Mon 2024.04.22

Cross-cultural Inspiration Detection and Analysis in Real and LLM-generated Social Media Data

Oana Ignat, Gayathri Ganesh Lakshmy, Rada Mihalcea

Inspiration is linked to various positive outcomes, such as increased creativity, productivity, and happiness. Although inspiration has great potential, there has been limited effort toward identifying content that is inspiring, as opposed to just engaging or positive. Additionally, most research has concentrated on Western data, with little attention paid to other cultures. This work is the first to study cross-cultural inspiration through machine learning methods. We aim to identify and analyze real and AI-generated cross-cultural inspiring posts. To this end, we compile and make publicly available the InspAIred dataset, which consists of 2,000 real inspiring posts, 2,000 real non-inspiring posts, and 2,000 generated inspiring posts evenly distributed across India and the UK. The real posts are sourced from Reddit, while the generated posts are created using the GPT-4 model. Using this dataset, we conduct extensive computational linguistic analyses to (1) compare inspiring content across cultures, (2) compare AI-generated inspiring posts to real inspiring posts, and (3) determine if detection models can accurately distinguish between inspiring content across cultures and data sources.

link: http://arxiv.org/abs/2404.12933v1

AuDaLa is Turing Complete

Tom T. P. Franken, Thomas Neele

AuDaLa is a recently introduced programming language that follows the new data autonomous paradigm. In this paradigm, small pieces of data execute functions autonomously. Considering the paradigm and the design choices of AuDaLa, it is interesting to determine the expressiveness of the language and to create verification methods for it. In this paper, we take our first steps to such a verification method by implementing Turing machines in AuDaLa and proving that implementation correct. This also proves that AuDaLa is Turing complete.

link: http://arxiv.org/abs/2404.12934v1

MAiDE-up: Multilingual Deception Detection of GPT-generated Hotel Reviews

Oana Ignat, Xiaomeng Xu, Rada Mihalcea

Deceptive reviews are becoming increasingly common, especially given the increase in performance and the prevalence of LLMs. While work to date has addressed the development of models to differentiate between truthful and deceptive human reviews, much less is known about the distinction between real reviews and AI-authored fake reviews. Moreover, most of the research so far has focused primarily on English, with very little work dedicated to other languages. In this paper, we compile and make publicly available the MAiDE-up dataset, consisting of 10,000 real and 10,000 AI-generated fake hotel reviews, balanced across ten languages. Using this dataset, we conduct extensive linguistic analyses to (1) compare the AI fake hotel reviews to real hotel reviews, and (2) identify the factors that influence the deception detection model performance. We explore the effectiveness of several models for deception detection in hotel reviews across three main dimensions: sentiment, location, and language. We find that these dimensions influence how well we can detect AI-generated fake reviews.

link: http://arxiv.org/abs/2404.12938v1

Neural Flow Diffusion Models: Learnable Forward Process for Improved Diffusion Modelling

Grigory Bartosh, Dmitry Vetrov, Christian A. Naesseth

Conventional diffusion models typically relies on a fixed forward process, which implicitly defines complex marginal distributions over latent variables. This can often complicate the reverse process' task in learning generative trajectories, and results in costly inference for diffusion models. To address these limitations, we introduce Neural Flow Diffusion Models (NFDM), a novel framework that enhances diffusion models by supporting a broader range of forward processes beyond the fixed linear Gaussian. We also propose a novel parameterization technique for learning the forward process. Our framework provides an end-to-end, simulation-free optimization objective, effectively minimizing a variational upper bound on the negative log-likelihood. Experimental results demonstrate NFDM's strong performance, evidenced by state-of-the-art likelihood estimation. Furthermore, we investigate NFDM's capacity for learning generative dynamics with specific characteristics, such as deterministic straight lines trajectories. This exploration underscores NFDM's versatility and its potential for a wide range of applications.

link: http://arxiv.org/abs/2404.12940v1

Purposer: Putting Human Motion Generation in Context

Nicolas Ugrinovic, Thomas Lucas, Fabien Baradel, Philippe Weinzaepfel, Gregory Rogez, Francesc Moreno-Noguer

We present a novel method to generate human motion to populate 3D indoor scenes. It can be controlled with various combinations of conditioning signals such as a path in a scene, target poses, past motions, and scenes represented as 3D point clouds. State-of-the-art methods are either models specialized to one single setting, require vast amounts of high-quality and diverse training data, or are unconditional models that do not integrate scene or other contextual information. As a consequence, they have limited applicability and rely on costly training data. To address these limitations, we propose a new method ,dubbed Purposer, based on neural discrete representation learning. Our model is capable of exploiting, in a flexible manner, different types of information already present in open access large-scale datasets such as AMASS. First, we encode unconditional human motion into a discrete latent space. Second, an autoregressive generative model, conditioned with key contextual information, either with prompting or additive tokens, and trained for next-step prediction in this space, synthesizes sequences of latent indices. We further design a novel conditioning block to handle future conditioning information in such a causal model by using a network with two branches to compute separate stacks of features. In this manner, Purposer can generate realistic motion sequences in diverse test scenes. Through exhaustive evaluation, we demonstrate that our multi-contextual solution outperforms existing specialized approaches for specific contextual information, both in terms of quality and diversity. Our model is trained with short sequences, but a byproduct of being able to use various conditioning signals is that at test time different combinations can be used to chain short sequences together and generate long motions within a context scene.

link: http://arxiv.org/abs/2404.12942v1

Next Generation Loss Function for Image Classification

Shakhnaz Akhmedova, Nils Körber

Neural networks are trained by minimizing a loss function that defines the discrepancy between the predicted model output and the target value. The selection of the loss function is crucial to achieve task-specific behaviour and highly influences the capability of the model. A variety of loss functions have been proposed for a wide range of tasks affecting training and model performance. For classification tasks, the cross entropy is the de-facto standard and usually the first choice. Here, we try to experimentally challenge the well-known loss functions, including cross entropy (CE) loss, by utilizing the genetic programming (GP) approach, a population-based evolutionary algorithm. GP constructs loss functions from a set of operators and leaf nodes and these functions are repeatedly recombined and mutated to find an optimal structure. Experiments were carried out on different small-sized datasets CIFAR-10, CIFAR-100 and Fashion-MNIST using an Inception model. The 5 best functions found were evaluated for different model architectures on a set of standard datasets ranging from 2 to 102 classes and very different sizes. One function, denoted as Next Generation Loss (NGL), clearly stood out showing same or better performance for all tested datasets compared

to CE. To evaluate the NGL function on a large-scale dataset, we tested its performance on the Imagenet-1k dataset where it showed improved top-1 accuracy compared to models trained with identical settings and other losses. Finally, the NGL was trained on a segmentation downstream task for Pascal VOC 2012 and COCO-Stuff164k datasets improving the underlying model performance.

link: http://arxiv.org/abs/2404.12948v1

Low-Depth Spatial Tree Algorithms

Yves Baumann, Tal Ben-Nun, Maciej Besta, Lukas Gianinazzi, Torsten Hoefler, Piotr Luczynski

Contemporary accelerator designs exhibit a high degree of spatial localization, wherein two-dimensional physical distance determines communication costs between processing elements. This situation presents considerable algorithmic challenges, particularly when managing sparse data, a pivotal component in progressing data science. The spatial computer model quantifies communication locality by weighting processor communication costs by distance, introducing a term named energy. Moreover, it integrates depth, a widely-utilized metric, to promote high parallelism. We propose and analyze a framework for efficient spatial tree algorithms within the spatial computer model. Our primary method constructs a spatial tree layout that optimizes the locality of the neighbors in the compute grid. This approach thereby enables locality-optimized messaging within the tree. Our layout achieves a polynomial factor improvement in energy compared to utilizing a PRAM approach. Using this layout, we develop energy-efficient treefix sum and lowest common ancestor algorithms, which are both fundamental building blocks for other graph algorithms. With high probability, our algorithms exhibit near-linear energy and poly-logarithmic depth. Our contributions augment a growing body of work demonstrating that computations can have both high spatial locality and low depth. Moreover, our work constitutes an advancement in the spatial layout of irregular and sparse computations.

link: http://arxiv.org/abs/2404.12953v1

Towards Reliable Latent Knowledge Estimation in LLMs: In-Context Learning vs. Prompting Based Factual Knowledge Extraction

Qinyuan Wu, Mohammad Aflah Khan, Soumi Das, Vedant Nanda, Bishwamittra Ghosh, Camila Kolling, Till Speicher, Laurent Bindschaedler, Krishna P. Gummadi, Evimaria Terzi

We propose an approach for estimating the latent knowledge embedded inside large language models (LLMs). We leverage the in-context learning (ICL) abilities of LLMs to estimate the extent to which an LLM knows the facts stored in a knowledge base. Our knowledge estimator avoids reliability concerns with previous prompting-based methods, is both conceptually simpler and easier to apply, and we demonstrate that it can surface more of the latent knowledge embedded in LLMs. We also investigate how different design choices affect the performance of ICL-based knowledge estimation. Using the proposed estimator, we perform a large-scale evaluation of the factual knowledge of a variety of open source LLMs, like OPT, Pythia, Llama(2), Mistral, Gemma, etc. over a large set of relations and facts from the Wikidata knowledge base. We observe differences in the factual knowledge between different model families and models of different sizes, that some relations are consistently better known than others but that models differ in the precise facts they know, and differences in the knowledge of base models and their finetuned counterparts.

link: http://arxiv.org/abs/2404.12957v1

Improving Pediatric Pneumonia Diagnosis with Adult Chest X-ray Images Utilizing Contrastive Learning and Embedding Similarity

Mohammad Zunaed, Anwarul Hasan, Taufiq Hasan

Despite the advancement of deep learning-based computer-aided diagnosis (CAD) methods for pneumonia from adult chest x-ray (CXR) images, the performance of CAD methods applied to pediatric images remains suboptimal, mainly due to the lack of large-scale annotated pediatric imaging datasets. Establishing a proper framework to leverage existing adult large-scale CXR datasets can thus enhance pediatric pneumonia detection performance. In this paper, we propose a

three-branch parallel path learning-based framework that utilizes both adult and pediatric datasets to improve the performance of deep learning models on pediatric test datasets. The paths are trained with pediatric only, adult only, and both types of CXRs, respectively. Our proposed framework utilizes the multi-positive contrastive loss to cluster the classwise embeddings and the embedding similarity loss among these three parallel paths to make the classwise embeddings as close as possible to reduce the effect of domain shift. Experimental evaluations on open-access adult and pediatric CXR datasets show that the proposed method achieves a superior AUROC score of 0.8464 compared to 0.8348 obtained using the conventional approach of join training on both datasets. The proposed approach thus paves the way for generalized CAD models that are effective for both adult and pediatric age groups.

link: http://arxiv.org/abs/2404.12958v1

Eyes Can Deceive: Benchmarking Counterfactual Reasoning Abilities of Multi-modal Large Language Models

Yian Li, Wentao Tian, Yang Jiao, Jingjing Chen, Yu-Gang Jiang

Counterfactual reasoning, as a crucial manifestation of human intelligence, refers to making presuppositions based on established facts and extrapolating potential outcomes. Existing multimodal large language models (MLLMs) have exhibited impressive cognitive and reasoning capabilities, which have been examined across a wide range of Visual Question Answering (VQA) benchmarks. Nevertheless, how will existing MLLMs perform when faced with counterfactual questions? To answer this question, we first curate a novel \textbf{C}ounter\textbf{F}actual \textbf{M}ulti\textbf{M}odal reasoning benchmark, abbreviated as \textbf{CFMM}, to systematically assess the counterfactual reasoning capabilities of MLLMs. Our CFMM comprises six challenging tasks, each including hundreds of carefully human-labeled counterfactual questions, to evaluate MLLM's counterfactual reasoning capabilities across diverse aspects. Through experiments, interestingly, we find that existing MLLMs prefer to believe what they see, but ignore the counterfactual presuppositions presented in the question, thereby leading to inaccurate responses. Furthermore, we evaluate a wide range of prevalent MLLMs on our proposed CFMM. The significant gap between their performance on our CFMM and that on several VQA benchmarks indicates that there is still considerable room for improvement in existing MLLMs toward approaching human-level intelligence. On the other hand, through boosting MLLMs performances on our CFMM in the future, potential avenues toward developing MLLMs with advanced intelligence can be explored.

link: http://arxiv.org/abs/2404.12966v1

Scalable Data Assimilation with Message Passing

Oscar Key, So Takao, Daniel Giles, Marc Peter Deisenroth

Data assimilation is a core component of numerical weather prediction systems. The large quantity of data processed during assimilation requires the computation to be distributed across increasingly many compute nodes, yet existing approaches suffer from synchronisation overhead in this setting. In this paper, we exploit the formulation of data assimilation as a Bayesian inference problem and apply a message-passing algorithm to solve the spatial inference problem. Since message passing is inherently based on local computations, this approach lends itself to parallel and distributed computation. In combination with a GPU-accelerated implementation, we can scale the algorithm to very large grid sizes while retaining good accuracy and compute and memory requirements.

link: http://arxiv.org/abs/2404.12968v1

Disentangling ID and Modality Effects for Session-based Recommendation

Xiaokun Zhang, Bo Xu, Zhaochun Ren, Xiaochen Wang, Hongfei Lin, Fenglong Ma

Session-based recommendation aims to predict intents of anonymous users based on their limited behaviors. Modeling user behaviors involves two distinct rationales: co-occurrence patterns reflected by item IDs, and fine-grained preferences represented by item modalities (e.g., text and images). However, existing methods typically entangle these causes, leading to their failure in

achieving accurate and explainable recommendations. To this end, we propose a novel framework DIMO to disentangle the effects of ID and modality in the task. At the item level, we introduce a co-occurrence representation schema to explicitly incorporate cooccurrence patterns into ID representations. Simultaneously, DIMO aligns different modalities into a unified semantic space to represent them uniformly. At the session level, we present a multi-view self-supervised disentanglement, including proxy mechanism and counterfactual inference, to disentangle ID and modality effects without supervised signals. Leveraging these disentangled causes, DIMO provides recommendations via causal inference and further creates two templates for generating explanations. Extensive experiments on multiple real-world datasets demonstrate the consistent superiority of DIMO over existing methods. Further analysis also confirms DIMO's effectiveness in generating explanations.

link: http://dx.doi.org/10.1145/3626772.3657748

FineRec:Exploring Fine-grained Sequential Recommendation

Xiaokun Zhang, Bo Xu, Youlin Wu, Yuan Zhong, Hongfei Lin, Fenglong Ma

Sequential recommendation is dedicated to offering items of interest for users based on their history behaviors. The attribute-opinion pairs, expressed by users in their reviews for items, provide the potentials to capture user preferences and item characteristics at a fine-grained level. To this end, we propose a novel framework FineRec that explores the attribute-opinion pairs of reviews to finely handle sequential recommendation. Specifically, we utilize a large language model to extract attribute-opinion pairs from reviews. For each attribute, a unique attribute-specific user-opinion-item graph is created, where corresponding opinions serve as the edges linking heterogeneous user and item nodes. To tackle the diversity of opinions, we devise a diversity-aware convolution operation to aggregate information within the graphs, enabling attribute-specific user and item representation learning. Ultimately, we present an interaction-driven fusion mechanism to integrate attribute-specific user/item representations across all attributes for generating recommendations. Extensive experiments conducted on several realworld datasets demonstrate the superiority of our FineRec over existing state-of-the-art methods. Further analysis also verifies the effectiveness of our fine-grained manner in handling the task.

link: http://dx.doi.org/10.1145/3626772.3657761

TRNet: Two-level Refinement Network leveraging Speech Enhancement for Noise Robust Speech Emotion Recognition

Chengxin Chen, Pengyuan Zhang

One persistent challenge in Speech Emotion Recognition (SER) is the ubiquitous environmental noise, which frequently results in diminished SER performance in practical use. In this paper, we introduce a Two-level Refinement Network, dubbed TRNet, to address this challenge. Specifically, a pre-trained speech enhancement module is employed for front-end noise reduction and noise level estimation. Later, we utilize clean speech spectrograms and their corresponding deep representations as reference signals to refine the spectrogram distortion and representation shift of enhanced speech during model training. Experimental results validate that the proposed TRNet substantially increases the system's robustness in both matched and unmatched noisy environments, without compromising its performance in clean environments.

link: http://arxiv.org/abs/2404.12979v1

Private Agent-Based Modeling

Ayush Chopra, Arnau Quera-Bofarull, Nurullah Giray-Kuru, Michael Wooldridge, Ramesh Raskar

The practical utility of agent-based models in decision-making relies on their capacity to accurately replicate populations while seamlessly integrating real-world data streams. Yet, the incorporation of such data poses significant challenges due to privacy concerns. To address this issue, we introduce a paradigm for private agent-based modeling wherein the simulation, calibration, and analysis of agent-based models can be achieved without centralizing the agents attributes or interactions. The key insight is to leverage techniques from secure multi-party computation to

design protocols for decentralized computation in agent-based models. This ensures the confidentiality of the simulated agents without compromising on simulation accuracy. We showcase our protocols on a case study with an epidemiological simulation comprising over 150,000 agents. We believe this is a critical step towards deploying agent-based models to real-world applications.

link: http://arxiv.org/abs/2404.12983v1