

**Mon 2024.07.15**

### **Turn-Level Empathy Prediction Using Psychological Indicators**

*Shaz Furniturewala, Kokil Jaidka*

**What's New:** This research introduces a novel turn-level empathy detection method that breaks down empathy into six psychological indicators.

**Technical Details:** The method involves text enrichment using a Large Language Model (LLM) and DeBERTA fine-tuning to improve Pearson Correlation Coefficient and F1 scores for empathy detection.

**Results:** The system developed based on this method ranked 7th at the CONV-turn track, showcasing the effectiveness of the proposed approach.

link: <http://arxiv.org/abs/2407.08607v1>

### **Tamil Language Computing: the Present and the Future**

*Kengatharaiyer Sarveswaran*

**What's New:** This research addresses the text processing elements of Language Computing and emphasizes the need for practical applications and advancements in technology for languages like Tamil.

**Technical Details:** The paper discusses the integration of linguistics, computer science, and cognitive psychology in language computing tasks like speech recognition and text summarization. It also focuses on recent deep learning advancements and the development of computational resources like data, dictionaries, and computational grammars for effective language processing.

**Results:** The research concludes that building practical applications for languages such as Tamil, digitizing historical texts, and enhancing digital usage are crucial for the comprehensive development of Tamil language processing and to improve global communication and access to digital services.

link: <http://arxiv.org/abs/2407.08618v1>

### **HACMan++: Spatially-Grounded Motion Primitives for Manipulation**

*Bowen Jiang, Yilin Wu, Wenxuan Zhou, Chris Paxton, David Held*

**What's New:** This research introduces spatially-grounded parameterized motion primitives in the method HACMan++ to improve policy generalization for robot manipulation.

**Technical Details:** The method proposes an action representation with three components: primitive type, grounding location, and execution details, creating a novel discrete-continuous action space. The approach enables robot agents to chain motion primitives and select parameters for long-horizon manipulation tasks.

**Results:** Through grounding primitives in the environment, the method achieves effective generalization across object shape and pose variations, outperforming existing methods in complex scenarios and demonstrating zero-shot sim-to-real transfer for real-world manipulation tasks.

link: <http://arxiv.org/abs/2407.08585v1>

### **On the Universal Truthfulness Hyperplane Inside LLMs**

*Junteng Liu, Shiqi Chen, Yu Cheng, Junxian He*

**What's New:** This research investigates the existence of a universal truthfulness hyperplane within large language models to distinguish factually correct and incorrect outputs.

**Technical Details:** The study scales up the number of training datasets to over 40 and evaluates the model's performance on cross-task, cross-domain, and in-domain generalization to determine

the impact of dataset diversity on the model's ability to adhere to facts.

**Results:** Increasing the diversity of training datasets significantly improves model performance in distinguishing factual outputs, suggesting that a universal truthfulness hyperplane may exist within the model and providing promising pathways for future research in improving factual awareness of language models.

link: <http://arxiv.org/abs/2407.08582v1>

### **A Review of Nine Physics Engines for Reinforcement Learning Research**

*Michael Kaup, Cornelius Wolff, Hyerim Hwang, Julius Mayer, Elia Bruni*

**What's New:** This research reviews popular simulation engines and frameworks used in reinforcement learning (RL) research, evaluating them based on popularity, features, quality, usability, and RL capabilities.

**Technical Details:** The study compared nine frameworks, including MuJoCo and Unity, highlighting their strengths and weaknesses for creating simulated physical environments for RL training setups.

**Results:** The findings designate MuJoCo as the top framework for performance and flexibility, while Unity stands out for ease of use despite scalability and fidelity limitations. The study emphasizes the need for improved usability and performance of simulation engines in RL research.

link: <http://arxiv.org/abs/2407.08590v1>

### **BiasPruner: Debaised Continual Learning for Medical Image Classification**

*Nourhan Bayasi, Jamil Fayyad, Alceu Bissoto, Ghassan Hamarneh, Rafeef Garbi*

**What's New:** This research introduces a new perspective in Continual Learning suggesting that intentional forgetting of spurious correlations could benefit the sequential learning process.

**Technical Details:** BiasPruner, a Continual Learning framework, uses a bias score to identify and prune units in a neural network that contribute to learning spurious features. The framework constructs a debaised subnetwork for each task, incorporating units from previous subnetworks to enhance adaptation and performance.

**Results:** Experimental results on medical datasets show that BiasPruner achieves improved classification performance and fairness compared to state-of-the-art Continual Learning methods, demonstrating the effectiveness of intentional forgetting in enhancing neural network adaptation during sequential learning tasks.

link: <http://arxiv.org/abs/2407.08609v1>

### **FlashAttention-3: Fast and Accurate Attention with Asynchrony and Low-precision**

*Jay Shah, Ganesh Bikshandi, Ying Zhang, Vijay Thakkar, Pradeep Ramani, Tri Dao*

**What's New:** This research introduces FlashAttention-3, a method to speed up attention on GPUs specifically designed for Hopper GPUs.

**Technical Details:** The method leverages three main techniques to speed up attention: exploiting asynchrony of the Tensor Cores and TMA, overlapping computation and data movement, and using block quantization and incoherent processing with hardware support for FP8 low-precision.

**Results:** FlashAttention-3 achieves a speedup on H100 GPUs by 1.5-2.0x with FP16, reaching up to 740 TFLOPs/s with 75% utilization, and by getting close to 1.2 PFLOPs/s with FP8. In addition, FP8 FlashAttention-3 achieves 2.6x lower numerical error than a baseline FP8 attention.

link: <http://arxiv.org/abs/2407.08608v1>

### **The Synergy between Data and Multi-Modal Large Language Models: A Survey from Co-Development Perspective**

*Zhen Qin, Daoyuan Chen, Wenhao Zhang, Liuyi Yao, Yilun Huang, Bolin Ding, Yaliang Li, Shuiguang Deng*

**What's New:** This research focuses on the interconnected development of multi-modal language models (MLLMs) and data, highlighting the importance of data in enhancing the performance of MLLMs.

**Technical Details:** The study examines how vaster and higher-quality data can improve MLLM performance, while MLLMs, in turn, can facilitate the development of data. It emphasizes the need to understand when data-centric approaches can enhance MLLM capabilities and how models can contribute to multi-modal data.

**Results:** The research concludes that the co-development of multi-modal data and MLLMs is crucial for advancing capabilities in this field. By systematically reviewing existing works, it aims to promote the symbiotic growth of data and models in the MLLM community.

link: <http://arxiv.org/abs/2407.08583v1>

### **Data-Locality-Aware Task Assignment and Scheduling for Distributed Job Executions**

*Hailiang Zhao, Xueyan Tang, Peng Chen, Jianwei Yin, Shuiguang Deng*

**What's New:** This research introduces an optimal balanced task assignment algorithm (OBTA) to minimize job completion times for distributed job executions without prior knowledge of future job arrivals.

**Technical Details:** The study narrows the search space of potential solutions to reduce OBTA's computational overhead and extends the water-filling (WF) algorithm, proving its approximation factor equals the number of task groups in the job assignment. A novel heuristic called replica-deletion (RD) is designed, surpassing the performance of WF.

**Results:** The research expands to include job reordering and evaluates the proposed algorithms through extensive trace-driven evaluations, validating their efficiency and effectiveness in reducing job completion times.

link: <http://arxiv.org/abs/2407.08584v2>

### **Semantic GUI Scene Learning and Video Alignment for Detecting Duplicate Video-based Bug Reports**

*Yanfu Yan, Nathan Cooper, Oscar Chaparro, Kevin Moran, Denys Poshyvanyk*

**What's New:** This research addresses the challenges of managing video-based bug reports by introducing a new approach called JANUS for duplicate detection in video-based bug reports.

**Technical Details:** JANUS utilizes the scene-learning capabilities of vision transformers to capture subtle visual and textual patterns on app UI screens to accurately detect duplicate bug reports. It also incorporates a video alignment technique that adapts the weighting of video frames to identify typical bug manifestation patterns.

**Results:** In a study with 7,290 duplicate detection tasks from 270 video-based bug reports, JANUS achieved an overall mRR/mAP of 89.8%/84.7%, outperforming previous methods by around 9% in most tasks. The research demonstrates the effectiveness of JANUS in leveraging scene-learning capabilities for improved bug report management.

link: <http://arxiv.org/abs/2407.08610v1>