# Department of Mathematics

May 26, 2025

## UNDERGRADUATE MATH SEMINAR – see you in the fall!

In deference to the welcoming reception for Union College's new incoming president, Elizabeth Kiss, which is being held Thursday, May 29 at Common Hour on the Reamer Back Patio, the same time as the math seminar is usually held, there will not be a seminar this week. And, since the term is winding down after that (or gearing up for finals), there will not be any more seminars this term. So, let's take a moment to thank this term's seminar organizers, **Professors Jeff Jauregui** and **Phanuel Mariano**.

## Two Pieces from Thesis: Audrey Benson and Janak Subedi

*Audrey wrote her thesis this fall and winter under the guidance of **Professor Phanuel Mariano.***

Working alongside my advisor **Professor Phanuel Mariano** this past winter, I submitted my senior thesis on the Lyapunov exponent for the product of random matrices related to an art collector model. For this project, we suppose that we are an art collector trying to expand our collection with two main goals. We want to optimize the monetary value of our collection, but also maintain an aesthetically pleasing one. We consider a model given by Rastegar, Roitershtein, Roytershteyn, and Seetharam (2024), which provides us with a way to represent the value of our collection as well as the amount of money available to put back into our collection, at time $n+1$, as a matrix. Focusing on an unbiased case of this matrix, our goal was to study the growth rate of the product of these random matrices. The exponential growth rate of the matrix product as time n approaches infinity is called the Lyapunov exponent.

It's very difficult, sometimes even impossible, to obtain an explicit expression for the Lyapunov exponent of random matrices. However, we were able to prove upper and lower bounds that converge to the Lyapunov exponent as $n$ approaches infinity. These bounds are given in terms of a multi-level recursion.

Following the submission of my thesis, Professor Mariano and I continued to work on this project. We were able to prove an explicit formula for the cumulative distribution function (CDF) of our random variable. Then, using the CDF, we were able to estimate upper and lower bounds for the Lyapunov exponent that are accurate up to 4 decimal places. As we finish up this project, we are compiling our results into a paper to submit for publication.

It has been an incredibly rewarding experience working with Professor Mariano. This is our second project together and I've learned so much from both experiences. Not only have I expanded my knowledge in probability and statistics, but I've also learned technical skills like coding in Mathematica and writing academic papers. I even developed skills that have nothing to do with math at all. I've learned to have confidence in my abilities, to be comfortable with not knowing everything, and the importance of trial and error. I think that's the best advice I could give to those about to begin their theses: don't be afraid to be wrong! I spent countless hours trying different ways to attack a problem only for none of them to work and be back at square one. You don't know what works until you try it, and if you fail, try again.

---

*Janak wrote his thesis this fall and winter under the guidance of **Professor Roger Hoerl***

In simple terms, machine learning methods are algorithms that can turn data into meaningful information. Thanks to the abundance of data, complex statistical methods, and advancements in computational power, machines can analyze large amounts of data in a relatively short amount of time. Leveraging statistical algorithms and mathematical optimization, machine learning algorithms can use data to understand patterns, make predictions or classifications, and provide useful insights through minimal instructions. However, these models are not truly intelligent; rather, they are a very effective and efficient tool for data processing and making informed decisions based on past information. Moreover, there are many types of machine learning models, each with its own strengths and limitations. Furthermore, real-world datasets can vary greatly in structure, size, and complexity. As a result, not all machine learning models work well with every kind of data.

(**continued on next page)**

So, which machine learning model works best for prediction and classification? This question is too broad to answer directly, as model performance often depends on the type of data and the specific task. To make the question more manageable, I focused my research on pricing datasets, where the goal is to predict a price or classify a payment outcome. These datasets typically have similar structures, with numerical and categorical features and a clear target variable related to value or risk. My thesis explores which machine learning models perform well on such pricing tasks, both in predicting continuous values and classifying default outcomes. Since most models can be improved by adjusting their parameters, I also investigated whether fine-tuning these models could lead to noticeably better performance.

For this study, I used three real-world datasets: a vehicle sales dataset from Kaggle, a real estate sales dataset from the Connecticut Office of Policy and Management, and a credit card default dataset from the UCI Machine Learning Repository. The first two are used for predicting the selling price of cars and properties, while the third is used for classification, to predict whether a customer will make their next credit card payment. To evaluate performance, I tested three popular machine learning models: Random Forest, XGBoost, and Deep Neural Networks (DNN).

Overall, I found that the two tree-based models, Random Forest and XGBoost, consistently performed well across the pricing datasets, especially when the data was structured and moderately sized. DNN, on the other hand, initially underperformed but showed significant improvement after tuning its parameters. Hence, hyperparameter tuning made a noticeable difference, particularly in datasets with high variability or imbalanced classes. While no single model was universally superior, each had strengths depending on the dataset and task. These results highlight that model selection should be guided by the characteristics of the data, and that tuning can lead to meaningful performance gains.

This thesis was a great learning experience from start to finish. The initial effort you put into reviewing existing literature and understanding your project greatly helps in laying a strong foundation. Furthermore, making an itinerary or timeline early on and trying to write sections of the thesis as you go helps in breaking the work into manageable steps. I am really grateful to my advisor, Professor Roger Hoerl, for his guidance and support. His feedback was invaluable whenever I ran into a roadblock.

## Fall Term Math Jobs Available: Calculus Help Center Tutor; Math 105 Coach

The Math Department is now accepting applications for vacant **Calculus Help Center (CHC) tutoring positions**. Tutors in the fall work in the CHC one fixed night per week, Sunday through Thursday, from 7:30-10:00pm.

Qualifications: Calculus through Math 115 with grades of no less than A-. Preference will be given to students who

- have also completed Math 117 (with a grade ≥ A-),
- are declared math majors,
- are considering becoming a math teacher or pursuing graduate work in mathematics, and
- have other tutoring experience.

To apply, send an email to **Professor Paul Friedman** (friedmap@union.edu) expressing your interest, listing your math background, *including coursework (term, professor, and grade)* and tutoring experience (if any), and discussing why you think you would be a good tutor.

**Deadline: Tuesday, June 3 at NOON**

**Math Coaches** attend and work with a section of MTH 105 in the fall to assist students with their understanding of course content. They assist the professor during class, circulating and working with students when the students are working on problems and worksheets. In addition, Math Coaches provide math tutoring sessions for the Math 105 students.

Apply

For more information, contact **Lesly Clay** (clayl@union.edu), Director of the Office of Student Success, or **Professor Paul Friedman** (friedmap@union.edu)

To apply, follow this link or the QR code shown.

**Deadline: Monday, June 2 at 9:00 am**