# Computer Science

## 43 - Evaluating the Reliability of SHAP Explanation Stability Under Feature Correlation

Tasrian Ahmed[1]

[1]University of Prince Edward Island

Explainable artificial intelligence methods such as SHapley Additive exPlanations (SHAP) are increasingly used to interpret machine learning predictions. However, explanations can vary when a model is retrained, raising questions about how reliably explanation stability can be estimated and how many model retrainings are required to obtain a sufficiently precise estimate. This ongoing study investigates the reliability of SHAP stability estimates under different levels of feature correlation. Synthetic binary classification datasets are generated with controlled correlation among informative features. Random Forest models are repeatedly trained on bootstrap samples, and SHAP values are computed on a fixed evaluation set to isolate variability arising from model retraining. The analysis considers three correlation conditions and evaluates explanation stability using three complementary measures: consistency of SHAP attribution magnitudes, reproducibility of feature rankings, and consistency of the highest-ranked features. Estimates obtained from different numbers of retrainings are compared with a 30-retraining pilot reference to assess how stability estimates change as additional model fits are included. Preliminary results indicate that variability in estimated SHAP stability decreases as the number of retrainings increases, with the largest improvements occurring when moving from small to moderate numbers of retrainings. The stability measures also exhibit different patterns, particularly under strong feature correlation, suggesting that no single metric fully characterizes explanation reproducibility. These findings motivate ongoing work to quantify the number of repeated model fits required for reliable stability estimation and to determine how feature correlation influences the precision and interpretation of SHAP stability measures across several synthetic data generation settings examined.

## 44 - Human Activity Recognition (HAR) using skeleton data

Omoroje Benjamin Alimele[1], Sazia Mahfuz[1]

[1]Acadia University

Skeleton-based human activity recognition (HAR) preserves privacy and is robust to lighting and background changes, but Kinect skeleton data is noisy and often contains spurious bodies caused by misdetected objects. Additionally, many top-performing models rely on heavy architectures that are costly to train and slow at inference. This work presents a lightweight model for skeleton-based HAR that aims to remain competitive in accuracy while substantially reducing model size and inference time. We develop a preprocessing pipeline that removes spurious bodies using frame length, aspect ratio, and motion energy, normalizes skeletons around the

main actor, and resamples sequences to a fixed length. The model combines spatial convolutional blocks, a bidirectional GRU, and a learned frame-attention layer, using position, velocity, acceleration, and bone-vector features with explicit handling of two-person actions. On NTU RGB+D, ablation studies show that motion features and two-actor handling provide the largest gains over a position-only baseline. These results suggest that careful preprocessing and feature design can close much of the performance gap with heavier models while maintaining a substantially smaller computational footprint.

## 19 - Persistent Initialization Susceptibility in a Hamiltonian Variational Eigensolver: Consequences for Reliable Benchmarking

Muhammad Umair A. B.[1]

[1]university of new brunswick

Variational quantum eigensolvers (VQEs), among the most studied algorithms for near-term quantum computers, estimate ground-state energies by tuning a quantum circuit with a classical optimizer, and the result can depend on where the optimizer starts. Benchmarks usually report random starts separately for each problem, so they cannot show whether a start that fails on one problem also fails on related ones. We studied this on the transverse-field Ising chain, a standard model of interacting quantum spins, using 16 qubits, a ten-layer Hamiltonian variational ansatz and exact noise-free simulation. Each random starting point was optimized independently at several magnetic field strengths. After 64 starting points revealed the pattern, we fixed the error threshold, field values and statistical tests before running 64 new ones. In the new runs, about two in five failed at the main test field, clearly separated from the accurate runs. Failure also tracked the starting point. Across seven field strengths, 42 of the 64 gave the same outcome every time, where fewer than three would be expected by chance. Starting all angles near zero removed every failure where tested, while a narrower random range and a capped L-BFGS-B optimizer did not. The failing runs we examined met the numerical criteria for suboptimal local minima. In this setting, runs on related problems were not independent, and treating them as such can overstate the evidence for a method. We recommend treating each starting point as the unit of analysis, reporting full outcome distributions, and comparing methods on matched random starts.

## 21 - Forecasting of Electric Vehicle Charging Behaviour: A Two-Stage Machine Learning Pipeline for Demand-Response Planning

Ditthi Chatterjee[1], Minming Ni[2], Dania Tamayo-Vera[1]

[1]University of Prince Edward Island, [2]BluWave-ai

Widespread electric vehicle adoption introduces a growing and highly variable source of demand on distribution grids. Utilities and demand-response programs therefore need more than an estimate of total electric vehicle load: they need to understand which vehicles are likely to charge, when charging is likely to occur, and whether these behaviours can be learned from historical data. This research develops an end-to-end machine learning pipeline for forecasting charging activity one day ahead for 568 vehicles collected from two telematics providers over a multi-year period.

Rather than modeling only aggregate demand, the proposed approach learns charging behaviour at the individual vehicle level. Historical charging records are used to capture recent activity, recurring temporal patterns, and per-device dependencies, while day-ahead weather forecasts are incorporated for vehicles with known regional information. The problem is structured in two stages. A binary classifier first predicts whether a vehicle will initiate charging the following day, after which downstream models estimate charging timing. Features are constructed using causal lag structures and cyclic time representations so that predictions rely only on information available at forecasting time.

Evaluation uses a chronologically held-out test set and walk-forward validation to prevent temporal leakage. Across eight tuned model families and 384 configurations, the best model achieved a score of 0.86 for predicting next-day charging occurrence, compared with a baseline decision tree score of 0.65. These results show that learning individual charging behaviour can provide a practical forecasting layer for demand-response and capacity-management planning by reliably identifying expected flexible demand before charging occurs.

## 23 - Automated Quality Assessment of Wood Products Using Computer Vision

Chhavi Chhavi[1]

[1]Acadia University

Quality assessment in wood manufacturing commonly relies on manual visual inspection, which can be time-consuming, subjective, and difficult to scale with increasing production demands. This research investigates the feasibility of using computer vision and deep learning to automate the quality assessment of manufactured wood products. The project focuses on developing an application-specific image classification system that learns visual characteristics associated with product quality and assigns quality grades in near real time.

Preliminary experiments evaluated the feasibility of convolutional neural networks (CNNs) for recognizing visual defects in wood surfaces. An experimental classifier achieved approximately 98% validation accuracy on a publicly available wood-defect dataset, demonstrating the potential of deep learning for distinguishing relevant visual characteristics. However, these results represent preliminary feasibility testing rather than the performance of the final grading system.

The next phase involves collecting and labelling application-specific images under representative manufacturing conditions to train and evaluate a custom CNN-based classification model. In parallel, an image-acquisition system is being developed to provide consistent image capture for

integration with the computer-vision workflow. The completed system will be evaluated based on classification accuracy, inference speed, and explainability, with its performance compared against an established computer-vision architecture. Ultimately, this research aims to determine whether an application-specific deep learning system can provide near-real-time wood grading with human-competitive accuracy and speed while supporting interpretable grading decisions.

## 11 - XGBoost for Network Intrusion Detection

Shekin David Raj[1]

[1]Department of Computer Science, Acadia University, Wolfville, Nova Scotia, Canada

Intrusion Detection Systems (IDS) play an important role in identifying malicious network activity, but class imbalance can make reliable attack detection challenging. This study investigates XGBoost for binary intrusion detection using the UNSW-NB15 dataset and compares its approach to a CNN-BiLSTM and SMOTE-based methodology presented in a previous work.

The dataset was preprocessed by removing identifier and attack-category fields and applying one-hot encoding to categorical network features. An improved XGBoost configuration achieved 87.36% accuracy, 82.16% precision, 98.42% attack recall, and an F1-score of 89.56% using the standard classification threshold. These results demonstrate strong sensitivity to malicious traffic while using a comparatively lightweight tree-based model.

The study provides a baseline for examining the differences between gradient-boosting and deep-learning approaches to intrusion detection. However, binary classification is not sufficient for identifying the specific types of attacks present in network traffic. Future work will therefore extend the study to multi-class classification, with the goal of distinguishing between different attack categories and evaluating how class imbalance affects the detection of individual attack types.

## 52 - Seizure detection using SeizeIT2 dataset

Dhriti Dhriti[1], Sazia Mahfuz[1], Esteve Hassan[1]

[1]Acadia University

Epilepsy is a neurological disorder marked by recurrent seizures, and timely, reliable seizure detection is essential for patient care. Continuous electroencephalogram (EEG) monitoring is the clinical standard, but it is largely confined to hospital settings, and EEG-only detectors face a persistent trade-off: tuning for high sensitivity produces many false alarms, while suppressing false alarms causes missed seizures. Wearable sensors that record movement and other physiological signals offer a complementary, independent source of information that may ease this trade-off. Our research work combines a focused review of recent machine learning approaches to seizure detection with experimental validation on the SeizeIT2 benchmark, which contains over 11,000 hours of multimodal recordings from 125 patients with focal epilepsy. We build a staged pipeline that begins with classical machine learning on handcrafted statistical and spectral EEG features, with feature relevance assessed by Information Gain and models

evaluated using cross-validation. Our research results show that a second, independent signal can make seizure detectors more useful in practice and they highlight current challenges and directions for wearable, real-world seizure monitoring.

## 12 - Workflow-Integrated Record Matching for Historical Ship Identification

Mohamed Elsabah[1]

[1]University of Prince Edward Island

Historical shipping records provide valuable information about trade, transportation, and economic activity to historians. Identifying and tracing the same vessel across large archival datasets becomes challenging when names are inconsistently transcribed or recorded. This project investigated how software could aid historians with reviewing ship records while preserving researcher control over uncertain matches. We developed our initial prototype in Python using Streamlit, fuzzy matching, similarity scoring, and a human-in-the-loop workflow. Although the prototype could identify duplicate or related vessel records, observing historians' manual workflow revealed an important limitation: Researchers work across spreadsheets, archival ledger images, and external ship registration sources to validate each record. Thus adding another application to their current workflow adds unnecessary workflow friction. Therefore, we have redesigned the system as an Excel tool embedded into the researchers' existing spreadsheet workflow. This prototype allows users to scan for candidate vessel records using vessel names, year ranges, and similarity thresholds. It also allows them to inspect relevant record information, select multiple candidate rows and apply changes while preserving the original source data. Usability features such as Undo and reviewer-tracking help prevent original information from being lost. The project highlights that effective research software should not depend solely on latest technology, but also on understanding the environment that the users work in. The design provides a foundation for more general spreadsheet-based tools supporting historical record linkage and annotation.

## 48 - Enhancing Alzheimer's Disease Detection Using MRI Preprocessing and Data Augmentation

Rosemary Etu-Efeotor[1]

[1]Dr. Lydia Bouzar-Benlabiod

Alzheimer's disease (AD) is a progressive neurodegenerative disorder for which early detection is essential to enable timely intervention and potentially slow cognitive decline. However, accurately identifying early-stage AD from structural magnetic resonance imaging (MRI) remains a significant challenge. This study investigates whether the performance of an existing 3D

convolutional neural network (CNN) can be improved through alternative MRI preprocessing, data augmentation, and class-balancing techniques.

The study will use a previously published CNN architecture as the baseline and evaluate its performance across different data preparation strategies. MRI data from the ADNI and PREVENT-AD datasets will be processed using different preprocessing and data augmentation approaches, with ADASYN applied to address class imbalance. The resulting models will be compared to determine how each strategy affects classification performance and which combination produces the best results. Model performance will be evaluated using established metrics, including accuracy and false-negative rate, to determine whether the proposed approaches provide measurable improvements over the baseline model.

In addition to predictive performance, this research considers the interpretability of deep learning models for clinical applications. Explainable AI techniques will be applied to identify neuroanatomical regions that influence model predictions and provide greater insight into the model's decision-making process. By combining improved data preparation with interpretability, this study aims to develop a more accurate and transparent approach to MRI-based Alzheimer's disease classification and contribute to the development of reliable and interpretable AI-assisted tools for early dementia detection.

## 4 - Formal Specification of Controlled Visual Dataflow Languages

Simon Gauvin[1]

[1]University of New Brunswick

In an important subclass of visual dataflow languages that includes many developed for industrial use, such as Prograph and LabView, programs consist of acyclic diagrams embedded in control structures of some form. We present here a formal specification of this class of languages, which we call Controlled Dataflow. To illustrate the formalism, we provide examples that shows how it captures conditional branching, iteration, procedural abstraction, recursion, polymorphism, exceptions, and concurrency. We report on use of this formalism as the basis for an implementation of new a CDL language workbench used to create new CDL languages through the declarative parameterization of graphical editors, debugger, compiler, and a dataflow virtual machine.

## 8 - Hydrology-Driven Leaching Risk Assessment through Remote Sensing: A Satellite-Only Generalised Additive Model Framework for Fertiliser Application Advisory

Samridhi Girdhar[1]

[1]University of New Brunswick Saint John

Poor fertiliser timing drives nutrient leaching in drainage water systems causing eutrophication, drinking water contamination and economic loss. Existing monitoring solutions rely either on expensive ground sensors limiting widespread deployment or static recommendations failing to reflect actual soil conditions. Remote sensing is a maturing technology and offering an alternative approach. However no validated framework integrates remote sensing into risk classification advisory for farmers. This study presents the Hydrology-driven Leaching Risk Assessment (HLRA) model, a five feature Generalised Additive Model (GAM) built using freely available MODIS and CHIRPS satellite data with no ground-sensor dependency. Each component of the model is a penalised spline function that maps to physically interpretable signal making it a transparent white-box model. Across 1,522 daily observations, the model achieved r=0.938, $R^2$=0.879 and Cohen's κ=0.840, with zero false positives. Temporal validation across unseen years demonstrated predictive generalisation, with hindcast r=0.917, $R^2$=0.839, κ=0.803 and forward-period assessment r=0.875, $R^2$=0.739, κ=0.847. Furthermore this study reports the first formal classification of leaching events into two types: firstly a hydrology-driven detectable from satellite and secondly stoichiometric driven events detectable only from ground sensing. It provides a key analytical contribution found by decomposing outputs from a prior ground sensor-based composite index (Nutrient Retention and Dilution Index). Overall, the results establish an interpretable, satellite-only framework delivering fertilizer application advisory to any temperate grassland farm providing a solution beyond controlled experimental farms into the hands of real farmers.

## 31 - Aligning and Steering Molecular Representations Across Graph Neural Networks and Large Language Models

Tushita Govindaraj[1]

[1]University of New Brunswick

Large language models (LLMs) are increasingly applied in chemistry, yet strong performance does not reveal what they learn about molecular structure. Graph neural networks (GNNs) treat molecules as graphs of atoms and bonds, whereas LLMs read linear strings such as SMILES. Multimodal architectures combining both achieve comparable accuracy, but evaluation has focused on prediction. Recent work maps GNN atom embeddings onto quantum-chemical quantities, but whether GNNs and LLMs arrive at related representations of the same molecule remains unclear.

We address two questions. First, to what extent are GNN and LLM embeddings of a molecule aligned, and at which layers and levels of granularity? Second, can property directions be identified in these spaces, transferred between them, and used to steer predictions controllably?

Phase one compares the embedding spaces of graph encoders and chemical LLMs such as MoLFormer on molecules, using representational similarity, cross-space mapping, neighbourhood preservation, and property probing at the molecule, functional-group, and atom levels, with untrained encoders and shuffled pairings as controls. Phase two, following contrastive activation addition, estimates directions for properties such as aqueous solubility and lipophilicity

from high- and low-value molecules, tests their stability across seeds and resamples, and adds them to held-out embeddings at varying strengths, measuring effects on target and unrelated properties. Where mapping is reliable, directions are transferred between spaces.

The anticipated contribution is a controlled account of how far graph and language representations agree and whether that agreement carries property directions between models. A negative result would also constrain multimodal model design.

## 42 - Girls STEM Up: Building Community and Visibility for Women in STEM Across Atlantic Canada

Tushita Govindaraj[1, 2]

[1]University of New Brunswick, [2]Girls STEM Up

Girls STEM Up (GSU) is a team of 40+ executive members across four universities in Atlantic Canada, working to make STEM more accessible for all girls, women, and underrepresented groups. Although more women are entering STEM fields, they still make up less than 30% of Canada's STEM workforce and are more likely than men to perceive STEM careers as difficult to succeed in. Visibility is also a barrier: one in four Canadian girls aged 15 to 17 reports not knowing a single female role model working in their dream job.

GSU addresses this gap through campus events, community initiatives, and an annual conference. The GSU Community, an online Slack space for female-identifying students and individuals in STEM, shares scholarships, bursaries, internships, research opportunities, jobs, conferences, and programs from across Canada and internationally. Each year, GSU hosts the only conference of its kind in Atlantic Canada, held in Fredericton. The next one, GSU 2027: Branch and Bloom, takes place March 13, 2027, and features keynote speakers, hands-on simulation labs, research poster presentations by high school and university students, professional breakout sessions, networking with sponsors and speakers, and bursaries and prizes for attendees.

This presentation will introduce GSU's mission and initiatives, show how students can get involved, and outline sponsorship and partnership opportunities for organizations interested in supporting women in STEM.

## 7 - A SMOTE-BASED APPROACH FOR MINORITY ATTACK DETECTION IN INTRUSION DETECTION SYSTEMS

Mohammod Haque[1]

[1]Acadia University

Class imbalance in network traffic datasets remains a fundamental obstacle to reliable intrusion detection, as models trained on skewed distributions tend to favour majority classes while failing

to detect rare but critical attack types. This thesis investigates whether a hybrid deep learning architecture can overcome this challenge on the UNSW-NB15 dataset over 2.5 million network flow records spanning ten classes with an imbalance ratio exceeding 12,750:1 between the dominant Normal class and the rarest attack category. A CNN-BiLSTM model is proposed, combining convolutional layers for spatial feature extraction with BiLSTM layers for sequential pattern modelling. The model is evaluated under four controlled configurations: a baseline with no resampling, standard SMOTE variants, per-class oversampling, and categorical-aware SMOTE-NC.The baseline achieved 98.00% overall accuracy but a Macro F1-score of only 0.5081, with near-zero detection rates for Analysis (2.39%) and DoS (5.08%)confirming that accuracy alone is insufficient for evaluating intrusion detection under imbalance. SMOTE-NC with a fixed target of 20,000 samples per minority class achieved the highest Macro F1-score of 0.5307 across all model families. Subsequent hyperparameter fine-tuning identified as the optimal configuration, improving Worms detection from 8.11% to 75.00% the largest single-class gain recorded while maintaining binary classification accuracy of 99.14%.These results demonstrate that categorical-aware oversampling combined with careful hyperparameter tuning meaningfully improves minority-class detection without sacrificing global model stability.

## 57 - Evaluating Multi-Agent Reinforcement Learning for C&C Beaconing Detection in Simulated Enterprise Networks

Daniel Hashemi Tehrani[1, 2]

[1]Acadia University, [2]Jodrey School of Computer Science

Autonomous cyber defence requires defensive agents capable of identifying malicious activity while operating under partial observability and coordinating responses across distributed network segments. This research investigates the use of Multi-Agent Reinforcement Learning (MARL) for detecting activity associated with the Command & Control (C&C or C2) phase of an Advanced Persistent Threat (APT) attack using the CybORG CAGE Challenge 4 environment. CAGE Challenge 4 provides a multi-agent enterprise network in which distributed defensive agents operate across separate network zones while responding to adversarial activity and maintaining benign services. This work will extend the environment with adversarial behaviours representing APT C&C beaconing activity and evaluate cooperative MARL defensive agents trained to identify and respond to these behaviours. Defensive performance will be evaluated under multiple attacker strategies, with experiments examining the effects of inter-agent communication, partial observability, and previously unseen adversarial behaviours. The research aims to determine whether distributed MARL agents can identify C&C beaconing activity while maintaining benign network operations and whether learned defensive policies remain effective as adversarial behaviour changes.

## 46 - Quadratic Kernel for Orchard Hybridization

Lathika Kathiresan[1], Norbert Zeh[1]

[1]Dalhousie University

Phylogenetic networks are generalizations of phylogenetic trees and allow us to describe non-tree-like evolutionary histories through reticulations, which are simply vertices with more than one parent. The *Hybridization Numbe*r problem takes as input a collection of phylogenetic trees and asks if there exists a phylogenetic network that contains an embedding of each of the input trees, while specifying no more than k reticulations. In this work, we study the hybridization number problem for the class of orchard networks. Orchard networks are biologically significant because their reticulations can be consistently interpreted as resulting from instantaneous horizontal gene transfers. Previous work has established that the problem is fixed-parameter tractable (FPT) for general networks and binary input trees, via a kernelization that produces a network with O(k^2) leaves, where k is the desired number of reticulations in the constructed network. We prove that the same result, with slightly larger constant factors, holds when restricting the constructed network to being orchard. The proof of this result is significantly more intricate than for the kernelization for general networks, and demonstrates that the *Orchard Hybridization Number* problem is FPT.

# 59 - Agentic Detection of Security Antipatterns in Microservices Architectures

Muhammad Shazil Khan[1], Francis Palma[1], Ahsen Tahir[1]

[1]University of New Brunswick

Security weaknesses in microservice systems frequently originate from architectural decisions, such as bypassing an API gateway or assigning excessive privileges to services, rather than from coding defects. Existing static analysis tools and evaluations of LLM-generated code predominantly address code-level security weaknesses (e.g., CWE-class vulnerabilities within individual functions), leaving design-level security antipatterns largely unexamined. This study investigates the capability of AI coding agents to detect security antipatterns in microservices and compares their detection performance. We first provide a catalog of nine security antipatterns for microservices, each specified with the detection criteria and not-applicable conditions, encoded as agent skill files. We then develop an experimental harness that executes three coding agents (Claude Code, OpenAI Codex CLI, and Google Antigravity CLI) through their command-line interfaces under a shared review protocol. For each microservice, the harness consolidates source and configuration files into a line-numbered representation, generates an architecture diagram, and initiates one session per antipattern. Each session returns a present, absent, or not-applicable verdict supported by file- and line-level evidence and a detailed rationale. The evaluation is conducted on 20 developer-built microservice repositories, annotated by two experts with a third reviewer adjudicating disagreements, forming the ground truth. As ongoing work, we are assessing the agent performance in terms of precision, recall, and F1-score. We plan to generate new microservice systems with these agents under varied prompting strategies, apply

the validated detector, and examine whether LLM-generated microservices exhibit design-level antipatterns and how prompt design influences their prevalence.

## 47 - Field-Scale Vegetation Anomaly Detection Using Integrated Sentinel-2, Crop Histories, and Hydroclimatic Data

Muhammad Maaz[1], Sara Sadri[1], Dania Tamayo-Vera[1]

[1]University of Prince Edward Island

Satellite vegetation anomalies can indicate crop stress, but low greenness can also reflect differences in crop development. We ask whether crop-rotation history during 2014-2024 helps explain potential stress in fields growing potatoes in Prince Edward Island in 2025, after accounting for regional rainfall and soil moisture.

We calculated monthly median normalized difference vegetation index (NDVI) from cloud-masked Sentinel-2 imagery at native 10 m resolution for May-September 2019-2025 and summarized observations within mapped field boundaries. Annual Crop Inventory maps supplied crop labels. The focused analysis will select fields growing potatoes in 2025 and characterize their preceding histories using potato frequency, time since the previous potato crop, previous crop identity, crop-change frequency, and crop diversity. We will assess 2025 vegetation conditions against calendar-month potato references from 2019-2024 and compare statistical models with and without crop-history indicators. Hydroclimatic inputs comprise HRDPA rainfall at 2.5 km and SMAP root-zone soil moisture at nominal 9 km resolution.

Existing analyses across crop types demonstrate the importance of crop identity: crop matching reduced the area-weighted mean May NDVI coefficient of variation from 0.42 to 0.22. Forage following potatoes fell below regional crop-specific thresholds in approximately 22% of field-months, compared with approximately 4.0% following forage.

The potato analysis will evaluate whether preceding rotation history adds explanatory information beyond regional hydroclimatic conditions. An interactive tool developed in Google Earth Engine supports field inspection and agronomic follow-up; NDVI anomalies indicate potential stress rather than confirmed crop failure.

## 10 - Heuristic Bandwidth Reduction Algorithms for Weighted Graphs

Alexander Matheson[1], Liam Keliher[1], Luis Varona[1]

[1]Mount Allison University

A layout of an (undirected, with positive edge weights) graph $G$ on $n$ nodes is an assignment of each of its nodes to positions 0, 1, …. $n - 1$. The unweighted bandwidth of a graph $G$ is the minimum integer $k$ such that there exists a layout of $G$ placing every pair of adjacent nodes at

most *k* positions apart. The weighted bandwidth of *G*, meanwhile, is the minimum real number *r* such that there exists a layout of *G* placing every pair of adjacent nodes at most *r* / *w* positions apart, where *w* is the weight of the edge between them. Reducing unweighted bandwidth has applications in fields like numerical analysis, computer data storage, and electronic design automation (EDA); thus, unweighted bandwidth reduction algorithms are well-studied in the literature.

On the other hand, very little work has been published on weighted bandwidth algorithms, even though they also have applications (albeit more niche) in EDA and quantum emulation. Although metaheuristics for unweighted bandwidth generalize fairly cleanly to the weighted variant, heuristics (which are often faster) are far more difficult to adapt. The present work is the first to explore nontrivial heuristic algorithms for weighted graph bandwidth. Benchmarking on VLSI circuits, weighted Erdős-Rényi graphs, weighted caterpillars, and more, we analyze both runtime and bandwidth obtained, finding that our methods generally outperform weighted metaheuristics.

## 5 - From Geometry to Function: A Multi-Task LiDAR Benchmark Dataset for Semantic and Functional Parsing of Urban Facades

Abdiaziz Muse[1]

[1]Saint Mary's University

Building facades play a central role in urban digital twins, heritage modelling, and sustainable city analysis. However, current LiDAR benchmarks focus only on geometry and semantic segmentation, leaving little room to reason about what each architectural element actually does. We present a large-scale urban facade benchmark spanning 26 buildings and roughly 492 million LiDAR points, captured across all four seasons, day and night, and a range of weather conditions. Each facade carries three tiers of annotation, semantic, instance, and per-instance functional attributes, making this, to our knowledge, the first facade dataset to combine geometric and semantic labels with functional ones. Labels are produced through a semi-automatic pipeline pairing geometry-guided heuristics with expert validation. Benchmarking a range of point-based, graph-based, convolutional, and transformer architectures, we find that semantic segmentation remains difficult under severe class imbalance, and that both separating repetitive instances and reasoning about function from geometry are considerably harder problems still. The benchmark lays new groundwork for function-aware 3D urban understanding.

## 1 - Demystifying the conventional choice of a key hyperparameter in the word2vec word embedding algorithm

Ufedo-Ojo Omale[1]

[1]University of Prince Edward Island

Word2vec is a foundational word embedding method. Its training relies on a negative sampling exponent hyperparameter that is conventionally set to 0.75 on empirical grounds. This preliminary investigation uses Kullback-Leibler divergence as a principled framework for studying the negative sampling exponent theoretically by comparing the empirical joint distribution of target-context word pairs with the distribution induced by negative sampling. We show that this Kullback-Leibler divergence is always minimized at a negative sampling exponent of 1 when the empirical joint distribution is used directly. In contrast, when the joint distribution is smoothed using Good-Turing estimation, the divergence-minimizing exponent shifts below 1. These findings establish a connection between negative sampling and smoothing, suggesting that the effectiveness of an exponent below 1 may be understood as compensating for properties of the unsmoothed empirical distribution. This work provides a starting point for a theoretical explanation of the conventional 0.75 exponent and motivates further analysis of the relationship between negative sampling and smoothing techniques used in natural language processing.

# 22 - Evaluating and Improving Large Language Model Accuracy for Occupational Cancer Information Retrieval using the ILO Diagnostic and Exposure Criteria

Jeremy Qiao[1]

[1]University of New Brunswick

Conventional Large language models (LLMs) can produce erroneous International Classification of Diseases (ICD-10) codes and carcinogen links when used by physicians for query-answer decision support, potentially leading to missed diagnoses and wrong compensation decisions for workers.

The Retrieval-augmented generation (RAG) approach helps by having the model refer also to a source document. This study presents an Ontology-Guided RAG system built on the occupational cancer section of the ILO Diagnostic and Exposure Criteria for Occupational Diseases (2022). An ontology built from the ILO document links 20 carcinogens to 15 cancer classes, 47 ICD-10 codes and National Occupational Classification (NOC) 2021 occupations, and was used to expand each question before retrieval.

In this study we used the Qwen2.5 model to answer a set of questions using only the ILO passages retrieved after ontology expansion. The performance was then compared with Illinois Chat, a document-grounded chatbot running the same model, against gold-standard answers from an occupational health specialist.

Across ten questions, ICD-10 precision rose from 75.21% to 88.61% using the Ontology-Guided RAG system, cutting the ICD hallucination rate from 24.79% to 11.39%. ICD F1 rose from 41.63% to 63.73%, source-page recall from 15.00% to 95.20%, and the mean composite score from 5.33 to 6.69/10. Overall, ontology guidance made LLM answers on occupational cancer more accurate and could improve decision support for experts in clinical settings.

# 51 - Reflection of Majority Choices with Experts' Preferences in LLM Alignment

Cameron Steeves[1], Peash Saha[1]

[1]University of New Brunswick, Saint John

Large Language Model (LLM) alignment with human values; the roundwise training of the models based on human feedback is an emerging research area. It gives AI-centric systems a better understanding of human preferences and makes them safer for society. Reinforcement Learning from Human Feedback (RLHF) the most established and widely used method to optimize human preferences in LLM alignment. However, recent work has discussed limitations of RLHF, such as producing misaligned outputs due to not understanding the underlying approaches of human preferences. One recent proposal replaces RLHF with maximal lotteries, a probabilistic Social Choice rule that reflects majority choices. This use of maximal lotteries, however, relies on an Anonymity axiom, which assumes that all voters carry equal weight. This assumption is problematic in certain domains such as healthcare, where non-expert feedback can outweigh expert consensus. The mechanism of maximal lottery fails to guarantee that expert feedback is preserved. This can be alarming for applications utilizing LLMs' capabilities, especially in domains where having expert feedback reflected in the majority choice is paramount. We propose a weighted, preference-based approach to ensure that expert preferences are prioritized over non-expert ones when maximal lotteries are applied in LLM alignment. Provided a sufficient proportion of the voters are experts, our algorithm guarantees that every result retains expert preferences to an acceptable extent. Experiments performed with synthetic instances suggest that the proposed algorithm is useful as a supplement to LLM alignment methods that reflect majority choices.

# 15 - VPNN Rotations as Fine-Tuning Adapters: What Helps, What Doesn't, and Where to Spend the Budget

Habiebatou Sam[1]

[1]UPEI

Large pretrained AI models can be adapted cheaply by freezing them and training only a small add-on, called an adapter. One family of adapters restricts each change to a rotation, which redirects what the model has learned without distorting it. We place the rotation layers of volume-preserving neural networks (VPNNs), which build large rotations from many inexpensive two-dimensional ones, inside a recent adapter, PSOFT, that rotates only each layer's most important directions and ask which parts of the combination matter. Across 452 training runs on five image tasks with a Vision Transformer, we report three findings. First, VPNN's rotation layers improve accuracy by up to 4.1 percentage points when replacing PSOFT's dense rotation on a task close to the model's original training data. This happens at budgets far below the smallest size of LoRA, the most widely used adapter. Second, the components that make VPNNs

volume-preserving do not help: a stretch-and-squeeze layer and a specialized activation function have no measurable effect. Third, how a fixed budget is spent matters. At equal cost, adding rotation layers rather than more rotations inside the same layer (width) lowers accuracy on familiar data by 2.4 points. On an unfamiliar task it appears to raise accuracy by 3.6 points, but training three times longer shrinks that advantage to 0.7. Extra depth therefore mainly speeds up learning rather than improving what is learned. These results show which parts of VPNN are worth keeping and how to allocate a small parameter budget.

## 50 - AI-Assisted Emergency Department Pre-Triage: A Multi-Agent Approach to Patient Prioritization and Wait-Time Management

Mithilesh Tumuganti[1], Lamija Kasumovic[1], Jeremy Qiao[1], Peash Ranjan Saha[1]

[1]University of New Brunswick, Saint John

Emergency department (ED) overcrowding is a major challenge in Atlantic Canada. In 2024-25, patients waited a median of 4 hours 28 minutes, and these long waits cause many patients to leave without being seen. Multi-agentic AI frameworks powered by large language models (LLMs) have shown promise for efficient solutions in healthcare. We propose a multi-agentic AI framework that reduces ED wait times by changing how the patient queue is managed. The model divides the triage workflow among specialized AI agents instead of relying on a single nurse or doctor. Each patient's case is assigned to an agent based on the information provided, such as symptoms, age, and recent medical history. The agent evaluates this information and historical records, assigns a Canadian Triage and Acuity Scale (CTAS) level, and estimates the patient's wait time. An outer-layer agent then analyzes all decisions from the inner-layer agents to produce an integrated queue with final CTAS levels and wait times for all patients. This lets agents collaboratively focus on specific clinical dimensions. Nurses and doctors remain the final decision makers, evaluating the outer-layer agent's decision and confirming or modifying it. A separate module explains the decision to patients and guides non-urgent patients toward alternative care options such as walk-in clinics. The system promises to lower the number of patients who leave without being seen and to let doctor/nurse focus on clinical work by reducing paperwork. Initial experiments with synthetic clinical cases support this claim. Such a framework can support resource-strained EDs across Atlantic Canada.

## 24 - Design and Optimization of Degree-Constrained Social Networks

Craig Brett[1], Luis Varona[1, 2, 3]

[1]Department of Economics, Mount Allison University, [2]Department of Mathematics & Computer Science, Mount Allison University, [3]Department of Politics & International Relations, Mount Allison University

Networks are seemingly ubiquitous in the economy, as social media, transportation systems, electrical grids, and supply chains bring people, goods, and ideas together. A noteworthy subfield in computational economics, therefore, is designing and optimizing such networks with metrics like efficiency and connectivity in mind. The present work considers a class of degree-constrained social networks wherein each agent invests some budget $X \in (0, 1)$ in a link to "communicate" with one other agent, forming a unicyclic functional digraph in which every node has indegree at most $D - 1$ for some degree cap $D \geq 2$. In this game-theoretic model, each agent derives a payoff equal to the sum of $X^d$ over the shortest-path distances $d$ to all other nodes in the network (where links, although directed, are traversable both ways).

For network design purposes, it is desirable to quickly identify all efficient network structures (i.e., maximizing the total sum of payoffs) given a fixed number of agents $n$, degree cap $D$, and link budget $X$. Utilizing structural theorems we have proven about optimal cycle size, pendant tree arrangement, and more, we present an $O(n^2 \cdot e^{O(\sqrt{n})} \cdot 2.9558^n)$ algorithm to do so. We then apply our algorithm to obtain extensive numerical results on efficient structures up to $n = 32$. Moreover, we identify various Nash equilibria, characterizing types of deviations and analyzing when Nash structures coincide with efficient ones.

## 38 - An Inter retable, Clinician-in-the-Loop AI Framework for Mobility Assessment and Fall-Risk Reporting

Sarun Weerakul[1]

[1]University of New Brunswick

Mobility assessment is central to evaluating balance, functional performance, and fall risk in older adults, but turning screening answers, test results, and population statistics into a clinical decision remains largely manual. Large language models (LLMs) can help summarize this information, yet in clinical settings they must explain their reasoning and remain under clinician control.

This research proposes an LLM-based framework that supports clinicians across three stages of mobility assessment. Before assessment, a retrieval-augmented generation (RAG) system draws on clinical fall-risk guidelines (e.g., STEADI), national fall-prevalence statistics, and test norms to produce an evidence-grounded risk profile and a recommended initial test battery, with each claim traceable to its source. During assessment, the system suggests test substitutions with a rationale when a patient shows floor or ceiling effects or safety concerns, and the clinician approves or rejects each suggestion. After assessment, it generates an evidence-linked clinical report, with patient-facing summaries released only after clinician review. The system is required to run fully locally (offline) to protect patient data, so all compared models are open-weight models that can be deployed on local hardware.

Rather than committing to a single design, we will compare existing RAG approaches (such as standard, graph-based, and agentic RAG) and several local language models against rule-based

(STEADI) and logistic regression baselines. Evaluation will cover factual accuracy, traceability of claims to evidence, and clinician usability and trust.

# Math

## 6 - Structure of Graph Polynomials

Damnada Abesiri Samarasekara[1]

[1]Acadia university

We studied properties of polynomials constructed from the upward closure of collections of minimal subsets. I studied how different structures of minimal subsets affect the resulting polynomial and identified conditions under which distinct collections produce identical polynomials. Additionally, I investigated the relationship between clutters and their blockers, including a polynomial relationship between the two, and explored double-dependent sets and properties such as symmetry and unimodality.

## 13 - Algebraic L-theory via Poincare Infinity Categories

Tanner Altenkirk[1]

[1]Dalhousie University

Algebraic L-theory was developed in the late 1960's by C.T.C. Wall and others as a means of classifying higher dimensional manifolds via surgery theory. Importantly, given a Poincare space of dimension greater than four, the theory defines a sequence of L-groups that encode the obstructions to surgery, which in turn tells us if the space is homotopic to a manifold. Ranicki formulated these L-groups using chain complexes and showed that the L-groups form a spectrum with four-periodicity.

More recent joint work by Baptiste Calmès, Markus Land, Thomas Nikolaus, and six others reformulates algebraic L-theory using the notion of a Poincare infinity category. The authors show that this reformulation recovers Ranicki's theory, and furthermore it shows that one can do L-theory on any Poincare infinity category, providing a substantial generalization of the theory.

In this talk, I will introduce Ranicki's chain complex formulation of L-theory and outline some challenges associated with it. I will then introduce the infinity categorical formulation of Calmès et al and demonstrate that this approach is not only more general than the original theory, but also more natural.

## 17 - Distinct Residues of the Gauss Factorial Modulo a Composite

Jonah Barrington[1]

[1]Dalhousie University

When we observe the residues of factorials modulo a prime p > 3, there can be at most p - 2 distinct non-zero values. However, it is still an open question whether there are any primes p > 5 with exactly p - 2 distinct factorials. In earlier work, this problem was extended to composites using a concept called the Gauss factorial. In this presentation, we first give an algorithm for detecting moduli with a maximal number of distinct Gauss factorials, alongside calculations using this algorithm. Then we present various restrictions on moduli that can have this property, and we show that such moduli must be square-free.

## 37 - Modelling Opioid Mortality during the COVID-19 Pandemic

Carter Clifton[1], Ryan Lukeman[1]

[1]St. Francis Xavier University

The opioid epidemic is a global public health crisis, which has had significant impacts across Canada, where over 50 thousand opioid related fatalities have occurred since 2016. During the COVID-19 pandemic, the impacts of this crisis continued to grow, with the number of opioid mortalities nearly doubling and remaining elevated for multiple years afterwards. To quantify the factors which led to this increase, we developed an individual-based model which allows us to track individual characteristics to quantify overdose and subsequent opioid mortality risk.

This model is first parametrized to match data from 2019 in Toronto. Then, using available data, we sequentially modified characteristics and parameters known to have changed during the pandemic to quantify the change in mortality associated with each.

We found that although these factors do show an increase in opioid mortality, they alone do not account for the large increase during the pandemic. Instead, we attribute this by a simultaneous increase in the presence of fentanyl and fentanyl-analogues in the drug supply. To quantify this increase in fentanyl, we use the limited data on the makeup of the drug supply along with methods from Bayesian statistics. This modelling work provides quantitative information on the state of fentanyl in the illicit drug supply which is valuable to guide drug policy decisions across Canada.

## 35 - Graph Burning: When Selection Does Not Matter

Gabe Denney-Martell[1]

[1]Acadia University

Graph burning is a discrete-time process that models the spread of information, influence, or a fire through a graph. At each round, fire spreads from previously burnt vertices and a new unburnt vertex is selected to burn. The burning number is the fewest number of rounds to burn every vertex, and a variant called graph cooling seeks to find the largest number of rounds. This presentation explores when these numbers are equal. That is, when the graph burns in the same number of rounds regardless of the choice of sources. Such graphs are called burning invariant. We characterize the burning invariant graphs with diameter 2, and show examples of larger diameter burning invariant graphs with particular attention toward vertex transitive graphs. It is not known whether there exist burning invariant graphs with diameter larger than 6. Applications to liminal burning, a variant where the choice of vertices is restricted, will also be mentioned.

## 2 - An Application of Graph Colouring and its Extension to List Colouring.

Katrina de Vera[1]

[1]n/a

A graph colouring is a lalbeling of the vertices of a graph with colours that are represented by integers. A list colouring is a graph colouring where each vertex is assigned a set of integers, called a list, in which each vertex receives a colour from its list. In this presentation I will present an application of graph colouring, called the scheduling problem and demonstrate how list colouring can be used to model the scheduling problem subject to additional constraints.

## 34 - Unsupervised Classification of Extreme Temperature Indices in New Brunswick: Contribution of Dimensionality Reduction (PCA, t-SNE, UMAP)

AIDA MARGOT DIOP[1]

[1]Université de Moncton

This study analyzes the evolution of temperature extremes in New Brunswick using CMIP6 climate projections (CanESM5) covering 1950-2100. The objective is to characterize the spatiotemporal structure of these extremes by identifying homogeneous climate regions.

The methodology relies on a set of fifteen extreme temperature indices (ETCCDI), standardized prior to analysis. Given the high dimensionality of these climate dataset, multivariate statistical analysis is used to model their interdependencies and reduce the dimensionaality. The originality of the project lies in the comparison of three dimension reduction methods, namely principal component analysis (PCA), t-SNE and UMAP, as a preliminary step to k-means clustering, in order to (1) classify historical and future thermal profiles and (2) identify spatiotemporal transition zones.

The results highlight a progressive shift toward warmer conditions, a decrease in regional diversity and an intensification of hot extremes across time horizons. This analytical framework provides a scientific basis to guide adaptation policies in response to temperature extremes in Atlantic Canada..

## 45 - Simultaneous Holotheticity and Bi-Hamiltonian Structures in Economic Growth Theory

Sarah Finkle[1], Roman Smirnov[1]

[1]Dalhousie University

This talk introduces a geometric framework applying bi-Hamiltonian structures and Nambu-Poisson mechanics to macroeconomic growth theory. Utilizing Ryuzo Sato's principle of simultaneous holotheticity, we demonstrate that economic production functions emerge natively as stable, time-independent geometric leaves of integrable phase flows. We establish a hierarchical taxonomy of three sequential growth regimes: the unconstrained Cobb-Douglas model, a resource-limited sigmoidal response, and a capacity-bounded "overshoot-and-collapse" regime. Finally, we address structural boundary crises where trajectories encounter resource ceilings. By deploying Filippov's convex reconstruction and a Dirac bracket reduction, we project the three-dimensional bi-Hamiltonian hierarchy onto a lower-dimensional constraint manifold, preserving algebraic compatibility and modeling systemic structural locks.

## 40 - Surface Codes and Factor Graphs

Melanie Gauthier[1]

[1]Dalhousie

A surface code arranges physical qubits on a two-dimensional patch and repeatedly measures local parity checks called plaquettes. Each plaquette checks whether an even or odd number of errors has appeared in its adjacent qubits. The pattern of plaquettes with errors is called the syndrome.

We are able to turn the surface code into an associated graph called a factor graph. These graphs have variable nodes which represent the qubits, and factor nodes which represent the plaquettes. We can use a process called belief propagation to pass messages along the edges of the graph and get back an estimate of each qubit's marginal probability of having an error. The main drawback is that the marginals of several variable nodes frequently tie. We implement this algorithm and show some results.

# 16 - Extensions of the Asymptotic Iteration Method for Linear Differential Equations and Coupled First-Order Systems

Mingyue Gloria Guo[1]

[1]University of Prince Edward Island

This presentation introduces the Asymptotic Iteration Method, usually abbreviated as AIM, and develops extensions of the method for systems of coupled first-order linear ordinary differential equations.

Our work begins with the classical scalar form of AIM for 2nd-order linear differential equations and then extends the main ideas to systems of two and three equations.

First, an extended AIM theorem is studied for 2nd-order linear differential equations. The result shows that the AIM solution formulas remain valid for a related modified differential equation even when the usual termination condition is not satisfied. The corresponding equation is also transformed into Liouville normal form, which gives a direct connection with Schrödinger-type equations.

For 2-dimensional coupled systems, four AIM coefficient sequences are introduced. When suitable ratios of these sequences stabilize, the system can be reduced and its general solution can be written in integral form. An extended AIM theorem is also obtained for the two-dimensional case. For constant coefficients, the stabilization conditions reduce to quadratic algebraic equations, and the resulting solutions agree with those obtained from the eigenvalue method.

The method is then extended to three-dimensional coupled systems. Repeated differentiation produces nine coefficient sequences, and the simultaneous stabilization of two ratios gives a linear relation among the three unknown functions. This relation reduces the original system to a nonhomogeneous two-dimensional system. The stabilization functions are also shown to satisfy a coupled Riccati-type system. Several examples with constant and variable coefficients are included to illustrate the method.

# 56 - Counting the Twist Isoclasses of the Irreducible Representations of M_5

Jase Kochhar[1]

[1]Cape Breton University

This talk will discuss research in counting the twist isoclasses of the irreducible representations of the nilpotent group of maximal class 5, M_5. The goal of this research is to construct, for exceptional primes p, the p-local representation zeta functions of M_5. These functions are conjectured to be equal to the corresponding functions for non-exceptional primes.

We will discuss the background and structure of M_5 and its irreducible representations, as well as explore examples of the calculation to count said representations. For M_5, the exceptional

primes are 2 and 3, as they are less than 5. The structure of the irreducible representations of M_5 with 2^N and 3^N dimensional matrices is much less uniform than for larger primes, and thus brings more complexity to the enumeration of all isoclasses.

## 26 - Determining the Winner of Some GraphDistance Games is PSPACE-complete

William Lawlor[1], Svenja Huntemann[1]

[1]MSVU

Combinatorial Game Theory studies 2-player games of perfect information. One such game is GraphDistance(D, S), which is played on a graph and a player may not play on a vertex at distance d ∈ D from another player's vertex or a distance s ∈ S from one of their own. We examine the computational complexity of determining whether the next player to move has a winning strategy. We show several families of these games to be PSPACE-complete by reducing from Bounded 2-player Constraint Logic. This is joint work with Svenja Huntemann (MSVU).

## 53 - Bounded Mean to Bounded Lower: Boundedness of the Maximal Operator from BMO to BLO

Luke MacKinnon[1]

[1]Cape Breton University

It is well-known that the Hardy-Littlewood maximal operator $M$ is bounded from $L^p$ to $L^p$ for $1

## 18 - Relating Wallpaper Groups Using Orbifold Coverings

Ana Mastnak[1]

[1]Dalhousie

Wallpaper groups are interesting geometric groups that can be viewed as patterns by an action of the group on the plane. Objects called orbifolds are created by quotienting the pattern by the group action, identifying points in the same orbit. We use these orbifolds and a modified version of topological covering maps to relate the wallpaper groups to each other, creating a hierarchy of who covers who. However, proving the lack of a covering is difficult, thus we introduce the concept of the Tom Dieck fundamental groupoid, and lead into future work.

# 29 - The Karlhede Bound in the case of Riemannian 3 Manifolds is Sharp

Asher Millett[1]

[1]Dalhousie University

A. Millett

How can you tell whether two Riemannian metrics describe the same geometry in different coordinates? The Cartan-Karlhede algorithm answers this. It works through the curvature tensor and its covariant derivatives one order at a time, and stops once no new invariant information appears. The order at which it stops, q, measures how much differentiation a geometry needs before it is fully classified. For Riemannian 3-manifolds a counting argument gives $q \leq 4$. Until now it was not known whether this bound is actually reached.

This talk shows that it is. We derive an explicit family of metrics

$g = 2(dt + 2z\,dy)^2 + (2/s^2)\,dy^2 + (dz - (n(y) + 4t)\,dy)^2$, where $s'^2 = Cs^4 - 2\lambda s^2$,

and show that every member achieves $q = 4$. The proof narrows the problem one step at a time.

Curvature: in three dimensions the Riemann tensor reduces to the Ricci tensor, which forces Ricci eigenvalues of type (2,1).

Lie theory: a U(1) charge decomposition separates the shear (charge 2) from the twist (charge 0). This shows that symmetry must be broken at first order.

Mathematical physics: gauge fixing, a Newman-Penrose-style set of structure equations, and a conserved quantity along a geodesic flow turn a second-order PDE into a solvable ODE.

Integrating the resulting frame gives the metric above. Computing scalar curvature invariants then confirms that the algorithm follows the unique sequence of steps allowing maximal order. This shows the Riemannian Karlhede bound is sharp, matching known Lorentzian results.

# 54 - BMO and the Tale of Two Jumps

Jesse Milley[1]

[1]Cape Breton University

BMO (Bounded Mean Oscillation) is a function space with a difficult norm to compute. Its norm is defined as the supremum of the mean oscillations over all intervals between [0,1]. This presentation offers an introductory look into this space with examples of norms that we can

currently calculate. We will end with a mention of ongoing research into calculating the norms of other functions.

## 39 - Symmetry Breaking of Wallpaper Patterns and its Relation to Orbifold Covering Maps

Evan Pelley[1]

[1]Dalhousie University

In classical topology, the fundamental group of a space can be described in terms of loops about a basepoint or as the group of deck transformations of the space's universal cover. Euclidean 2-orbifolds are obtained as quotients of the plane by wallpaper groups. For these orbifolds the universal covering space is the plane. Using this, we construct the orbifold fundamental group from deck transformations.

It turns out that for Euclidean 2-orbifolds, the orbifold fundamental group is precisely the wallpaper group. This allows us to study how the 17 Euclidean 2-orbifolds cover each other, leading us to an intuitive method for visualizing coverings of wallpaper patterns.

## 36 - Travelling Clusters in Bacterial Aggregation

Timothy Power[1], Hildur Knutsdottir[2], Leah Edelstein-Keshet[2], Theodore Kolokolnikov[1]

[1]Dalhousie University, [2]University of British Columbia

Cyanobacteria aggregate into localized clusters of high density. An agent-based model of this has been coarse-grained to a fourth-order nonlinear PDE possessing an exact stationary cluster, but real colonies are transported by fluid flow and biased by light. We add directed motion in two ways: an intrinsic cell velocity, and a bias in each cell's choice of neighbour. In both cases we derive the corresponding PDE. Seeking a travelling cluster, together with the scaling symmetry of the equation, reduces the problem to a first-order Abel equation, which a quadratic ansatz closes exactly. This determines both the cluster profile, an exponentially tilted power of the hyperbolic secant, and its group velocity in closed form. Existence and stability thresholds separate as soon as the cluster moves, producing a parameter gap in one model and bistability in the other.

## 28 - Cops and Robber Reliability Polynomial

Gage Richard[1], Josh Robar[1], Danielle Cox[1], Denys Bulavka[2]

[1]Mount Saint Vincent University, [2]Dalhousie University

In this talk, we will introduce a new polynomial, the Cops and Robber Reliability Polynomial. We will begin with an introduction to the game of Cops and Robber

and network reliability. Then, we will present results on the design of optimal graphs, that is graphs where there is a high probability that the cop will capture the robber. This is joint work with Josh Robar (MSVU), Denys Bulavka (Dalhousie), and Danielle Cox (MSVU).

## 20 - Bounding the Average Order of Dominating Sets on Tournaments

Colin Roberts[1]

[1]Acadia University

For a given oriented graph G and a vertex subset S, we call S a “dominating set” if the closed out-neighbourhood of S is the entire vertex set of G. All oriented graphs with at least one arc have multiple such dominating sets. The average cardinality of all these sets is called the “average order of dominating sets of G”, and may be studied as its own graph parameter. In this presentation, we find bounds for the parameter across the family of oriented complete graphs, also known as tournaments.

## 49 - Reflections on Rational Orthogonal Matrices

Rory Schepp[1]

[1]Dalhousie University

This talk concerns the group of orthogonal matrices with rational entries. Many standard results from the real setting carry over, including the decomposition of orthogonal matrices into reflections. We discuss generators and relations for this group, with particular attention to generators that act on only a small number of coordinates. Some of these questions are purely linear-algebraic, while others depend on arithmetic properties of the rationals. Related questions and possible generalizations will also be mentioned.

## 3 - Optimizing the Location and Capacity of Primary Care Clinics in Prince Edward Island Using Capacitated p-Median Models

Rupinder Singh[1]

[1]University of Prince Edward Island

Many Prince Edward Islanders face limited access to primary care. As of July 31, 2026, more than 35,000 residents were registered as waiting to be connected to a primary-care provider. This

study examines how clinic location and service capacity jointly affect geographic access across the province. We formulate a series of capacitated p-median facility-location models using 2021 population data for Prince Edward Island's 319 dissemination areas and road-network travel distances. The models examine the effects of the number of clinics, uniform clinic capacities, and different rural and urban capacity levels. We also compare the locations and modeled capacity of the province's existing 13-site primary-care network with optimized alternatives. The results show diminishing reductions in travel as additional clinics are introduced and demonstrate that capacity constraints can substantially increase travel, particularly for rural residents. The existing primary-care network has an average travel distance only 11.2% above that of a travel-optimal 13-site configuration, suggesting that its broad geographic placement is reasonably efficient. However, under the staffing and panel-size scenarios considered, the network has sufficient modelled capacity to serve only about 44% of the provincial population. These findings indicate that insufficient service capacity, rather than clinic placement alone, is a major barrier to primary-care access in Prince Edward Island and illustrate how optimization can support healthcare location and capacity-planning decisions.

## 30 - An Introduction to Symmetry, Wallpaper Patterns, and Orbifolds

Madeleine Taylor[1]

[1]Dorette Pronk

Symmetry surrounds us in nature, art, architecture, science, and mathematics. This talk will begin with a guide to enumerating symmetric patterns, particularly "wallpaper patterns" (repeating patterns on the plane), using the "orbifold signature." It will then transition into an intuitive introduction to orbifolds and how these topological and geometric objects relate to wallpaper patterns.

## 14 - Geometry in Random Graphs

Noah Wayne[1]

[1]Dalhousie University

Much research has been done in the area of randomly generated graphs, but there is a distinct lack of research on these graphs in high dimension. It is known that in high dimension, the "geometry" of the random geometric graph (RGG) dissapears, and the graph morphs into the non-geometric Erdos-Reniy graph (ERG). My work this summer focused on, experimentally, showing this result and proposing a similar relationship between the Spatial Preferential Attachment model and the Preferential Attachment model.

# Statistics

## 9 - An Age Grouping Framework for Multi-Population Mortality Modeling

Cezar Campeanu[1]

[1]University of Prince Edward Island

This study extends existing mortality prediction frameworks by incorporating information borrowed from population-gender-age subgroups that exhibit similar mortality patterns. The borrowed information is integrated into classical mortality models to improve the accuracy of future mortality rate forecasts. To capture structural similarities among mortality trajectories, several distance measures are evaluated in combination with four linkage methods, particularly when each subgroup comprises multiple age-specific mortality trajectories. Extensive empirical analyses using data from the Human Mortality Database demonstrate the superior predictive performance of the proposed approach.

## 58 - Causally guided predictive models can lead to better ecological forecasts.

Noah Coleman[1], Anthony Lecoure[1], Suchinta Arif[1]

[1]University of New Brunswick Saint John

Predictive models are ubiquitous across ocean science but can often fail to generalize across space and time, particularly when environmental conditions change from the training data. Incorporating causal knowledge into a predictive model's structure and learning process offers a promising way forward for improving predictive performance under environmental change. This talk will synthesize contemporary cross-disciplinary advances for applying causally guided predictive models, bridging insights from causal inference and AI. Simulated ecological examples will be used to illustrate when and why causal approaches outperform traditional correlation-based models. Ultimately, causally guided predictive models can improve the transferability of species distribution models, population forecasting, and other predictive tools used to support ocean conservation and management under a dynamic and changing world, and can be a widely applied by ocean researchers, stewards and practitioners.

## 27 - Partial Identification of Causal Effects using a financial analogy

Emmanuel Mindia[1]

[1]University of Prince Edward Island

Sometimes the causal effect of a treatment cannot be determined exactly from the available data. In these cases, the honest answer consists of upper and lower bounds rather than a single value. We consider models in which the effect can be determined exactly if the probability distribution of the observed data is fully known. When its parameters are known only to lie within intervals, we obtain the causal bounds using a backward induction algorithm on a tree. The method mirrors the computation of bid and ask prices in an incomplete financial market. For finite samples, we use individual confidence regions for each separate parameter to obtain a confidence interval for the effect through a Bonferroni-type approach. Our method preserves their overall confidence level, although the resulting interval can be conservative. In a simulation study, our method favourably with the bootstrap in terms of coverage when there are positivity or near-positivity violations.

## 55 - Investigating Methods of Fatty Acid Sub-selection which Maximize Separation of Prey Groups

Keely Mueller-Classen[1]

[1]University of New Brunswick, Department of Mathematics and Statistics

Quantitative Fatty Acid Signature Analysis (QFASA) is a method of diet estimation which minimizes the distance between the predator's fatty acid signature and the mean of the prey signatures. Commonly, real-life data used for QFASA have substantial overlap in prey groups due to similarities in their fatty acid compositions. Due to the overlap, QFASA may have issues with identifying the proportions of similar prey species in a predator's diet. Selecting a subset of fatty acids that maximize separation among prey groups may improve QFASA estimation of diet proportions. In this talk, a method of fatty acid selection will be explored. This method uses Linear Discriminant Analysis (LDA), after accounting for the compositional nature of the data, for maximizing prey separation, followed by a fatty acid selection procedure developed by Neubauer (2015). A simulation study as well as a comparison using real-life grey seal data between QFASA diet estimates, based on the conventional set of fatty acids, and those using only a subset are conducted to determine the estimation ability of the fatty acid sub-selection methodology.

## 32 - Gaussian process emulation of spatio-temporal outputs of a 2D hydrodynamic model

Asma Oussar[1]

[1]Université de Moncton

Spring floods on the Saint John River (New Brunswick, Canada) pose a growing risk to riverside communities, and their assessment under climate change requires a large number of hydrodynamic simulations. Two-dimensional models such as HEC-RAS provide detailed representations of flood extent and depth, but their high computational cost limits the number of feasible runs and makes uncertainty analysis difficult. This study aims to develop a Bayesian

emulator based on Gaussian processes, a fast statistical surrogate that reproduces HEC-RAS outputs while quantifying the uncertainty of its predictions. Since standard Gaussian processes become computationally prohibitive for high-dimensional spatiotemporal outputs, the objective is to compare scalable approaches, namely dimension reduction combined with Gaussian processes, Low-Rank approximations, and Nearest-Neighbour Gaussian Processes (NNGP), in order to identify the best compromise between accuracy, uncertainty quantification, and computational efficiency.

## 60 - How weather affects daily catch in LFA 27

Morgan Taylor Rose[1]

[1]St. Francis Xavier University

Lobster fishing is a significant part of Nova Scotia's economy. Daily catch rates can be inconsistent. Fishermen have to rely on experience and weather forecasts to anticipate what their catch will be. This research investigates whether historical weather data can help explain inconsistencies in daily lobster catch in northern Cape Breton's LFA 27. Catch records were collected from 12 local fishermen, covering the 2015, 2022, 2023, and 2024 fishing seasons. These records were manually entered into a spreadsheet, cleaned, and used to calculate median daily catch values. Historical weather data, including maximum wind speed, precipitation, and daily temperature, were obtained through the Open-Meteo API and merged with the catch records. Statistical analysis and exploratory data visualizations were used to examine seasonal catch distributions and relationships between weather conditions and catch rates. Results showed significant differences in median catches between years, with 2024 recording the highest median catch. However, weather variables demonstrated weak relationships with catch rates. Temperature showed a slight positive trend, while wind speed and precipitation showed a slight negative trend. This suggests that the examined weather variables alone cannot reliably predict daily catch. Differences in trap depth, locations, and other harvesting practices may better contribute to the variations in catch. Despite the analysis not showing strong predictive relationships, it demonstrates how statistical methods can combine real fishing records with environmental data to investigate local fisheries. Further research incorporating larger data sets and more precise weather data may help improve our understanding of catch variations.

## 33 - Modelling Opioid Dispensing patterns in Nova Scotia following the March 2020 CDSA Subsection56(1) Exemption

David Simonov[1]

[1]Acadia University

Opioids related harms involving misuse and overdose have been a topic of concern over the years, with the number of incidents being on a sharp incline in recent years. Following the COVID-19 pandemic, a new discussion arose regarding the accessibility of first and subsequent

prescriptions due to reduced social interaction between health providers and patients. These concerns prompted the Controlled Drugs and Substances Act (CDSA) subsection 56(1) policy intervention, which granted a temporary exemption for controlled substance prescriptions, allowing direct patient care pharmacists to prescribe, extend and renew opioid treatment. In collaboration with Dr. Sanyal and Dr. Zhang, the current study examined prescription patterns among health providers using data from the Nova Scotia Prescription Monitoring Program (NSPMP) over 26 months following the CDSA exemption.

This presentation walks through the analysis from raw claims to a fitted model. We cover data cleaning, descriptive statistics comparing first and subsequent fills, and model selection. The presentation discusses variance-to-mean ratio imbalances and how monthly counts led to a negative binomial model with a patient-count offset and a random intercept. We close by interpreting the results, with attention to how covariates were chosen and handled.

## 41 - Averaging and Transforming Repeated Measurements in Longitudinal Data: Does the Order of Operations Matter?

Patrick Talbot[1]

[1]University of New Brunswick, Saint John

Repeated measurements are commonplace in longitudinal surveys, where multiple observations are made on each individual. One way in which the repeated measurements of an individual can be combined is by taking the mean.  In many cases, the data are further transformed to another scale for analysis. When doing so, a choice must be made regarding the order in which averaging and transforming take place: should we average the repeated measurements of an individual and then transform the mean, or should we perform these two operations in the reverse order? In this talk, we attempt to answer this question for certain transformations, with our notion of the “correct” order of operations resulting in the estimator with smaller variance. Our focus rests on a transformation used in nutritional epidemiology. For a given component of this transformation, we approximate it with a smooth function and assess the variance of the above two estimators using series expansions. Results on the error in truncating such series expansions are given. An alternative approach for computing the variances of the estimators using Gauss-Hermite quadrature is also presented. We illustrate our results using data from the nutrition component of the 2015 Canadian Community Health Survey.

# Author Index


Abesiri Samarasekara, Damnada 6
Ahmed, Tasrian 43
Alimele, Omoroje Benjamin 44
Altenkirk, Tanner 13
B., Muhammad Umair A. 19
Barrington, Jonah 17
Campeanu, Cezar 9
Chatterjee, Ditthi 21
Chhavi, Chhavi 23
Clifton, Carter 37
Coleman, Noah 58
David Raj, Shekin 11
de Vera, Katrina 2
Denney-Martell, Gabe 35
Dhriti, Dhriti 52
DIOP, AIDA MARGOT 34
Elsabah, Mohamed 12
Etu-Efeotor, Rosemary 48
Finkle, Sarah 45
Gauthier, Melanie 40
Gauvin, Simon 4
Girdhar, Samridhi 8
Govindaraj, Tushita 31
Govindaraj, Tushita 42
Guo, Mingyue Gloria 16
Haque, Mohammod 7
Hashemi Tehrani, Daniel 57
Kasumovic, Lamija 50
Kathiresan, Lathika 46
Khan, Muhammad Shazil 59
Kochhar, Jase 56
Lawlor, William 26
Maaz, Muhammad 47
MacKinnon, Luke 53
Mastnak, Ana 18
Matheson, Alexander 10
Millett, Asher 29
Milley, Jesse 54
Mindia, Emmanuel 27
Mueller-Classen, Keely 55

| | |
|---|---|
| Muse, Abdiaziz | 5 |
| Omale, Ufedo-Ojo | 1 |
| Oussar, Asma | 32 |
| Pelley, Evan | 39 |
| Power, Timothy | 36 |
| Qiao, Jeremy | 22 |
| Richard, Gage | 28 |
| Roberts, Colin | 20 |
| Rose, Morgan Taylor | 60 |
| Saha, Peash | 51 |
| Sam, Habiebatou | 15 |
| Schepp, Rory | 49 |
| Simonov, David | 33 |
| Singh, Rupinder | 3 |
| Steeves, Cameron | 51 |
| Talbot, Patrick | 41 |
| Taylor, Madeleine | 30 |
| Tumuganti, Mithilesh | 50 |
| Varona, Luis | 24 |
| Wayne, Noah | 14 |
| Weerakul, Sarun | 38 |