Tue 2024.03.26

Using Large Language Models for OntoClean-based Ontology Refinement

Yihang Zhao, Neil Vetter, Kaveh Aryan

This paper explores the integration of Large Language Models (LLMs) such as GPT-3.5 and GPT-4 into the ontology refinement process, specifically focusing on the OntoClean methodology. OntoClean, critical for assessing the metaphysical quality of ontologies, involves a two-step process of assigning meta-properties to classes and verifying a set of constraints. Manually conducting the first step proves difficult in practice, due to the need for philosophical expertise and lack of consensus among ontologists. By employing LLMs with two prompting strategies, the study demonstrates that high accuracy in the labelling process can be achieved. The findings suggest the potential for LLMs to enhance ontology refinement, proposing the development of plugin software for ontology tools to facilitate this integration.

link: http://arxiv.org/abs/2403.15864v1

RAAMove: A Corpus for Analyzing Moves in Research Article Abstracts

Hongzheng Li, Ruojin Wang, Ge Shi, Xing Lv, Lei Lei, Chong Feng, Fang Liu, Jinkun Lin, Yangguang Mei, Lingnan Xu

Move structures have been studied in English for Specific Purposes (ESP) and English for Academic Purposes (EAP) for decades. However, there are few move annotation corpora for Research Article (RA) abstracts. In this paper, we introduce RAAMove, a comprehensive multi-domain corpus dedicated to the annotation of move structures in RA abstracts. The primary objective of RAAMove is to facilitate move analysis and automatic move identification. This paper provides a thorough discussion of the corpus construction process, including the scheme, data collection, annotation guidelines, and annotation procedures. The corpus is constructed through two stages: initially, expert annotators manually annotate high-quality data; subsequently, based on the human-annotated data, a BERT-based model is employed for automatic annotation with the help of experts' modification. The result is a large-scale and high-quality corpus comprising 33,988 annotated instances. We also conduct preliminary move identification experiments using the BERT-based model to verify the effectiveness of the proposed corpus and model. The annotated corpus is available for academic research purposes and can serve as essential resources for move analysis, English language teaching and writing, as well as move/discourse-related tasks in Natural Language Processing (NLP).

link: http://arxiv.org/abs/2403.15872v1

LAMPER: LanguAge Model and Prompt EngineeRing for zero-shot time series classification

Zhicheng Du, Zhaotian Xie, Yan Tong, Peiwu Qin

This study constructs the LanguAge Model with Prompt EngineeRing (LAMPER) framework, designed to systematically evaluate the adaptability of pre-trained language models (PLMs) in accommodating diverse prompts and their integration in zero-shot time series (TS) classification. We deploy LAMPER in experimental assessments using 128 univariate TS datasets sourced from the UCR archive. Our findings indicate that the feature representation capacity of LAMPER is influenced by the maximum input token threshold imposed by PLMs.

link: http://arxiv.org/abs/2403.15875v1

Cognitive resilience: Unraveling the proficiency of image-captioning models to interpret masked visual content

Zhicheng Du, Zhaotian Xie, Huazhang Ying, Likun Zhang, Peiwu Qin

This study explores the ability of Image Captioning (IC) models to decode masked visual content sourced from diverse datasets. Our findings reveal the IC model's capability to generate captions from masked images, closely resembling the original content. Notably, even in the presence of masks, the model adeptly crafts descriptive textual information that goes beyond what is observable in the original image-generated captions. While the decoding performance of the IC model experiences a decline with an increase in the masked region's area, the model still performs well when important regions of the image are not masked at high coverage.

link: http://arxiv.org/abs/2403.15876v1

Diffusion-based Aesthetic QR Code Generation via Scanning-Robust Perceptual Guidance

Jia-Wei Liao, Winston Wang, Tzu-Sian Wang, Li-Xuan Peng, Cheng-Fu Chou, Jun-Cheng Chen

QR codes, prevalent in daily applications, lack visual appeal due to their conventional black-and-white design. Integrating aesthetics while maintaining scannability poses a challenge. In this paper, we introduce a novel diffusion-model-based aesthetic QR code generation pipeline, utilizing pre-trained ControlNet and guided iterative refinement via a novel classifier guidance (SRG) based on the proposed Scanning-Robust Loss (SRL) tailored with QR code mechanisms, which ensures both aesthetics and scannability. To further improve the scannability while preserving aesthetics, we propose a two-stage pipeline with Scanning-Robust Perceptual Guidance (SRPG). Moreover, we can further enhance the scannability of the generated QR code by post-processing it through the proposed Scanning-Robust Projected Gradient Descent (SRPGD) post-processing technique based on SRL with proven convergence. With extensive quantitative, qualitative, and subjective experiments, the results demonstrate that the proposed approach can generate diverse aesthetic QR codes with flexibility in detail. In addition, our pipelines outperforming existing models in terms of Scanning Success Rate (SSR) 86.67% (+40%) with comparable aesthetic scores. The pipeline combined with SRPGD further achieves 96.67% (+50%). Our code will be available https://github.com/jwliao1209/DiffQRCode.

link: http://arxiv.org/abs/2403.15878v1

TrustSQL: A Reliability Benchmark for Text-to-SQL Models with Diverse Unanswerable Questions

Gyubok Lee, Woosog Chay, Seonhee Cho, Edward Choi

Recent advances in large language models (LLMs) have led to significant improvements in translating natural language questions into SQL queries. While achieving high accuracy in SQL generation is crucial, little is known about the extent to which these text-to-SQL models can reliably handle diverse types of questions encountered during real-world deployment, including unanswerable ones. To explore this aspect, we present TrustSQL, a new benchmark designed to assess the reliability of text-to-SQL models in both single-database and cross-database settings. The benchmark tasks models with providing one of two outcomes: 1) SQL prediction; or 2) abstention from making a prediction, either when there is a potential error in the generated SQL or when faced with unanswerable questions. For model evaluation, we explore various modeling approaches specifically designed for this task. These include: 1) optimizing separate models for answerability detection, SQL generation, and error detection, which are then integrated into a single pipeline; and 2) developing a unified approach that optimizes a single model to address the proposed task. Experimental results using our new reliability score show that addressing this challenge involves many different areas of research and opens new avenues for model development. Nonetheless, none of the methods surpass the reliability performance of the naive baseline, which abstains from answering all questions.

link: http://arxiv.org/abs/2403.15879v1

Fast and Unified Path Gradient Estimators for Normalizing Flows

Lorenz Vaitl, Ludwig Winkler, Lorenz Richter, Pan Kessel

Recent work shows that path gradient estimators for normalizing flows have lower variance compared to standard estimators for variational inference, resulting in improved training. However, they are often prohibitively more expensive from a computational point of view and cannot be applied to maximum likelihood training in a scalable manner, which severely hinders their widespread adoption. In this work, we overcome these crucial limitations. Specifically, we propose a fast path gradient estimator which improves computational efficiency significantly and works for all normalizing flow architectures of practical relevance. We then show that this estimator can also be applied to maximum likelihood training for which it has a regularizing effect as it can take the form of a given target energy function into account. We empirically establish its superior performance and reduced variance for several natural sciences applications.

link: http://arxiv.org/abs/2403.15881v1

VLUE: A New Benchmark and Multi-task Knowledge Transfer Learning for Vietnamese Natural Language Understanding

Phong Nguyen-Thuan Do, Son Quoc Tran, Phu Gia Hoang, Kiet Van Nguyen, Ngan Luu-Thuy Nguyen

The success of Natural Language Understanding (NLU) benchmarks in various languages, such as GLUE for English, CLUE for Chinese, KLUE for Korean, and IndoNLU for Indonesian, has facilitated the evaluation of new NLU models across a wide range of tasks. To establish a standardized set of benchmarks for Vietnamese NLU, we introduce the first Vietnamese Language Understanding Evaluation (VLUE) benchmark. The VLUE benchmark encompasses five datasets covering different NLU tasks, including text classification, span extraction, and natural language understanding. To provide an insightful overview of the current state of Vietnamese NLU, we then evaluate seven state-of-the-art pre-trained models, including both multilingual and Vietnamese monolingual models, on our proposed VLUE benchmark. Furthermore, we present CafeBERT, a new state-of-the-art pre-trained model that achieves superior results across all tasks in the VLUE benchmark. Our model combines the proficiency of a multilingual pre-trained model with Vietnamese linguistic knowledge. CafeBERT is developed based on the XLM-RoBERTa model, with an additional pretraining step utilizing a significant amount of Vietnamese textual data to enhance its adaptation to the Vietnamese language. For the purpose of future research, CafeBERT is made publicly available for research purposes.

link: http://arxiv.org/abs/2403.15882v1

UPSS: a User-centric Private Storage System with its applications

Arastoo Bozorgi, Mahya Soleimani Jadidi, Jonathan Anderson

Strong confidentiality, integrity, user control, reliability and performance are critical requirements in privacy-sensitive applications. Such applications would benefit from a data storage and sharing infrastructure that provides these properties even in decentralized topologies with untrusted storage backends, but users today are forced to choose between systemic security properties and system reliability or performance. As an alternative to this status quo we present UPSS: the user-centric private sharing system, a cryptographic storage system that can be used as a conventional filesystem or as the foundation for security-sensitive applications such as redaction with integrity and private revision control. We demonstrate that both the security and performance properties of UPSS exceed that of existing cryptographic filesystems and that its performance is comparable to mature conventional filesystems - in some cases, even superior. Whether used directly via its Rust API or as a conventional filesystem, UPSS provides strong security and practical performance on untrusted storage.

link: http://dx.doi.org/10.5220/0012306600003648

STEntConv: Predicting Disagreement with Stance Detection and a Signed Graph Convolutional Network

Isabelle Lorge, Li Zhang, Xiaowen Dong, Janet B. Pierrehumbert

The rise of social media platforms has led to an increase in polarised online discussions, especially on political and socio-cultural topics such as elections and climate change. We propose a simple and novel unsupervised method to predict whether the authors of two posts agree or disagree, leveraging user stances about named entities obtained from their posts. We present STEntConv, a model which builds a graph of users and named entities weighted by stance and trains a Signed Graph Convolutional Network (SGCN) to detect disagreement between comment and reply posts. We run experiments and ablation studies and show that including this information improves disagreement detection performance on a dataset of Reddit posts for a range of controversial subreddit topics, without the need for platform-specific features or user history.

link: http://arxiv.org/abs/2403.15885v1

Leveraging Zero-Shot Prompting for Efficient Language Model Distillation

Lukas Vöge, Vincent Gurgul, Stefan Lessmann

This paper introduces a novel approach for efficiently distilling LLMs into smaller, application-specific models, significantly reducing operational costs and manual labor. Addressing the challenge of deploying computationally intensive LLMs in specific applications or edge devices, this technique utilizes LLMs' reasoning capabilities to generate labels and natural language rationales for unlabeled data. Our approach enhances both finetuning and distillation by employing a multi-task training framework where student models mimic these rationales alongside teacher predictions. Key contributions include the employment of zero-shot prompting to elicit teacher model rationales, reducing the necessity for handcrafted few-shot examples and lowering the overall token count required, which directly translates to cost savings given the pay-per-token billing model of major tech companies' LLM APIs. Additionally, the paper investigates the impact of explanation properties on distillation efficiency, demonstrating that minimal performance loss occurs even when rationale augmentation is not applied across the entire dataset, facilitating further reductions of tokens. This research marks a step toward the efficient training of task-specific models with minimal human intervention, offering substantial cost-savings while maintaining, or even enhancing, performance.

link: http://arxiv.org/abs/2403.15886v1

Human Motion Prediction under Unexpected Perturbation

Jiangbei Yue, Baiyi Li, Julien Pettré, Armin Seyfried, He Wang

We investigate a new task in human motion prediction, which is predicting motions under unexpected physical perturbation potentially involving multiple people. Compared with existing research, this task involves predicting less controlled, unpremeditated and pure reactive motions in response to external impact and how such motions can propagate through people. It brings new challenges such as data scarcity and predicting complex interactions. To this end, we propose a new method capitalizing differential physics and deep neural networks, leading to an explicit Latent Differential Physics (LDP) model. Through experiments, we demonstrate that LDP has high data efficiency, outstanding prediction accuracy, strong generalizability and good explainability. Since there is no similar research, a comprehensive comparison with 11 adapted baselines from several relevant domains is conducted, showing LDP outperforming existing research both quantitatively and qualitatively, improving prediction accuracy by as much as 70%, and demonstrating significantly stronger generalization.

link: http://arxiv.org/abs/2403.15891v1

MatchSeg: Towards Better Segmentation via Reference Image Matching

Ruiqiang Xiao, Jiayu Huo, Haotian Zheng, Yang Liu, Sebastien Ourselin, Rachel Sparks

Recently, automated medical image segmentation methods based on deep learning have achieved great success. However, they heavily rely on large annotated datasets, which are costly and time-consuming to acquire. Few-shot learning aims to overcome the need for annotated data by using a small labeled dataset, known as a support set, to guide predicting labels for new, unlabeled images, known as the query set. Inspired by this paradigm, we introduce MatchSeg, a novel

framework that enhances medical image segmentation through strategic reference image matching. We leverage contrastive language-image pre-training (CLIP) to select highly relevant samples when defining the support set. Additionally, we design a joint attention module to strengthen the interaction between support and query features, facilitating a more effective knowledge transfer between support and query sets. We validated our method across four public datasets. Experimental results demonstrate superior segmentation performance and powerful domain generalization ability of MatchSeg against existing methods for domain-specific and cross-domain segmentation tasks. Our code is made available at https://github.com/keeplearning-again/MatchSeg

link: http://arxiv.org/abs/2403.15901v1

Towards Low-Energy Adaptive Personalization for Resource-Constrained Devices

Yushan Huang, Josh Millar, Yuxuan Long, Yuchen Zhao, Hamed Hadaddi

The personalization of machine learning (ML) models to address data drift is a significant challenge in the context of Internet of Things (IoT) applications. Presently, most approaches focus on fine-tuning either the full base model or its last few layers to adapt to new data, while often neglecting energy costs. However, various types of data drift exist, and fine-tuning the full base model or the last few layers may not result in optimal performance in certain scenarios. We propose Target Block Fine-Tuning (TBFT), a low-energy adaptive personalization framework designed for resource-constrained devices. We categorize data drift and personalization into three types: input-level, feature-level, and output-level. For each type, we fine-tune different blocks of the model to achieve optimal performance with reduced energy costs. Specifically, input-, feature-, and output-level correspond to fine-tuning the front, middle, and rear blocks of the model. We evaluate TBFT on a ResNet model, three datasets, three different training sizes, and a Raspberry Pi. Compared with the $Block Avg$, where each block is fine-tuned individually and their performance improvements are averaged, TBFT exhibits an improvement in model accuracy by an average of 15.30% whilst saving 41.57% energy consumption on average compared with full fine-tuning.

link: http://arxiv.org/abs/2403.15905v1

Deep Gaussian Covariance Network with Trajectory Sampling for Data-Efficient Policy Search

Can Bogoclu, Robert Vosshall, Kevin Cremanns, Dirk Roos

Probabilistic world models increase data efficiency of model-based reinforcement learning (MBRL) by guiding the policy with their epistemic uncertainty to improve exploration and acquire new samples. Moreover, the uncertainty-aware learning procedures in probabilistic approaches lead to robust policies that are less sensitive to noisy observations compared to uncertainty unaware solutions. We propose to combine trajectory sampling and deep Gaussian covariance network (DGCN) for a data-efficient solution to MBRL problems in an optimal control setting. We compare trajectory sampling with density-based approximation for uncertainty propagation using three different probabilistic world models; Gaussian processes, Bayesian neural networks, and DGCNs. We provide empirical evidence using four different well-known test environments, that our method improves the sample-efficiency over other combinations of uncertainty propagation methods and probabilistic models. During our tests, we place particular emphasis on the robustness of the learned policies with respect to noisy initial states.

link: http://dx.doi.org/10.1109/ACDSA59508.2024.10467448