ARIA

APPLIED RESEARCH IN ACTION

2022 INTERNSHIP PROJECTS

MScAC Master of Science in Applied Computing

Computer Science
UNIVERSITY OF TORONTO

THANK YOU TO OUR SPONSORS

- SOTI
- Roche
- Walmart
- AMD and Entrepreneur First

THANK YOU TO OUR PARTNERS

- Advanced Micro Devices (AMD)
- Arima
- Canada Pension Plan Investment Board
- CIBC
- Cedience
- Conflux
- Crossing Minds
- Deloitte
- Dewpoint Therapeutics
- Double Negative (DNEG)
- Fudan University, Electronic Detection Lab
- Geotab
- Google
- Hypatia System
- ICICI Bank
- Layer 6
- LG Electronics Canada
- Meta
- Metabob
- Microsoft
- Modiface
- NerveX Neurotechnologies
- Nexxt Intelligence
- Nureva
- Pearson Canada
- Pelmorex
- Princess Margaret Cancer Centre
- Robert Bosch
- Samsung AI Center
- Samsung Electronics
- Scotiabank
- Scrawler Development
- Snowflake
- SOTI
- SRA Staffing Solutions
- Surgical Safety Technologies
- Talka AI
- Unity Health Toronto
- University Health Network (UHN)
- Vanguard
- Wealthsimple
- XLScout

A MESSAGE FROM THE MSCAC PROGRAM DIRECTOR

The Master of Science in Applied Computing (MScAC) program is proud to be the training home of ground-breaking technological innovators. Our industry-academic collaborations continue to defy gravity, pushing the limits towards ensuring the workforce that is effectively solving real-world problems. The MScAC program opens its doors once again, inviting the world to our Applied Research in Action (ARIA) showcase – the centre stage for experiencing what is the most innovative and cutting-edge research being conducted in the world.

ARIA is as one of Toronto's foremost tech innovation showcases run by the academic sector taking place in North America's third largest tech hub. The University of Toronto is recognized for fuelling the boom in Toronto's tech-innovation ecosystem and it is our MScAC program that continues to play a major role in this ever-growing technological boom.

This year ARIA is back in person and bigger than ever, with attendees from around the world. At ARIA you can network with a rich cross-section of industrial researchers and see first-hand technological innovations being showcased by industry partners. You will also meet the exceptional research talent coming out of our many partner departments: Computer Science, Statistical Sciences, Mathematics, Physics, Electrical and Computer Engineering, and Mechanical and Industrial Engineering.

Our sincerest thanks to everyone who made ARIA 2022 possible and who make the MScAC program one of the world's best Applied Computing graduate programs.

A special congratulations to our 2021/22 cohort who we welcome to our MScAC alumni family!

ARVIND GUPTA
Professor
Academic Director, Professional Programs
Department of Computer Science
University of Toronto

2022 ARIA AWARD WINNERS

STUDENT INNOVATION AWARD

Aslesha Pokhrel and Siyuan Shang

Nominated by an industry supervisor, this award is presented to a student in the internship phase of the MScAC program who has demonstrated innovation, dedication, and creativity in the completion of their internship project.

FACULTY RECOGNITION AWARD

Vardan Papyan and Kirill Serkh

The Faculty Recognition Award recognizes an academic supervisor who has shown exceptional commitment, time, and skilled attention to developing the talents of their MScAC students during the internship phase of the program.

INDUSTRY ICON AWARD

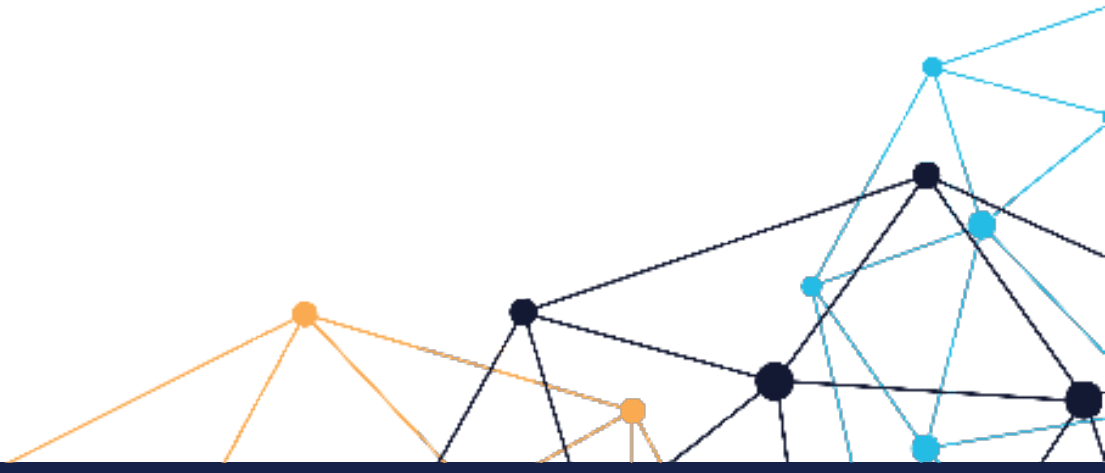
Alex Levinshtein

The Industry Icon Award recognizes an industry partner who has demonstrated exceptional commitment, time, mentorship, and skilled attention to developing the talents of their MScAC student during the internship phase of the program.

ARIA SPOTLIGHT AWARD

Nathan Taback

The ARIA Spotlight Award is awarded to an individual who has had a tremendous impact on the MScAC program.



2022 RESEARCH INTERNSHIP PROJECTS

Enhancing Web-based Applications and Generating Content with GPT-3 Model

- **Student:** Afzal, Amaar
- **Academic:** Dehan Kong
- **Industry:** Konstantin Eletskiy
- **Company:** ArcadeJolt

The OpenAI Generative Pre-trained Transformer 3 (GPT-3) model is the currently widely used language model for generating human-like text. Our research aims to answer if it is possible for a machine to have human-like thinking and creativity in producing automatically generated content with a GPT-3 model. The research has four main steps: creating the dataset, fine tuning GPT-3, modeling the trained data, and then generating the unique content.

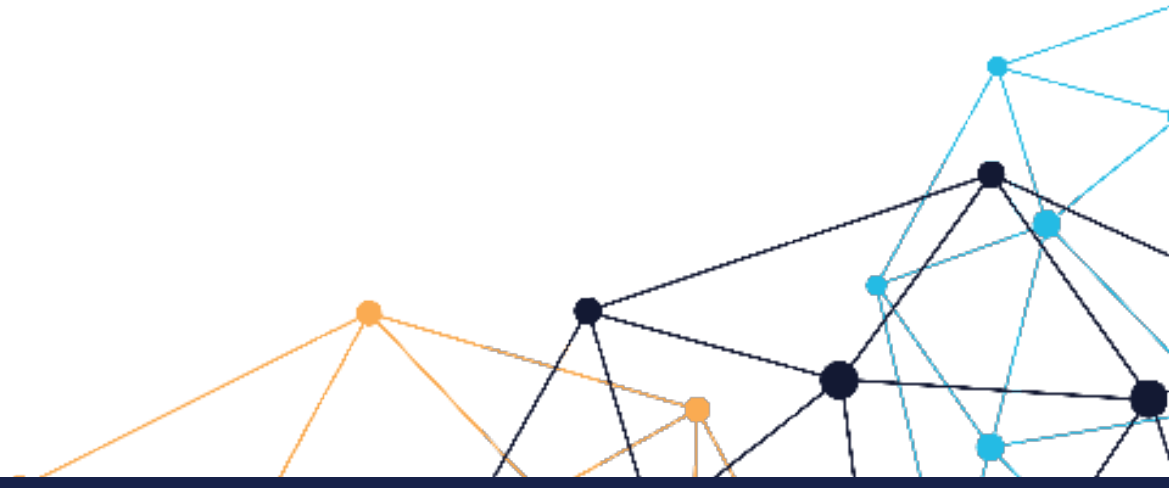
We present a novel approach of creating the dataset for the GPT-3 model through web scraping content. The web scraping aspect tries to solve the problem of information extraction from repetitive blocks in the structure of web pages, by utilizing both the visual features and the semantic information to extract informative blocks, and then use this information in the generative model. This allows us to generate sentences and long-form articles that are on-topic and relevant. We then perform an analysis of similarity to existing content on related web pages to determine if it is possible to produce unique and human-like content from the generative model.

The findings show that the proposed model is capable of generating human-like content after training the model on relevant informative blocks. The research shows that it is possible to use this in production systems to generate articles, summaries, and long form unique content in web pages.

Using single-cell CITE-seq data, to find strange connections in lung cancer patients

- **Student:** Azarmina, Arvin
- **Academic:** Prof. Kieran Campbell
- **Industry:** Azin Sayad
- **Company:** Ohashi Lab – Immuno Oncology Group (PMH)

Human cells can be sequenced in many ways. One approach is CITE-seq single cell RNA sequencing, which contains information about thousands of genes found in each cell taken as a sample from a patient. Given lung cancer patients' overall cells as a study, we can extract the cells of a tumor, which then will be divided into multiple groups, the main two categories of which being the cancerous cells, and the non-cancerous ones. In this analysis, using multiple methods of visualization, analysis, clustering, and the cell environment "transcription factor" detection, we try to clean the cell data, recognize each cell type, categorize cell types, and find the important and impressionable cells in that environment. Future work can be done with another type of cell recognition method called CyTOF, with a similar analysis on larger but narrower data, with fewer features for each data point. Additional future work includes the connection between the CITE-seq and CyTOF set of cells in the same patient, and the use of that connection to understand and analyze the evolution of cells in different time points.



Quantifying and modulating popularity bias in very-large recommendation datasets

- **Student:** Banthia, Shreyansh
- **Academic:** Prof. Nisarg Shah
- **Industry:** Garrin McGoldrick
- **Company:** Crossing Minds

Recommender systems often suffer from popularity bias, over-recommending popular items while under-recommending less popular but more suitable items. One popular approach to address popularity bias is to make all items appear approximately the same number of times (statistical parity) while ignoring the suitability of new recommendations to the users. The preferred solution to popularity bias is instead to tailor the bias modulation to each user's past item interactions (equal opportunity).

This study explores how user-side (uPB) and item-side (iPB) popularity-opportunity bias metrics can be calculated for large real-world datasets with commodity hardware. We also explore debiasing methods which re-rank recommendations in a way which modulates the resulting uPB and iPB. We show that this debiasing method, on average, reduces uPB by 2603.8% and iPB by 2089.2% while having a minimal impact on recommender system performance metrics (~0.3% average reduction in Normalized Discounted Cumulative Gain scores). We plan to further extend this work by introducing the debiasing logic as a regularizing parameter in model training and using subsampling to make the metrics and the method work for even larger (1 billion+ interactions) datasets.

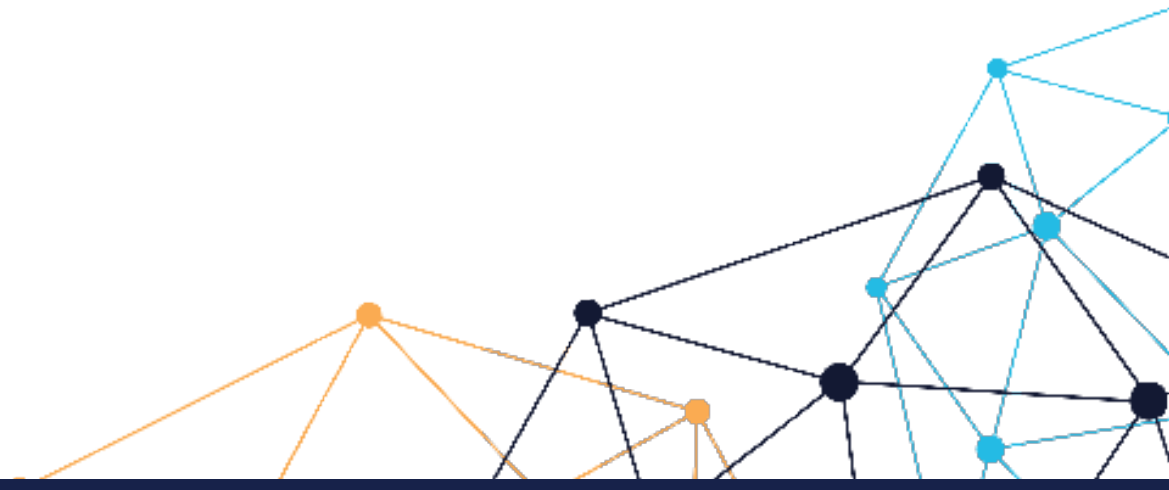
A Quantum-Hybrid Algorithm to Solve the Vehicle Routing Problem with Time Windows

- **Student:** Cai, Qiyuan
- **Academic:** Christopher Beck
- **Industry:** Aida Ahmadzadegan
- **Company:** ForeQast

Last-mile routing for package delivery refers to the process of packages being delivered from the last distribution point, such as a warehouse or a distribution center, to the recipients. The nature of the problem is susceptible to the use of a quantum computer.

In this project, we explored the application of quantum annealing in solving a variation of the vehicle routing problem, for which the objective is to maximize the number of orders served within their respective time windows, while guaranteeing visits to a subset of the orders. We designed a two-stage algorithm that first sorts the orders into clusters assigned to each vehicle, then solves a single-vehicle routing problem for each cluster. We then attempted to solve each stage of the problem by formulating it as a quadratic unconstrained binary optimization model (QUBO) and solving it with D-Wave's quantum hardware.

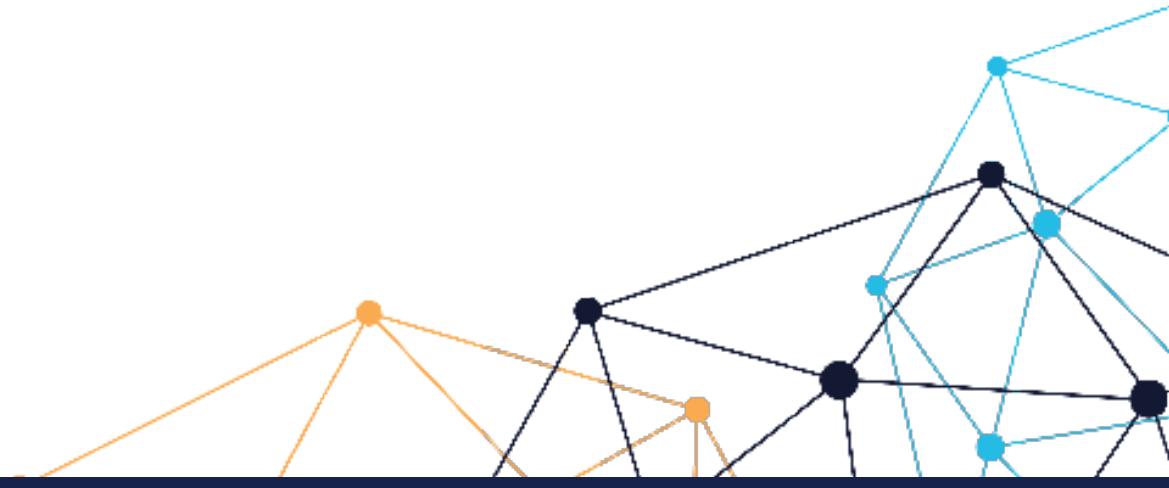
We found that the computational power of current quantum devices is limited in its capability to solve the vehicle routing problem with time windows. The current implementation is a hybrid algorithm where the clustering stage is solved partially with quantum annealing, and the routing is performed classically using dynamic programming.



Meta-Learning Gaussian Processes for Faster Hyperparameter Optimization

- **Student:** Chaiwachirasak, Teerapat
- **Academic:** Prof. Scott Sanner
- **Industry:** Garrin McGoldrick
- **Company:** Crossing Minds

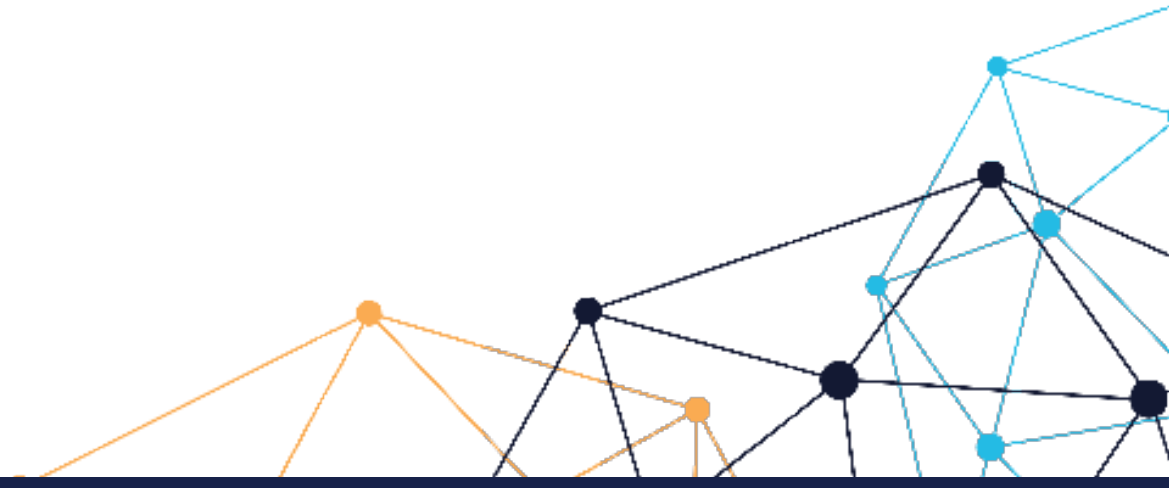
Hyperparameter Optimization (HPO) is the process of finding a set of hyperparameters which allow for a learning algorithm to reach an optimal loss value on a given validation dataset. The optimization process can be time-consuming, especially when dealing with large datasets or complex models. A common approach uses GP to approximate the mapping of the hyperparameter values to the algorithm's resulting validation loss on a run of observations taken from the algorithm being optimized. Given a run of observations taken from a similar algorithm and dataset, it would be desirable for the GP to utilize this information. We propose to extend Gaussian process-based Bayesian optimization using a rank-weighted Gaussian process ensemble (RGPE). The method estimates the target function as the weighted sum of the Gaussian processes of the current and previous optimization runs. The weight of each GP is proportional to its ability to correctly rank hyperparameters according to their performance on the algorithm currently being optimized. RGPE outperforms random search and traditional GP-based HPO on 12 public recommendation datasets with an average reduction in a wall clock time of 75%. This project demonstrates the transferability of HPO tasks across many recommendation datasets and proposes RGPE to speed up HPO for recommender systems.



A Color Classifier for Virtual Hair Product Try-On

- **Student:** Chau, Lydia
- **Academic:** Lueder Kahrs
- **Industry:** Vicky Yu, Irene Jiang
- **Company:** ModiFace

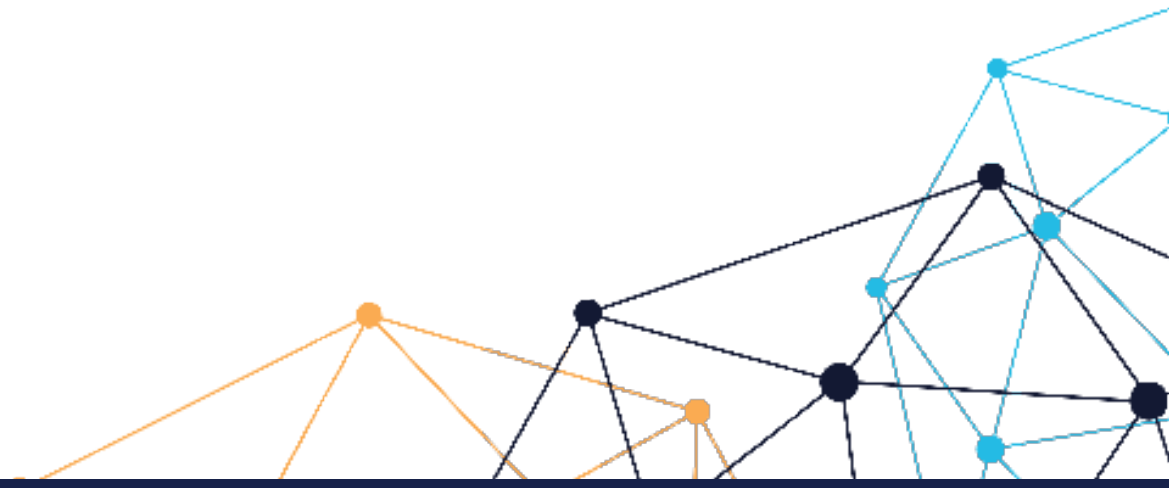
Due to increasing popularity in online shopping, e-commerce companies have been exploring ways to improve customer experience. Virtual try-on technology is developed to assist customers in searching for products that best suit them. In this project, we developed a model that identifies hair color in head-shot photos which can be used for virtual hair dye try-on. There are multiple challenges when developing the hair color classification model, for example, color labels from hair experts are subject to human basis and labeled data are expensive to obtain. To handle these problems, we built a classifier model through a two-step process. To alleviate the impact of human bias, we separated the modeling of the real label distribution and annotators' biases through incorporating annotator confusion matrices to the base model; to further improve the model performance leveraging unlabeled data, we trained the model using a consistent-based semi-supervised learning framework. With the use of only 1000 labeled data, the final classifier model achieves classification accuracy that is 20% higher than a human professional annotator. The trained model can be used for a wide range of downstream tasks, including being used as a classifier to train generative models for virtual hair color try-on.



Neural Networks for Observable Market Data Validation

- **Student:** Danisa, Siphelele
- **Academic:** Prof. Kirill Serkh, Prof. Vardan Papyan
- **Industry:** Simeon Spencer
- **Company:** CIBC

Observable Market Data is critical for effective valuation of trades for risk management. Banks collect and store risk-factors for thousands of assets on a daily basis, with validation of this data being a costly process. Statistical methods are typically used to identify potentially suspect incoming data for validation, but these have known drawbacks, and data-filling can be a complex task. The goal of our project was to develop deep learning models that improve the false positive rate of extant anomaly detection methods by at least 10%, and to examine the capacity of these models as a source of comparable data. Our best-performing Transformer model reduces the anomaly detection false positive rate by roughly 65%. Moreover, evaluated on a carefully tailored dataset (a subproblem), 50% of its predictions for missing data are within 3% of the real data, the next 20% are within 11%, and the final 30% are within 25%, where this percentage is the relative disparity between model predictions and actual data. The extension of these results to the full problem is underway. Similarly, the extension of both applications to other data classes, like Bond Yields and Commodity Future Prices.



A Real-time Semantic Based Human Tracking and Identification System for CCTV

- **Students:** Deng, Wenyue & Pan, Zixuan
- **Academic:** Igor Gilitschenski, Babak Taati
- **Industry:** Sunil Jacob
- **Company:** SOTI

Multi-Object Tracking (MOT) and Re-Identification (ReID) are essential tasks in information safety. As a leading character in mobile device management, SOTI Inc. aims to integrate MOT and ReID to provide an intelligent pipeline for closed-circuit television (CCTV). In this work, we combine these two techniques and present a smart CCTV surveillance solution for high-security environments.

Our solution comprises a customized version of one of the most popular MOT methods, StrongSORT, and our compact novel ReID for video streaming. To retrieve a given person's track among video streaming captured by security cameras, an instance-segmentation-based person re-identification module has been implemented to improve the tracking performance. We also perform parallel computing to speed up the inference of the system and obtain a real-time result.

To evaluate our solution, several continuous video streams have been recorded in a real office environment. We feed these test sequences into both the original StrongSORT and our solution. As a result, our system provides us with a significant identification capability that accurately detects intruders and provides activity summaries when needed.

Dialogue generation for AI assistants using web search

- **Students:** Dige, Omkar Anant
- **Academic:** Prof. Frank Rudzicz
- **Industry:** Peter Potash
- **Company:** Microsoft

Recent advances in natural language generation have helped AI chat assistants achieve human-like conversational abilities. Despite these improvements, they often produce factually incorrect and outdated information ([1],[2]). To mitigate misrepresentation of information, new methods ([3],[4]) allow the generated dialogue to dynamically leverage content from the internet. The approach consists of generating a search query based on the conversational context, fetching passages using a search engine API and generating a response conditional on these passages from the web.

The pipeline for grounded response generation consists of three modules: Search query generation, passage retrieval and, response generation. We leverage the prompting ability of large language models and use suitable prompts with few shot examples to generate both the search query and the grounded response. The Bing search engine is used for passage retrieval. Additionally, since the AI assistant is expected to respond to messages which are either information seeking or chitchat in nature, web grounded responses may not be required for every turn. Hence, different than the approach used in [5], we allow the model to always search and retrieve the passage but condition the response on the passage only if necessary. Human evaluation results show marginal increase in credibility and a significant increase in specificity and interestingness scores.

References

[1] Roller, Stephen, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu et al. "Recipes for building an open-domain chatbot." arXiv preprint arXiv:2004.13637 (2020).

[2] Shuster, Kurt, Spencer Poff, Moya Chen, Douwe Kiela, and Jason Weston. "Retrieval augmentation reduces hallucination in conversation." arXiv preprint arXiv:2104.07567 (2021).

[3] Komeili, Mojtaba, Kurt Shuster, and Jason Weston. "Internet-augmented dialogue generation." arXiv preprint arXiv:2107.07566 (2021).

[4] Shuster, Kurt, Mojtaba Komeili, Leonard Adolphs, Stephen Roller, Arthur Szlam, and Jason Weston. "Language models that seek for knowledge: Modular search & generation for dialogue and prompt completion." arXiv preprint arXiv:2203.13224 (2022).

[5] Thoppilan, Romal, Daniel De Freitas, Jamie Hall, Noam Shazeer, Apoorv Kulshreshtha, Heng-Tze Cheng, Alicia Jin et al. "Lamda: Language models for dialog applications." arXiv preprint arXiv:2201.08239 (2022).

Extracting images from patent PDF files: Distinguishing text from images with pretrained deep neural network

- **Student:** Ding, Hongyi
- **Academic:** Prof. Eric Yu
- **Industry:** Khushwant Rai
- **Company:** xlscout

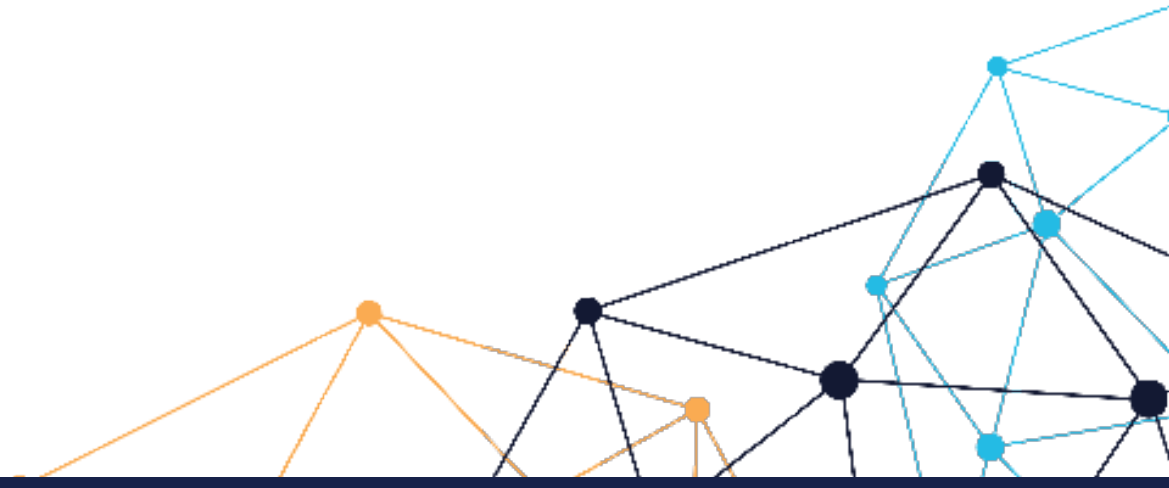
Images are an important component of patent documents and make it easy for the viewers to understand the contents of a patent document. Some existing python modules may extract the whole page as an image and cannot always extract images from a pdf document accurately. The goal of this project is to improve the outputs of existing modules and propose a machine-learning approach to automatically distinguish text blocks from image blocks.

We applied statistical-based methods to get the correct outputs from existing modules. A pretrained deep neural network was improved to fit our cases through transfer learning. Finally, we achieved 98% accuracy of the input with the new model that demonstrated the feasibility of using our approach to extract images from patent PDFs. The next step of the project is to deploy models into other types of PDf files.

Immune-modulatory and T cell landscapes of Melanoma, Lung, Ovarian and Breast cancers, and their association with tumor genotype

- **Student:** Ding, Yuyi
- **Academic:** Prof. Nancy Reid
- **Industry:** Azin Sayad
- **Company:** University Health Network (UHN)

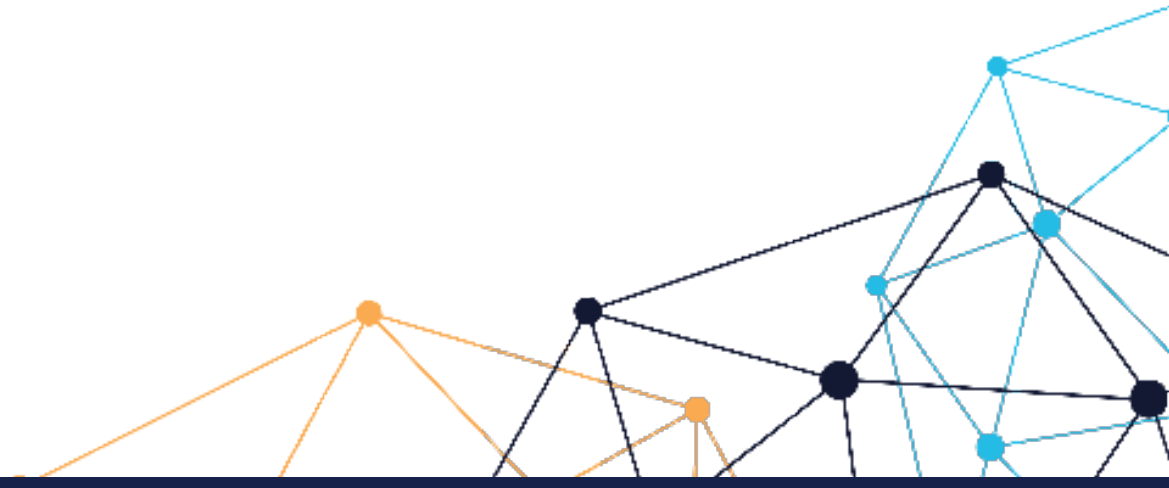
Immunotherapies that restore T cell anti-tumor immunity have revolutionized cancer treatment in the last decade. While many cancer types respond well to immunotherapies, long-term response rates remain poor in some malignancies. A major goal of current immuno-oncology research is to catalogue the cellular and molecular factors that regulate T cell anti-tumor immunity in different tumors and indications, to better determine patient response to immunotherapy and inform future clinical trial designs. In this project, we aim to characterize immuno-regulatory and T cell landscapes in Melanoma, Non-Small Cell Lung Cancer, notably Ovarian, and Breast Cancer based on tumor genotype, patient recurrence and response to immunotherapy status. Applying state of the art single-cell high-throughput immune profiling technologies, we produce an in-depth, multimodal compendium of molecular data. Combined with applied modeling and computational approaches, we aim to extract key insights into the mechanisms of tumor immunotherapy response. By exploring differences in immuno-regulatory contexts across immunotherapy responsive and non-responsive cancers, we aim to expand our limited understanding of the factors involved in divergent patient response to immunotherapy and use this understanding to direct future tumor-specific immunotherapy trials at our hospital.



Continual Learning via User Mediation for Decentralized Agent Aggregation

- **Student:** Duan, Yihan
- **Academic:** Prof. Anthony Bonner
- **Industry:** David Marques
- **Company:** Google, Meta

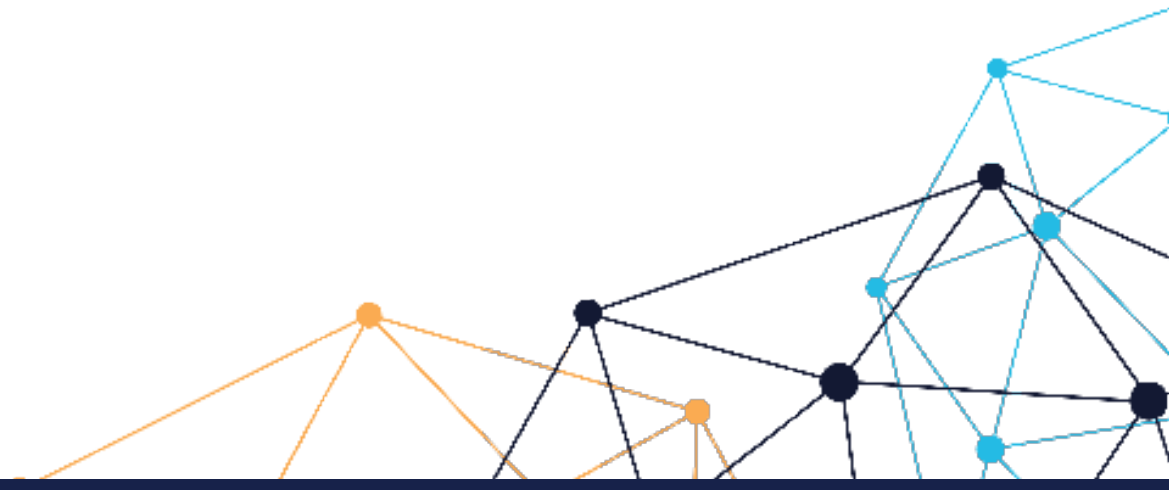
Recent improvements in computing power and ambient technologies have opened up new perspectives in ambient intelligence and proactive software systems. The need emerges for an ambient and supportive system that deeply understands the user's needs and preferences under the current context. The new system should leverage multimodal AI agents to understand the user and proactively make recommendations for the next action. In this work, we propose a recommendation system that aggregates and understands the current situation and continuously learns from the user's decision history. To leverage a rich pool of contextual data, the system dynamically changes the strategy and scope for contextualization and encodes multimodal data into unified embeddings. A contextual similarity database consisting of these embeddings is leveraged for finding, ranking and comparing scenarios. The proposed recommender is intended to fit into a decentralized multi-agent system and addresses challenges in agent communication dynamics, user engagement, user privacy and scalability.



Autonomous Navigation for Small UAVs in Indoor GPS-denied Environments

- **Student:** Farghadani, Soroush
- **Academic:** Prof. Igor Gilitschenski
- **Industry:** Sunil Jacob
- **Company:** SOTI

One of the difficult issues for unmanned aerial vehicles is autonomous navigation in a new or uncharted area (UAVs). High level control systems that are complex and capable of collision detection and adaptation are required to meet this challenge. As unmanned aerial vehicles (UAVs) have advanced in recent years, an increasing number of researchers are focusing on the application study of the inside environment. Among other indoor uses, there are warehouse inventory management, search and rescue, and industrial facility inspection. UAV deployment in these applications, however, need ongoing high-accuracy location and posture information, as well as an effective obstacle avoidance algorithm. To solve the aforementioned issue, we have developed an algorithm that can: (1) perceive the immediate surroundings; (2) manage the attitude and position of UAVs; and (3) steer UAVs over obstacles. This is done through (1) Visual Odometry, (2) Motion planning, and (3) Manipulation of UAVs.



Taking the First Steps Towards Connecting Providers and Consumers in Cloud Services

- **Student:** Feng, Yutian
- **Academic:** Prof. Eyal de Lara
- **Industry:** Xiaodi Ke
- **Company:** Snowflake

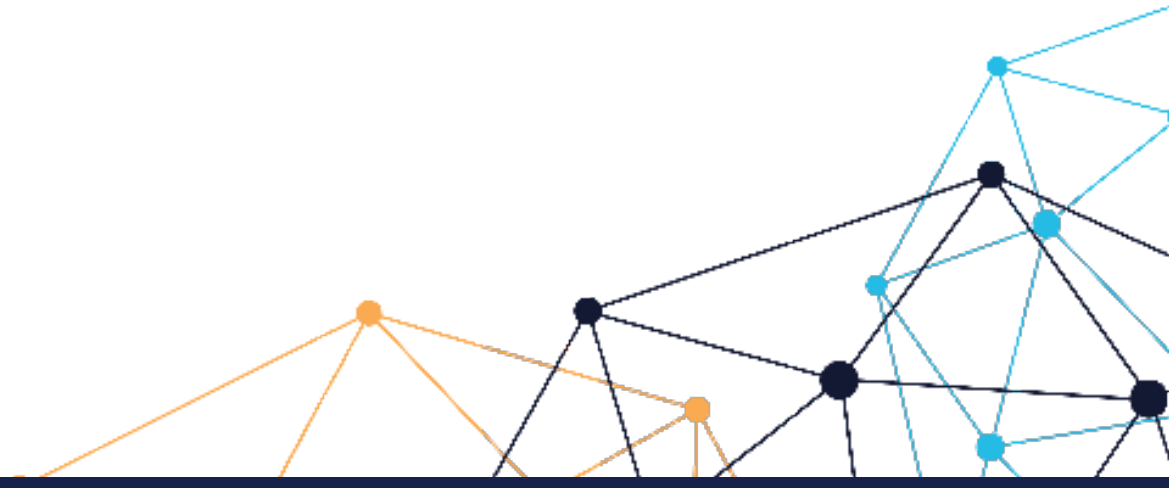
Snowflake's mission is to enable every organization to mobilize data. The Snowflake Native Application Framework provides the ability for developers to develop and publish apps to the Marketplace. Consumers can go to the Marketplace, get the app, and install it securely in their Snowflake account, keeping their sensitive data in their account instead of exporting it to SaaS applications. However, with the growing number of providers and consumers, there is a need for a framework that connects them so that consumers can share application log files with providers when the application fails.

This project takes the first steps toward building a framework to connect providers and consumers, allowing consumers to share log files with providers on an as-needed basis. First, we research and analyze the challenges and pain points of building such a framework, especially in the world of highly distributed cloud applications. Second, we propose two possible approaches and study their pros and cons. Third, we implement a proof of concept model to prove the technical feasibility. As part of future work, we plan to integrate the model into the current system. Further investigation is required in terms of monitoring and monetization.

Weight Detection of Heavy-duty Vehicles

- **Student:** Ge, Yifeng
- **Academic:** Andrei Badescu, Prof. Xiaodong Lin
- **Industry:** Willem Petersen
- **Company:** Geotab Inc.

Transportation is a necessary piece that connects people around the world. The increasing demand of driving generates a large amount of geospatial data that can be collected by telematics devices and further processed through machine learning models to support different use cases. One of the big use cases is weight detection. With vehicle weight information provided, many downstream tasks like overweight monitoring, evaluation of traffic bottlenecks, risk evaluation of driving, etc., becomes feasible. The objective of this project was to complete binary classification of cargo weight (loaded/unloaded) on heavy-duty (class 7 or 8) vehicles, which could be correlated to safety benchmarking such as harsh driving events. To classify vehicle weight, geospatial data and vehicle engine data like latitude, longitude, elevation, torque, RPM, etc. were collected during acceleration maneuvers to build one supervised learning model and two semi-supervised learning models. They were Random Forest (RF), DeepTab, and AutoEncoder (AE)+RF respectively. With ideal inputs, DeepTab model outperformed other models with an accuracy of 89.4%, because it learned salient feature interaction during pre-training and used multi-head self-attention to achieve state of the art classification. In the future, the goal is to incorporate weight into a feature set that can be used to conduct safety evaluation.



Financial Arbitrage Opportunity Identification using Machine Learning-based Framework

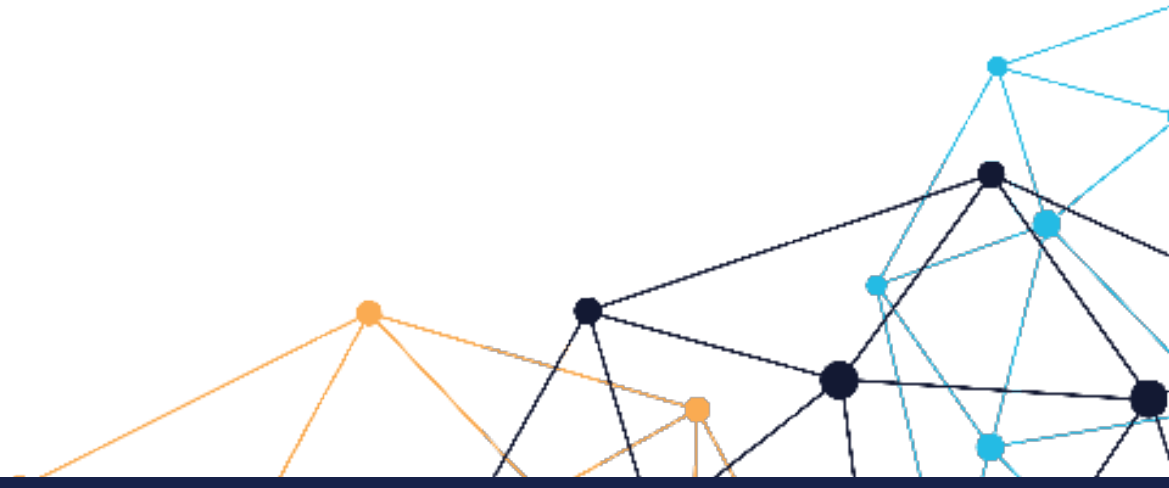
- **Student:** Gill, Deepkamal
- **Academic:** Sushant Sachdeva
- **Industry:** Jithin Pradeep, Tina Wang
- **Company:** Vanguard

Due to volatility in supply and demand of an Exchange-Traded Fund (ETF) and market inefficiency, the price of an ETF may deviate from its Net Asset Value (NAV). This creates arbitrage opportunities which can be leveraged to effectively prepare for future creation and redemption activities. However, it is theoretically hard to predict arbitrage because of noisy and partially observable market conditions and existing research using pure machine learning techniques is limited. We propose a data-driven approach to detect arbitrage using historical data. We conduct a comprehensive study to identify 135 macro-economic and fund characteristic features that are drivers contributing to this deviation. The first experiment is designed to identify top 10% deviations since those represent significant arbitrage. The second experiment aims to classify the sign of deviation (positive or negative). After testing the hypotheses on various classical and deep learning-based classifiers, we obtain an F1 score of 0.62 using LightGBM for detecting arbitrage and 0.82 for identifying the deviation sign. Thus, our method presents a novel machine learning-based framework for financial arbitrage identification and forms a strong baseline for future work in the domain.

Reconstructing 3D Meshes from Neural Radiance Fields

- **Student:** Guo, Yijie
- **Academic:** David Swart, David B. Lindell
- **Industry:** Martine Bertrand
- **Company:** Double Negative (DNEG)

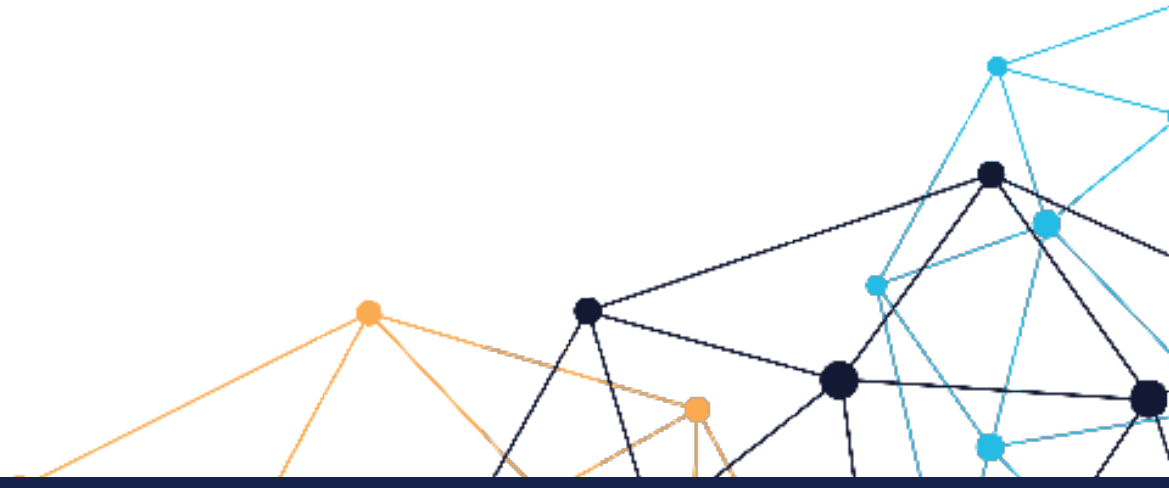
3D reconstruction from multi-camera inputs is useful for many applications, such as robotics, animations, visual effects (VFX), AR/VR, etc. Recently, neural radiance fields (NeRF) have emerged for rendering photorealistic novel view synthesis from multiview imagery. In this project, we aim to compare NeRF variants for 3D mesh reconstructions. We propose to extract point clouds with accurate normal estimates to reconstruct 3D meshes that tightly fit the surface of NeRF's implicit volume representations. Thus, the output mesh can be directly used in downstream applications, reducing manual efforts in 3D modeling. Additionally, we show that self-supervised estimates of normals allow the model to learn a better regularized density field, leading to more accurate volume reconstruction, and improvements in pixel color reconstructions for better image rendering. Thus, we gained a comprehensive understanding of NeRF approaches for surface mesh reconstructions. Future work will build on this to make scenes editable and relightable for VFX.



Multi-modal machine learning for business-critical insights in video conversations

- **Student:** Guo, Zhongkang
- **Academic:** Prof. Anthony Bonner
- **Industry:** Ling Wang
- **Company:** Talka AI

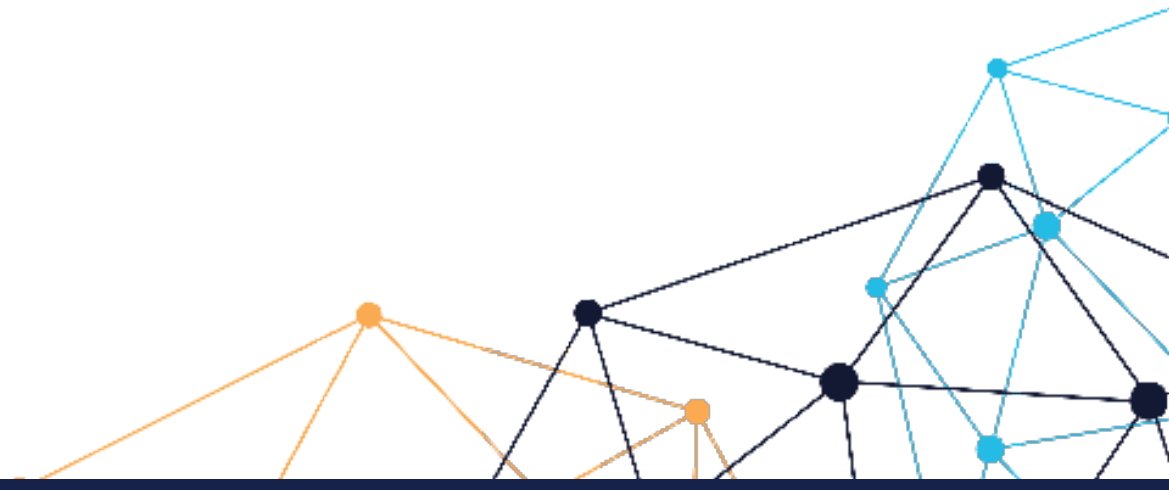
Conversation is the bedrock of solid business relationships. After the pandemic, online meetings are becoming more important in industrial fields, and contain different types of features for human communication. In this project, we use machine learning methods to extract features from online meetings and predict the deal success probability. Based on a dataset consisting of tens of thousands of recorded sales video, we extract different types of features, including text feature, audio feature, and visual feature, and further build neutral network models to combine the multi-modal features together to predict the success probability. By using different network structures, we achieve an accuracy of prediction of 80%, indicating that different types of features can be integrated to identify the key points of business meetings and allow product managers to know their products and customers better.



UAV-based 3D Multi-Object Detection for Indoor Applications using DCNN

- **Student:** Gupta, Pranav
- **Academic:** Steven Waslander
- **Industry:** Sunil Jacob
- **Company:** SOTI

Scene Understanding is essential for indoor autonomous navigation. While 2D object detection is a well-researched problem, there have been fewer advancements in the 3D space tackling indoor object detection. In this research project, we address the problem of estimating information of indoor objects from the environment and representing it in a 3D map. Th The objective is to develop an optimal algorithm that produces a three-dimensional detection in real-time on a UAV for further downstream tasks, such as obstacle avoidance, tracking and scene understanding. We optimize the FCAF3D neural network architecture to achieve real-time performance on an edge-compute device (Jetson Xavier NX). A robust data pipeline is developed to ingest the input obtained from any stereo and depth cameras, providing flexibility and scope for further advancement with improved hardware capabilities. While few datasets for indoor 3D object detection exist, we use SUNRGBD for training and benchmarking since it is open-source and contains RGB-Depth images obtained from real-world indoor environments, using stereo and depth cameras. Empirical results obtained using Intel RealSense camera show near real-time performance, with an of 63.8 on the SUNRGBD dataset.



Interpretable Market Segmentation with Time-Series Clustering

- **Student:** Hafez, Sarah
- **Academic:** Prof. Murat Erdogdu
- **Industry:** Angie Huang
- **Company:** Deloitte

Identifying customer segments is a key part of market growth as it helps businesses better identify and target customer needs. However, most valuable customer data is in time-series form, which complicates the segmentation process. While the prior literature suggests algorithms that can efficiently segment temporal data, it doesn't offer a framework on how to interpret its clusters. This project offers a methodology to efficiently segment time-series data in an interpretable manner.

We focus on a new representation of time-series data which extracts 41 features from each series; in doing so, it eliminates the problematic highly dimensional time-element of the data and replaces it with an interpretable description of its features (Baldán & Benítez, 2021). We nickname the resulting dataset as the flattened dataset. We make use of this flattened dataset in two experiments. Firstly, a state-of-the-art time-series clustering model is used to build the initial segments while the flattened dataset is later used in a classification model to find the cluster's important features. Secondly, the flattened dataset is used to drive the full clustering and feature importance process. We show that using the flattened dataset produces clear segmentation rules with definite improvements in interpretability.

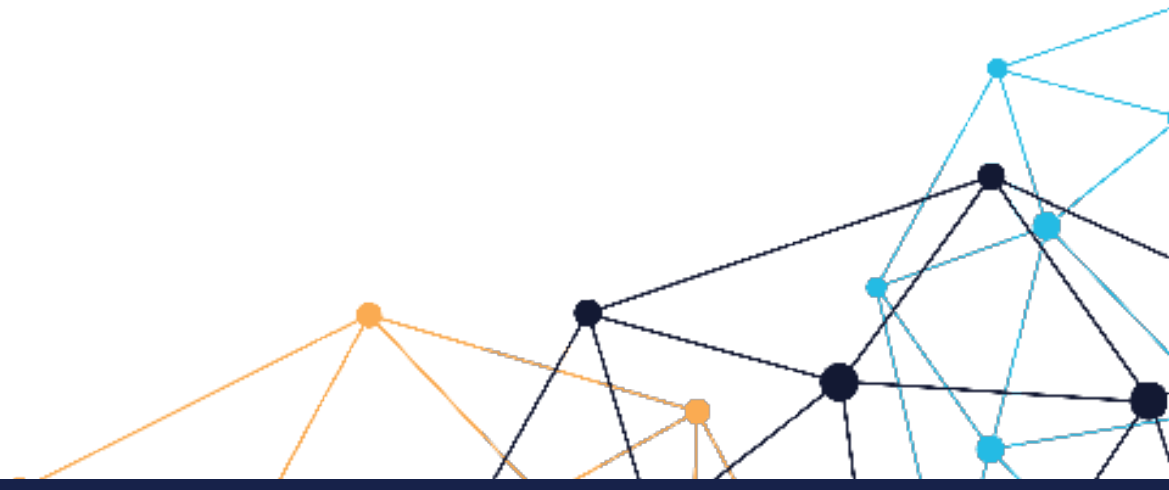
Lifelong Learning Strategies for Mobile Service Robots

- **Student:** Hu, Tianyu
- **Academic:** Prof. Animesh Garg
- **Industry:** Sharath Gopal
- **Company:** Bosch

Traditionally, software for home products could not be changed once shipped. Recent advancements in cloud infrastructure allow such products to collect a large amount of data and continuously update the software system throughout their lifetime.

In this project, we explored the use cases of lifelong learning in a mobile service robot for home applications. We implemented proof-of-concept systems that aim to improve user experience through improved performance and adaptation to user preferences.

Specifically, We implemented an Image segmentation model to identify clutter causing longer mission times using a LiDAR map robot using a heat map that accumulates the robot's visit frequency. In general, the average mission time can be reduced by 7.4% if the top 1/5 of clutter that the robot spends the most time around are removed. In addition, we are also working on a continual object detection strategy based on experience replay to allow the robot to learn new classes without the need for retraining on previous datasets. By saving a portion of images from previous tasks in memory, the model can robustly expand the range of objects it can detect while avoiding catastrophic forgetting.



Speech Enhancement and Recognition with Generative Adversarial Network

- **Student:** Huang, Jinda
- **Academic:** Prof. Gerald Penn
- **Industry:** Richard Wang
- **Company:** Pearson Canada Inc.

In order to improve speech enhancement and recognition, we measure the effectiveness of the NU-GAN, a generative adversarial network model for neural audio upsampling, for generating good-quality 16 KHz audio signals from inputs of 8 KHz audios. Since most of the automatic speech recognition (ASR) systems are trained on 16kHz sampling rate audios and above, and it's hard to collect more data from customers due to increasing privacy and cost concerns, we conduct the study on an 8kHz Pearson internal audio dataset, trying to boost the amount of data and enhance the quality of existing data, from a lower frequency(8 KHz) to a higher frequency(16 KHz). NU-GAN is a useful end-to-end deep learning tool, with the original audio spectrogram as inputs and targets. Inspired by [1], first, we use a similar NU-GAN architecture and train models on the DAPS dataset [2], we evaluate the performance of our model by the log-spectral distance (LSD) score. Then we apply the best NU-GAN model on the 8 KHz Pearson internal audio dataset and generate 16 KHz audio signals as training data for our modified Kaldi ASR pipeline [3]. Finally, we investigate the word error rate (WER) when measuring the performance of ASR model and we experimentally demonstrate that NU-GAN performs 1-2 % better than the baseline upsampling technique (sinc interpolation) on WER. The ASR model will help us to generate more accurate transcripts for inputs of various quality, and benefit the downstream tasks like scoring. More datasets and architectures will be introduced to further improve the NU-GAN performance in the future.

References

[1] Kumar, R., Kumar, K., Anand, V., Bengio, Y., & Courville, A. (2020). NU-GAN: High resolution neural upsampling with GAN.

[2] Gautham J. Mysore, "Can We Automatically Transform Speech Recorded on Common Consumer Devices in Real-World Environments into Professional Production Quality Speech? - A Dataset, Insights, and Challenges", in the IEEE Signal Processing Letters, Vol. 22, No. 8, August 2015. https://ccrma.stanford.edu/~gautham/Site/daps.html

[3] Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N.K., Hannemann, M., Motlícek, P., Qian, Y., Schwarz, P., Silovský, J., Stemmer, G., & Veselý, K. (2011). The Kaldi Speech Recognition Toolkit.

Explainer Divergence Score: Evaluating Neural Network Explanations Using Information Theory

- **Student:** Montero, Islas; Gabriel, Jose & Cardozo, Shea
- **Academic:** Prof. Anthony Bonner, Prof. Murat A. Erdogdu
- **Industry:** Dmitry Kazhdan
- **Company:** Tenyks

The widespread usage of deep neural networks (DNNs) in many sensitive settings has given rise to new methods to explain their predictions. 'Post-Hoc Explainers' are a class of technique that aims to explain a DNN with minimal assumptions about its specific architecture and training procedure. While many post-hoc explainers have been proposed, there has yet to be a systematic method for comparing the performance of different explainers when identifying common problems with DNNs.

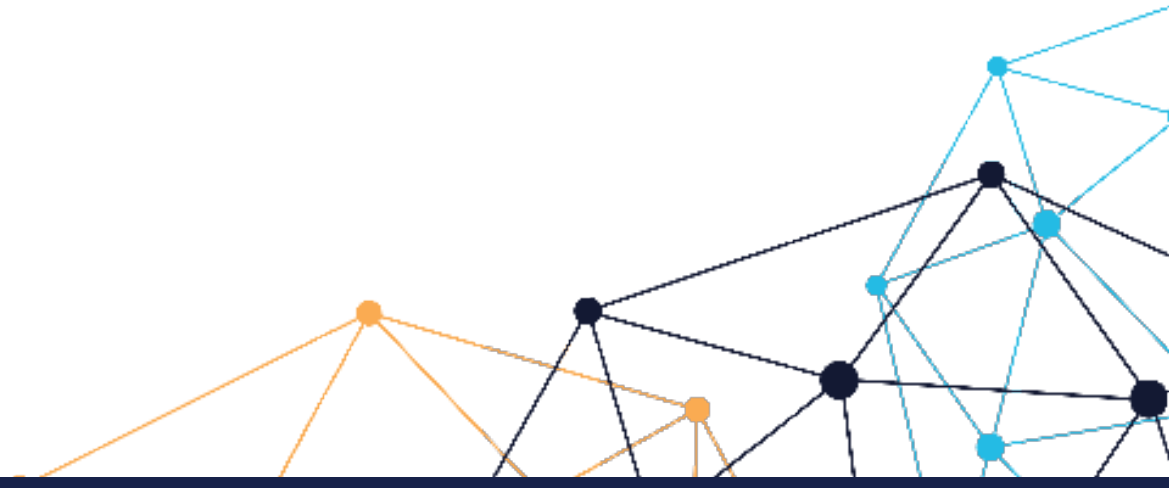
We propose a new evaluation procedure, Explainer Divergence Scores (EDS), that quantitatively evaluates a post-hoc explainer's ability to detect a common defect in DNNs - spurious correlations. Defined as the performance of a classifier trained to distinguish between explanations generated by defective and non-defective models, it is easy to interpret, naturally comparable, and grounded in an information theory analysis between explainers. We show EDS' effectiveness in a variety of settings, and find post-hoc explainers containing substantial amounts of 'hidden' information which can diagnose defects in DNN's but are imperceptible to a human user.

Design and Implementation of a Deployment Platform for Machine Learning models in a Distributed Manner

- **Student:** Kan, Boyang
- **Academic:** Prof. Ben Liang
- **Industry:** Mengliao Wang
- **Company:** Geotab Inc.

As more machine learning models are developed by data scientists, ways of using and serving these models to solve real time problems and generate predictions becomes important. The purpose of our project is to design and implement a machine learning model deployment platform, which supports deploying and serving models to a cluster of distributed servers. Users are able to specify the location of their models and the platform will be able to load the model from the given source and serve it as an application to receive different formats of prediction requests. It is able to correctly and efficiently handle all server failures, restarts, and autoscaling. The deployed applications should maintain reliability, consistency and correctness regardless of all kinds of network and server failures.

Our project can be separated into three phases: research, design, and implementation. Research includes investigating the related technologies and tools, such as databases and threading mechanisms. The design phase includes designing the workflow and structure that can maintain consistency throughout under high concurrency and load balance. Finally, implementation brings about the design and makes it scalable. The resulting solution uses google kubernetes engine as the servers and Bigquery as database to support both synchronous and asynchronous requests.



Calculating Similarity Between Bipartite Graphs for Auto-ML in Recommender Systems

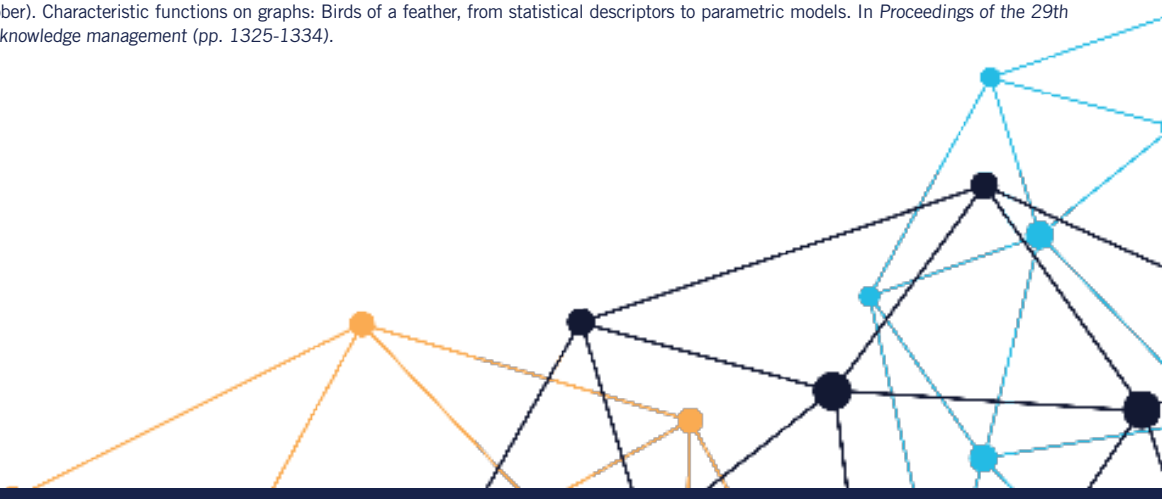
- **Student:** Kharosekar, Aditya
- **Academic:** Prof. Eldan Cohen
- **Industry:** Garrin McGoldrick
- **Company:** Crossing Minds

Automatic machine learning (Auto-ML) is the process of automating the iterative and time-consuming steps of the ML pipeline. One common Auto-ML approach is to transfer knowledge and configurations from related datasets to new ones. This requires a measure of similarity between datasets but existing research into dataset similarity is not naturally applicable to bipartite graphs, which are a major source of data for recommender systems.

In this study, we explore methods based on meta-features and graph embeddings to calculate similarity between bipartite graphs of user-item interactions. Such a similarity measure will speed up ML decision-making in Crossing Minds by allowing for easier transfer of knowledge between engineers and ultimately even a systematic transfer of ML configurations from one dataset to another. Using custom experiments that involve perturbing varying proportions of customer interaction graphs, we observe that hand-crafted dataset meta-features as well as certain whole-graph embedding methods such as FEATHER-G [1] can efficiently capture similarity in the structure of bipartite graphs. Future work will involve calibration and further analysis to ensure the similarity metric reflects human judgment.

References

[1] Rozemberczki, B., & Sarkar, R. (2020, October). Characteristic functions on graphs: Birds of a feather, from statistical descriptors to parametric models. In *Proceedings of the 29th ACM international conference on information & knowledge management (pp. 1325-1334).*



An Interpretable Unsupervised Clustering Algorithm for Understanding Consumer Utterances

- **Student:** Kolagati, Santosh
- **Academic:** Prof. En-Shiun Annie Lee
- **Industry:** Josh Seltzer
- **Company:** Nexxt Intelligence Inc.

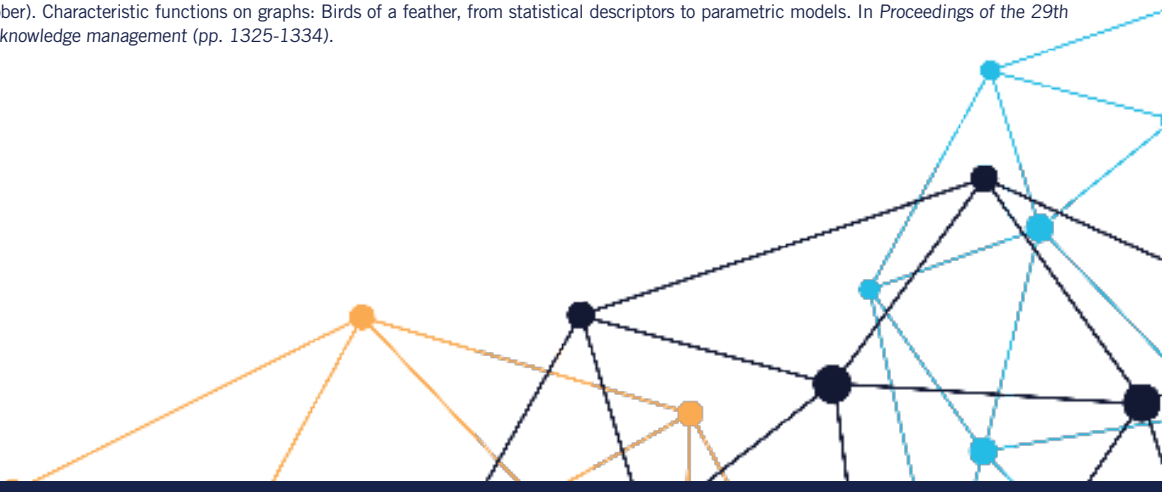
Market research allows companies to research what consumers think about their products by gathering information about consumers' needs and preferences. Within the market research industry, consumer language constitutes a wide variety of user-generated sentences such as opinions, feedback, complaints. Understanding and deriving insights from such utterances constitute a challenging problem. The contextual meaning of these utterances can vary broadly and may be specific to the domain, making it difficult to create a classification model that generalizes well. In this research project, we approach the problem of understanding consumer utterances by modelling an unsupervised clustering algorithm.

Using state-of-the-art techniques in natural language processing, particularly in semantic text similarity, named entity recognition, sentence segmentation, and one-shot learning, we devise a robust clustering algorithm that churns valuable insights to further the understanding of consumer language. Our algorithm produces sensible contextual themes and provides interpretable results by leveraging natural language features of interest to market researchers, marketing executives, and many others.

Our unique clustering model produces relevant clusters with automatically generated labels and significantly reduces the number of outliers by at least 30%. Furthermore, our one-shot classification technique provides 65% more perceived utility in gathering insights as opposed to purely unsupervised classification techniques.

References

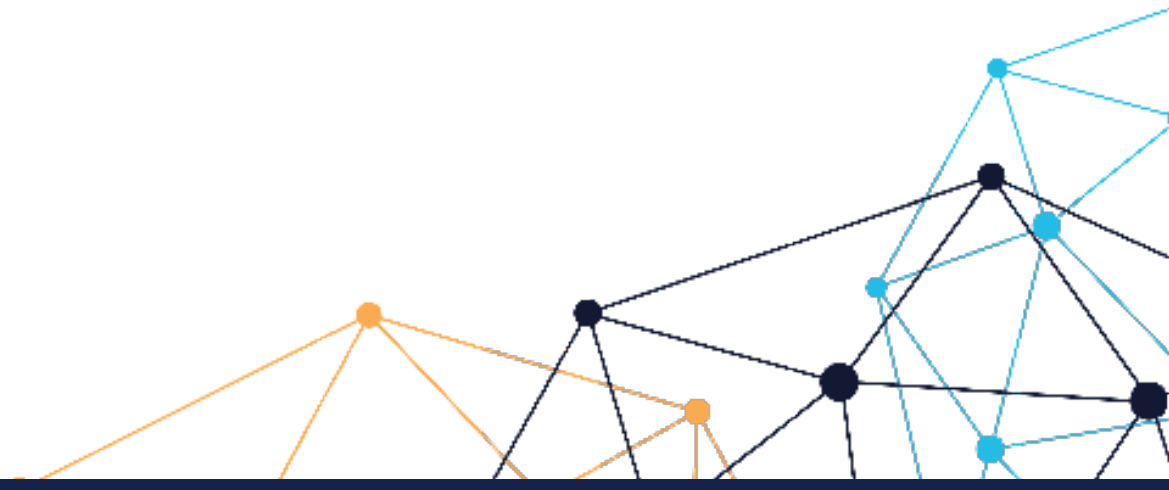
[1] Rozemberczki, B., & Sarkar, R. (2020, October). Characteristic functions on graphs: Birds of a feather, from statistical descriptors to parametric models. In *Proceedings of the 29th ACM international conference on information & knowledge management (pp. 1325-1334).*



Modelling the Evolution of Economic Discourse in the Business Outlook Survey

- **Student:** Koya, Charita
- **Academic:** Prof. Marsha Chechik
- **Industry:** Dave Campbell
- **Company:** Bank of Canada

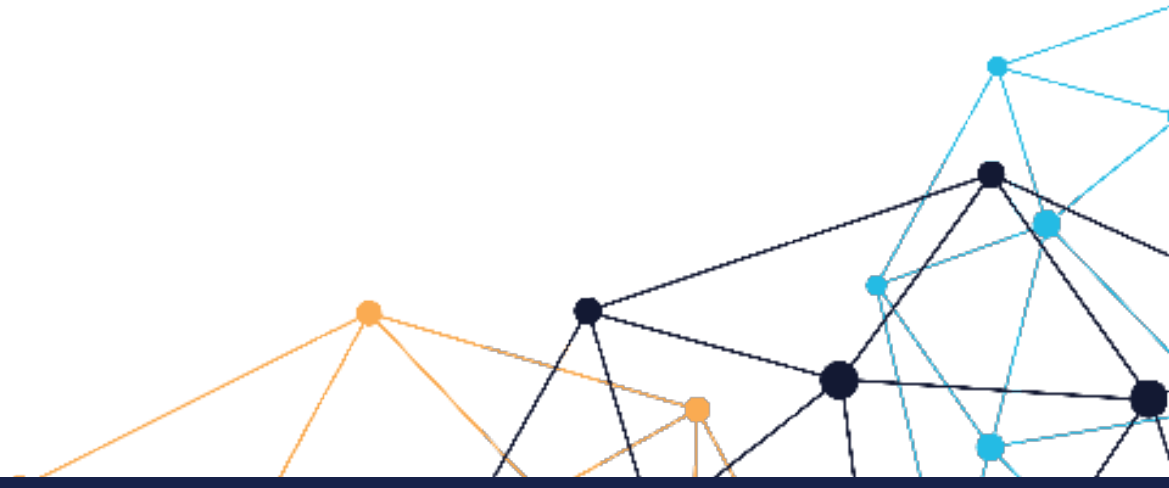
The Bank of Canada's quarterly Business Outlook Survey provides a timely narrative on the economic perspectives of firms across Canada. The raw data contains nearly 150 000 text responses from thousands of firms over a 13-year span, making it difficult to manually process and analyze. The goal of our project is to produce a meaningful summary of the responses and assess the evolution in economic discourse over time. This can help track changes in businesses' views on the economy, deepening the Bank's understanding of economic trends. Our experimental approach involved categorizing the survey questions by topic via non-negative matrix factorization with anchor words. This allows us to estimate covariate effects on topics of responses to establish any change in discourse over time. We validated our insights by looking at the economics of the past decade. Our results show meaningful trends in language about economic activity over time and over the COVID-19 pandemic. To our knowledge, this is the first time this approach has been effectively used on an economics dataset and in a time-series context. Future work will aim to improve accuracy by extending the analysis from a single topic word to n-gram anchor words.



Enhancing Private Market Data to Improve Performance of Predicting Startup Success

- **Student:** LeMesurier, Carey
- **Academic:** Prof. Ting-Kam Leonard Wong
- **Industry:** Brendon Freeman
- **Company:** CPP Investments

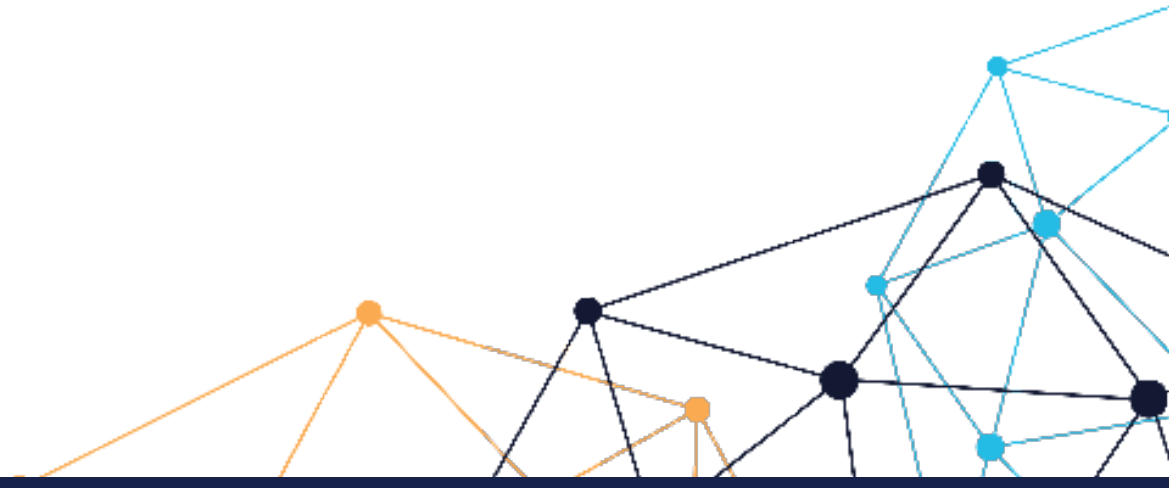
Predicting the future success of private companies is a non-trivial task, due to limited available data and complicated market dynamics. This research builds upon an existing baseline model which predicts success, defined by the target of having an Initial Public Offering (IPO) within 5 years. The goal of this work is to improve the performance of this model and gain insights into the challenges of the current approach. A significant challenge was found to be the poor data quality and coverage in private markets. Data Imputation was used to improve the applicability of the data. While the best performing method showed significant skill in predicting the features themselves, it led to very little improvement in the performance of the IPO model. Additionally, the significant target imbalance in the data makes our prediction task challenging. To account for this, it was found that carefully tuning model hyperparameters could improve the classifier's ability to accurately predict more positive classes, while maintaining overall performance metrics. Lastly, many new features were built to describe factors such as the macro environment, investor skill, and CEO experience. These features proved to be additive to the model, increasing the Area Under the Precision-Recall Curve (PR-AUC) by 12%.



Bridging the Gap Between Robustness and Accuracy with Partial Robustness

- **Student:** Li, Cassandra
- **Academic:** Prof. Anthony Bonner
- **Industry:** Ihab Amer
- **Company:** Advanced Micro Devices (AMD)

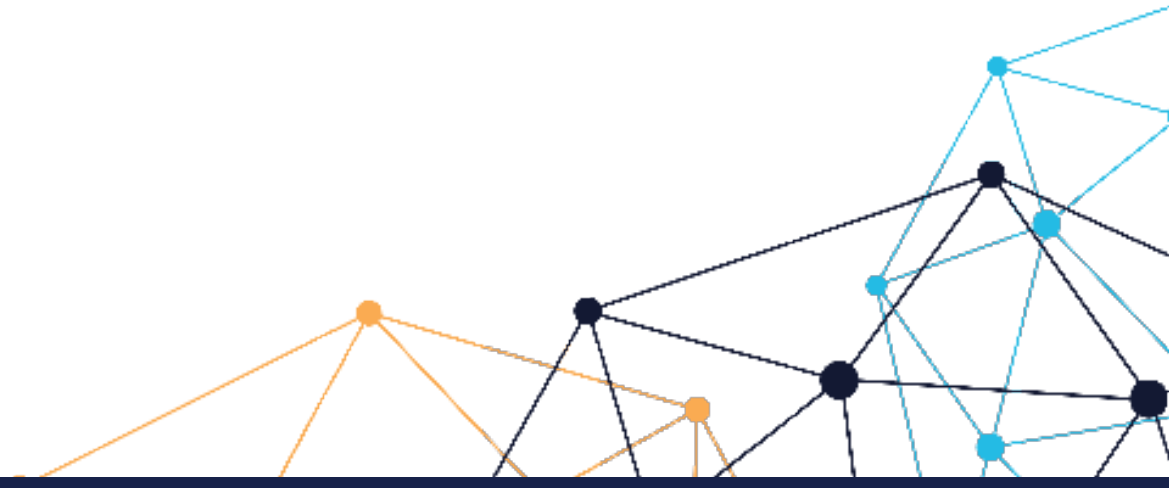
Deep neural networks (DNNs) have been applied widely in security-critical applications, from self-driving cars to identity-recognition technology. It has been shown that DNNs are vulnerable to adversarial examples – maliciously perturbed data samples that are perceptually indistinguishable from natural data to humans but are misclassified by DNNs. The state-of-the-art solution is to train the network on adversarial examples, which causes performance degradation on benign images compared to non-adversarial training. Notably, adversarial training improves the robustness of the model against adversarial attacks on all image classes. However, in some applications, such as image filtering, preventing an adversary from manipulating images from classes representing non-offensive content is unimportant. This leads to the question: "Can we increase the accuracy on natural images by sacrificing the robustness of non-critical classes?" To answer this question, we propose Partial Robustness – improve the robustness of the model only towards images in a set of protected classes in order to minimize the performance degradation on clean images. We address the problem of partial robustness by training a neural network on a combination of adversarial images from the protected classes and unmodified images from the unprotected classes. Our method achieves 80% robust accuracy on the protected class and ~9% increased accuracy on natural images when compared with an adversarially trained counterpart. These results demonstrate that Partially Robust models achieve higher clean accuracy than models trained with adversarial training without sacrificing the robustness of the protected classes. One open question to be addressed in further research is whether network pruning can be leveraged to increase the partial robustness of a network by removing the network connections exploited by adversarial attacks.



Exploration Guided Nonrigid-Body Manipulations

- **Student:** Li, Junjiang
- **Academic:** Prof. Animesh Garg
- **Industry:** Karime Pereida
- **Company:** Kindred

State-of-the-art research results in robotics and controls enable standard industrial robotic arms to perform general manipulation tasks of rigid objects with super-human efficiency and accuracy. However, when constraints on motions are involved, or when objects change during manipulation, one often must derive task-specific control schemes which requires significant time and effort to tune. In this work, we aim to improve the manipulation capabilities of standard industrial robotic arms by proposing a meta-algorithm for non-rigid body manipulation. Mimicking how humans interact with unknown objects, our approach first performs strategic in-hand exploration to gather information about the objects that are relevant to the task, and then acts on this information to achieve the desired outcome. We train the pipeline using reinforcement learning without providing prior knowledge of the manipulation task. We evaluate our approach on several manipulation tasks in which the non-rigidity of the object has a strong influence on the optimal control policy and demonstrate that the proposed approach can successfully perform the stated tasks in which context-unaware agents would fail.

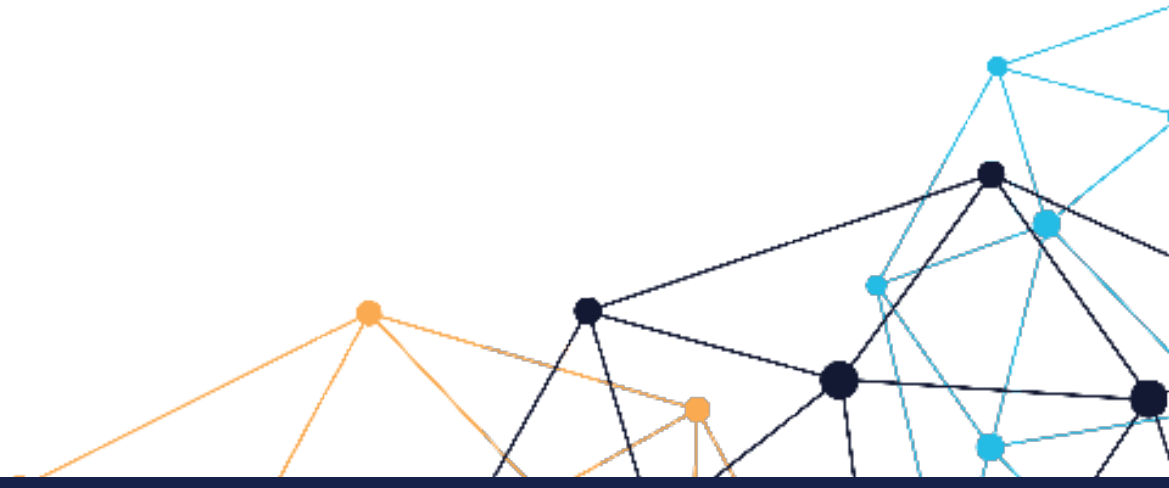


Acquisition AI: Applying Machine learning and optimization algorithms to automatically sort postal codes for potential profitability

- **Student:** Li, Wen
- **Academic:** Prof. Elias Khalil
- **Industry:** William Ives
- **Company:** Deloitte

Acquisition AI works to help banks allocate resources by postal code to different regions. Within a province a bank has a finite amount to spend on potential customers, thus they need to know how much to spend where based on postal code. To obtain this ranking, it is necessary to predict the acquisition rate of a new customer per postal code and optimize the monetary allocation by the acquisition rate. We have developed a predictive model for this task.

The core of our work is to build evaluation metrics and train predictive models based on designed targets. Due to the large number of 0s, we failed to directly predict the proportion of acquisition per postal code, so we introduced some new targets and then performed mathematical and statistical related transformations from predicted results. Ultimately, our approach obtained an average precision score of 70% and AUC ROC of 0.8 from different classification models(SVM, weighted logistic regression). We would continue to improve the performance of models and apply prediction results to score postal codes.



Exploring Three Methods of Bolt Looseness and Fracture Detection: Capacitive Sensing, Acoustic Emission Blind Separation, and Computer Vision

- **Student:** Lin, Zhimao
- **Academic:** Prof. Fae Azhari
- **Industry:** Tieliang Lv
- **Company:** Electronic Detection Lab at Fudan University

Bolted joints are widely used in hydropower plants. A failure of the bolted joint, especially the ones on the pump-turbine head cover, can cause catastrophic disasters such as flooding and power shortage. Currently, manual inspection during maintenance is the most common method adopted by hydro-power plants to detect the looseness and fracture of bolt joints. However, this method is prone to human errors and resource-intensive. Therefore, an online real-time bolt looseness and fracture monitor is in demand in the market for hydro generator condition monitoring. This project aims to explore the following three methods of detecting bolt looseness and fracture, which builds a foundation for further product development.

Capacitive sensing

The charge transfer principle is used to measure the capacitance at the level of picofarads. According to a simulation of the sensor circuit, we found it could measure a length change in the range of 0-3 millimeters with an accuracy of 3 micrometers.

Acoustic emission blind separation

A blind separation algorithm was evaluated by separating two human voices. As a result, the signal noise ratio (SNR) of the two separated signals are 15.63 dB and 7.61 dB respectively. Once the signal source is separated, further parametric and waveform analysis can be conducted.

Computer vision

An algorithm using IP cameras is developed to read the dial reading of the depth gauge. With this algorithm, a range of 0-3 millimeters depth change can be detected with a resolution of 1 micrometer.

Research and Implementation of Data Quality Improvement Algorithms for Vehicle Streaming Data Analytics

- **Student:** Liu, Mengyang
- **Academic:** Prof. Ben Liang
- **Industry:** Amy Zhang
- **Company:** Geotab Inc.

IoT devices have infiltrated every corner of our daily life. Many of those devices collect TBs or even PBs of IoT data and send it to a backend server for data analytics in a real-time manner. This kind of task belongs to stream data processing which is featured with low latency and unbounded incoming data. Apache Spark[1], Apache Flink[2], Apache Storm[3] are popular frameworks for stream data processing. However, each tool has its own niche speciality and mastery of all is difficult. To make the development easier, we leverage the power of Apache Beam[4] to implement a cross-language cross-platform stream data processing engine. Developers only need to implement the code once and run it over Apache Spark and Apache Flink. Furthermore, we proposed a useful duplication removal algorithm for stream data processing and applied it in vehicle collision detection problems.

References

[1] S. Salloum, R. Dautov, X. Chen, P. X. Peng, and J. Z. Huang, "Big data analytics on apache spark," International Journal of Data Science and Analytics, vol. 1, no. 3, pp. 145–164, 2016.

[2] P. Carbone, A. Katsifodimos, S. Ewen, V. Markl, S. Haridi, and K. Tzoumas, "Apache flink: Stream and batch processing in a single engine," Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, vol. 36, no. 4, 2015.

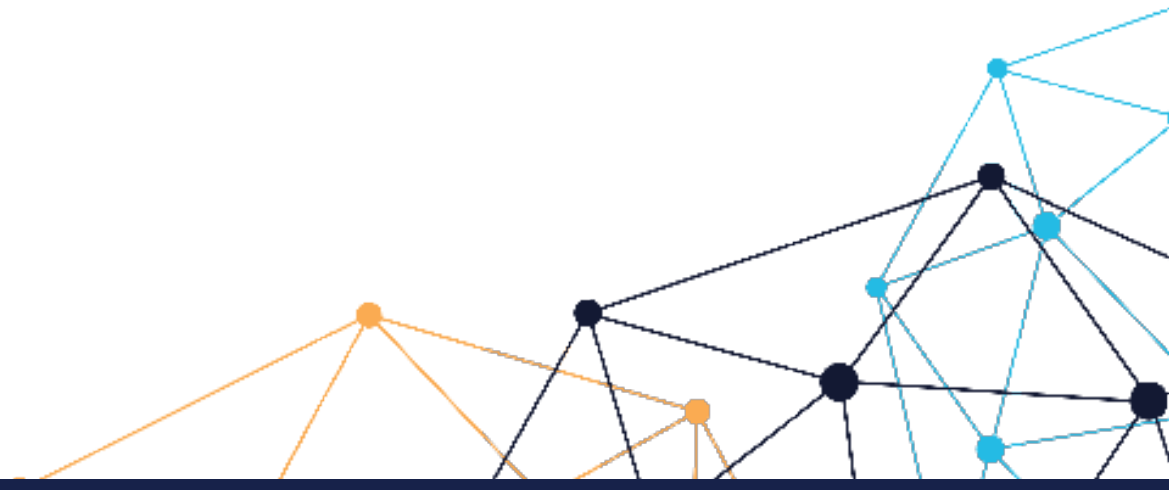
[3] M. H. Iqbal, T. R. Soomro, et al., "Big data analysis: Apache storm perspective," International journal of computer trends and technology, vol. 19, no. 1, pp. 9–14, 2015.

[4] T. Akidau, R. Bradshaw, C. Chambers, S. Chernyak, R. J. Fernández-Moctezuma, R. Lax, S. McVeety, D. Mills, F. Perry, E. Schmidt, et al., "The dataflow model: a practical approach to balancing correctness, latency, and cost in massive-scale, unbounded, out-of-order data processing," 2015.

Multi Sources and Sinks: An Adaptive Stream/Batch Big Data Ingestion Solution

- **Student:** Liu, Shanning
- **Academic:** Prof. Ben Liang
- **Industry:** Wenyang Liu
- **Company:** Geotab Inc.

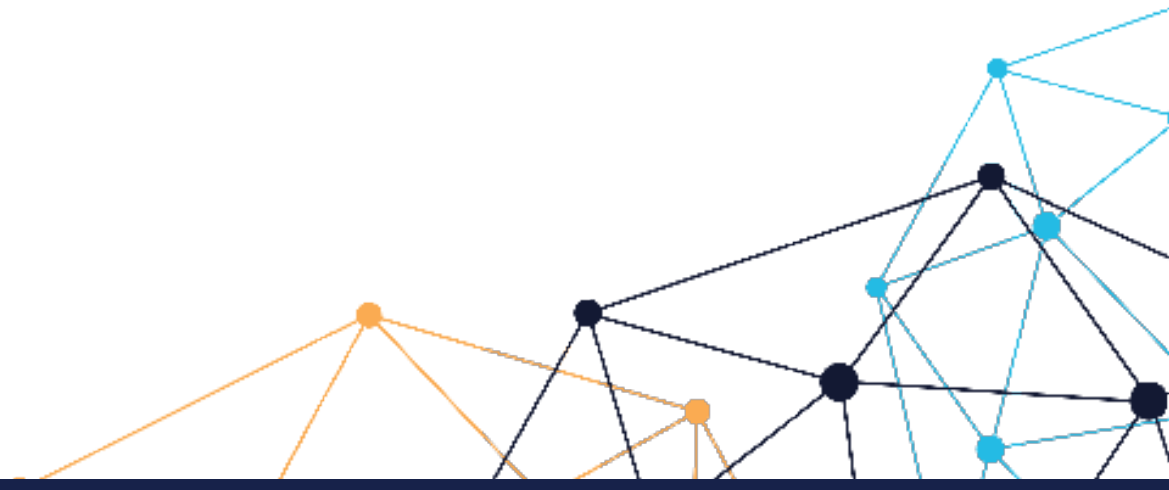
Tremendous amounts of data are generated by software and IoT devices making it important for data scientists to leverage this data to deliver meaningful insights. Considering the nature of IoT devices, the traffic they generate will fluctuate as they will come from many sources and sinks. In view of delivering this kind of data to data scientists, the ingestion infrastructure needs to be scalable and extensible in the production, therefore, the capability to adapt to different data sources/sinks and compatible batch and streaming processing simultaneously is required absolutely. In this report, we propose a cloud-based adaptive big data ingestion solution that can be applied to the problems by utilizing dynamic connector extensions to accommodate multiple data sources and data sinks together with different ingestion methodologies. Furthermore, we provide procedures for implementing and introducing the extensions into the solution. In addition, several validation scenarios and corresponding results of the solution's big data ingestion capability have been described. To be concluded, the solution is production-level stable, and functional for multiple sources/sinks and streaming/batch ingestion scenarios.



Zero-shot Text-to-Image Generation via Amortizing GAN Inversion

- **Student:** Liu, Zhaoyan
- **Academic:** Prof. Jimmy Ba
- **Industry:** Gabriel Loaiza
- **Company:** Layer 6

Recent work in text-to-image generation has achieved remarkable success, making headlines even outside of the machine learning community. Training these models requires internet-scale datasets and correspondingly massive amounts of compute. In this work we show that text-to-image models can be obtained for much cheaper by leveraging pre-trained models, even when these were not originally designed for this task. Our work combines pretrained CLIP, any caption model, and any GAN, and has three key steps: (1) Generate a synthetic dataset by sampling from the GAN; (2) Train a translator network to map from CLIP to GAN latent space using curriculum learning, essentially amortizing GAN inversion; (3) at sample time, we use Langevin dynamics to further improve upon the output of the translator network. Importantly, our entire procedure only trains the relatively simple translator network, and does not require access to any pre-existing dataset since the used data is synthetically generated through the pre-trained models. Our method achieves 25.8 FID score and 0.35 augmented CLIP score on MS-COCO without training on this dataset, faster sampling than optimization-based methods, and better out-of-distribution generation than competing alternatives.



Machine Learning Techniques for Speech Enhancement in Audio Conferencing Systems

- **Student:** Ma, Boyi
- **Academic:** Prof. Nick Koudas
- **Industry:** Paul Ferry
- **Company:** Nureva Inc.

Audio conferencing systems must address three main technical problems in all meeting spaces: reverberation, background noise, and echo. Traditional signal processing techniques have been developed for decades to try to solve these problems, but they all have limitations, especially with non-stationary noises. This project aims to explore the capabilities and limitations of ML techniques for speech enhancement and how they can be integrated with Nureva's audio technology.

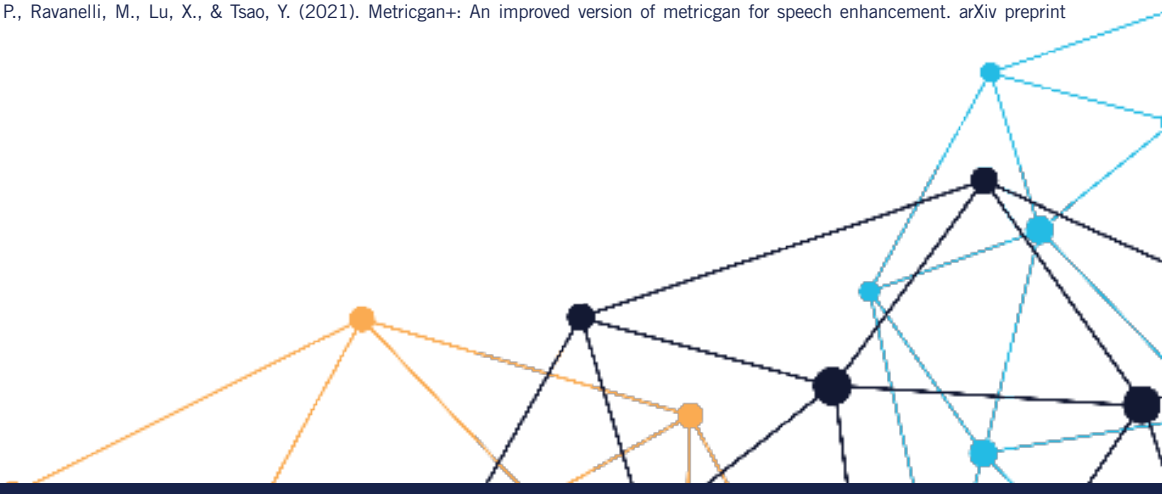
We started by evaluating open-source architectures using DNSMOS [1], which predicts a model's noise suppression performance, and we found significant degradations in real-world scenarios. To tackle that, we fine-tuned MetricGAN+ [2] with more realistic data, and the results show improved DNSMOS noise scores by over 0.5. However, we also observed that the built-in denoising features of Unified Communications and Collaboration (UC&C) clients would interfere with the audio and generate worse results overall if our model's denoised input was fed in.

We conclude that general ML denoising research faces the challenge of degraded performance when put into real-world applications. This can be improved by utilizing more realistic data in training, but interoperability with UC&C clients shall also be investigated.

References

[1] Reddy, C. K., Gopal, V., & Cutler, R. (2022, May). DNSMOS P. 835: A non-intrusive perceptual objective speech quality metric to evaluate noise suppressors. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 886-890). IEEE.

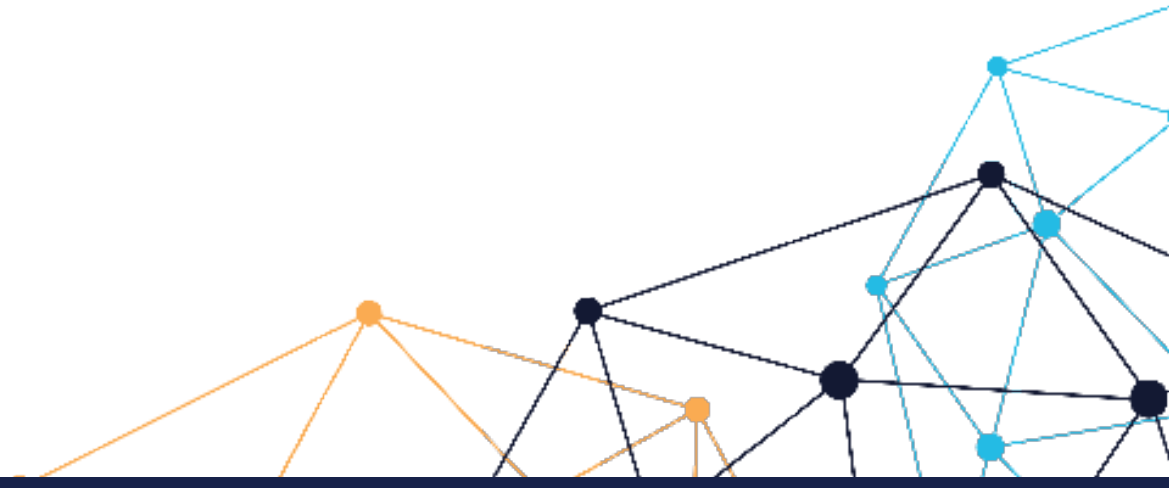
[2] Fu, S. W., Yu, C., Hsieh, T. A., Plantinga, P., Ravanelli, M., Lu, X., & Tsao, Y. (2021). Metricgan+: An improved version of metricgan for speech enhancement. arXiv preprint arXiv:2104.03538



Mobile Seizure Detection

- **Student:** Miscevic, Filip
- **Academic:** Prof. Alex Mariakakis
- **Industry:** Gerard O'Leary
- **Company:** NerveX Neurotechnologies

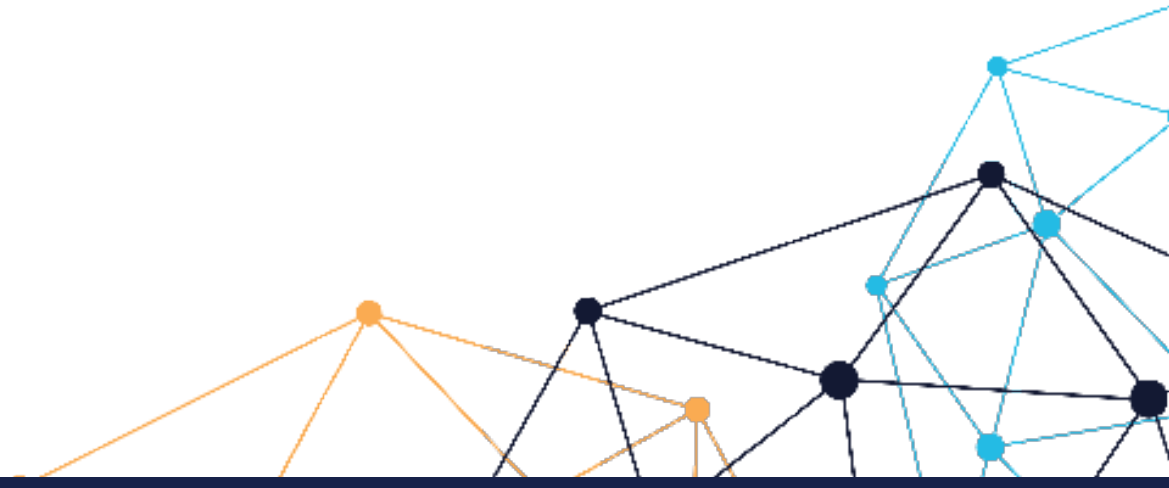
Anti-seizure medications are ineffective for approximately one third of people with epilepsy, but electrical stimulation of the vagus nerve offers an effective alternative. NerveX Neurotechnologies is developing medical devices suited for this purpose. This involves the development of an automated machine learning pipeline for personalized seizure detection. An unsupervised approach is initially used to learn individual patient biomarkers and identify potential seizures. Once these are corroborated with the patient's seizure log, a supervised machine learning model that runs efficiently on battery-powered hardware is trained and transferred into the device to detect seizures on the fly and apply vagus nerve stimulation. Applying this approach on the CHB-MIT seizure dataset, we achieve a clinically acceptable seizure sensitivity of 0.90 with a false positive rate of 0.1. The development of this pipeline represents a major milestone in leveraging machine learning for personalized epilepsy treatment with medical devices.



Multimodal learning with Transformer for sparse and irregular data

- **Student:** Pokhrel, Aslesha
- **Academic:** Prof. Rahul G. Krishnan
- **Industry:** Saba Zuberi
- **Company:** Layer 6

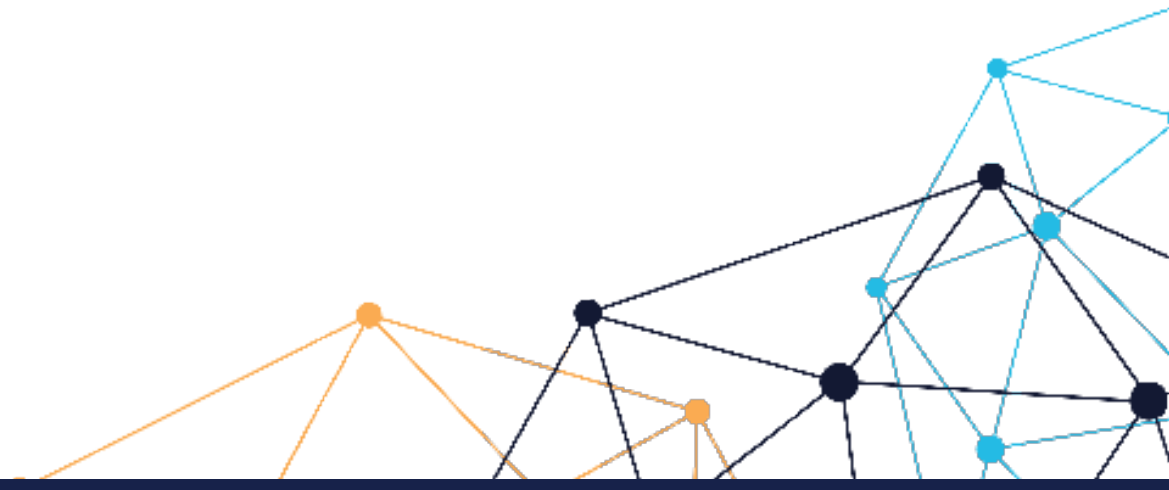
Real-world health and finance data are multimodal in nature, including text, images, time series, and tabular data, and characterized by high sparsity and irregular observations. Transformer-based models have demonstrated unparalleled performance in their application to computer vision, natural language processing and speech recognition tasks. However, their use in modelling sparse irregular time series data with other modalities such as images, text and tabular data has not been widely explored. In this work, we introduce a self-supervised Transformer-based architecture designed to extract robust representations from multimodal electronic health record (EHR) data. As Transformers are limited in their input sequence length due to the quadratic scaling of the self-attention layers, we propose an input binning scheme for the time series that compresses highly correlated data into small time bins. We also introduce multiple self-supervised prediction tasks that provide rich and informative signals for model pre-training. We show that our model learns useful representations and outperforms state-of-the-art deep learning models on the MIMIC-IV mortality prediction task.



Multimodal Procedure Understanding

- **Student:** Abdollahpouri Hosseini, Seyed Ahmad
- **Academic:** Prof. Frank Rudzicz
- **Industry:** Afsaneh Fazly
- **Company:** Samsung AI Center - Toronto

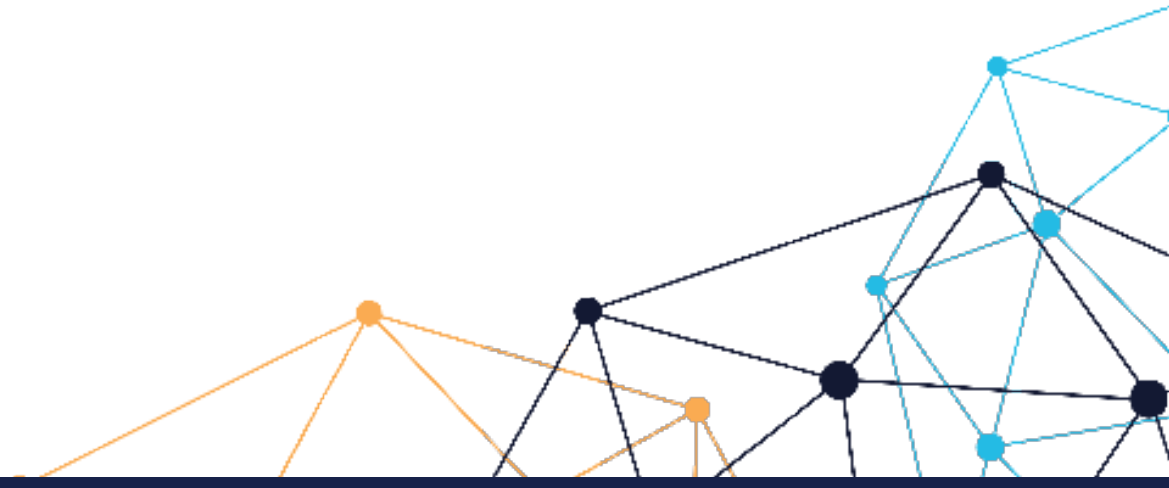
Deep learning has been very successful in Natural Language Processing, but the scarcity of annotated data poses significant challenges for supervised models, especially in niche domains, such as cooking recipes. The most recent method for parsing flow-graphs from recipes involves a Conditional Random Field (CRF) tagger and a biaffine dependency parser that uses the tagger's output. This model uses a small, annotated dataset, resulting in low generalizability. To address this, we begin with error analysis, revealing interesting error patterns and highlighting some problems with the annotation scheme. We then propose a solution combining ideas from multi-task learning, silver data usage, and pre-trained large language models. We merge the tagging and parsing stages into a single pipeline that simultaneously does both. This model is then used to generate labels on a large unannotated dataset, which we use to train better models. We achieve 1.35% and 0.99% f1-score improvement for labelled and unlabelled edge prediction, respectively. However, due to significant gaps between the validation and test set performances, we propose a more robust evaluation framework, which is not limited to a single validation/test split. Under this framework, our method no longer shows improvements over the baseline, suggesting that the improvements might have been side effects of that particular data split. We suggest future work adopt a robust evaluation strategy, such as ours.



Automated Code Fix Suggestions

- **Student:** Prakash, Yash
- **Academic:** Prof. Shurui Zhou
- **Industry:** Yonas Bedasso
- **Company:** Advanced Micro Devices (AMD)

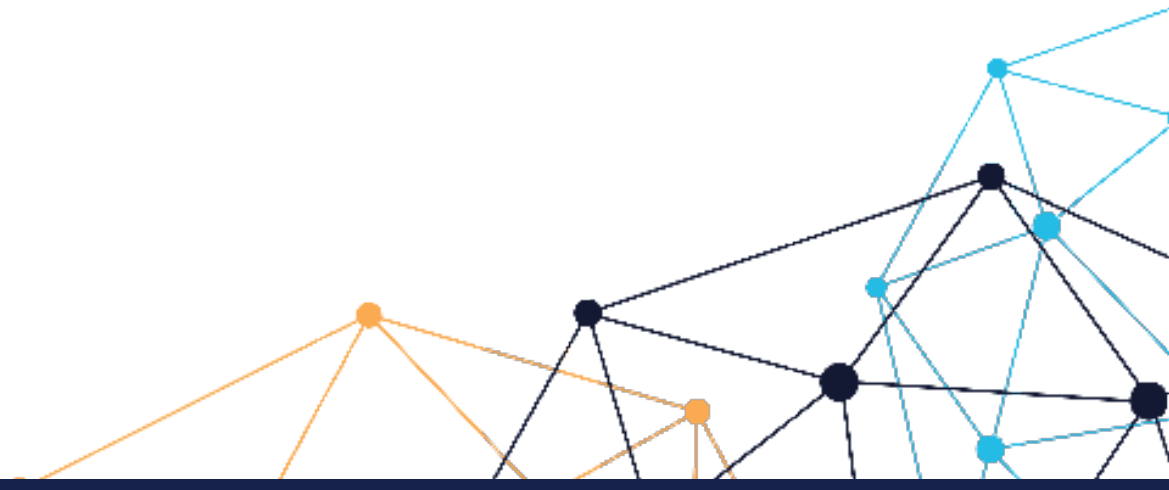
Mature software systems, such as AMD's driver codebase, regularly undergo a high rate of code changes, typically from a combination of dedicated developers and testers. As such, early detection of issues is necessary for maintaining software quality. Since bug detection and fixing is such a time and labor-intensive process, we sought to develop an automated solution to address the difficulty. Using careful extraction of changed code blocks and usage of syntax tree transformations, we pinpoint structural issues in the changed code and automatically detect similar issues in a set of applicable release configurations for the project. This successful identification of the recurring issue also allows us to generate necessary fixes to resolve those issues in those corresponding releases. We are aiming to qualitatively evaluate the process in our software development lifecycle to show its significant time efficiency for detecting and producing fixes for issues while reducing manual intervention and being computationally efficient.



Inventory availability forecasting for PC Express

- **Student:** Qu, Yang
- **Academic:** Prof. Murat A. Erdogdu
- **Industry:** Emir Kavurmacioglu
- **Company:** Loblaw Digital

At Loblaw Digital, one of our missions is to optimize customers' online shopping experience. One avenue is to mimic in-store shopping experiences where customers have visibility on the quantity of items on-shelf directly when they are building their shopping carts. PC Express is an online shopping service that allows customers to build their online carts and select a pick up timeslot and store. There is a possibility that at the time when their items are picked by the in-store pickers, the items may be sold out which would lead to a substitution or shortage in the items received by the customers. To avoid having customers disappointed at not being able to receive what they ordered, we use Machine Learning to build a model so that customers can know how likely the items they ordered will be available ahead of time. More specifically, we use the transaction logs, fulfillment events, replenishment events, etc as signals around the on-shelf availability and build a Hidden Markov Model to predict the probability of finding more than x items on-shelf at a given timeslot and store. By enabling this capability, we also open the doors for improving the fulfillment process in a cost-effective manner.



Data-Efficient Self-supervised Learning for Video Action Recognition

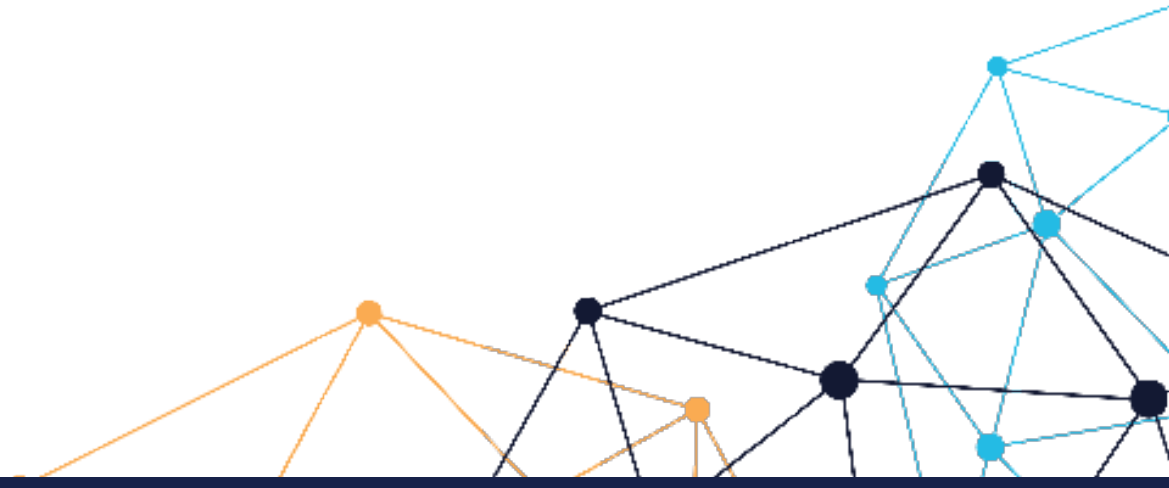
- **Student:** Ren, Weiming
- **Academic:** Prof. Animesh Garg
- **Industry:** Iqbal Mohomed
- **Company:** Samsung AI Center - Toronto

State-of-the-art methods for video action recognition require large-scale pretraining datasets and show improved generalizability on a larger finetuning dataset. However, access to large-scale video training data can be restricted in real-world applications due to technical issues and legal concerns, such as storage limitations and commercialization issues. We study the problem of building robust video action recognizers in a low data regime, where we aim to minimize the requirement for both pre-training and finetuning data. To this end, we propose a data-efficient self-supervised learning (SSL) method for action recognition, which aligns the representations of two different clips from semantically similar videos to learn spatiotemporal-persistent representations that generalize well on downstream action recognition tasks. Our experiments on a real-world indoor action dataset and several public action recognition benchmarks suggest that our method achieves better generalizability than models without pretraining, achieving +7.1% mAP and +10.2% top-1 accuracy on ActivityNet action recognition. However, we still observe a relatively large gap comparing our method with large-scale supervised pretraining, suggesting further improvements to our SSL framework. As a next step, we plan to investigate other video understanding tasks that can be benefited from our SSL framework, as well as understand what characteristics constitute a good dataset for efficient few-shot video understanding.

Automatic Optical Character Recognition (OCR) Pre-Processing and Recognition for Custom Gameplay Text

- **Student:** Roy, Subhayan
- **Academic:** Prof. Babak Taati
- **Industry:** Yonas Bedasso
- **Company:** Advanced Micro Devices (AMD)

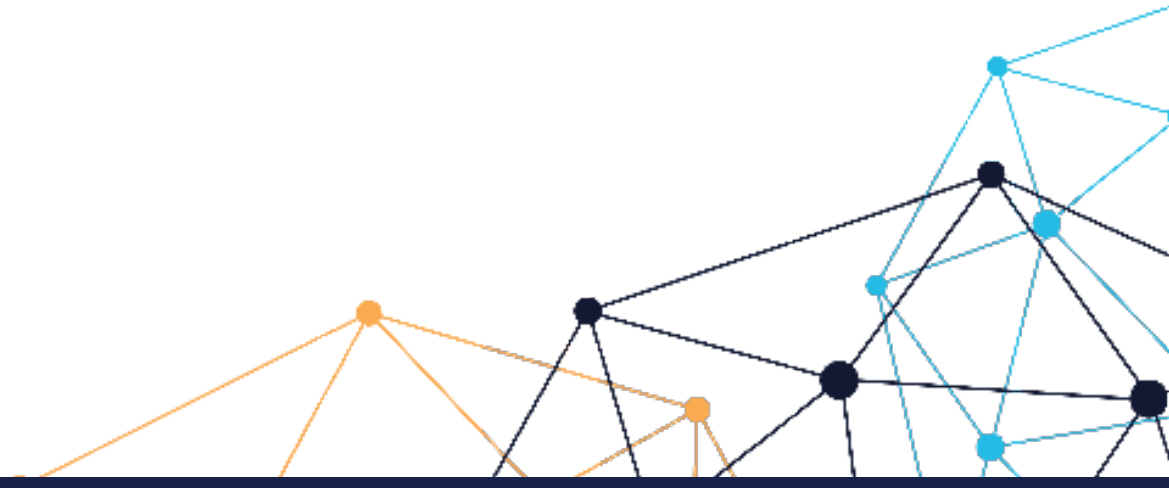
Computer Games are one of the key use cases for AMD Graphics cards. AMD performs extensive testing of its graphics hardware and software. However, majority of gameplay testing is manual and requires significant efforts. Text plays a significant role in automating gameplay testing as it allows to navigate through game menus and scenes. Recognition of such heavily styled and colored instances of text found in games with varied backgrounds is a challenging task which currently requires manual tuning of thresholds. The primary goal of this project was to implement a machine-learning solution that can automatically pre-process any game scene to produce a black/white output of text and background. We achieved this goal by implementing a U-Net based architecture for segmentation of background and text. We achieved a Validation Jaccard similarity score of 95% and text recognition accuracy of 82.5%. The secondary goal of the project was to develop a fully autonomous OCR solution that can handle text recognition on its own. A fully convolutional architecture (EAST) was adopted to do real-time text detection which was followed by making use of convolutional recurrent neural networks for text recognition. Thus, this project eliminated the need for manual tuning of thresholds for producing the optimal binarized image and streamlined the OCR process.



Evaluating Impact of Data-Centric Approaches on Robustness

- **Student:** Saini, Parul
- **Academic:** Prof. Kelly Lyons
- **Industry:** Hossein Taghinejad
- **Company:** SOTI

Text-to-SQL systems have the potential to empower users to get insights from their data without having to be adept at query languages. The reliability of such systems vastly depends on the robustness of the models in the presence of noisy and out-of-distribution data. This research focuses on two data-centric approaches for achieving robustness: (1) Natural Language Diversification generated using T5 by modifying the fine-tuning strategy and opting for a diversity-promoting decoding strategy during inference. Filters based on N-gram overlap and semantic correctness are leveraged to circumvent the quality-diversity tradeoff; (2) Data Augmentation to supplement the data with perturbed samples that imitate real-world noise. Previous work has explored these two approaches independently to improve the performance and robustness of machine learning models. We combine the two techniques to evaluate the impact of augmenting data with perturbed diversified samples in a Text-to-SQL system. To analyze the robustness accuracy, stress tests are created for a given dataset consisting of questions with lexical variations and synthetic perturbations. Experiments using different in-house datasets have shown that models trained using our approach are more robust and show accuracy improvement ranging from 2-10% on the perturbed set.

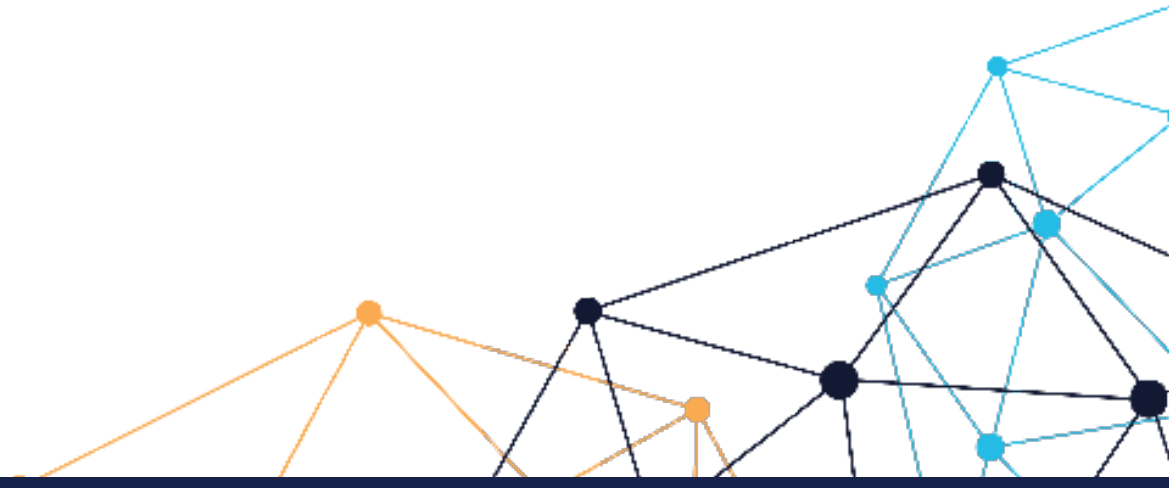


Forecasting Patient Flow Pressures with Machine Learning Models at St. Michael's Hospital

- **Student:** Shang, Siyuan
- **Academic:** Prof. Igor Jurisica
- **Industry:** Derek Beaton
- **Company:** Unity Health Toronto

In any hospital, having insight into patient flow throughout the hospital is essential to resource planning and operational efficiency. When patients are admitted, discharged, or transferred, several actions need to be taken to ensure patients receive timely care and resources do not become backlogged. Improving patient flow reduces wait times, improves operational efficiency, and ultimately improves care. However, due to reasons like the epidemic outbreak, traditional methods of controlling the flow of patients are no longer effective. The goal of this project is to seek a better way to control patient flow, efficiently by utilizing state-of-the-art machine learning models. Specifically, these models can forecast various indicators that are important in managing hospital pressures.

Our solution utilizes multiple data sets for training, which minimize and reduce the risk of bias and eliminate the data shift influence caused by the pandemic. We used both robust and advanced models like lasso regression, XGBoost, and neural network in an ensemble way to make optimal predictions. With the help of retrospective models training workflow, we determine the best hyper-parameters and obtained a high AUC score (0.85-0.9). The result demonstrated machine learning technique is a feasible way to improve health care system efficiency. The next step of this project is to build and deploy a real-time pipeline with a self-auditing function to monitor its performance in daily use.



Characterizing cellular phenotypes with self-supervised learning

- **Student:** Sharma, Vasudev
- **Academic:** Prof. Rahul G. Krishnan
- **Industry:** Oren Kraus
- **Company:** Recursion Pharmaceuticals

High-Content screening (HCS) systems that combine fluorescence microscopy with high throughput biotechnology have become potent systems for studying cell biology and drug discovery. Accurately quantifying the biological meaningful effects of millions of chemical and genetic perturbations on cells is important to the scientists and remains an open research problem. Earlier approaches, such as features-based or supervised methods, suffer from several limitations, including pre/post-processing or the need for annotations for learning robust representations.

To circumvent these limitations, we investigate self-distillation-based self-supervised models, specifically DINO (Caron et al., 2021), on HCS datasets. We use one of the publicly available datasets, BBBC021 (Broad Institute), consisting of human breast cancer cells. To evaluate our DINO model, we see how well treatments are assigned to the correct Mechanism of Action (MOA) based on a 1 nearest neighbor search. On the evaluation, we obtain 94% and 88% MOA accuracy on NSC (not-same-compound) and NSCB (not-same-compound-and-batch) respectively trained on patch-level BBBC021 images. Interestingly, even without post-processing (Typical Variation Normalization), we do not see a considerable drop in the performance, unlike earlier supervised and weakly supervised approaches. The preliminary results showcase self-supervised methods as a promising alternative compared to weakly supervised methods for learning cellular representations.

References

Mathilde Caron et al. "Emerging Properties in Self-Supervised Vision Transformers". In: arXiv (2021). eprint: 2104.14294.

Human MCF7 cells – compound-profiling experiment | Broad Bioimage Benchmark Collection. https://bbbc.broadinstitute.org/BBBC021.

Pawlowski, N., Caicedo, J. C., Singh, S., Carpenter, A. E., and Storkey, A. (2016). Automating Morphological Profiling with Generic Deep Convolutional Networks. bioRxiv, page 085118.

Perakis, A., Gorji, A., Jain, S., Chaitanya, K., Rizza, S., and Konukoglu, E. (2021). Con- trastive Learning of Single-Cell Phenotypic Representations for Treatment Classification. arXiv.

Singh, S., Bray, M.-A., Jones, T., and Carpenter, A. (2014). Pipeline for illumination correction of images for high-throughput microscopy. Journal of microscopy, 256(3):231– 104.

Ljosa, V., Sokolnicki, K. L., and Carpenter, A. E. (2012). Annotated high-throughput microscopy image sets for validation. Nature Methods, 9(7):637–637

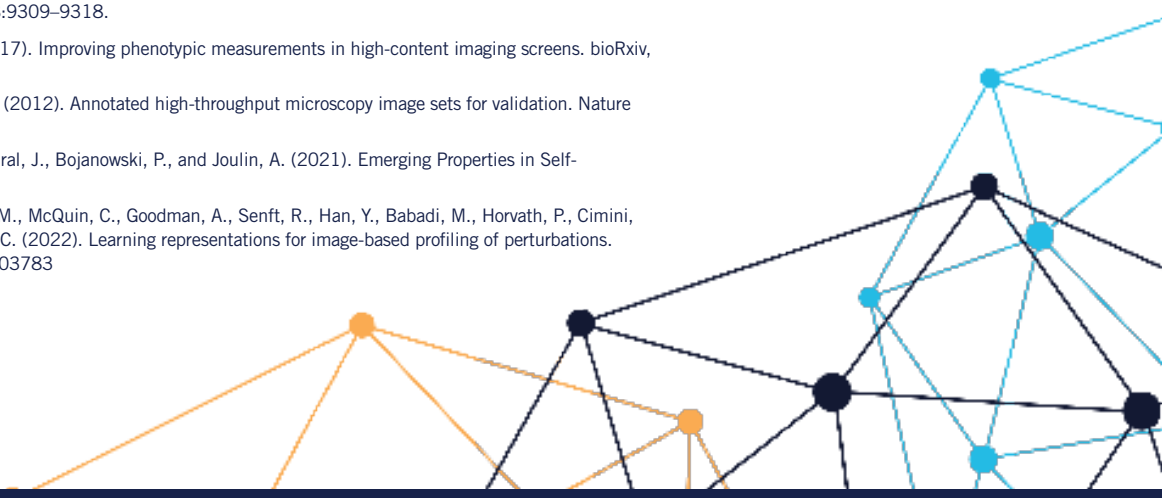
Caicedo, J. C., McQuin, C., Goodman, A., Singh, S., and Carpenter, A. E. (2018). Weakly Supervised Learning of Single-Cell Feature Embeddings. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018:9309–9318.

Ando, D. M., McLean, C. Y., and Berndl, M. (2017). Improving phenotypic measurements in high-content imaging screens. bioRxiv, page 161422.

Ljosa, V., Sokolnicki, K. L., and Carpenter, A. E. (2012). Annotated high-throughput microscopy image sets for validation. Nature Methods, 9(7):637–637

Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., and Joulin, A. (2021). Emerging Properties in Self-Supervised Vision Transformers. arXiv.

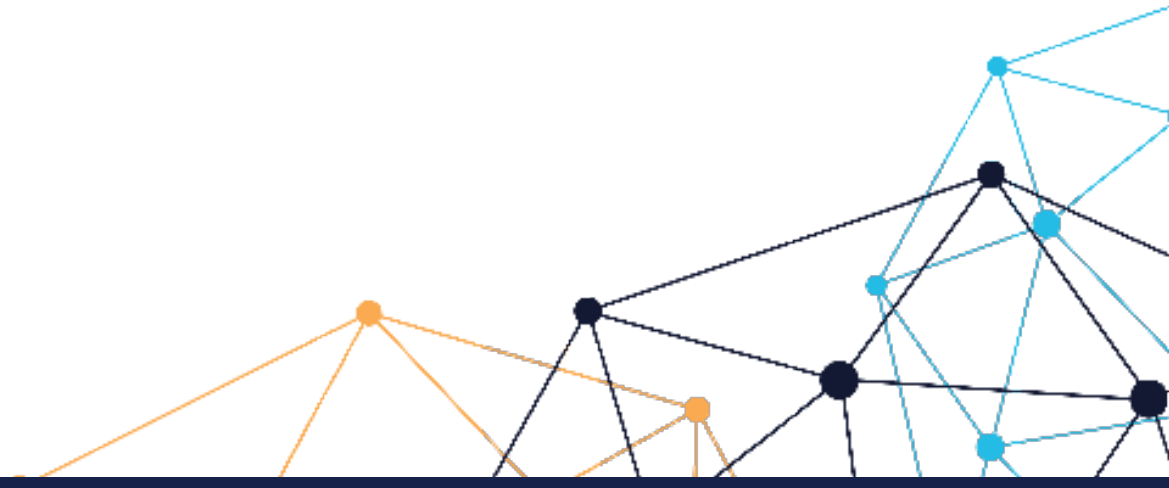
Moshkov, N., Bornholdt, M., Benoit, S., Smith, M., McQuin, C., Goodman, A., Senft, R., Han, Y., Babadi, M., Horvath, P., Cimini, B. A., Carpenter, A. E., Singh, S., & Caicedo, J. C. (2022). Learning representations for image-based profiling of perturbations. BioRxiv. https://doi.org/10.1101/2022.08.12.503783



Exploring faulty code with artificial intelligence

- **Student:** Shi, Haoxuan
- **Academic:** Prof. Shurui Zhou
- **Industry:** Benjamin Reaves
- **Company:** Metabob

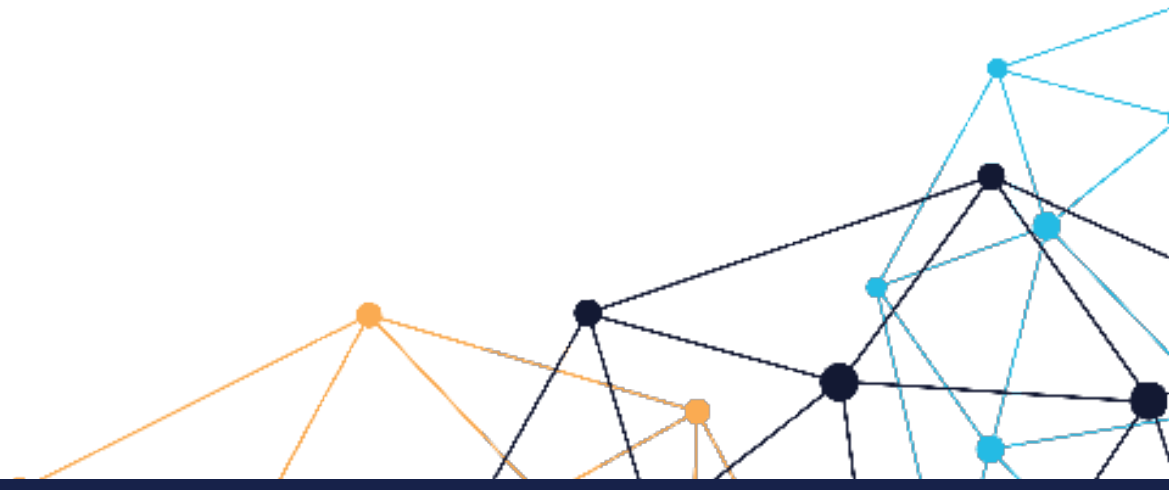
Software developers spend ~50% of their programming time dealing with software bugs (a.k.a., faulty software code). Once a software bug is encountered, developers attempt to find its original location in the code and then understand its erroneous behaviours, which is a prerequisite to correcting any bug. Although there has been significant research to find the bugs automatically, only a little attempt has been made to explain the faults in the code that trigger these bugs. In this project, we designed and develop a novel, intelligent framework to find topics and relationships between faulty software code with one major objective: organize, explain, understand and summarize faulty software code. First, we designed a topic modelling module by training a latent dirichlet allocation model using millions of pull requests and issue texts from GitHub that contain only faulty code as it's unsupervised learning. Second, apply various models of topic modelling with different focuses to better fit the dataset. For example, biterm topic model has greater metrics when facing a short-text dataset compared to latent dirichlet allocation. We evaluated our designed solutions both empirically using historical data from GitHub and qualitatively involving professional developers. Unlike rule-based, non-scalable alternatives (e.g., static analysis tools), this project offers data-driven, scalable, intelligent solutions that can learn from millions of faulty codes and their corresponding explanations stored on GitHub. It will not only advance the current state of research for bug understanding but also will produce methods and tools that will be adopted by the software industry and facilitate programmer efficiency.



Document Verification Using Visual Question Answering

- **Student:** Singh, Diljot
- **Academic:** Prof. Frank Rudzicz
- **Industry:** Andrew Kujtan
- **Company:** Scotiabank

Scotiabank's existing document understanding system verifies trade contracts using machine learning and text analysis. Given the size and variety of financial documents that must be processed, it is labor-intensive to train separate extraction models for various document-types. Therefore, the system is now being enhanced to develop a task-agnostic general document understanding model capable of responding to natural language queries across the broad spectrum of financial documents. In this project, we aim to leverage the current state-of-the-art multi-model architecture on document-understanding to create a document verification system that, given a pdf document, identifies the correct document type and queries for information accordingly. For the classifier, we fine-tune the LayoutLMv2[1] model on the document classification task to achieve an overall 96% test accuracy. Then, using the same architecture, we train a question-answering model that, when trained on both the internal and DocVQA datasets, produces an 11% absolute improvement in ANLS[6] score compared to a score of 0.51 when trained only on the DocVQA dataset. The model performs quite well with text-based documents, but suffers in performance when dealing with documents incorporating tabular data. Therefore, for the question-answering task, we continue to train and evaluate the performance of other document-understanding models, especially DocFormer [3], Donut [4], and Dessurt [5].

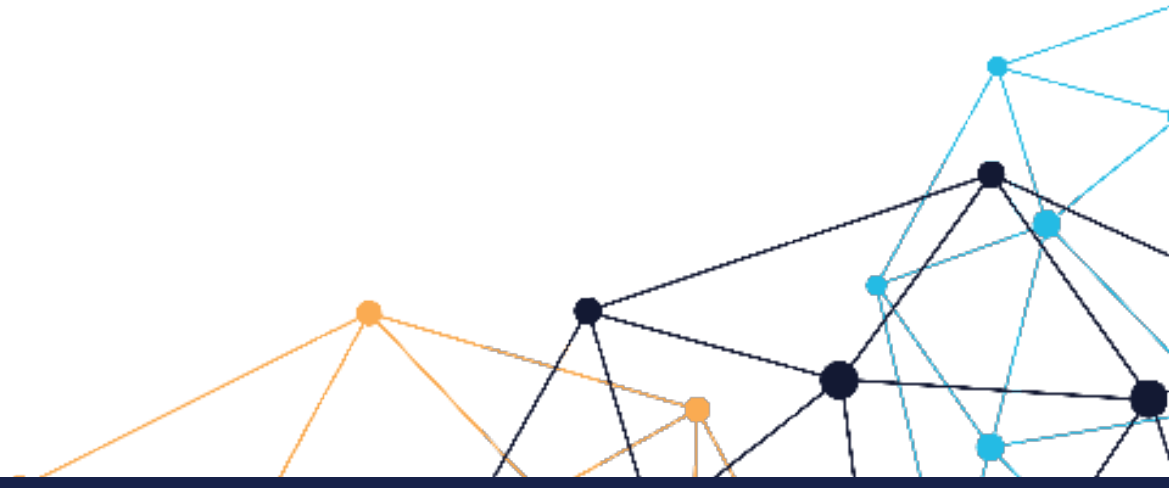


WYSIWYG Overlay tool for generating custom PDF templates

- **Student:** Singh, Mallika
- **Academic:** Prof. Daniel Wigdor
- **Industry:** Sheldon Davis
- **Company:** SOTI

Businesses today struggle with manual operations such as paper-based processes and copying and pasting information from one system to another. These manual operations are costing companies hours of unnecessary labor, which trickles down to customers that are expecting on-demand services. SOTI SNAP is a no-code application format for cross-platform mobile application development. In this project, we developed a robust PDF generation tool that allows users to upload existing templates of invoices, etc., or create new templates and overlay data collected from the SNAP applications to populate the PDFs on end-user submission. We designed and prototyped the complete user experience to allow SNAP developers to create and manage PDF templates, and provide the functionality to autogenerate the code to populate templates with SNAP data via SOTI's proprietary 'what you see is what you get' (WYSIWYG) application development environment.

Throughout the project, we conducted various interviews with the Sales team, Product Owners, etc., within the organization to understand the needs of the customers and gaps within the currently used PDF-generation products. Based on this feedback, we determined a set of use cases, designed user flows and prototyped 2 user flows for low-fidelity wireframes and then organized a round of AB Testing. Participants emphasized the need for ease of use, security, consistency with related products, simple navigation, familiar terminology for a non-technical audience, and effective onboarding. Following the implementation, we plan to conduct a remote usability study with three phases: exploratory testing, usability testing with simulated tasks, and a post-study questionnaire.

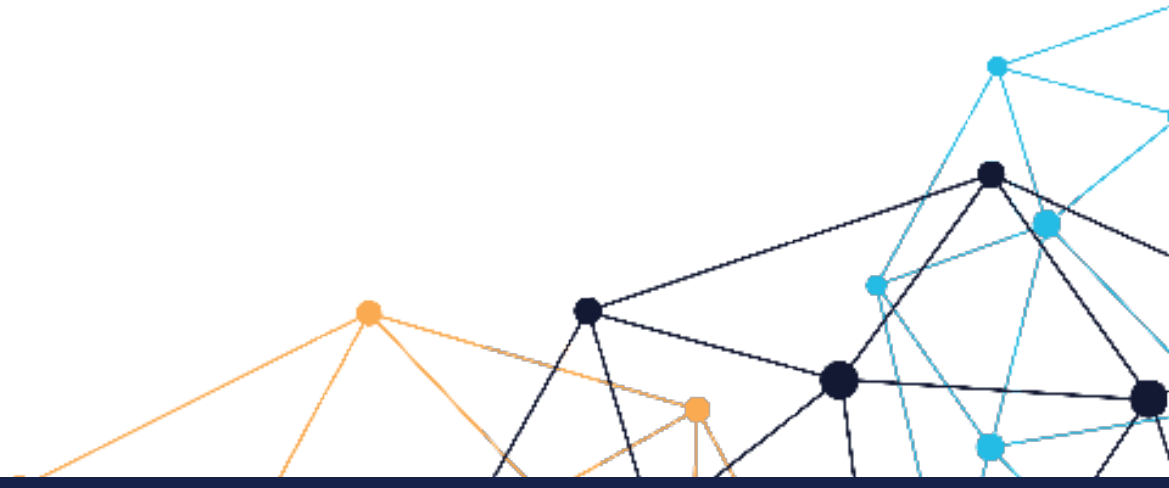


Defaulter Prediction with Alternate Data

- **Student:** Sun, Kaihua
- **Academic:** Prof. Sheldon Lin, Prof. Andrei Badescu
- **Industry:** Aditya Garg
- **Company:** ICICI Bank

While traditional credit scores play a key role in determining who is trustworthy and can get loans, there are many people in the market who have no credit history or minimal credit footprint. This puts these people at a disadvantage of not being approved for loans, as banks cannot verify their creditworthiness. As the competition in today's financial market becomes increasingly fierce, discovering trustworthy customers with limited credit history can increase profit and market share. The goal of this research project is to explore alternative data sources which can be used to predict defaulters and non-defaulters.

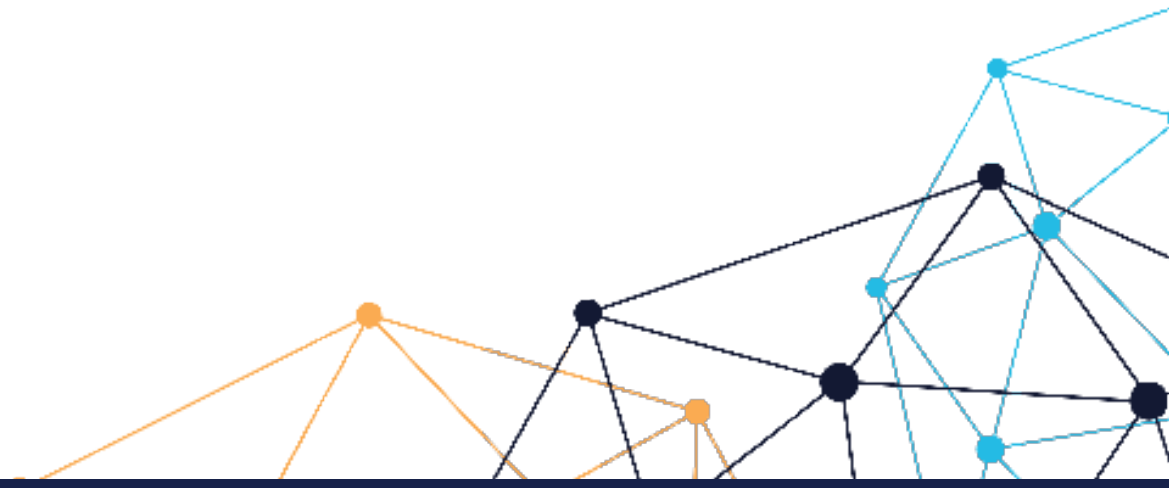
Our solution uses the UPI dataset as the alternative data source. UPI is an instant real-time payment system that records peer-to-peer and person-to-merchant transactions. An exploratory data analysis was conducted to extract useful features and a few different models were tried, like logistic regression, random forest, and Adaboost. Since the dataset is quite unbalanced (1% of data is defaulter), The SMOTE method was applied to up-sample the number of defaulters and used the Edit Nearest Neighbors method to down-sample non-defaulters. The dataset is more balanced before entering the training part. In doing so, my random forest model achieves a very high area-under-operator-curve(~0.9) by using 5-fold cross-validation.



Discourse Representation Framework for Consumer NLP

- **Student:** Sun, Shawn (Yuxiao)
- **Academic:** Prof. Shurui Zhou
- **Industry:** Josh Seltzer
- **Company:** Nexxt Intelligence

Representing pertinent information expressed during a natural language conversation, and generating sensible utterances and questions is a fundamental challenge in information extraction and natural language processing. Although some techniques have seen success in generating plausible and on-topic natural language utterances to engage with humans, these methods tend to employ large language models and are generally unable to extract structured information in a meaningful way. Our project seeks to approach the problem of natural language conversations through the research and development of a discourse representation framework which addresses both of those lacunae: namely, this framework will allow for the simultaneous extraction of structured information in a manner that is meaningful to market researchers and generalizable across conversations, as well as the further elicitation of information which is pertinent to ad-hoc market research objectives. With extracted keywords, an opinion graph will be implemented which allows structured information retrieval. The downstream task will focus on question-asking tasks that are based on human-annotated data and question templates. Future improvement of the framework will lie in controllable interfaces and robust deep learning capabilities.

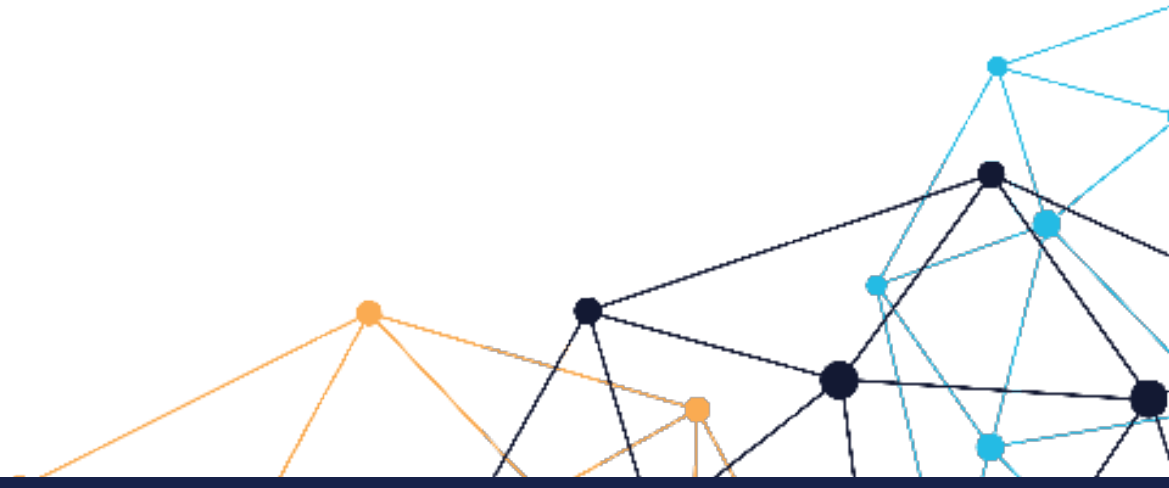


Ionian Web3 Storage: Study of Arweave's mining algorithm to design a decentralized storage system

- **Student:** Tan, Zhanwen
- **Academic:** Prof. Fan Long
- **Industry:** Ming Wu
- **Company:** Conflux Technology

Limit on storage capacity and extremely high storage fees have been prominent problems in some existing blockchain systems. For example, users can't store data that's larger than 1MB in a single block on Bitcoin network due to its limit on block size. Also, it takes around 40K USD to store 1MB of data on Ethereum. These problems deter users from storing large files on existing blockchain systems. Our project, Ionian, aims at designing and implementing a decentralized storage system that provides permanent and unlimited storage with low cost. Also, it aims at providing high availability, high throughput, great scalability and flexibility.

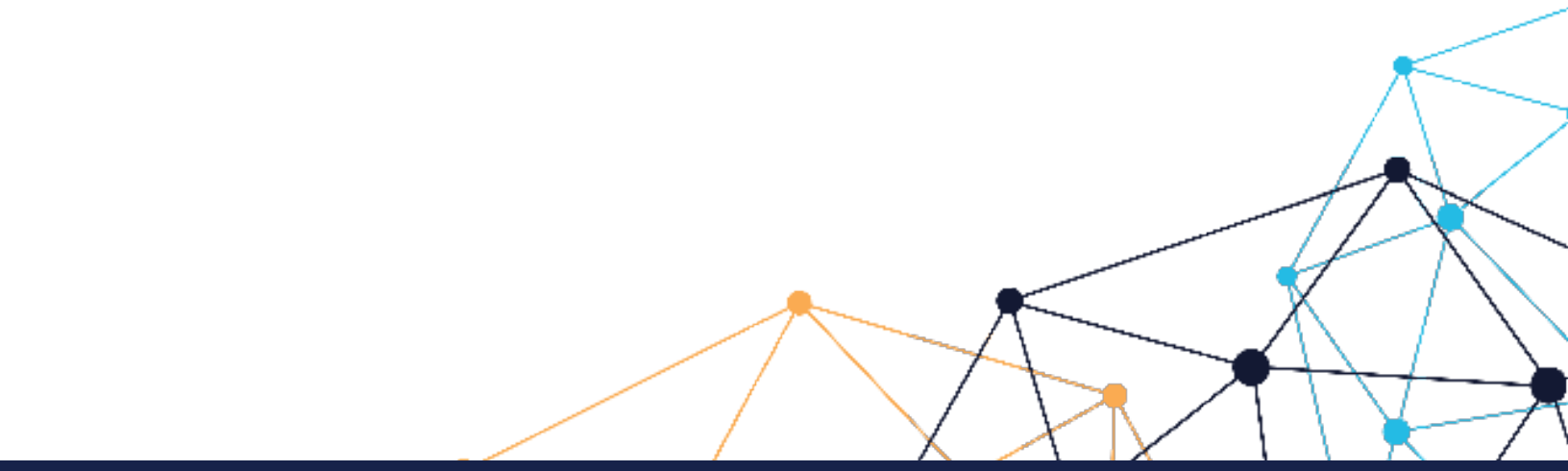
To learn from existing solutions, we analyzed Arweave, one of the existing decentralized storage systems. The mining algorithm of Arweave was studied and two vulnerabilities of the system were evaluated. Our results show that system attack on one of the vulnerabilities is feasible but not profitable, while the attack on the other vulnerability is not feasible. These findings shed light on the design of our own mining algorithm. We have implemented a prototype of Ionian's decentralized storage system. In the prototype, we improved the system performance and usability by developing features to allow users to upload file concurrently and cache small files.



Data Communication Optimization between Mobile Devices and Servers

- **Student:** Tang, Yilai
- **Academic:** Prof. Eyal de Lara
- **Industry:** Ashley Pereira
- **Company:** SOTI

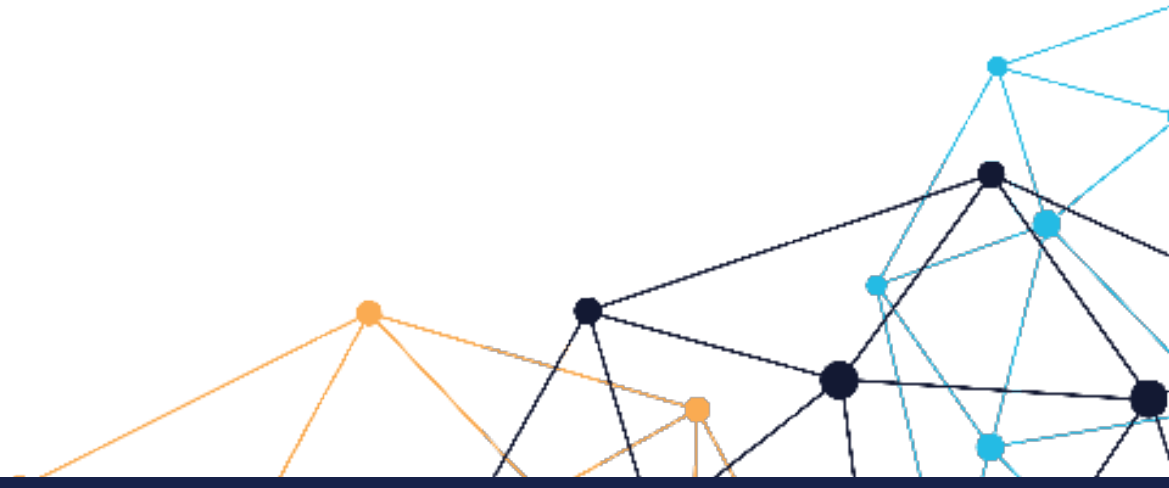
With data being the new digital currency, its consumption is growing at an exponential rate. Data communication between mobile devices and servers generates a massive cost to businesses because of the high throughput. Transmission of data becomes expensive when it happens within millions of devices, and every “bit” matters. One possible way is to reduce the size of messages being transmitted for the current communication protocol, named COMM. We established the baseline for the current protocol and re-implemented it by using a serialization tool, which is Google’s Protocol Buffer, to generate a tighter message. Finally, we created a new communication protocol named COMM2 and it has advantages over both COMM and protocol buffer. We found the message size generated by the protocol buffer is only 75% of that generated by COMM, but it takes nearly 430% more time to do serialization and deserialization on average. While applying COMM2, the message size is only 50% of that generated by COMM, and it only takes 90% of the time required by COMM to complete serialization and deserialization. This new communication protocol would be integrated into SOTI products to facilitate faster File Transfer/ Synchronization, Application Deployment and Remote Control of mobile devices.



Study of biases in artificial intelligence for IT recruitment

- **Student:** Usha, Rowshni Tasneem
- **Academic:** Prof. Robert Soden
- **Industry:** Roy Hussein
- **Company:** SRA Staffing Solutions

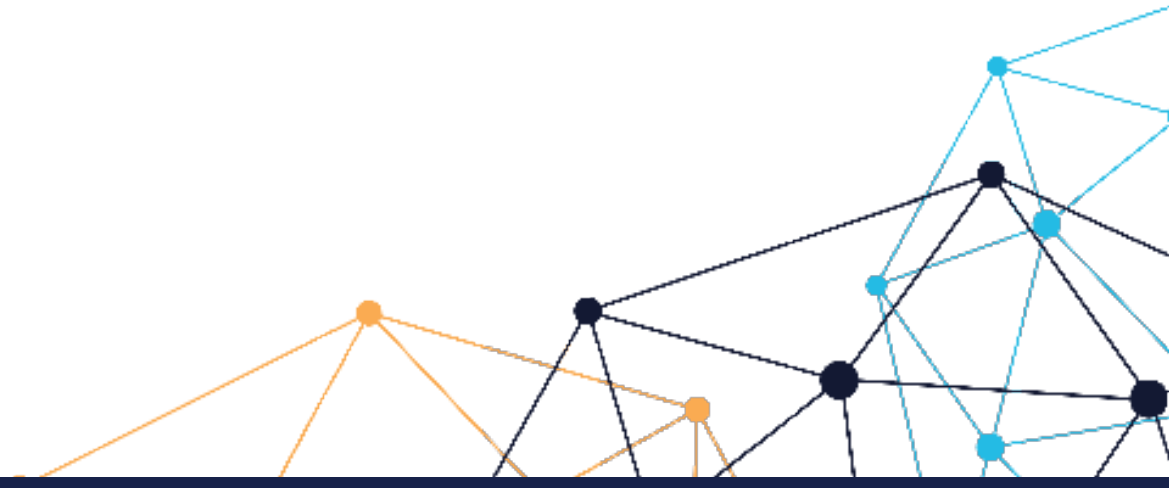
Traditional recruiting processes are time-consuming for both hiring managers and applicants. Some common steps in any hiring method are matching, verifying, rating, and shortlisting applicants based on their resumes. The proposed research intends to develop an application to aid hiring organizations in searching for outstanding candidates. Our application will focus on these redundant tasks, lowering barriers to seeking potential applicants, and reducing wait time. The technology will give clients access to interviews and the ability to find people based on their knowledge, skills, and experiences. Our goal is to have pre-screened individuals available for recruiters, which can benefit both parties by quickly cutting down the time needed for searching and making decisions. We used synthetic data to mirror the confidential hiring information and analyzed it to design and tune the hiring algorithm as part of the research to have potentially unbiased predictions. We examine how the system handles the issue of fairness balanced with client needs. We report on the findings of auditing the applicant screening tool, focusing on subjective and familiarity biases. For each, we identify potential biases in the algorithm, their sources and viable solutions to reduce them. Overall, we improve the hiring process by enabling fair sourcing of applicants and efficient pre-screening.



Multi-turn Image Manipulation using Text guidance

- **Student:** Verma, Nikhil
- **Academic:** Prof. Scott Sanne
- **Industry:** Kevin Ferreira
- **Company:** LG Electronics Canada, Inc.

Image manipulation using text is an interesting Cross-modal learning task. The task in multi-turn image manipulation is to iteratively update the background image using a sequence of modification text. The aim of this study is to generate a compositional image following a sequence of modification text. Techniques used previously include generative adversarial training and counterfactual reasoning for accomplishing this task. These approaches are limited with object generation in the image guided by text instructions decreasing the overall performance of object detection. We propose an Image manipulation architecture using diffusion modeling technique guided using the fused information of both modification text and previous image in the iteration. It starts with a pattern of random dots and slowly converts that pattern into an image using the textual information fused with the base image content. Diffusion models have recently beat other generative models in image generation tasks and cross-modality learning using guiding information. The guiding features were obtained using representation learning via contrastive language-image pair (CLIP) model. The experiments were performed on benchmark datasets for this task, namely CoDraw and iCLEVR to achieve the performance using the proposed diffusion model.



Automated Blendshape Generation for Facial Animation

- **Student:** Wang, Ellery
- **Academic:** Prof. Eitan Grinspun
- **Industry:** Martine Bertrand
- **Company:** Double Negative (DNEG)

Blendshape animation is an industry standard facial animation system where each facial expression is represented as a linear combination of a basis of face shapes referred to as Blendshapes. Blendshape animation is preferred by artists as it offers a high level of control. However, generating the required Blendshape assets is a challenging process.

To improve efficiency, we investigated the viability of two automated Blendshape generation procedures: a linear regression model which optimises for the target Blendshapes by minimizing the reconstruction loss compared to captured complex expressions of the target, and an image-to-image translation network which generates the target Blendshapes from the geometry image of the neutral mesh using a 2D convolutional encoder-decoder network. Both methods are shown to successfully create personalized Blendshapes in our experiments with a synthetic dataset. We provide an analysis of the efficacy, advantages, and limitations of each of the methods.

A transformer-based classification model for the intent of user utterances in text-to-SQL settings

- **Student:** Wang, Weiqing
- **Academic:** Kelly Lyons
- **Industry:** Hossein Taghinejad
- **Company:** SOTI

Most decision makers rely on business analysts to query data stored in a database to answer questions before making business decisions. Recently, the development of Natural Language Processing (NLP) technologies has made querying databases simpler – questions raised in decision-making processes can be answered by NLP-powered text-to-SQL engines, which generate SQL queries to databases that respond to a decision maker's utterance. However, most existing solutions can only handle self-contained context-independent utterances (such as Q1: "what was the total sales in the first quarter of 2022?") and perform poorly on context-dependent utterances (e.g. following Q1 with "what was it for the second quarter?"), frequently giving wrong answers. Moreover, until today, no context-dependent text-to-SQL pipeline has been developed with an accuracy that could be put into production.

To prevent the context-dependent question and other utterances (such as greetings) from being processed randomly by the existing text-to-SQL engines, an intent-classifier could be used to ensure only the answerable and context-independent questions are being passed onto the text-to-SQL engine. This study proposes a multi-modal transformer-based classification model that serves this purpose.

We tested the accuracy of this framework on the CoSQL dataset, a state-of-the-art context-dependent text-to-SQL benchmark, where it reached an accuracy of over 80%. We are currently experimenting with SOTI's datasets to compare our framework against its standard for productization. This framework will help existing text-to-SQL engines deliver more informative responses to user utterances that could previously not be handled.

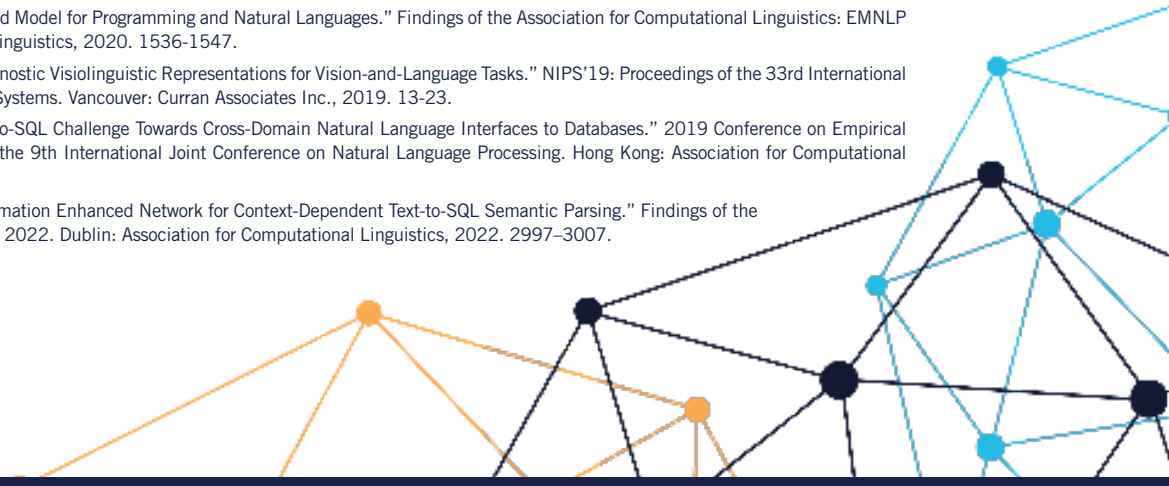
References

Feng, Zhangyin, et al. "CodeBERT: A Pre-Trained Model for Programming and Natural Languages." Findings of the Association for Computational Linguistics: EMNLP 2020. Online: Association for Computational Linguistics, 2020. 1536-1547.

Lu, Jiasen, et al. "ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks." NIPS'19: Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2019. 13-23.

Yu, Tao, et al. "CoSQL: A Conversational Text-to-SQL Challenge Towards Cross-Domain Natural Language Interfaces to Databases." 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 1962-1979.

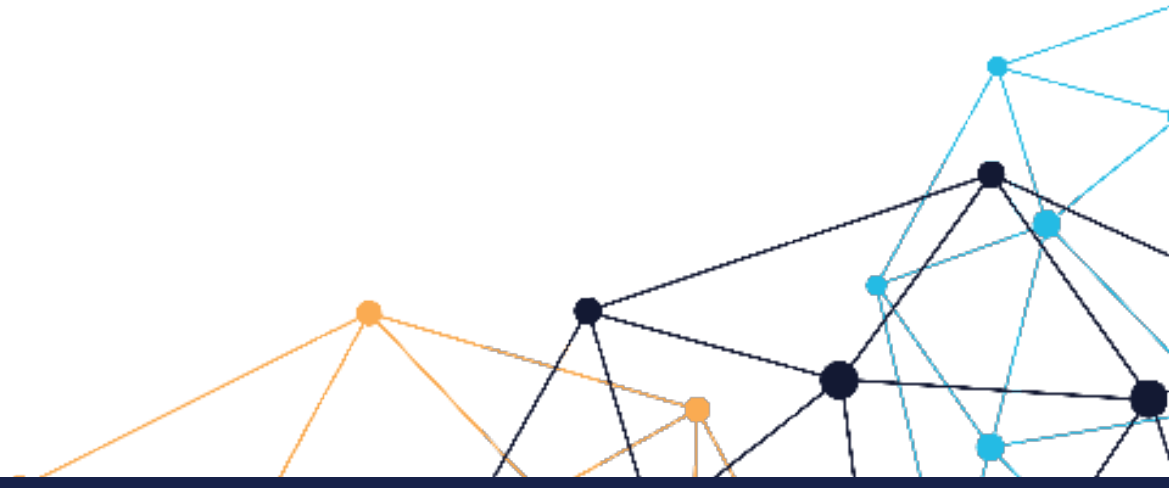
Zheng, Yanzhao, et al. "HIE-SQL: History Information Enhanced Network for Context-Dependent Text-to-SQL Semantic Parsing." Findings of the Association for Computational Linguistics: ACL 2022. Dublin: Association for Computational Linguistics, 2022. 2997–3007.



Learning Ephemeral-Sparse Attention for Efficient Neural Networks

- **Student:** Wei, Cong
- **Academic:** Prof. Florian Shkurti
- **Industry:** Brendan Duke
- **Company:** ModiFace

Vision Transformers (ViTs) have been one of the most successful deep-learning models for Computer vision. One of their core limitations is the quadratic dependency on the number of image patches due to their full self-attention mechanism. Fixed sparse attention mask patterns have previously accelerated operations on images by orders of magnitude. These attention masks take the form of a HxW binary mask for each patch in an image with HxW spatial dimensions. Rather than use fixed attention masks, we go a step further and propose learning attention masks. We propose Sparsifiner, a learnable sparse attention mechanism that reduces the quadratic dependency to linear. Our contributions are in two folds: First, We show that learned unstructured attention sparsity produces a superior Pareto-optimal trade-off between FLOPs and top-1 accuracy (78.8% with -40% FLOPs) on ImageNet compared to token sparsity. Second, we show that the sparse attention pattern introduces additional interpretability to Vision Transformer models.



Multi-modal Machine Learning for Business-critical Insights in Video Conversations

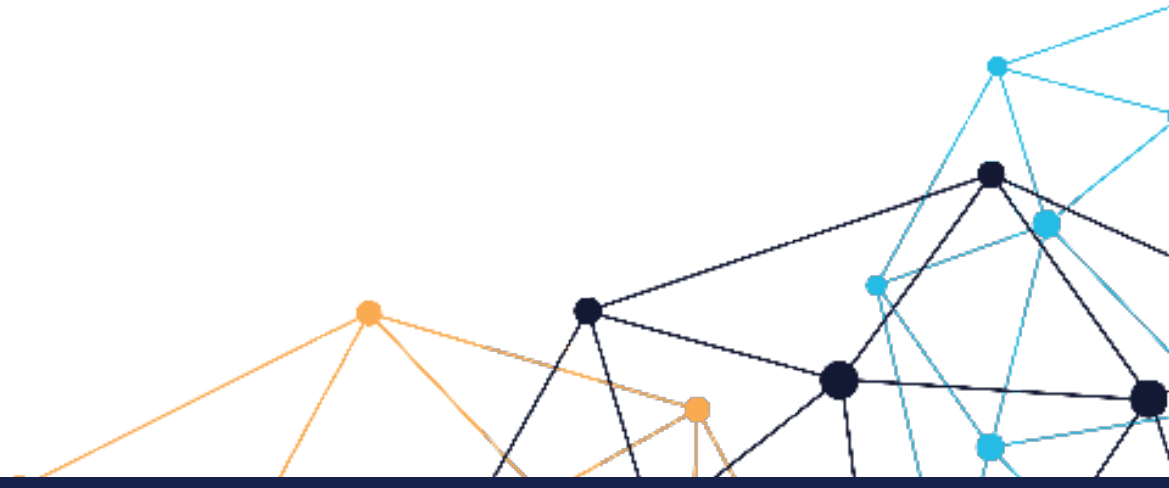
- **Student:** Wu, Tongzi
- **Academic:** Prof. Scott Sanner
- **Industry:** Henry Zhou
- **Company:** Talka AI Canada, Inc.

Gaining insights in online meetings can help identify moments that may impact a companies future product features, revenue and customer satisfaction. This project first focuses on developing an interruption detection system and a smile classifier. Although the intersecting work of computer vision, natural language understanding and audio signal processing has rapidly developed, no existing work focuses specifically on interruption or smile detection in video format, which make it challenging to build and deploy working models in production. With the abundance of sales video recordings inside of our company, we first annotated and constructed the two datasets. Then, we leverage Wav2Vec 2.0 model, OPT model and CLIP model for feature extraction. We used a random model as our baseline, and applied Transformer and MLPMixer models to prove the feasibility and reliability of our dataset. The contribution of this report is three-fold: (1) we built a dataset on the task of video interruption detection and a dataset on the task of smile detection; (2) we evaluated the validity of 2 models on our datasets and provide evidence of the improvement brought by different modalities; 3) We also created a thorough ablation study to understand which perceptual modality play a more important role in each task.

Temporal Neural Networks with Association Rules for Retail Demand Forecasting

- **Student:** Xu, Ziyue
- **Academic:** Prof. Murat A. Erdogdu
- **Industry:** Joe Roussy
- **Company:** Deloitte

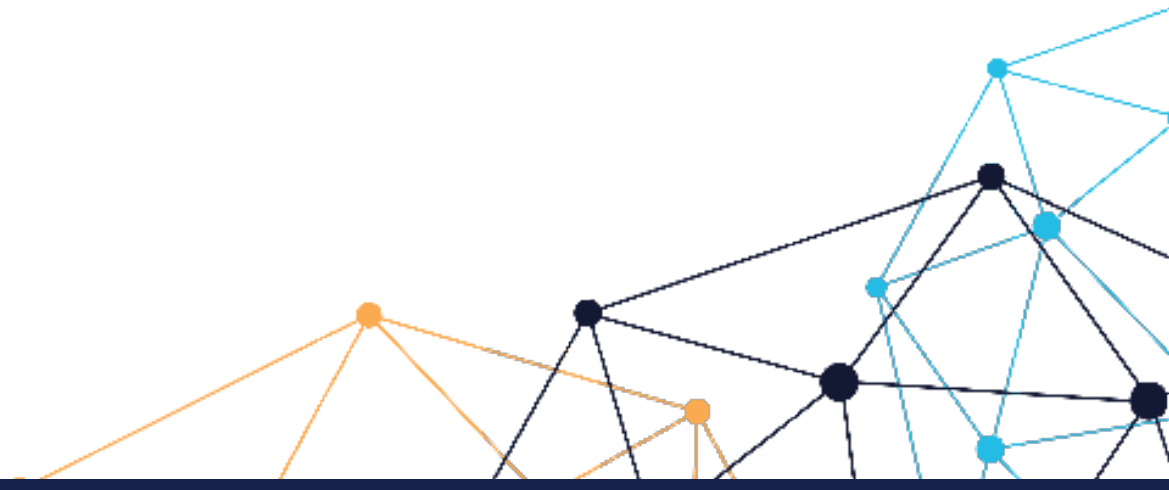
Temporal neural networks (TCN) are widely used for time series forecasting, especially demand forecasting in retail, due to their lower training time compared to recurrent neural networks and lower feature engineering costs than multilayer perceptrons. Our current version of the demand forecasting TCN model handles each SKU individually, and no relationships between SKUs are considered. However, the sales of one product could influence the sales of some other products. Therefore, the proposed new TCN model takes the historical sales data of the products with a high association, mined using the Market Basket Analysis method, as input, and the sales of both products are predicted simultaneously. Trained using the same data, this model incorporating product associations lifts the retail accuracy of the consequent products by 4.2%. Future work will be adding price promotion features to the model to achieve higher retail accuracy.



Channel-Agnostic Pretraining for Microscopy

- **Student:** Yan, Shujun (Diana)
- **Academic:** Prof. Benjamin Haibe-Kains
- **Industry:** Kshitij Gupta
- **Company:** Dewpoint Therapeutics

Experiments involving the examination of microscopy images are crucial in the drug discovery process. With the improvement of computer vision models, we wonder if it would be beneficial to apply pre-trained image models to microscopy images. More specifically, our primary goal is to see if pre-trained ImageNet models can improve the performance of existing pipelines or accelerate the training process. Besides that, we also aim to validate whether a model pre-trained with microscopy images can outperform ImageNet pre-trained models. In this project, we have compared the results from CytoImageNet models with ImageNet models on classifying the mechanism of actions of various chemical compounds. On the other hand, there is less correlation among channels of microscopy images when compared to channels of regular RGB images. As this difference could potentially lead to less than optimal results if we directly apply pre-trained ImageNet models, we have explored a few ways to handle the channels when it comes to creating embedding space for microscopy images.

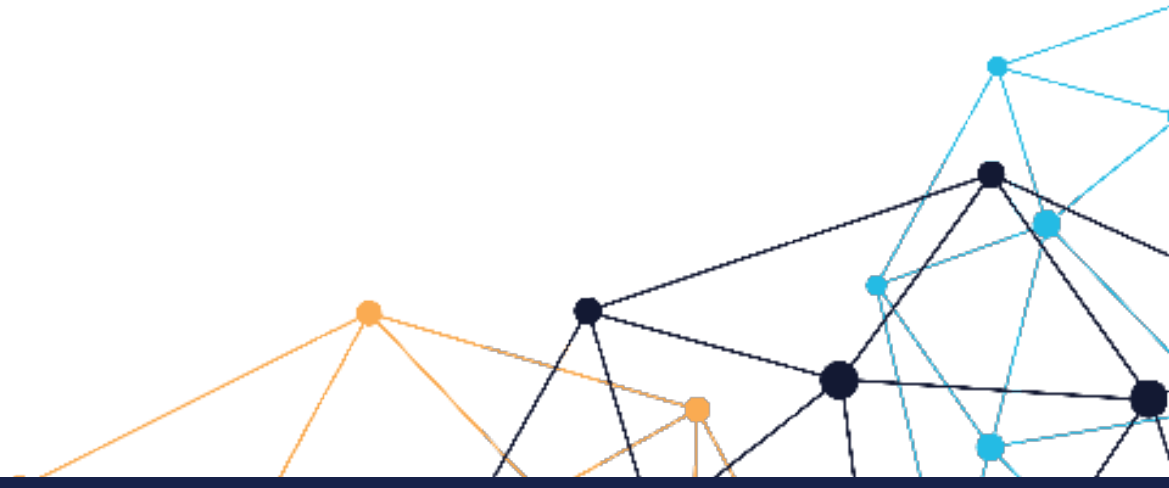


High-Throughput Linguistic Content Comparison System

- **Student:** Yang, Bowen
- **Academic:** Prof. Nick Koudas
- **Industry:** Jonathan Shahen
- **Company:** Scrawlr Development Inc.

Monitoring copyright infringement is one of the responsibilities of social media. An effective copy detection mechanism will protect users' originality and encourage them to create. But the high volume of new posts generated every second makes it a vexing issue. On an average day, hundreds of million tweets are posted and shared by all users on Twitter, which requires the detection system to handle thousands of tasks per second. This project aims to explore this problem.

In practical experiments, we encountered that the amount of incoming data was larger than we expected. Therefore, it is necessary to develop a more efficient data processing pipeline to improve the throughput of the entire system. Potential problems include the processing capability of larger corpora, data redundancy and storage optimization, the trade-off between model performance and accuracy, etc. We have used several profiling tools to measure the performance of each component. Severe bottlenecks have been recognized and improved after rounds of testing. Different techniques have been integrated to support multiple detection metrics. At the current stage, we could achieve 3,500 posts per second processing throughput on the experimental server cluster. And we are aiming to up the processing speed to a target of 6,000 posts per second.



An In-Memory Database Solution for Industrial Telematics Data At Large-Scale

- **Student:** Yang, Yichao
- **Academic:** Prof. Ben Liang
- **Industry:** Mengliao (Mike) Wang
- **Company:** Geotab Inc.

As the Internet of Things (IoT) market has grown, for many modern companies a large of amount of data needs to be exposed daily to different users including data scientists, software engineers and customer companies. Thus, A robust, scalable, and high-performance data infrastructure is necessary for data processing and analytics. In this scenario, in order to improve the data querying performance between users and cloud data warehouses, and cost-saving on cloud services, such as BigQuery, an In-Memory Database service (IMDB) is introduced.

In this paper, we present the overall architecture and workflow of our In-Memory Database service, several common use cases, and the behavior tests based on real-world data. Finally, we discuss some challenges and future work that we need to continuously address.

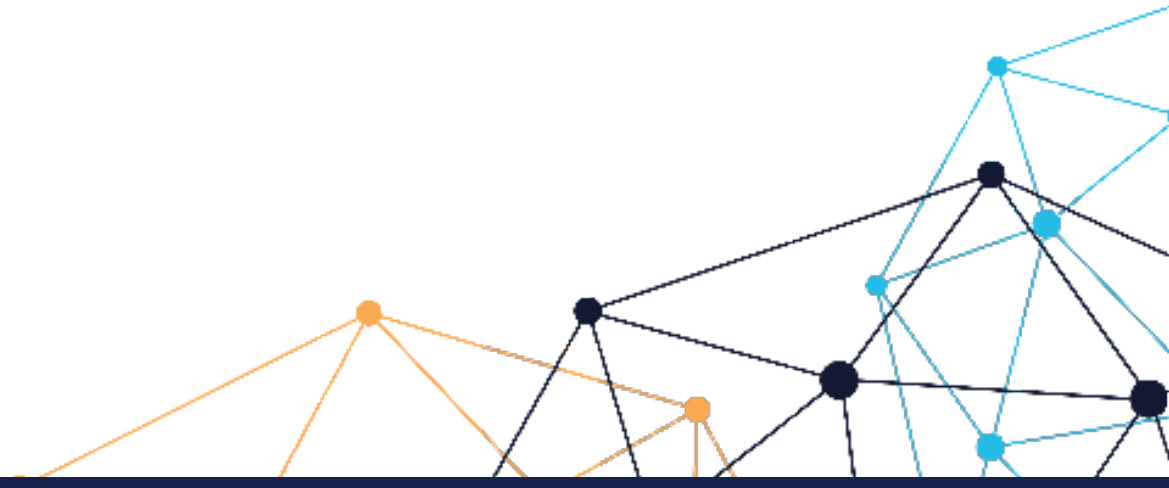
Our solution relies on Apache Ignite for the key areas (basic distributed storage layer), we also add improvements and customizations to make the solutions fit in the Geotab use scenario. The analysis of querying performance is mainly measured by response time (in milliseconds).

With the help of the In-Memory Database service, our query response time gets reduced by half on average, compared with directly querying data from cloud data warehouses. In the future, we may migrate key-value caching and on-spot loading to further improve its functionality and performance.

An In-Memory Database Solution for Industrial Telematics Data At Large-Scale

- **Student:** Yu, Qingyang
- **Academic:** Prof. Samin Aref
- **Industry:** Sam Talasila
- **Company:** Wealthsimple

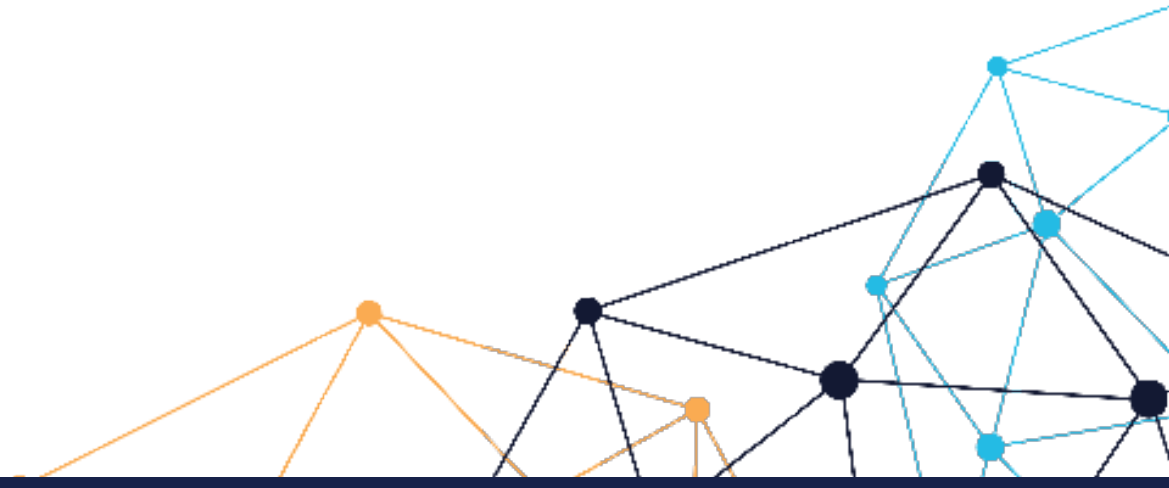
One of the missions of Wealthsimple is to help clients manage money in a simple and affordable way. To achieve this goal and grow the user base, Wealthsimple would like to perform smart cross-selling among products. In traditional cross-selling between Trade and Crypto app, mass campaigns are distributed through emails and in-app pushes. As Crypto investment generally has higher risk than stock trading, Wealthsimple wants to target the cross-selling of Crypto app to the users with the suitable level of risk tolerance and high acceptance probability. In this project, graph neural networks (GNN) are applied to learn the representation of the graph and build users' stock trading portfolio. Stock holdings, account registration, in-app/ contact book connection, and users' personal information are collected to construct the graph. We applied GraphSage to learn the embedding of bipartite use-stocks graph and predict the probability score of the downstream cross-selling task. The graph neural networks show strong capability in learning underlying information of the graph and improve the performance of downstream tasks.



Predicting Indicated Diseases from Drug Labels

- **Student:** Yuan, Jingming
- **Academic:** Prof. Kieran Campbell
- **Industry:** Krish Perumal
- **Company:** Cedience

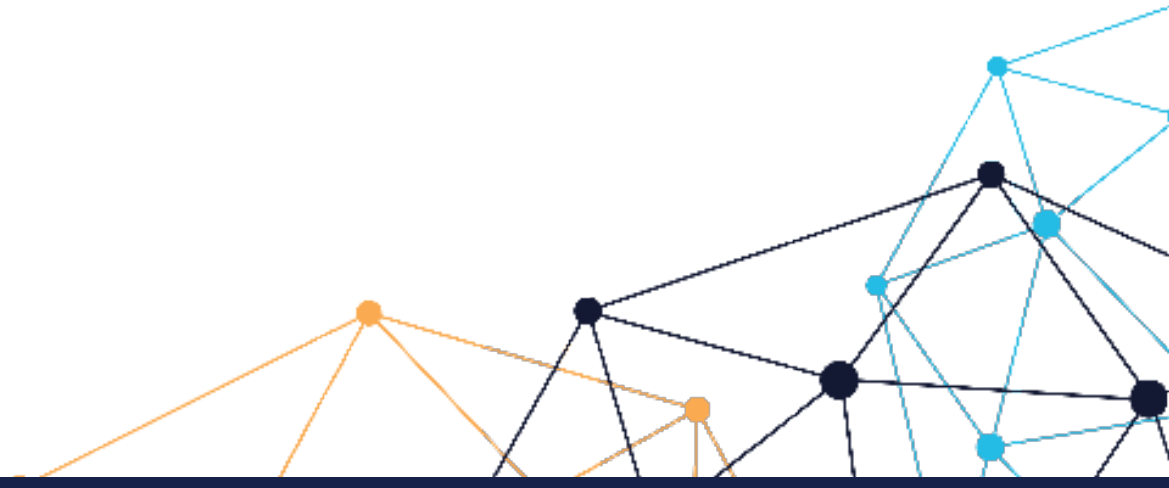
The US FDA publishes detailed label documents for each approved drug specifying its indication, dosage, precautions etc. For medical regulatory professionals, it is instrumental to be able to read and process these documents using an AI algorithm to extract the diseases associated with particular drugs, such that one can easily navigate through thousands of label documents without having to read them individually. The main challenges involved are that indication statements and disease names in the labels can have many variations and noises in their wording and format, making conventional text processing approaches difficult. The main goal of this project is to predict the diseases indicated for a drug given its FDA label document. We proposed a three-step solution: First, leveraging fined-tuned BIOBERT token classifier, indication statements in the drug label documents were extracted to remove noise. Then, AWS Comprehend Medical API was run on these statements to identify the diseases indicated (as a Named-Entity Recognition task). Finally, as the task required almost perfect accuracy, rule-based manual filtering was applied to the result to remove problematic predictions. The first token classification step resulted in very high accuracy (93%) and almost always extracted indication statements correctly. The second step resulted in more noisy predictions, but after a final stage of manual filtering, accurate indications were predicted for more than 40000 drugs with high precision (though lower recall). In conclusion, a machine-learning pipeline was constructed to accurately extract indicated diseases from raw FDA documents.



Removing Confounders in Neural Networks for Microscopy

- **Student:** Zeng, Beiqin
- **Academic:** Prof. Jessica Gronsbell
- **Industry:** Grant Watson
- **Company:** Dewpoint Therapeutic

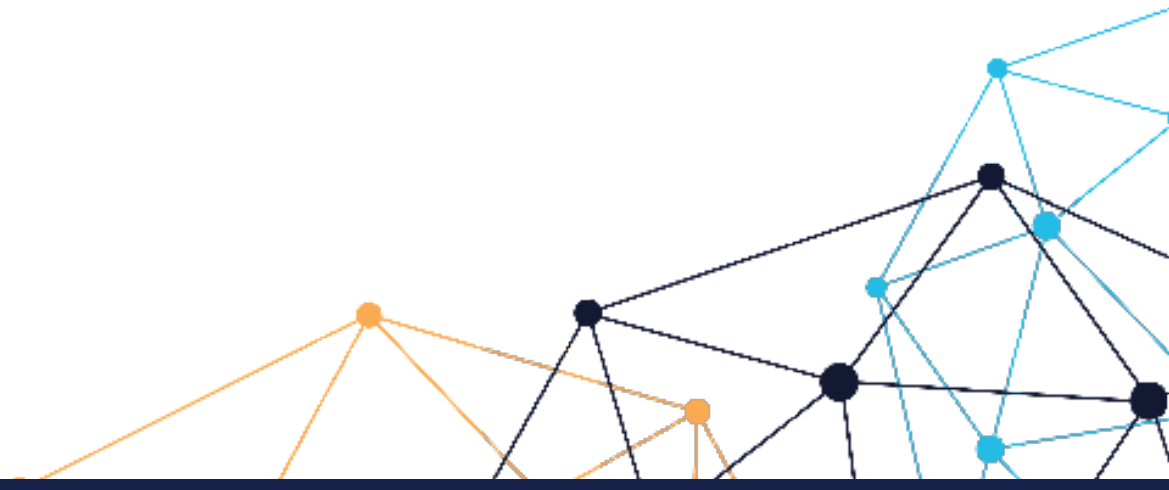
Deep learning models have achieved great performance on phenotypic profiling tasks. However, when performing biology experiments at scale, it is difficult to dissociate biological effects of interest from irrelevant confounders created by the experimental protocol itself. In this project, our goal is to remove confounders in microscopy images. We specifically investigate the batch effect, a known confounder that has a big influence on biological data analysis. Our work is mostly benchmarked on BBBC021 dataset, a public dataset of phenotypic profiling that has strong batch effect. First, we explore an existing batch effect correction technique called Typical Variation Normalization (TVN) and discuss the strength and limitations of the technique. Second, we investigate and compare different metrics that can be used to evaluate the batch effect. Third, we test a contrast enhancement technique that can potentially force the model to ignore background noise, which can lead to the removal of confounding effects.



Leveraging Few-shot Learning and Interactive Workflow for Controllable Text-to-image Generation

- **Student:** Zeng, Ruijing
- **Academic:** Prof. Frank Rudzicz
- **Industry:** Jerry Chen, David Landsman
- **Company:** Vanguard Investments Canada

Diffusion models have recently been shown to generate high-quality images and achieve state-of-the-art results on text-to-image tasks. Although existing models can condition on arbitrary text and are trained on large-scale datasets to cover diverse outputs, it remains a challenge for users to control the model to generate images in a specific style. We hypothesize three potential reasons for the challenge: (1) natural language is not expressive enough to describe specific styles; (2) models are not trained on a specific-domain dataset; (3) one single generation alone is insufficient to incorporate users' ideas. Inspired by the work on controllable image generation in GAN, we explore few-shot learning capabilities of diffusion models and leverage Textual Inversion to learn the text embedding of a certain style in a few-shot fashion. We further experiment Textual Inversion on Stable Diffusion and propose a novel training prompt design to facilitate the training quality. Following previous work, reconstructability and editability are evaluated using CLIP-space distances, which reveal that our method performs well on both metrics compared to the baselines. In addition, we combine other sampling methods and prototype a human-interactive workflow to allow for not only image generation but also user-defined image operations including variation creation and image inpainting.



Efficient Flow-Guided Multi-frame Fence Removal

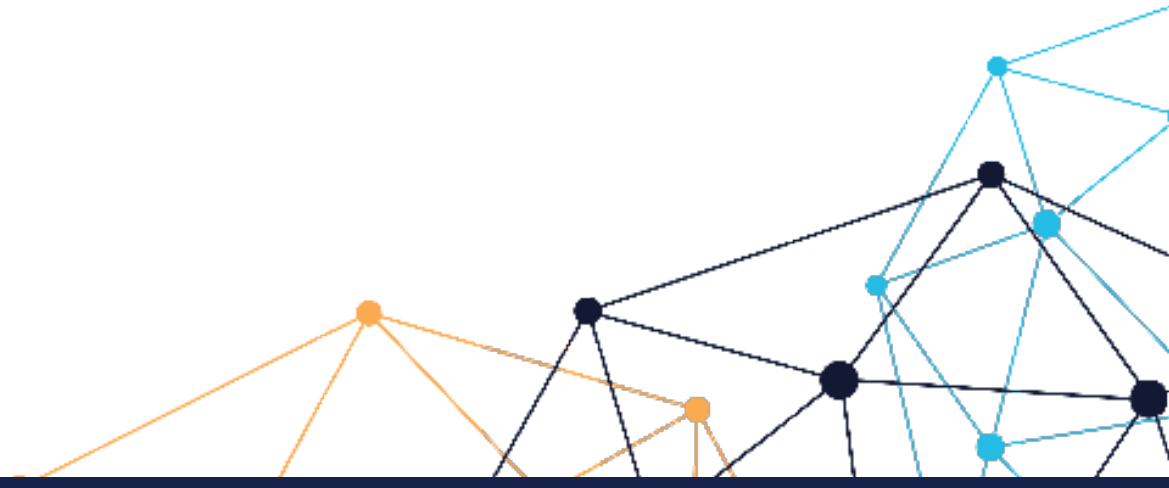
- **Student:** Zhang, Fengjia
- **Academic:** Prof. Kyros Kutulakos
- **Industry:** Alex Levinshtein
- **Company:** Samsung AI Center - Toronto

Taking photographs is often hindered by fence obstructions standing between the camera user and the scene of interest, which are hard or impossible to avoid. De-fencing is the algorithmic process of automatically removing such obstructions from images, revealing the invisible parts of the scene. While this problem can be formulated as a combination of fence segmentation and image inpainting, this often leads to implausible hallucinations of the occluded regions. Existing multi-frame approaches rely on propagating information to a selected keyframe from its temporal neighbors, but they are often inefficient and struggle with alignment of severely obstructed images. In this work we draw inspiration from the video completion literature, and develop a simplified framework for multi-frame de-fencing that computes high quality flow maps directly from obstructed frames, and uses them to accurately align frames. Our primary focus is efficiency and practicality in a real world setting: the input to our algorithm is a (5-frame) image burst -- a data modality commonly available in modern smartphones -- and the output is a single reconstructed keyframe, with the fence removed. Our approach leverages simple yet effective CNN modules, trained on carefully generated synthetic data, and outperforms more complicated alternatives on real bursts, both quantitatively and qualitatively, while running real-time.

Improving Federated Learning on Heterogeneous Data with Adaptive Local Weight Interpolation

- **Student:** Zhang, Runqing
- **Academic:** Nandita Vijaykumar
- **Industry:** Ian Colbert, Thomas Perry
- **Company:** Advanced Micro Devices (AMD)

In applications where data privacy is of paramount importance, federated learning (FL) has emerged as a popular collaborative learning paradigm in which a federation of models is trained across a collection of decentralized data silos without exchanging or transferring data. Standard FL techniques orchestrate learning a global consensus model across a network of distributed devices using a central server. The heterogeneity of such decentralized data silos often voids the standard machine learning assumption of independent identically distributed (IID) datasets, which can result in slower convergence rates and lower model accuracy across the federation of models. In this study, we introduce Adaptive Local Weight Interpolation (ALoWI) which combines state-of-the-art (SOTA) parameter decoupling and model interpolation algorithms to optimize FL model performance over heterogeneous data silos (i.e., non-IID data). We performed a comprehensive evaluation of the performance of both client local model and server global model on both local and global data distributions. We find that Fed-Rod yields robust performance under different training settings (e.g., client participation ratios, batch sizes, etc.), especially in the presence of strong non-IID data. We also show that ALoWI improves the performance of Fed-Rod under the extreme training settings (i.e., less client participation ratios, larger batch sizes, and extreme non-IID data distributions).

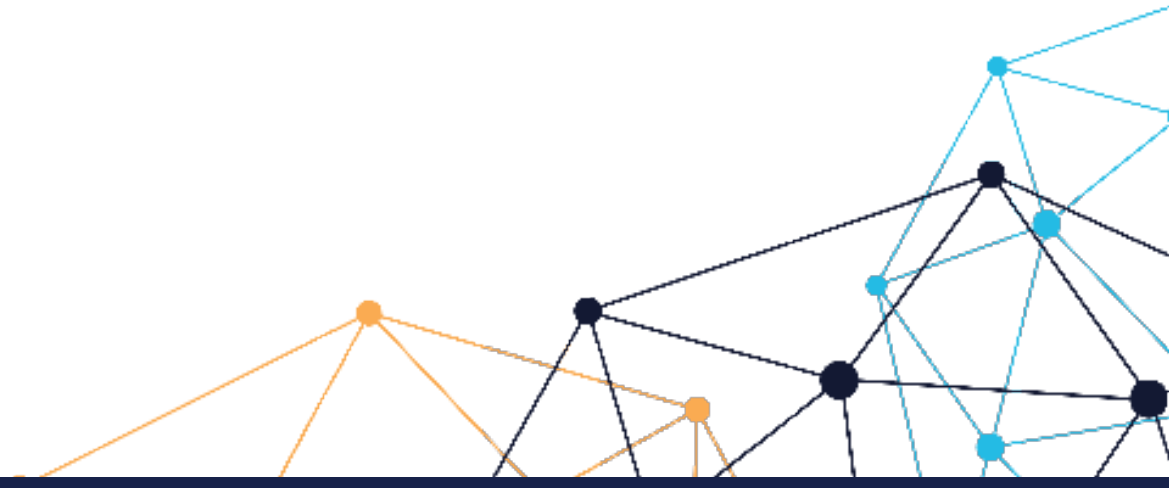


Deep Learning Based Approaches to Large-scale Synthetic Data Generation

- **Student:** Zhao, Weiwen
- **Academic:** Prof. Periklis Andritsos
- **Industry:** Winston Li
- **Company:** Arima

As a fundamental step for many data science tasks, creating a single dataset from multiple sources when direct collection is difficult or costly. In applications where large-scale data collection involves manual surveys, or when the collected data is highly sensitive and cannot be fully released to the public, synthetically generated datasets become an ideal substitute.

This project aims to achieve the goal of developing a scalable, data-rich Synthetic Global Population for research and commercial uses. In this project, we study a specific synthetic data generation task called downscaling, a procedure to infer high-resolution, harder-to-collect information, from many low-resolution, easy-to-collect sources. On top of the current Synthetic Data Generation algorithm in Arima, we developed and published a scalable software package for academic and industrial uses, providing users with the ability to generate the data in-house. The library published in the python pypl repository allows users to locally generate synthetic data, which preserves data privacy and extends customizability. In addition, the input flexibility for various application scenarios, and compatibility with different data formats are achieved by us, well incorporating more data types into the Synthetic Population.

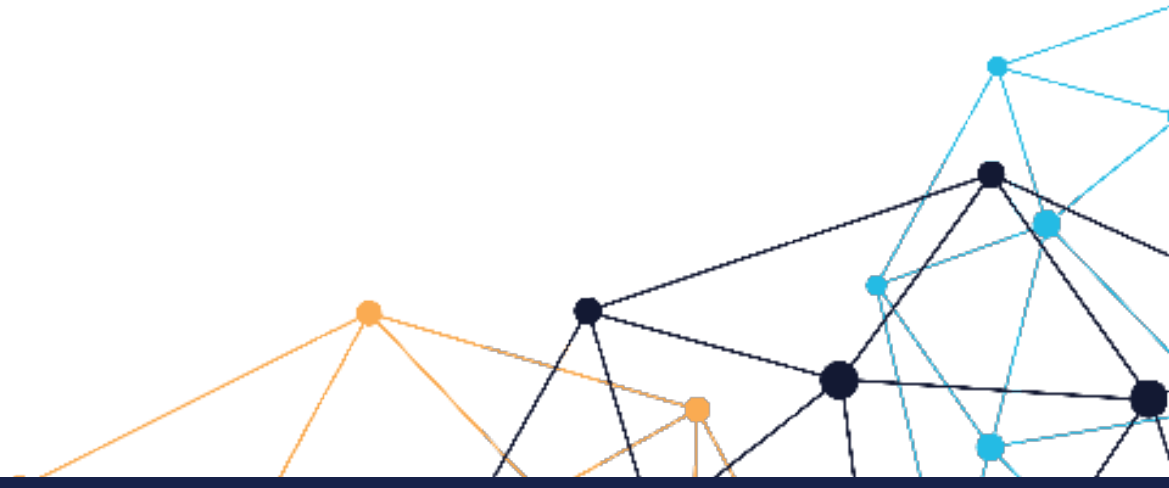


Hypatia-Learn: A State-of-the-Art Mathematics Learning and Tutoring System

- **Student:** Zhao, Yi
- **Academic:** Prof. Azadeh Farzan
- **Industry:** Ladislav Stacho
- **Company:** Hypatia System

Hypatia is now building Hypatia-Learn, a state-of-the-art mathematics ITS for Grades 8 to 12. Hypatia-Learn includes a powerful "Math Checker" that can auto-check students' work on a variety of topics. In this work, we aim to enhance the "Math Checker" by improving coverage of problems Hypatia Learn handles and build a system for autogenerating problems.

We extended the existing parser, designed a relation knowledge network of math expressions, and drew insights from this network before then designing a graph neural network for various learning predictions based on client (learners) to items (solution graph & hints) relations. The evaluation results show the graph neural network approach is able to enhance the capabilities of Hypatia Learn to improve coverage of problems Hypatia Learn handles by 20%.

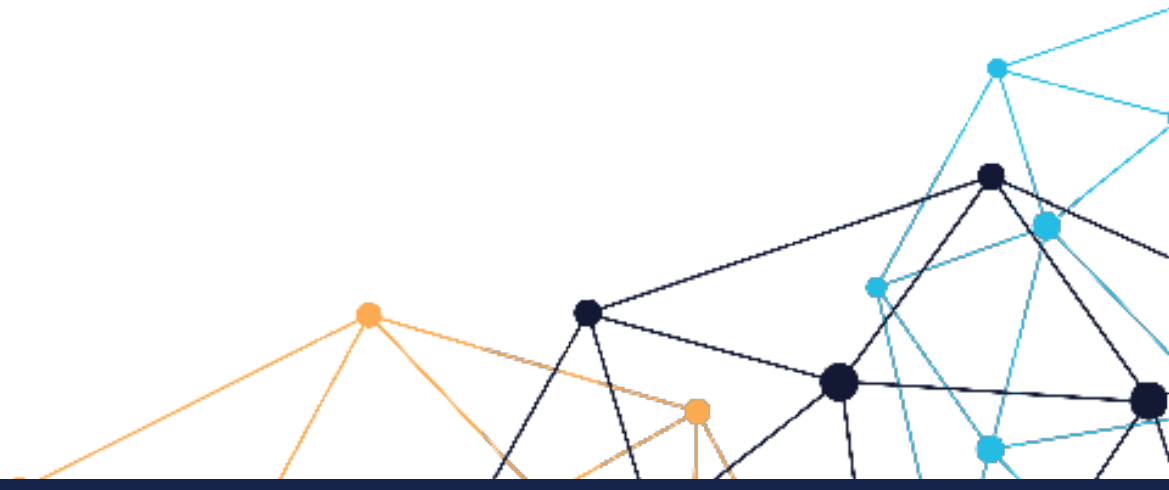


Overlapping Gradient GEMMs in PyTorch Backward Pass with Multi-Stream Execution

- **Student:** Zhou, Yun Fan (John)
- **Academic:** Prof. Gennady Pekhimenko
- **Industry:** Abhinav Vishnu
- **Company:** Advanced Micro Devices (AMD)

Deep Learning (DL) models are trained using a series of forward and backward passes on the models with mini-batches of input dataset. The backward pass – which is computationally expensive – comprises of a series of general matrix-multiplications (GEMMs), specifically that provide gradients on activation and the model weights. By default, these GEMMs are serialized, yet as the GPUs – the mainstays in DL learning – are becoming faster, there are significant opportunities to improve the performance by overlapping these GEMMs.

This project studies the performance improvement of overlapping gradient computations in the ROCm PyTorch framework – AMD software stack for MI Instinct GPUs. We create a custom PyTorch linear module with a CUDA/C++ extension, which during the backward pass, schedules the weight gradient GEMMs onto a separate stream while keeping the activation gradient GEMMs on the default stream; this allows us to overlap GEMM kernels while maintaining the correct order of execution of the GEMM kernels. Our approach achieves up to 1.1x speedup in GEMM kernel durations, and up to 1.1x speedup in the backward runtime of a network consisting of seven linear layers with ReLU activations. Future steps of the project are to study the tile-size selection of concurrent GEMM kernels to further optimize the performance gain from overlapping, and eventually to upstream the module with PyTorch.



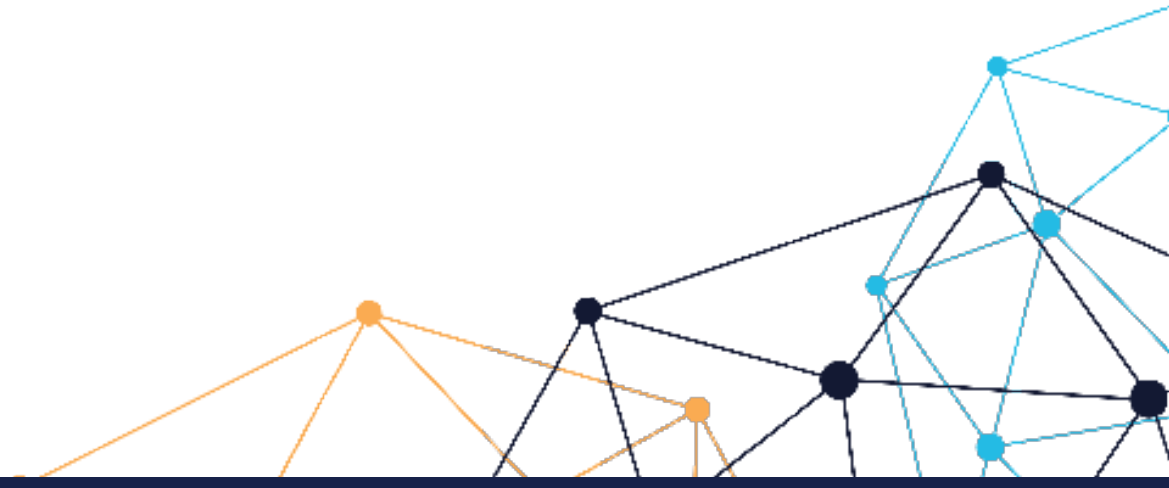
Contextual Targeting for Online Advertising

- **Student:** Zhu, Bingzhang
- **Academic:** Prof. Annie Lee
- **Industry:** Vincent Tembo
- **Company:** Pelmorex Corp

Users increasing care about the privacy of their data, motivating industry attempts to serve relevant ads to users without compromising their privacy.

Contextual targeting is a method which targets users based on the content they are consuming rather than history of their online behavior. This task is challenging because the ad-bidding process typically has a strict time limit, under 100ms. Additionally, since brands do not want to show their ads on irrelevant websites, a low false positive should be kept for the web categorization. Therefore, we attempted to categorize websites based on only URLs. Our work proposes a hybrid method to predict the industry standard categories for websites using TF-IDF and GloVe as the embedding feature, and Logistic Regression as the model.

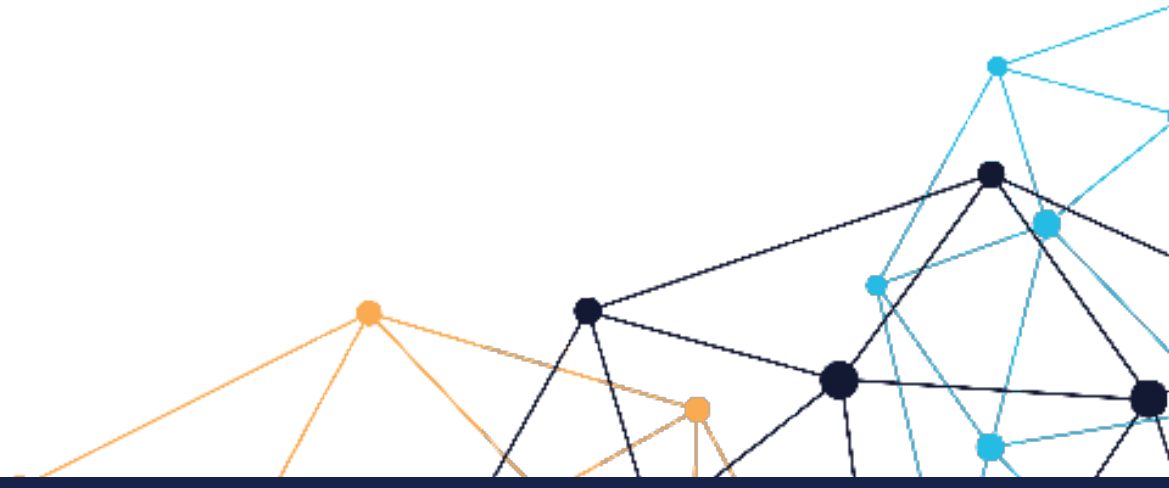
As a result, on average, 65.37% of the URLs in one day's ad-bidding requests could be categorized with the proposed method. And the F1-score for the categorization is 0.85. The categories here will be used in the context-based CTR prediction used by the ad-bidding process to get relevant ads to users.



Recursive LDA Topic Modelling for 'Similar' Short Texts

- **Student:** Zhu, Emily
- **Academic:** Prof. Linbo Wang
- **Industry:** Emily Ho
- **Company:** Fusion Analytics

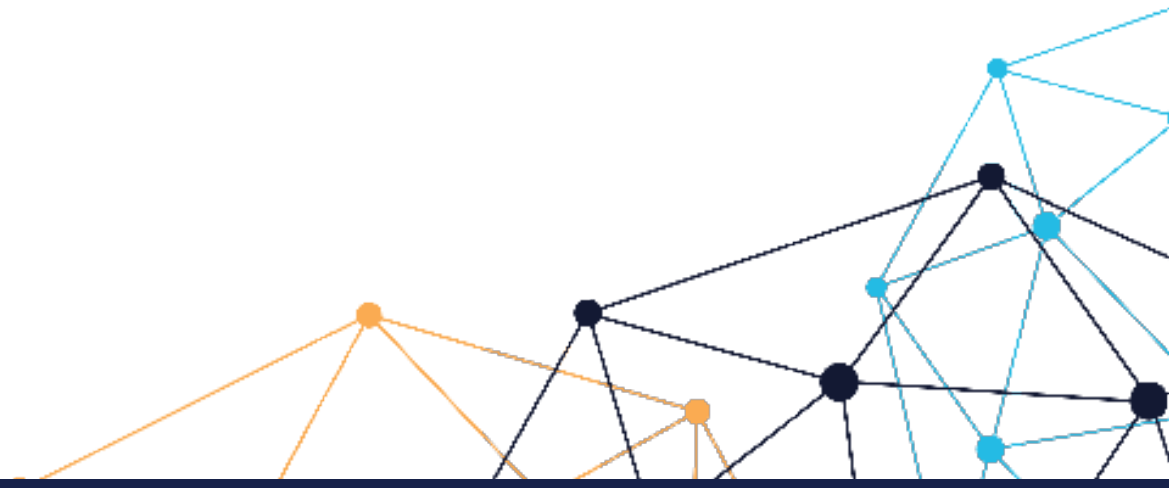
When it comes to extracting information from short, user-generated text, current topic modeling techniques have considerable space for improvement. They especially fall short when the text pool also includes different layers of abstraction, where the goal is to extract a hierarchy of subtopics. Currently, hybrid LDA models that focus in this area include hLDA and LF-DMM, but each are subject to constraints and limitations in their implementation. This paper proposes a hybrid LDA topic model called recursive LDA (rLDA), which is aimed to mine and construct a two-level hierarchy on small data pools of short unstructured text. Fundamentally, rLDA employs a recursive top-down approach when constructing the hierarchy, using word2vec to standardize words of similar semantic distance, and c-TF-IDF and coherence scoring algorithms to choose and split topics. In this paper, rLDA is implemented on eight pools of survey response data to extract the main messaging in the survey responses, and the results are compared against the popular topic models BERTopic and regular LDA.



A Transformer Based Approach to Automated Identification of Surgical Safety Checklists

- **Student:** Zhu, Tianshu
- **Academic:** Prof. Igor Jurisica, Prof. Frank Rudzicz
- **Industry:** Peter Grantcharov
- **Company:** Surgical Safety Technologies

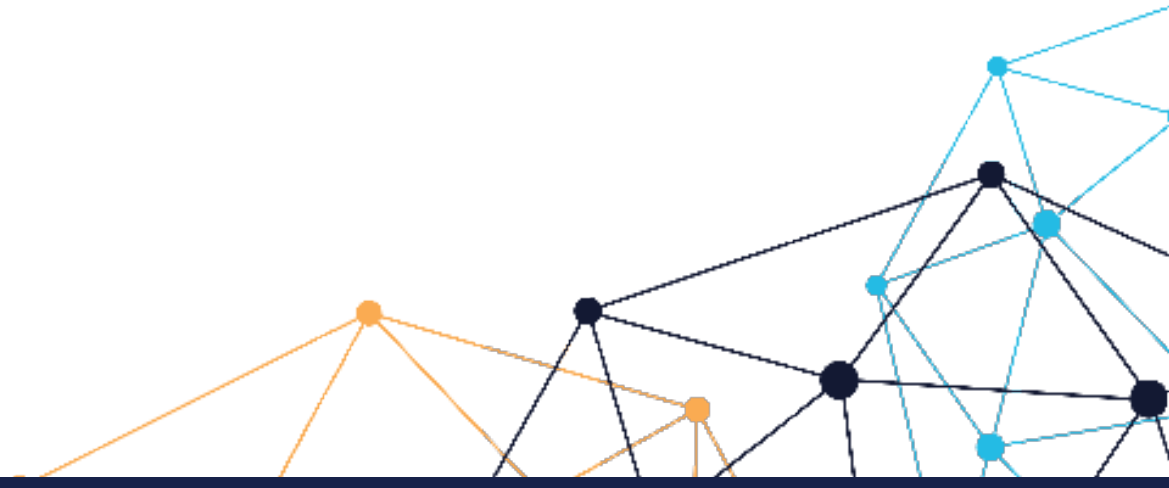
Automated identification of surgical safety checklist (SSC) aims to do Temporal Action Localization (TAL) on surgery case videos. To segment videos with different SSC steps (including briefing, timeout, debriefing, etc.) delineated by start and end times. Traditionally, analysts had to spend a lot of time watching surgical videos and manually annotating SSC steps. Because the surgical video is usually several hours or even dozens of hours long, we experimented with a transformer-based TAL model ActionFormer to better capture the global context of the video. This model is in a single shot, without using action proposals or relying on pre-defined anchor windows. It combines a multi-scale feature representation with local self-attention and uses a light-weighted decoder to classify every moment in time and estimate the corresponding action boundaries. Our model is tested on an internal dataset with over 400 cases and could potentially save between 50% and 80% of analysts' time for segmenting each case for SSC. We hope this work can set up a baseline for future research on not only the automated identification of SSC but any problem that involves TAL.



Calculation and prediction of battery capacity using data-driven models based on large-scale real-world data

- **Student:** Zhu, Yining
- **Academic:** Prof. Bo Wang
- **Industry:** Hossein Taghinejad
- **Company:** SOTI

Lithium batteries are widely used in mobile phones, electric vehicles and other scenarios because of their low manufacturing price and high energy density. However, due to side reactions, the capacity of the battery will gradually decrease. When the battery capacity is 70% of the rated capacity, it may not only make the device unable to complete the work but also pose a safety concern. Therefore, this research hopes to solve the following two problems: use the battery charge and discharge data to calculate the current battery capacity, and use the time series model to predict future battery degradation. The study uses more than six million data points from the real world in total. The existing papers mainly study the data obtained under laboratory conditions, so this research is a more extensive application of existing battery capacity algorithms. The research uses machine learning and deep learning algorithms including feature engineering based on the physical and chemical properties of batteries, sequence models of deep learning, and traditional models represented by Bayesian regression. In this study, we compared the accuracy produced by different algorithms. The accuracy of battery capacity prediction and the measurement of battery capacity are both above 98%.



Master of Science in Applied Computing (MScAC) Program University of Toronto

@UofTMScAC #MScAC

700 University Ave. 9th Floor
Toronto, ON M5G 1Z5

mscac@cs.toronto.edu
mscac.utoronto.ca
+1-416-978-5180