NUS COMPUTING

AI-X: INDUSTRY DAY

in partnership with IMDA's Technical Sharing Session

Poster and Demo Sessions

1	**A Plug-and-Play AI Box for Just-in-Time Decision Intelligence** **Abstract:** *Just-in-Time.AI offers a plug-and-play AI box for just-in-time decision intelligence, addressing the critical need for early anomaly detection, a common yet challenging requirement across manufacturing, finance, cybersecurity, and healthcare, complicated further by scarce data on rare or special cases. Combining patented anomaly detection with controllable and interpretable data generation, it empowers organizations to detect risks early, simulate edge scenarios, clearly explain model decisions, and proactively respond in real-time, all within a lightweight, edge-ready deployment.*
2	**AI + Professional Sport: Case Study in Professional Football** Amirhassan Monajemi, Saumya Shah, Ojas Surana **Abstract:** *This project, AI + Professional Sport: Case Study in Professional Football, investigates advanced artificial intelligence methods to enhance tactical decision-making and enrich both player and fan experiences. We focus on two emerging directions: Reinforcement Learning (RL) and Generative Artificial Intelligence (Gen-AI).* *Together, these two approaches demonstrate how state-of-the-art AI can transform professional football from tactical optimization to immersive fan experience, bridging analytics, creativity, and strategy.*
3	**Broaden your SCOPE! Efficient Multi-turn Conversation Planning for LLMs with Semantic Space** Zhiliang Chen, Xinyuan Niu, Chuan-Sheng Foo, Bryan Kian Hsiang Low **Abstract:** *Can you imagine an AI chatbot that plans its responses not just for the next turn, but for an entire conversation—optimizing engagement, safety, and quality? Traditional LLM conversation planning can take up to half an hour per turn, simulating thousands of future responses to figure out what works best. This is too expensive and slow for any real-time applications. Our demo introduces SCOPE, an approach that allows any off-the-shelf LLMs (e.g., ChatGPT, Llama, Mistral) to select good responses in just a few seconds by planning fully in semantic space rather than simulating full dialogues. SCOPE delivers higher-quality, more engaging multi-turn conversations 70-100x faster and at a fraction of the cost, making real-time intelligent assistants a reality.*

4	**CAESARS: Modernizing Credit Rating Reporting with AI** Asian Institute of Digital Finance **Abstract:** *CAESARS (Credit Analytics Engine & Strategic Automation Reporting System) is an AI solution designed to modernize the credit rating reporting process by enhancing efficiency, ensuring reliability, and providing scalability. Moreover, it serves as a dynamic platform that showcases R&D outputs. CAESARS utilizes AI technologies to identify firm-specific data from diverse sources, including annual reports, call scripts, and news articles. It features an automated report-writing capability that delivers timely updates, ensuring stakeholders always have access to the latest financial insights.*
5	**CARE-ACD Model for Chronic Disease Patients in Acute Care**
6	**Cloud Service Crises in Incubation**
7	**CogniCare Agent: Conversational Psychological Diagnosis & Support** Abstract: *This project features a Large Language Model (LLM)-powered conversational AI designed for discreet psychological diagnosis and therapeutic support. It implicitly assesses user mental states by subtly probing for symptoms aligned with PHQ-9 & GAD-7 scales, ensuring the process is seamless and non-intrusive. Through natural dialogue, the AI answers user questions, asks insightful follow-up queries, and provides timely empathy and emotional comfort, fostering a supportive environment for mental well-being.*
8	**Continual Multimodal Contrastive Learning** Xiaohao Liu, Xiaobo Xia, See-Kiong Ng, Tat-Seng Chua **Abstract:** *Multimodal Contrastive Learning (MCL) advances in aligning different modalities and generating multimodal representations in a joint space. By leveraging contrastive learning across diverse modalities, large-scale multimodal data enhances representational quality. However, a critical yet often overlooked challenge remains: multimodal data is rarely collected in a single process, and training from scratch is computationally expensive. Instead, emergent multimodal data can be used to optimize existing models gradually, i.e., models are trained on a sequence of modality pair data. We define this problem as Continual Multimodal Contrastive Learning (CMCL), an underexplored yet crucial research direction at the intersection of multimodal and continual learning. In this paper, we formulate CMCL through two specialized principles of stability and plasticity. We theoretically derive a novel optimization-based method, which projects updated gradients from dual sides onto subspaces where any gradient is prevented from interfering with the previously learned knowledge. Two upper bounds provide theoretical insights on both stability and plasticity in our solution. Beyond our theoretical contributions, we conduct experiments on multiple datasets by comparing our method against advanced continual learning baselines. The empirical results further support our claims and demonstrate the efficacy of our method.*
9	**DIXIT Bench**

10	**Discursive Circuits: How Do Language Models Understand Discourse Relations?** Yisong Miao, Min-Yen Kan **Abstract:** *It puts a "microscope" into language models' "brain" and decipher how they understand sentence logics.*
11	**DRONEAUDIOSET: An Audio Dataset for Drone-based Search and Rescue** **Abstract**: *Unmanned Aerial Vehicles (UAVs) or drones are increasingly used in search and rescue missions to detect human presence. Existing systems primarily leverage vision-based methods which are prone to fail under low-visibility or occlusion. Drone-based audio perception offers promise but suffers from extreme ego-noise that masks sounds indicating human presence. Existing datasets are either limited in diversity or synthetic, lacking real acoustic interactions, and there are no standardized setups for drone audition.* *To this end, we present DroneAudioset, a comprehensive drone audition dataset featuring 23.5 hours of annotated recordings, covering a wide range of signal-to-noise ratios (SNRs) from -60 dB to 0 dB, across various drone types, throttles, microphone configurations as well as environments. The dataset enables development and systematic evaluation of noise suppression and classification methods for human-presence detection under challenging conditions, while also informing practical design considerations for drone audition systems, such as microphone placement trade-offs, and development of drone noise-aware audio processing. This dataset is an important step towards enabling design and deployment of drone-audition systems.*
12	**Guide Genius: AI-powered Itinerary Planning for Effortless Travel** In this demo, we will show how to improve the visitors' visiting experience in attractions like Mandai Zoo or Sentosa. Empowered by AI, we provider visitors a more engaging, interactive, and intelligent experience when they visit a place.
13	**Helpful or Harmful Data? Fine-tuning-free Shapley Attribution for Explaining Language Model Predictions** Jingtan Wang, Xiaoqiang Lin, Rui Qiao, Chuan-Sheng Foo, Bryan Kian Hsiang Low **Abstract:** *The demo showcases the FreeSHAP algorithm, designed to robustly and efficiently explain test predictions. For any given test instance, it identifies both helpful and harmful training examples.* *- Helpful examples are those whose inclusion most improves the accuracy of the test prediction.* *- Harmful examples are those that negatively impact the prediction.* *This allows for a clearer understanding of how specific training data influences model decisions.*

14	**IGD: Token Decisiveness Modeling via Information Gain in LLMs for Personalized Recommendation** Zijie Lin, Yang Zhang, Xiaoyan Zhao, Fengbin Zhu, Fuli Feng, Tat-Seng Chua **Abstract:** *Large Language Models (LLMs) have shown strong potential for recommendation by framing item prediction as a token-by-token language generation task. However, existing methods treat all item tokens equally, simply pursuing likelihood maximization during both optimization and decoding. This overlooks crucial token-level differences in decisiveness-many tokens contribute little to item discrimination yet can dominate optimization or decoding. To quantify token decisiveness, we propose a novel perspective that models item generation as a decision process, measuring token decisiveness by the Information Gain (IG) each token provides in reducing uncertainty about the generated item. Our empirical analysis reveals that most tokens have low IG but often correspond to high logits, disproportionately influencing training loss and decoding, which may impair model performance. Building on these insights, we introduce an Information Gain-based Decisiveness-aware Token handling (IGD) strategy that integrates token decisiveness into both tuning and decoding. Specifically, IGD downweights low-IG tokens during tuning and rebalances decoding to emphasize tokens with high IG. In this way, IGD moves beyond pure likelihood maximization, effectively prioritizing high-decisiveness tokens. Extensive experiments on four benchmark datasets with two LLM backbones demonstrate that IGD consistently improves recommendation accuracy, achieving significant gains on widely used ranking metrics compared to strong baselines.*
15	**KAHAN: Knowledge-Augmented Hierarchical Analysis and Narration for Financial Data Narration** Yajing Yang, Tony Deng, Min-Yen Kan **Abstract:** *This work is about generating financial market reports from tabular data by extracting data insights through a hierarchical analysis pipeline guided with LLM generated domain knowledge.*
16	**KALEIDO: Diverse News Retrieval with Interpretable Embedding** Anthony Tung, Yiqun Sun, Yixuan Tang, Qiang Huang, Yuanyuan Shi, Yingchaojie Feng **Abstract:** *In today's information landscape, users are often exposed to repetitive or biased news coverage, making it difficult to form a well-rounded understanding of events. We tackle this challenge by combining large-scale news indexing with a two-stage diverse news retrieval framework: first retrieving relevant articles, then re-ranking them to promote diversity across perspectives. We leverage sentence-level semantic analysis and clustering to identify distinct narratives and reduce redundancy. To make this accessible in real time, we deploy our model as a browser plug-in that surfaces diverse news views directly within users' browsing experience.*
17	**Koditsu** AI Centre for Educational Technologies

18	**Neural Dueling Bandits: Preference-Based Optimization with Human Feedback** Arun Verma, Zhongxiang Dai, Xiaoqiang Lin, Patrick Jaillet, Bryan Kian Hsiang Low **Abstract:** *Many real-world applications, such as conversational agents and online recommendation systems, require AI systems that adapt dynamically to human preferences as they arrive sequentially. We propose algorithms that enable efficient learning from human preference feedback while providing rigorous theoretical guarantees. These guarantees ensure that the AI system improves adaptively while incurring only a limited mistakes over time.*
19	**Protein–Protein Interaction, Binding Affinity, and Interface Contact Prediction Using a Pairwise Language Model** Jun Liu, Hungyu Chen, Yang Zhang **Abstract:** *Protein–protein interactions (PPIs) form the basis of virtually all cellular processes, from signal transduction to immune recognition, and represent critical targets for therapeutic discovery. Despite their importance, predicting PPIs, including whether two proteins interact, how strongly they bind, and which residues form their interface, remains a fundamental and unsolved challenge. Existing protein language models have transformed sequence-based protein analysis, but their focus on individual protein chains limits their ability to capture the cooperative patterns intrinsic to protein complexes. To address this gap, we developed a pairwise protein–protein language model that directly learns joint representations of interacting proteins. Building upon this foundation, we design predictors for interaction, binding affinity, and interface contact recognition, offering a unified framework for advancing PPI research. This approach underscores the importance of modeling proteins in their natural interactive context and opens new opportunities for understanding molecular mechanisms and guiding therapeutic design.*
20	**ScholAIStic** AI Centre for Educational Technologies **Abstract:** *With ScholAIstic, we can bring an AI tutor to life to enhance students learning – either as a facilitator to guide and challenge students to go further in their thinking through Socratic questioning, or by role-playing so that students practice and refine valuable human skills through their inter;actions with a chatbot.*
21	**Self-Improvement Towards Pareto Optimality: Mitigating Preference Conflicts in Multi-Objective Alignment** Moxin Li, Yuantao Zhang, Wenjie Wang, Wentao Shi, Zhuo Liu, Fuli Feng, Tat-Seng Chua **Abstract:** *Multi-Objective Alignment (MOA) aims to align LLMs' responses with multiple human preference objectives, with Direct Preference Optimization (DPO) emerging as a prominent approach. However, we find that DPO-based MOA approaches suffer from widespread preference conflicts in the data, where different objectives favor different responses. This results in conflicting optimization directions, hindering the optimization on the Pareto Front. To address this, we propose to construct Pareto-optimal responses to resolve preference conflicts. To efficiently obtain and utilize such responses, we*

	propose a self-improving DPO framework that enables LLMs to self-generate and select Pareto-optimal responses for self-supervised preference alignment. Extensive experiments on two datasets demonstrate the superior Pareto Front achieved by our framework compared to various baselines.
22	**Streamlining Contract Life Cycle Management with Advanced AI** **Abstract:** *We exploit the power of GenAI to manage contracts more easily and smartly at scale. In our demo, we will show that we can chat with contracts, auto fill reports, review the contracts, and highlight any legal risks, etc.*
23	**TETRIS: Optimal Draft Token Selection for Batch Speculative Decoding** Zhaoxuan Wu, Zijian Zhou, Arun Verma, Alok Prakash, Daniela Rus, and Bryan Kian Hsiang Low **Abstract:** *Tetris is a smarter way to speed up large language models like ChatGPT. Instead of having the "expert" model generate every word one by one itself, an "assistant" model quickly drafts possible next words and the "expert" checks them. This already saves time, but when many users ask questions at once, blindly checking every draft can slow things down. Tetris acts like a traffic controller, strategically picking the most promising words to verify first, so the system delivers faster, smoother responses to more users at the same time.*
24	**Understanding Code-Mixed Queries in Multilingual Dense Retrieval** **Abstract:** *Code-mixing, when two or more languages are used within the same sentence (e.g., "Dey, wo men paktor always makan at kopitiam one")—is common in multilingual contexts such as Singapore. Yet, code-mixed retrieval remains under-explored. This project examines how Dense Passage Retrieval (DPR) performs when queries are code-mixed at different ratios across language pairs. Using DPR-based multilingual retrieval on standard benchmarks, we systematically vary language pairs, code-mixing ratios, and also probe representation-level mixing directly in the embedding space. Our hypothesis is that moderate code-mixing can improve retrieval by leveraging complementary semantics across languages. Preliminary results support this: on MSMARCO with an Indonesian–Chinese pair, a 70% Indonesian / 30% Chinese code-mixed query achieves a 1.2% relative improvement over a pure Indonesian query and 2.5% relative over a pure Chinese query, indicating that controlled code-mixing can outperform monolingual baselines and suggesting practical avenues for more effective cross-lingual search and query representations.*
25	**WaterDrum: Watermarking for Data-centric Unlearning Metric** Xinyang Lu, Xinyuan Niu, Gregory Kang Ruey Lau, Bui Thi Cam Nhung, Rachael Hwee Ling Sim, Fanyu Wen, Chuan-Sheng Foo, See-Kiong Ng, Bryan Kian Hsiang Low **Abstract:** *Sometimes, we need an LLM to "forget" certain information, such as private/copyrighted data or harmful content. We propose WaterDrum as a metric for efficiently evaluating the success of removing the influence of such unwanted data (that has been requested to be forgotten) from a trained LLM. Unlike other evaluation metrics that are often impractical, we show that WaterDrum can be effectively and efficiently applied to real-world applications.*

26	**Waterfall: Scalable Framework for Robust Text Watermarking and Provenance for LLMs** Gregory Kang Ruey Lau, Xinyuan Niu, Hieu Dao, Jiangwei Chen, Chuan-Sheng Foo, Bryan Kian Hsiang Low **Abstract:** *This demo shows how text watermarking is performed with Waterfall. Given an original piece of text, Waterfall uses a Large Language Model (LLM) to generate an alternative watermarked version of the text. This watermarked text contains a unique watermark signal only detectable to the owner, while still retaining the same meaning and content as the original. Waterfall allows data owners to detect unauthorized usage of Waterfall-watermarked text, such as through plagiarism or use for training LLMs.*
27	**Your Company, your Benchmark!** **Abstract:** *The current landscape of public benchmarks often suffers from benchmark saturation, where many state-of-the-art LLMs achieve similar, near-perfect scores (e.g., above 95%). This saturation makes it difficult for companies to differentiate between models and select the one best suited for their specific needs and budget. As a result, many organizations default to selecting the most advanced and often most expensive LLM, which may not be the optimal solution for their unique use case or proprietary knowledge domain. Our work proposes a new methodology for curating unique, company-specific benchmarks that LLMs have not been trained on. This approach aims to provide a more accurate evaluation of a model's performance on tasks relevant to a company's specific domain, moving beyond generic public metrics to enable more informed and effective LLM selection.*