Sat 2024.05.11

QFMTS: Generating Query-Focused Summaries over Multi-Table Inputs

Weijia Zhang, Vaishali Pal, Jia-Hong Huang, Evangelos Kanoulas, Maarten de Rijke

Table summarization is a crucial task aimed at condensing information from tabular data into concise and comprehensible textual summaries. However, existing approaches often fall short of adequately meeting users' information and quality requirements and tend to overlook the complexities of real-world queries. In this paper, we propose a novel method to address these limitations by introducing query-focused multi-table summarization. Our approach, which comprises a table serialization module, a summarization controller, and a large language model (LLM), utilizes textual queries and multiple tables to generate query-dependent table summaries tailored to users' information needs. To facilitate research in this area, we present a comprehensive dataset specifically tailored for this task, consisting of 4909 query-summary pairs, each associated with multiple tables. Through extensive experiments using our curated dataset, we demonstrate the effectiveness of our proposed method compared to baseline approaches. Our findings offer insights into the challenges of complex table reasoning for precise summarization, contributing to the advancement of research in query-focused multi-table summarization.

link: http://arxiv.org/abs/2405.05109v1

XAMPLER: Learning to Retrieve Cross-Lingual In-Context Examples

Peiqin Lin, André F. T. Martins, Hinrich Schütze

Recent studies have shown that leveraging off-the-shelf or fine-tuned retrievers, capable of retrieving high-quality in-context examples, significantly improves in-context learning of English. However, adapting these methods to other languages, especially low-resource ones, presents challenges due to the scarcity of available cross-lingual retrievers and annotated data. In this paper, we introduce XAMPLER: Cross-Lingual Example Retrieval, a method tailored to tackle the challenge of cross-lingual in-context learning using only annotated English data. XAMPLER first trains a retriever with positive/negative English samples, which are constructed based on the predictions of the multilingual large language model for in-context learning. Then, the trained retriever is directly employed to retrieve English examples as few-shot examples for in-context learning of target languages. Experiments on the massively multilingual text classification benchmark of SIB200 with 176 languages demonstrate that XAMPLER substantially improves the in-context learning performance across languages. Our code is available at https://github.com/cisnlp/XAMPLER.

link: http://arxiv.org/abs/2405.05116v1

Full Version: (De/Re)-Composition of Data-Parallel Computations via Multi-Dimensional Homomorphisms

Ari Rasch

We formally introduce a systematic (de/re)-composition approach, based on the algebraic formalism of "Multi-Dimensional Homomorphisms (MDHs)". Our approach is designed as general enough to be applicable to a wide range of data-parallel computations and for various kinds of target parallel architectures. To efficiently target the deep and complex memory and core hierarchies of contemporary architectures, we exploit our introduced (de/re)-composition approach for a correct-by-construction, parametrized cache blocking and parallelization strategy. We show that our approach is powerful enough to express, in the same formalism, the (de/re)-composition strategies of different classes of state-of-the-art approaches (scheduling-based, polyhedral, etc), and we demonstrate that the parameters of our strategies enable systematically generating code that can be fully automatically optimized (auto-tuned) for the particular target architecture and characteristics of the input and output data (e.g., their sizes and memory layouts). Particularly, our experiments confirm that via auto-tuning, we achieve higher performance than state-of-the-art approaches, including hand-optimized solutions provided by vendors (such as NVIDIA

cuBLAS/cuDNN and Intel oneMKL/oneDNN), on real-world data sets and for a variety of data-parallel computations, including: linear algebra routines, stencil and quantum chemistry computations, data mining algorithms, and computations that recently gained high attention due to their relevance for deep learning.

link: http://arxiv.org/abs/2405.05118v1

Web Intelligence Journal in perspective: an analysis of its two decades trajectory

Diogenes Ademir Domingos, Victor Emanuel Santos Moura, Antonio Fernando Lavareda Jacob Junior, Fabio Manoel Franca Lobato

The evolution of a thematic area undergoes various changes of perspective and adopts new theoretical approaches that arise from the interactions of the community and a wide range of social needs. The advent of digital technologies, such as social networks, underlines this factor by spreading knowledge and forging links between different communities. Web intelligence is now on the verge of raising questions that broaden the understanding of how artificial intelligence impacts the Web of People, Data, and Things, among other factors. To the best of our knowledge, there is no study that has conducted a longitudinal analysis of the evolution of this community. Thus, we investigate in this paper how Web intelligence has evolved in the last twenty years by carrying out a literature review and bibliometric analysis. Concerning the impact of this research study, increasing attention is devoted to determining which are the most influential papers in the community by referring to citation networks and discovering the most popular and pressing topics through a co-citation analysis and the keywords co-occurrence. The results obtained can guide the direction of new research projects in the area and update the scope and places of interest found in current trends and the relevant journals.

link: http://arxiv.org/abs/2405.05129v1

Multi-scale Bottleneck Transformer for Weakly Supervised Multimodal Violence Detection

Shengyang Sun, Xiaojin Gong

Weakly supervised multimodal violence detection aims to learn a violence detection model by leveraging multiple modalities such as RGB, optical flow, and audio, while only video-level annotations are available. In the pursuit of effective multimodal violence detection (MVD), information redundancy, modality imbalance, and modality asynchrony are identified as three key challenges. In this work, we propose a new weakly supervised MVD method that explicitly addresses these challenges. Specifically, we introduce a multi-scale bottleneck transformer (MSBT) based fusion module that employs a reduced number of bottleneck tokens to gradually condense information and fuse each pair of modalities and utilizes a bottleneck token-based weighting scheme to highlight more important fused features. Furthermore, we propose a temporal consistency contrast loss to semantically align pairwise fused features. Experiments on the largest-scale XD-Violence dataset demonstrate that the proposed method achieves state-of-the-art performance. Code is available at https://github.com/shengyangsun/MSBT.

link: http://arxiv.org/abs/2405.05130v1

Low-Distortion Clustering in Bounded Growth Graphs

Yi-Jun Chang, Varsha Dani, Thomas P. Hayes

The well-known clustering algorithm of Miller, Peng, and Xu (SPAA 2013) is useful for many applications, including low-diameter decomposition and low-energy distributed algorithms. One nice property of their clustering, shown in previous work by Chang, Dani, Hayes, and Pettie (PODC 2020), is that distances in the cluster graph are rescaled versions of distances in the original graph, up to an $O(\log n)$ distortion factor and rounding issues. Minimizing this distortion factor is important for efficiency in computing the clustering, as well as in other applications. We prove that there exist graphs for which an $\Omega((\log n)^{1/3})$ distortion factor is necessary for any clustering. We also consider a class of nice graphs which we call uniformly bounded independence graphs. These include, for example, paths, lattice graphs, and "dense" unit disk graphs. For these

graphs, we prove that clusterings of distortion $O(1)$ always exist, and moreover, we give new efficient distributed algorithms to construct them. This clustering is based on Voronoi cells centered at the vertices of a maximal independent set in a suitable power graph. Applications include low-energy simulation of distributed algorithms in the LOCAL, CONGEST, and RADIO-CONGEST models and efficient approximate solutions to distributed combinatorial optimization problems. We also investigate related lower bounds.

link: http://arxiv.org/abs/2405.05132v1

Identifying every building's function in large-scale urban areas with multi-modality remote-sensing data

Zhuohong Li, Wei He, Jiepan Li, Hongyan Zhang

Buildings, as fundamental man-made structures in urban environments, serve as crucial indicators for understanding various city function zones. Rapid urbanization has raised an urgent need for efficiently surveying building footprints and functions. In this study, we proposed a semi-supervised framework to identify every building's function in large-scale urban areas with multi-modality remote-sensing data. In detail, optical images, building height, and nighttime-light data are collected to describe the morphological attributes of buildings. Then, the area of interest (AOI) and building masks from the volunteered geographic information (VGI) data are collected to form sparsely labeled samples. Furthermore, the multi-modality data and weak labels are utilized to train a segmentation model with a semi-supervised strategy. Finally, results are evaluated by 20,000 validation points and statistical survey reports from the government. The evaluations reveal that the produced function maps achieve an OA of 82% and Kappa of 71% among 1,616,796 buildings in Shanghai, China. This study has the potential to support large-scale urban management and sustainable urban development. All collected data and produced maps are open access at https://github.com/LiZhuoHong/BuildingMap.

link: http://arxiv.org/abs/2405.05133v1

Dynamic Size Counting in the Population Protocol Model

Dominik Kaaser, Maximilian Lohmann

The population protocol model describes collections of distributed agents that interact in pairs to solve a common task. We consider a dynamic variant of this prominent model, where we assume that an adversary may change the population size at an arbitrary point in time. In this model we tackle the problem of counting the population size: in the dynamic size counting problem the goal is to design an algorithm that computes an approximation of $\log n$. This estimate can be used to turn static, non-uniform population protocols, i.e., protocols that depend on the population size n, into dynamic and loosely-stabilizing protocols. Our contributions in this paper are three-fold. Starting from an arbitrary initial configuration, we first prove that the agents converge quickly to a valid configuration where each agent has a constant-factor approximation of $\log n$, and once the agents reach such a valid configuration, they stay in it for a polynomial number of time steps. Second, we show how to use our protocol to define a uniform and loosely-stabilizing phase clock for the population protocol model. Finally, we support our theoretical findings by empirical simulations that show that our protocols work well in practice.

link: http://arxiv.org/abs/2405.05137v1

Hybrid Convolutional Neural Networks with Reliability Guarantee

Hans Dermot Doran, Suzana Veljanovska

Making AI safe and dependable requires the generation of dependable models and dependable execution of those models. We propose redundant execution as a well-known technique that can be used to ensure reliable execution of the AI model. This generic technique will extend the application scope of AI-accelerators that do not feature well-documented safety or dependability properties. Typical redundancy techniques incur at least double or triple the computational expense of the original. We adopt a co-design approach, integrating reliable model execution with non-reliable execution, focusing that additional computational expense only where it is strictly

necessary. We describe the design, implementation and some preliminary results of a hybrid CNN.

link: http://arxiv.org/abs/2405.05146v2

The Potential and Implications of Generative AI on HCI Education

Ahmed Kharrufa, Ian G Johnson

Generative AI (GAI) is impacting teaching and learning directly or indirectly across a range of subjects and disciplines. As educators, we need to understand the potential and limitations of AI in HCI education and ensure our graduating HCI students are aware of the potential and limitations of AI in HCI. In this paper, we report on the main pedagogical insights gained from the inclusion of generative AI into a 10 week undergraduate module. We designed the module to encourage student experimentation with GAI models as part of the design brief requirement and planned practical sessions and discussions. Our insights are based on replies to a survey sent out to the students after completing the module. Our key findings, for HCI educators, report on the use of AI as a persona for developing project ideas and creating resources for design, and AI as a mirror for reflecting students' understanding of key concepts and ideas and highlighting knowledge gaps. We also discuss potential pitfalls that should be considered and the need to assess students' literacies and assumptions of GAIs as pedagogical tools. Finally, we put forward the case for educators to take the opportunities GAI presents as an educational tool and be experimental, creative, and courageous in their practice. We end with a discussion of our findings in relation to the TPACK framework in HCI.

link: http://dx.doi.org/10.1145/3658619.3658627

Selective Classification Under Distribution Shifts

Hengyue Liang, Le Peng, Ju Sun

In selective classification (SC), a classifier abstains from making predictions that are likely to be wrong to avoid excessive errors. To deploy imperfect classifiers -- imperfect either due to intrinsic statistical noise of data or for robustness issue of the classifier or beyond -- in high-stakes scenarios, SC appears to be an attractive and necessary path to follow. Despite decades of research in SC, most previous SC methods still focus on the ideal statistical setting only, i.e., the data distribution at deployment is the same as that of training, although practical data can come from the wild. To bridge this gap, in this paper, we propose an SC framework that takes into account distribution shifts, termed generalized selective classification, that covers label-shifted (or out-of-distribution) and covariate-shifted samples, in addition to typical in-distribution samples, the first of its kind in the SC literature. We focus on non-training-based confidence-score functions for generalized SC on deep learning (DL) classifiers and propose two novel margin-based score functions. Through extensive analysis and experiments, we show that our proposed score functions are more effective and reliable than the existing ones for generalized SC on a variety of classification tasks and DL classifiers.

link: http://arxiv.org/abs/2405.05160v1

ProbRadarM3F: mmWave Radar based Human Skeletal Pose Estimation with Probability Map Guided Multi-Format Feature Fusion

Bing Zhu, Zixin He, Weiyi Xiong, Guanhua Ding, Jianan Liu, Tao Huang, Wei Chen, Wei Xiang

Millimetre wave (mmWave) radar is a non-intrusive privacy and relatively convenient and inexpensive device, which has been demonstrated to be applicable in place of RGB cameras in human indoor pose estimation tasks. However, mmWave radar relies on the collection of reflected signals from the target, and the radar signals containing information is difficult to be fully applied. This has been a long-standing hindrance to the improvement of pose estimation accuracy. To address this major challenge, this paper introduces a probability map guided multi-format feature fusion model, ProbRadarM3F. This is a novel radar feature extraction framework using a traditional FFT method in parallel with a probability map based positional encoding method. ProbRadarM3F fuses the traditional heatmap features and the positional features, then effectively achieves the estimation of 14 keypoints of the human body. Experimental evaluation on the HuPR dataset

proves the effectiveness of the model proposed in this paper, outperforming other methods experimented on this dataset with an AP of 69.9 %. The emphasis of our study is focusing on the position information that is not exploited before in radar singal. This provides direction to investigate other potential non-redundant information from mmWave rader.

link: http://arxiv.org/abs/2405.05164v1

Picking watermarks from noise (PWFN): an improved robust watermarking model against intensive distortions

Sijing Xie, Chengxin Zhao, Nan Sun, Wei Li, Hefei Ling

Digital watermarking is the process of embedding secret information by altering images in a way that is undetectable to the human eye. To increase the robustness of the model, many deep learning-based watermarking methods use the encoder-decoder architecture by adding different noises to the noise layer. The decoder then extracts the watermarked information from the distorted image. However, this method can only resist weak noise attacks. To improve the robustness of the algorithm against stronger noise, this paper proposes to introduce a denoise module between the noise layer and the decoder. The module is aimed at reducing noise and recovering some of the information lost during an attack. Additionally, the paper introduces the SE module to fuse the watermarking information pixel-wise and channel dimensions-wise, improving the encoder's efficiency. Experimental results show that our proposed method is comparable to existing models and outperforms state-of-the-art under different noise intensities. In addition, ablation experiments show the superiority of our proposed module.

link: http://arxiv.org/abs/2405.05170v1

Custom Gradient Estimators are Straight-Through Estimators in Disguise

Matt Schoenbauer, Daniele Moro, Lukasz Lew, Andrew Howard

Quantization-aware training comes with a fundamental challenge: the derivative of quantization functions such as rounding are zero almost everywhere and nonexistent elsewhere. Various differentiable approximations of quantization functions have been proposed to address this issue. In this paper, we prove that when the learning rate is sufficiently small, a large class of weight gradient estimators is equivalent with the straight through estimator (STE). Specifically, after swapping in the STE and adjusting both the weight initialization and the learning rate in SGD, the model will train in almost exactly the same way as it did with the original gradient estimator. Moreover, we show that for adaptive learning rate algorithms like Adam, the same result can be seen without any modifications to the weight initialization and learning rate. We experimentally show that these results hold for both a small convolutional model trained on the MNIST dataset and for a ResNet50 model trained on ImageNet.

link: http://arxiv.org/abs/2405.05171v1

A Survey on Occupancy Perception for Autonomous Driving: The Information Fusion Perspective

Huaiyuan Xu, Junliang Chen, Shiyu Meng, Yi Wang, Lap-Pui Chau

3D occupancy perception technology aims to observe and understand dense 3D environments for autonomous vehicles. Owing to its comprehensive perception capability, this technology is emerging as a trend in autonomous driving perception systems, and is attracting significant attention from both industry and academia. Similar to traditional bird's-eye view (BEV) perception, 3D occupancy perception has the nature of multi-source input and the necessity for information fusion. However, the difference is that it captures vertical structures that are ignored by 2D BEV. In this survey, we review the most recent works on 3D occupancy perception, and provide in-depth analyses of methodologies with various input modalities. Specifically, we summarize general network pipelines, highlight information fusion techniques, and discuss effective network training. We evaluate and analyze the occupancy perception performance of the state-of-the-art on the most popular datasets. Furthermore, challenges and future research directions are discussed. We hope this report will inspire the community and encourage more research work on 3D occupancy

perception. A comprehensive list of studies in this survey is available in an active repository that continuously collects the latest work: https://github.com/HuaiyuanXu/3D-Occupancy-Perception.

link: http://arxiv.org/abs/2405.05173v1