-training for Treatment Effect Estimation by Synergizing Patient Data with Knowledge Graphs

Ruoqi Liu, Lingfei Wu, Ping Zhang

Treatment effect estimation (TEE) is the task of determining the impact of various treatments on patient outcomes. Current TEE methods fall short due to reliance on limited labeled data and challenges posed by sparse and high-dimensional observational patient data. To address the challenges, we introduce a novel pre-training and fine-tuning framework, KG-TREAT, which synergizes large-scale observational patient data with biomedical knowledge graphs (KGs) to enhance TEE. Unlike previous approaches, KG-TREAT constructs dual-focus KGs and integrates a deep bi-level attention synergy method for in-depth information fusion, enabling distinct encoding of treatment-covariate and outcome-covariate relationships. KG-TREAT also incorporates two pre-training tasks to ensure a thorough grounding and contextualization of patient data and KGs. Evaluation on four downstream TEE tasks shows KG-TREAT's superiority over existing methods, with an average improvement of 7% in Area under the ROC Curve (AUC) and 9% in Influence Function-based Precision of Estimating Heterogeneous Effects (IF-PEHE). The effectiveness of our estimated treatment effects is further affirmed by alignment with established randomized clinical trial findings.

link: http://arxiv.org/abs/2403.03791v1

Neural Exec: Learning (and Learning from) Execution Triggers for Prompt Injection Attacks

Dario Pasquini, Martin Strohmeier, Carmela Troncoso

We introduce a new family of prompt injection attacks, termed Neural Exec. Unlike known attacks that rely on handcrafted strings (e.g., "Ignore previous instructions and..."), we show that it is possible to conceptualize the creation of execution triggers as a differentiable search problem and use learning-based methods to autonomously generate them. Our results demonstrate that a motivated adversary can forge triggers that are not only drastically more effective than current handcrafted ones but also exhibit inherent flexibility in shape, properties, and functionality. In this direction, we show that an attacker can design and generate Neural Execs capable of persisting through multi-stage preprocessing pipelines, such as in the case of Retrieval-Augmented Generation (RAG)-based applications. More critically, our findings show that attackers can produce triggers that deviate markedly in form and shape from any known attack, sidestepping existing blacklist-based detection and sanitation approaches.

link: http://arxiv.org/abs/2403.03792v1

A Precision Drone Landing System using Visual and IR Fiducial Markers and a Multi-Payload Camera

Joshua Springer, Gylfi Þór Guðmundsson, Marcel Kyas

We propose a method for autonomous precision drone landing with fiducial markers and a gimbal-mounted, multi-payload camera with wide-angle, zoom, and IR sensors. The method has minimal data requirements; it depends primarily on the direction from the drone to the landing pad, enabling it to switch dynamically between the camera's different sensors and zoom factors, and minimizing auxiliary sensor requirements. It eliminates the need for data such as altitude above ground level, straight-line distance to the landing pad, fiducial marker size, and 6 DoF marker pose (of which the orientation is problematic). We leverage the zoom and wide-angle cameras, as well as visual April Tag fiducial markers to conduct successful precision landings from much longer distances than in previous work (168m horizontal distance, 102m altitude). We use two types of April Tags in the IR spectrum - active and passive - for precision landing both at daytime and nighttime, instead of simple IR beacons used in most previous work. The active IR landing pad is heated; the novel, passive one is unpowered, at ambient temperature, and depends on its high reflectivity and an IR differential between the ground and the sky. Finally, we propose a high-level control policy to manage initial search for the landing pad and subsequent searches if it is lost - not

addressed in previous work. The method demonstrates successful landings with the landing skids at least touching the landing pad, achieving an average error of 0.19m. It also demonstrates successful recovery and landing when the landing pad is temporarily obscured.

link: http://arxiv.org/abs/2403.03806v1

Confidence-Aware Decision-Making and Control for Tool Selection

Ajith Anil Meera, Pablo Lanillos

Self-reflecting about our performance (e.g., how confident we are) before doing a task is essential for decision making, such as selecting the most suitable tool or choosing the best route to drive. While this form of awareness -- thinking about our performance or metacognitive performance -- is well-known in humans, robots still lack this cognitive ability. This reflective monitoring can enhance their embodied decision power, robustness and safety. Here, we take a step in this direction by introducing a mathematical framework that allows robots to use their control self-confidence to make better-informed decisions. We derive a mathematical closed-form expression for control confidence for dynamic systems (i.e., the posterior inverse covariance of the control action). This control confidence seamlessly integrates within an objective function for decision making, that balances the: i) performance for task completion, ii) control effort, and iii) self-confidence. To evaluate our theoretical account, we framed the decision-making within the tool selection problem, where the agent has to select the best robot arm for a particular control task. The statistical analysis of the numerical simulations with randomized 2DOF arms shows that using control confidence during tool selection improves both real task performance, and the reliability of the tool for performance under unmodelled perturbations (e.g., external forces). Furthermore, our results indicate that control confidence is an early indicator of performance and thus, it can be used as a heuristic for making decisions when computation power is restricted or decision-making is intractable. Overall, we show the advantages of using confidence-aware decision-making and control scheme for dynamic systems.

link: http://arxiv.org/abs/2403.03808v1

Incentivized Learning in Principal-Agent Bandit Games

Antoine Scheid, Daniil Tiapkin, Etienne Boursier, Aymeric Capitaine, El Mahdi El Mhamdi, Eric Moulines, Michael I. Jordan, Alain Durmus

This work considers a repeated principal-agent bandit game, where the principal can only interact with her environment through the agent. The principal and the agent have misaligned objectives and the choice of action is only left to the agent. However, the principal can influence the agent's decisions by offering incentives which add up to his rewards. The principal aims to iteratively learn an incentive policy to maximize her own total utility. This framework extends usual bandit problems and is motivated by several practical applications, such as healthcare or ecological taxation, where traditionally used mechanism design theories often overlook the learning aspect of the problem. We present nearly optimal (with respect to a horizon T) learning algorithms for the principal's regret in both multi-armed and linear contextual settings. Finally, we support our theoretical guarantees through numerical experiments.

link: http://arxiv.org/abs/2403.03811v1

ProbSAINT: Probabilistic Tabular Regression for Used Car Pricing

Kiran Madhusudhanan, Gunnar Behrens, Maximilian Stubbemann, Lars Schmidt-Thieme

Used car pricing is a critical aspect of the automotive industry, influenced by many economic factors and market dynamics. With the recent surge in online marketplaces and increased demand for used cars, accurate pricing would benefit both buyers and sellers by ensuring fair transactions. However, the transition towards automated pricing algorithms using machine learning necessitates the comprehension of model uncertainties, specifically the ability to flag predictions that the model is unsure about. Although recent literature proposes the use of boosting algorithms or nearest neighbor-based approaches for swift and precise price predictions, encapsulating model uncertainties with such algorithms presents a complex challenge. We introduce ProbSAINT, a

model that offers a principled approach for uncertainty quantification of its price predictions, along with accurate point predictions that are comparable to state-of-the-art boosting techniques. Furthermore, acknowledging that the business prefers pricing used cars based on the number of days the vehicle was listed for sale, we show how ProbSAINT can be used as a dynamic forecasting model for predicting price probabilities for different expected offer duration. Our experiments further indicate that ProbSAINT is especially accurate on instances where it is highly certain. This proves the applicability of its probabilistic predictions in real-world scenarios where trustworthiness is crucial.

link: http://arxiv.org/abs/2403.03812v1

Evaluating the Elementary Multilingual Capabilities of Large Language Models with MultiQ

Carolin Holtermann, Paul Röttger, Timm Dill, Anne Lauscher

Large language models (LLMs) need to serve everyone, including a global majority of non-English speakers. However, most LLMs today, and open LLMs in particular, are often intended for use in just English (e.g. Llama2, Mistral) or a small handful of high-resource languages (e.g. Mixtral, Qwen). Recent research shows that, despite limits in their intended use, people prompt LLMs in many different languages. Therefore, in this paper, we investigate the basic multilingual capabilities of state-of-the-art open LLMs beyond their intended use. For this purpose, we introduce MultiQ, a new silver standard benchmark for basic open-ended question answering with 27.4k test questions across a typologically diverse set of 137 languages. With MultiQ, we evaluate language fidelity, i.e.\ whether models respond in the prompted language, and question answering accuracy. All LLMs we test respond faithfully and/or accurately for at least some languages beyond their intended use. Most models are more accurate when they respond faithfully. However, differences across models are large, and there is a long tail of languages where models are neither accurate nor faithful. We explore differences in tokenization as a potential explanation for our findings, identifying possible correlations that warrant further investigation.

link: http://arxiv.org/abs/2403.03814v1