



27TH ACADEMIC SYMPOSIUM

WEDNESDAY, APRIL 29, 2026

THE DUNCAN FAMILY SKY ROOM

MAC MAHON STUDENT CENTER

Office of the Vice President for Academic Affairs

Welcome to the Annual Academic Symposium - a celebration of student accomplishments. The Symposium represents a longstanding tradition that highlights the research and scholarly achievements of Saint Peter's students who have worked in collaboration with dedicated faculty members throughout the year. It began as "Poster Day" in April 1998 with Laura Twersky, Ph.D. as the advisor and Amira Roess as the first President of TriBeta, in consultation with Virginia "Ginny" Bender, Ph.D. '78.

These collaborations represent a model for integrating teaching with scholarship outside the classroom, to promote the intellectual development and professional growth of our students. The Symposium also provides an opportunity for the campus community to gain a greater appreciation of the collective accomplishments of our faculty and students, as well as to celebrate their achievements.

I hope that you enjoy and benefit from viewing the posters, hearing the presentations and speaking with the authors of these impressive works. Congratulations to all the presenters!

Sincerely,

WeiDong Zhu, Ph.D.

Interim Vice President for Academic Affairs

OVERVIEW

UNDERGRADUATE
STUDENT SYMPOSIUM

Wednesday, April 29, 2026
11 a.m. - 1 p.m.

Extinct Potential: Investigating Pediatric Organ Discard Rates and Underutilization in America

Ajani Peck-Wekunda

Advisor: Brandy Garrett Kluthe, Ph.D.

Food Desert Strategies For Growing Medicinal Plants in Hydroponics Vs. Traditional Soil Environments

Martina Awad

Advisor: Brandy Garrett Kluthe, Ph.D.

Synthesis of an Abundant Metal Catalyst Based on a Cobalt-Bipyridine Complex for Sustainable Energy

Allyssia Merisier

Advisor: Yosra M. Badiei, Ph.D.

AI -Assisted Analysis of BPC-157 Dose-Dependent Effects on Fibroblast Wound Closure *in vitro*

James Ellis

Advisor: Karen Wydner, Ph.D., Leonard Sciorra, Ph.D., & Laura Twersky, Ph.D.

Polymer-Catalyst Assemblies for Electrochemical Oxidation of Organosulfur Compounds

Yessica Ramirez, Arham Muhammad

Advisor: Yosra Badiei, Ph.D.

Questioning the “Belongingness Gap” at a Small, Diverse Liberal Arts University

Emily Perez

Advisor: Brittany Hanson Ph.D., & Daniel Wisneski, Ph.D.

The Effects of Aging on Cardiovascular Adaptation to Athletic Training

Keyon Tims

Advisor: Joshua Williams, DHSc

Facial Resemblance and Its Effect on Helping Behavior

Valeria Ramirez

Advisor: Joshua Feinberg, Ph.D.

Periodization Models for Sprint Performance in Collegiate Track and Field Athletes

Karilynn Jimenez

Advisor: Joshua Williams, DHSc

The Possible Preventative Effects of Thymoquinone and Kaempferol, Individually and in Combination, on Bisphenol A-Induced Impacts in *Xenopus laevis* (African Clawed Frog)

Ruqya Elouarraq

Advisor: Laura Twersky, Ph.D.

The Role of Aquatic Rehabilitation In Improving Flexibility and Reducing Joint Stiffness After Musculoskeletal Injury

Yana Piletskaya

Advisor: Joshua Williams, DHSc

The Effects of Detraining in Sprint Athletes Vs Endurance Athletes

Madison McKinney

Advisors: Joshua Williams, DHSc

Muscle Development Interventions in Paralysis Rehabilitation: Approaches for Paraplegic and Disabled Populations

Jacob Turingan

Advisor: Joshua Williams, DHSc

Factors Influencing Testosterone Levels: Effects of Exercise, Supplementation, and Performance-Enhancing Drugs

David Canales

Advisor: Joshua Williams, DHSc

Physiological Responses to Extreme Temperatures and Their Impact on Endurance Performance

Audrey Franco

Advisor: Joshua Williams, DHSc

Why Some People Build Muscle Faster Than Others: Genetic Influences on Muscle Mass and Bone Density of Microplastics/Bisphenols on Blood

Ciara Brown

Advisor: Joshua Williams, DHSc

Central Nervous System Regulation of Running: Neural Signaling, Motor Control, and Fatigue

Luisana Flores

Advisor: Joshua Williams, DHSc

The Effectiveness of High-Intensity Exercise in Cardiac Rehabilitation

Kayla McKinnon

Advisor: Joshua Williams, DHSc

Comparing Intrinsic & Extrinsic Motivation in Resistance Training & Goal Achievement

Lamine Diop

Advisor: Joshua Williams, DHSc

How Social Media Data Shapes Public Opinion

Emanuel Almanzar, Aidan Forrest

Advisor: Glenda Guerrero, Ed.D.

Do the Benefits of Modern Major Technological Advancement Outweigh the Negative Effects on America's Small Cities?

Lyna Baal, Olwethu Mona

Advisor: Glenda Guerrero, Ed.D.

How Data Shapes Business Decisions in the Age of AI

Justin Rodriguez, Godtime Onokpasa
Advisors: Glenda Guerrero, Ed.D.

Data Spill: When Information Goes Where It Shouldn't

Victoria Johnson, Valeria Rodriguez
Advisor: Glenda Guerrero, Ed.D.

Predicting Protein Drug Complexes for Crohn's Disease-Associated Targets

Emily Hernandez
Advisor: Hugo Guterres, Ph.D.

Do moral and pragmatic judgments differ in their physiological correlates?

Keily Barrera, Marcus Bell, Jaylen Hargrove
Advisor: Daniel Wisneski, Ph.D. & Brittany Hanson Ph.D.

The Effects of High- vs. Low-Intensity Exercise and Its Effects on Brain Chemistry and Focus in Individuals with ADD/ADHD

Mia Dominguez
Advisor: Joshua Williams, DHSc

Predicting the "protein-drug" complex structure for Crohn's Disease-associated targets

Misheel Myagmarjav
Advisor: Hugo Guterres, Ph.D.

The Bizarreness Effect is Inconsistent in Free Recall

Maria Medrano
Advisor: Maryellen Hamilton, Ph.D.

Resistance and Climbing-Specific Training for Injury Prevention in Youth Climbers

Noah Santos
Advisor: Joshua Williams, DHSc

Putative Protective Effects of Ginkgolide B and Folic Acid Individually, and in Combination, Against Polystyrene-Induced Developmental Dysregulation in *Xenopus laevis* (African Clawed Frog)

Darwin Breton

Advisor: Laura Twersky, Ph.D.

The Great Backyard Bird Count 2026: Joining Together to Count Winter Birds

Alexandra Ramirez, Suleimy Gomez Colon, Phoenix Campbell

Advisor: Katherine Wydner, Ph.D.

The Twelfth Season of FeederWatch: Fostering a Diverse Avian Community Amidst a Historic Winter!

Darwin Breton, Alexandra Ramirez, Phoenix Campbell, Genesis Serpas, Kaidy Abreu, and Suleimy Gomez Colon

Advisor: Katherine Wydner, Ph.D.

Predicting Comets' Trajectories: Integrating Data Science with Physics

Agha Sibte E Ali Qazalbash

Advisor: Gulhan Bizel, Ph.D.

Interest Rates and Government Bond Yields across Countries: Comparative and Predictive Analysis

Hadeed Khan

Advisor: Gulhan Bizel, Ph.D.

"The Sonic Evolution of Hip-Hop: A Data-Driven Analysis of Audio Features Over Time"

David Dandridge

Advisor: Gulhan Bizel, Ph.D.

Vitamin D Deficiency in Athletes: A Literature Review of Performance and Health Consequences

Justin Stotz

Advisor: Joshua Williams, DHSc

Visualizing the Impact of COVID-19

Alberth Rojas, Dev Patel & Sraboni Haque
Advisor: Glenda Guerrero, Ed.D.

Cold Case Map

Ashley Ramcharran
Advisor: Gulhan Bizel, Ph.D.

Credit Card Fraud Detection Using Data Mining Techniques

Sneha Gurampally Malkan, Roystin Ronaldo Felix
Advisor: Vijay Kumar Reddy Voddi, MSDS

Network Intrusion Detection Using Machine Learning

Gurpreet Kaur
Advisors: Vijay Kumar Reddy Voddi, MSDS

Detection of Bank Account Fraud: Securing Digital Onboarding through Machine Learning and Resampling Techniques

Mohammed Aroui
Advisors: Vijay Kumar Reddy Voddi, MSDS

Global Warming Patterns: A Data Mining Study

Biyue Wang, Chander Rekha Sehgal
Advisors: Vijay Kumar Reddy Voddi, MSDS

Airline Delay Prediction using Data Mining

Tejaswini Konakanchi, Vinesh Nangunoori
Advisors: Vijay Kumar Reddy Voddi, MSDS

Doctor's Handwritten Medical Prescription Recognition Using Deep Learning

Ganesh Chilukuri
Advisors: Vijay Kumar Reddy Voddi, MSDS

Heart Disease Risk Prediction using Machine Learning

Irfan shaik, Abhi Sekhar Rao Talluri
Advisors: Vijay Kumar Reddy Voddi, MSDS

Detecting Depression Signals from Social Media Using Data Mining Techniques

Nikhila Boppana
Advisors: Vijay Kumar Reddy Voddi, MSDS

Cardiovascular Integrated Risk Analytics System

Manish Lakshmipuram, Md Alamgir Shahidullah
Advisor: Vijay Kumar Reddy Voddi, MSDS

S&P 500 Stock Prediction & Clustering

Praveen Kumar Inti
Advisor: Vijay Kumar Reddy Voddi, MSDS

Hospital Inpatient Length of Stay

Bethlehem Seid, Srijana Lama
Advisor: Vijay Kumar Reddy Voddi, MSDS

Clinical Pattern Discovery & Mortality Risk Modeling

Deepali Poudel
Advisors: Vijay Kumar Reddy Voddi, MSDS

Predicting and Interpreting Residential Rent Patterns : A Modeling and Interactive Decision-Support Approach

Yuqiu Gan, Yujie Liu
Advisors: Vijay Kumar Reddy Voddi, MSDS

AI-Driven Predictive Modeling and Time Series Forecasting of Employee Burnout Using Behavioral, Productivity, and Lifestyle Analytics

Rishabh Perewar, Dakshayani Uppu
Advisors: Vijay Kumar Reddy Voddi, MSDS

Delhi AQI Prediction

Sagar Maru, Pankaj Sharma, Garima Chaudhary
Advisors: Vijay Kumar Reddy Voddi, MSDS

GateKeep: An Agentic AI Platform for Personal Financial Intelligence

Fnu Maliha, Sree Udhai Raj
Advisors: Gulhan Bizel, Ph.D.

Recidivism Prediction & Algorithmic Fairness

Selin Ozguler, Hasan Mert Sungur, Umut Baris Basol
Advisors: Vijay Kumar Reddy Voddi, MSDS

Modeling Fatal Outcomes of Motor Vehicle Collisions Using Random Forest

Sonali Gupta
Advisors: Vijay Kumar Reddy Voddi, MSDS

Credit Card Fraud Detection Using Data Mining

Afreen Tabasum begum Fnu
Advisors: Vijay Kumar Reddy Voddi, MSDS

The Global Heartbeat: A Computational Analysis of Rhythmic Universals and Cross-Cultural Neural Resonances

Gaurab Shrestha, Bishal Gupta
Advisors: Vijay Kumar Reddy Voddi, MSDS

Predicting Colorectal Cancer Survival with Machine Learning

Michael Uhler, Suman Lamsal
Advisors: Gulhan Bizel, Ph.D.

Language Development & Evolution in Children

Angelica Amegashie
Advisors: Dong Ryeol Lee, Ph.D.

AI-Driven Market Intelligence for Identifying High-Growth Startup Opportunities: Insights from the Pet Nutrition Industry

Laura Becerra, Dari Dambaeva, Jaykumar Panchal
Advisors: Gulhan Bizel, Ph.D.

Smart Closet

Nihaar Sirikonda

Advisors: Gulhan Bizel, Ph.D.

Clinical Intelligence System

Fnu Maliha, Sree Udhai Raj

Advisors: Gulhan Bizel, Ph.D.

M.A.R.K AI: Agentic AI for Job Hunt

Arsalan Hassan, Samir Sengupta, Afreen Tabasum Begum Nln

Advisors: Gulhan Bizel, Ph.D.

Credit Card Fraud Detection

Haritha Renuguntla, Jahnavi Somu

Advisors: Vijay Kumar Reddy Voddi, MSDS

Credit Card Fraud Detection using Data Mining Techniques with Societal Impact Analysis

Venkata Kasi Praveen Yarru, James Davis
Advisors: Vijay Kumar Reddy Voddi, MSDS

Predicting Breast Cancer Diagnosis Using Machine Learning: A Data Driven Approach Using the Wisconsin Diagnostic Dataset

Joseph Liggio, Udumaga Onyeukwu
Advisors: Vijay Kumar Reddy Voddi, MSDS

AEGIS: A Three-Layer Security Proxy for Healthcare AI Agents Using Confidence-Gated Classification

Andrew Espira, Umut Baris
Advisors: Gulhan Bizel, Ph.D.

Using Large Language Models to Predict Analyst Questions from Earnings Call Content

Natasha Kapoor
Advisors: Dong Ryeol Lee, Ph.D.

Investigating Microplastic Interactions with Biological Molecules

Andres Zuniga
Advisors: Hugo Guterres, Ph.D.

Measuring ESG Commitment and Credibility of Tech Companies Through NLP

Moustapha Gebai
Advisors: Dong Ryeol Lee, Ph.D.

Comparing Zero-Shot Prompting, Few-Shot Prompting, and Traditional ML Approaches for Email Intent Classification

Chen Zhou
Advisors: Dong Ryeol Lee, Ph.D.

"Lose Yourself" in the Data

Nnamdi Okoli
Advisors: Dong Ryeol Lee, Ph.D.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Extinct Potential: Investigating Pediatric Organ Discard Rates and Underutilization in America

Ajani Peck-Wekunda

Advisor: Brandy Garrett Kluthe, Ph.D.

Pediatric organ donation requires careful management and recovery methods to maximize the placement and utilization of donated organs. This study reviews discard rates and contributing factors among pediatric donors. A systematic review of 14 peer-reviewed articles published between 2000 and 2025 was conducted, focusing on discard rates for pediatric hearts, livers, kidneys, and pancreata.

Heart discard rates ranged from 29.3% to 56.6% in infant donors. Contributing factors included donor-recipient size mismatch, suboptimal graft function, and assessments of donor quality. Liver discard rates were highly variable, ranging from 7.4% to 94.5% in donors under six years old, with the highest rates observed among Donation after Circulatory Death (DCD) donors and Public Health Service (PHS) increased risk donors. Discard was often due to organ quality or size mismatch. Kidney discard rates ranged from 15.9% to 89.7%, especially in donors weighing less than 10 kg.

Common challenges included vascular anatomy, small donor size, and limited procurement experience. Pancreas discard rates reached up to 69%, particularly in infant donors, largely due to concerns about low islet mass.

Despite high rates, many unused organs had normal radiologic and laboratory findings. PHS increased risk organs yield outcomes comparable to standard donors. Pediatric donors remain an underutilized resource, and further research is warranted.

Food Desert Strategies For Growing Medicinal Plants in Hydroponics Vs. Traditional Soil Environments

Martina Awad

Advisor: Brandy Garrett Kluthe, Ph.D.

In urban environments it is sometimes difficult to find fresh medicinal and culturally relevant plants. Developing a growing method that maximizes the ability of an individual to grow their own plants is important. This study looked at two different types of medicinal herbs: *Rosmarinus officinalis* commonly known as rosemary seeds and *Hibiscus sabdariffa* commonly known as roselle. Rosemary seeds have antioxidant and anti-inflammatory properties, which helps protect our cells from damage or harm.

Roselle's a medicinal herb that has many health benefits, such as lowering blood pressure and cholesterol, and weight management. The growth of roselle and rosemary seeds was monitored in soil and hydroponics systems. A total of 96 seeds were used, 48 replicates of each seed type. Over 7 weeks roselle and rosemary seeds were measured every Monday from the surface of the growth medium. The results indicated that soil for both plants had a lot more root production, However, the hydroponics had a better germination and better yield overall.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Synthesis of an Abundant Metal Catalyst Based on a Cobalt-Bipyridine Complex for Sustainable Energy

Allyssia Merisier

Advisor: Yosra M. Badiei, Ph.D.

The quest for sustainable, carbon-neutral fuels, particularly hydrogen, necessitates the development of efficient water-splitting catalysts. While precious metal complexes, such as the Ru-based polypyridine "Blue dimer," are typically used for water oxidation, this study focuses on more affordable and abundant metal catalysts. Specifically, we have synthesized and will comparatively analyze the catalytic activity of two cobalt-based compounds in water splitting: a mononuclear cobalt(II) bipyridine aqua ligated complex, $[\text{Co}(\text{bpy})_2(\text{H}_2\text{O})]^{2+}$ with 2,2'-bipyridine and a dinuclear $[\text{Co}(\text{bpy})_2(\text{en})(\text{H}_2\text{O})_2]^{2+}$. The electrocatalytic activity of these complexes will be evaluated by constructing an electrochemical cell for anodic water oxidation and performing cyclic voltammetry studies. This work advances sustainable fuel development by examining viable cobalt-based alternatives.

AI-Assisted Analysis of BPC-157 Dose-Dependent Effects on Fibroblast Wound Closure *in vitro*

James Ellis

Advisor: Karen Wydner, Ph.D., Leonard Sciorra, Ph.D., & Laura Twersky, Ph.D

Fibroblasts are the primary connective tissue cells that are responsible for wound healing. Enhancing fibroblast activity is one way improved outcomes in wound healing research may be achieved. BPC-157, a synthetic gastric pentadecapeptide, has quickly gained attention for its ability to accelerate regenerative properties seen in sports medicine. Although preclinical studies show promising results in tendon healing, there is little *in vitro* evidence examining its effect on dermal fibroblasts. BPC-157 has been proposed to influence cell migration through pathways that involve nitric oxide signaling, cytoskeletal organization, and focal adhesion dynamics. This study investigates whether BPC-157 accelerates dermal fibroblast wound closure in a dose-dependent manner using a scratch assay. Vehicle controls and four BPC-157 concentrations (0.01 μM , 0.1 μM , 1 μM , and 10 μM) are used in fibroblast basal media, which contains fibroblast growth factor (FGF2). Wound closure is measured using an AI-assisted image analysis pipeline that detects wound edges and quantifies percent wound closure across time points with greater precision than manual measurement. This study represents one of the first applications of artificial intelligence to quantify scratch assays while investigating if BPC-157 positively affects wound closure in dermal fibroblasts. This work may provide insight for sports medicine applications. Presented at the ICUNJ Undergraduate Research Symposium, Bell Works, March 23, 2026

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Polymer-Catalyst Assemblies for Electrochemical Oxidation of Organosulfur Compounds

Yessica Ramirez, Arham Muhammad
Advisor: Yosra Badiei, Ph.D.

Sustainable processes to eliminate sulfur-containing compounds from hydrocarbon fuels (i.e., desulfurization) are needed to prevent sulfur dioxide emissions, which can contribute to atmospheric pollution, acid rain, and adverse health effects. Electrochemical oxidation can be a green alternative to chemical oxidant reagents (e.g., hydrogen peroxide), in which water molecules can serve as a source of oxygen atoms under an applied oxidative potential. For this purpose, cationic Ru-based polypyridine complexes were immobilized onto conductive polymer films (poly(acrylic acid)) and Nafion on glassy carbon (GC) electrodes. Ru-polymer catalyst molecular assemblies will be tested for electrocatalytic S-oxidation of benzothiophenes, compared with the background reactivity of the electrodes, using cyclic voltammetry and bulk electrolysis. Factors that drive catalysis will be better understood by studying the inner-sphere coordination environment within these Ru-polymer assemblies through careful measurements of the electrochemical potentials of the immobilized complexes and spectroscopic studies.

Questioning the “Belongingness Gap” at a Small, Diverse Liberal Arts University

Emily Perez
Advisor: Brittany Hanson Ph.D., & Daniel Wisneski, Ph.D.

The current research finds that a tailored STEM belongingness intervention increased feelings of being prepared for college but did not impact GPA or retention, both at the university and within a STEM major, during the first year of college for low-income and Hispanic students. A follow-up study tests whether low-income and Hispanic students actually differ in baseline belongingness compared to non-low-income and non-Hispanic students in this specific university population.

The Effects of Aging on Cardiovascular Adaptation to Athletic Training

Keyon Tims
Advisor: Joshua Williams, DHS

This literature review focuses on the effects of aging on cardiovascular adaptation to athletic training, with an emphasis on the underlying physiological mechanisms and their impact on performance. Regular exercise significantly improves cardiac output, capillary density, stroke volume, and overall aerobic efficiency in younger athletes. In contrast, aging is associated with both structural and functional changes, such as decreased ventricular compliance, increased arterial stiffness, and lower maximal heart rate, which all impact cardiovascular performance. Age-related changes in key hormones, such as growth hormone, estrogen, and testosterone, can impair endurance and metabolic performance.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

The Effects of Aging on Cardiovascular Adaptation to Athletic Training (cont.)

Despite these changes, cardiovascular health in older individuals can still improve with appropriate exercise. While younger individuals tend to maintain adaptations more effectively, exercises such as moderate-intensity aerobic training, resistance training, and high-intensity interval training (HIIT) have been shown to improve oxygen utilization, vascular function, and muscle strength across age groups. Additionally, to optimize results, exercise frequency, intensity, and recovery are crucial for preventing overtraining. Overall, consistent physical activity is essential for maintaining functional capacity, supporting cardiovascular health, and promoting longevity and quality of life.

Facial Resemblance and Its Effect on Helping Behavior

Valeria Ramirez

Advisor: Joshua Feinberg, Ph.D.

This study was designed to investigate whether an individual's judgment in prosocial situations varies depending on whether the stranger in need of help resembles a person in the individual's life. The decision-making process of engaging in helping behavior can be instant and influenced by our emotions, empathy, closeness, and familiarity. We hypothesize that participants will be more likely to engage in helping behavior towards strangers who share the facial resemblance to someone they already know, compared to those who do not. In this study, there were three images of the same individual varying in levels of familiarity.

Participants were provided with one of these images and told to imagine them as the one in need of help for the following three prosocial scenarios. Each scenario varied in the degree of required sacrifice. They were then asked questions to measure elements such as their willingness and self-concept to determine whether facial resemblance can strengthen perceived connection, thereby increasing altruism.

Periodization Models for Sprint Performance in Collegiate Track and Field Athletes

Karilynn Jimenez

Advisor: Joshua Williams, DHS

This literature review examines how different structures of strength and power training impact sprint performance in collegiate sprinters. To improve sprinting performance, resistance training must develop high levels of strength, power, and speed while accommodating for fatigue management and injury prevention. Periodization is the predominant technique used to organize training by progressively varying intensity, volume, and recovery time periods. The purpose of this literature review is to explore how each of the three types of periodization – linear, nonlinear (undulating), and block – affects the sprinting performance of collegiate athletes. Each of the three types of periodization will be analyzed based on its impact on key performance variables – acceleration, maximum velocity, force output or power, and rate of force development – as well as its role in fatigue management and injury risk throughout a competitive season. While all three types of periodization can positively affect performance, research indicates that block and nonlinear periodized programs may produce greater benefits in trained athletes

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Periodization Models for Sprint Performance in Collegiate Track and Field Athletes (cont.)

due to their emphasis on specific attributes and their greater variability in training. Overall, no single periodization model is optimal for every athlete, and effective sprint training depends on individual needs, training experience, and the balance between performance development and adequate recovery.

The Possible Preventative Effects of Thymoquinone and Kaempferol, Individually and in Combination, on Bisphenol A-Induced Impacts in *Xenopus laevis* (African Clawed Frog)

Ruqya Elouarraq

Advisor: Laura Twersky, Ph.D.

Bisphenol A (BPA) is an endocrine disruptor that exerts anti-estrogen, anti-thyroid, and DNA-disrupting effects in human and animal models. While eliminating bisphenol A is impossible, holistic approaches can be used to avoid or minimize BPA-induced impacts. This study seeks to examine if thymoquinone, the main ingredient in the herb *Nigella sativa*, and kaempferol, a natural flavonoid, both used traditionally in ancient Middle Eastern medicine, can mitigate teratogenic effects of BPA on *Xenopus laevis* (clawed frog) development. *Xenopus* tailbuds at Nieuwkoop and Faber stages 40-41 are incubated in 1µg/mL solutions of thymoquinone, kaempferol, BPA, ethanol, thymoquinone and BPA, kaempferol and BPA, and thymoquinone, kaempferol, and BPA combined. As the *Xenopus* tailbuds develop, a stereomicroscope and iPhone 14 camera are used to record developmental rates and any abnormal morphology. Heart rate and motility rates in specimens are also observed visually, recorded, and averaged per group.

Preliminary results revealed significant cardiovascular and morphological malformations in tadpoles exposed to BPA alone, with improved developmental rates, heart rate, morphology, and motility when thymoquinone and kaempferol were introduced individually and in combination. Analysis of these results suggest that thymoquinone and kaempferol may potentially protect against structural malformations caused upon exposure to bisphenol A.

The Role of Aquatic Rehabilitation In Improving Flexibility and Reducing Joint Stiffness After Musculoskeletal Injury

Yana Piletskaya

Advisor: Joshua Williams, DHS

During our lifetime, there is a chance that we may experience at least one orthopedic trauma, not to mention those whose occupations or active lifestyles carry a higher risk of injury. Even though medicine and therapy have advanced greatly over time, there are still many cases in which rehabilitation outcomes could be more favorable than they actually are. Nowadays, healthcare offers us a wide range of rehabilitation techniques and methods that have indisputably proven their worth over time; however, one is often overlooked. Aquatic rehabilitation can be a great alternative to traditional land-based therapy, and its combined efforts can lead to improved recovery outcomes for musculoskeletal injuries. Moreover, water-based therapy is profoundly valued for its universal water properties, which make this type of recovery comfortable and enable the recovery process to start as early as possible, thereby reducing the harmful effects of the post-mobilization period.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

The Role of Aquatic Rehabilitation In Improving Flexibility and Reducing Joint Stiffness After Musculoskeletal Injury (cont.)

Musculoskeletal traumas are characterized by significant loss of flexibility and possible joint stiffness, which, if not treated promptly, can develop from prolonged rehabilitation challenges to long-term well-being issues. Hence, this literature review focuses on the importance of aquatic rehabilitation and compares it to its counterpart, with emphasis on how effective the former can be in addressing the problem of post-trauma flexibility and stiffness.

The Effects of Detraining in Sprint Athletes Vs Endurance Athletes

Madison McKinney

Advisors: Joshua Williams, DHS

This literature review explores the effects of detraining in sprint and endurance athletes. Detraining is an important topic in sport due to the frequency of injuries, illnesses, and breaks in play. The purpose of this literature review is to explore the aspects of detraining and the physiological changes each type of athlete experiences during these periods. In this review, some studies focus on a variety of athletes, including runners, soccer players, cyclists, and swimmers. Key outcomes in this review include cardiovascular function, metabolic adaptations, muscular strength, power output, and VO_2 max. The findings indicate that detraining leads to declines in multiple physiological systems. Endurance athletes experience declines in aerobic capacity and cardiovascular efficiency; on the other hand, sprint athletes maintain their strength and power, but experience declines in neuromuscular coordination and explosiveness. Many themes were found in this research, and they included training history and the duration of detraining.

Overall, both groups see a performance drop off with detraining, but the endurance group experiences changes more rapidly than the sprinters, which emphasizes the importance of maintaining training continuity in athletic populations.

Muscle Development Interventions in Paralysis Rehabilitation: Approaches for Paraplegic and Disabled Populations

Jacob Turingan

Advisor: Joshua Williams, DHS

Paralysis is the partial or complete loss of muscle function, movement, or sensation caused by disrupted nerve signals between the brain and the muscles, usually due to an underlying condition such as a spinal cord injury or stroke. Muscle growth and development rehabilitation aims to re-stimulate nerve sensations and prevent muscle atrophy. This literature review examines different types of interventions used to treat paralysis and their effectiveness in enhancing muscle development in order to improve quality of life. Evidence was gathered from PubMed and Google Scholar, focusing on studies that involved muscle development and growth in paralysis, which resulted in different interventions being used. Key variables included type of injury, type of treatment used, length of treatment, and data collected post-treatment. Across the interventions, overall results were significantly positive, indicating increased muscle development and neural re-stimulation through rehabilitation. Other benefits include an increase in mobility, stability and coordination, and a better quality of life. This review covers the differentiation between interventions and demonstrates that no single approach was consistently shown to be superior.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Factors Influencing Testosterone Levels: Effects of Exercise, Supplementation, and Performance-Enhancing Drugs

David Canales

Advisor: Joshua Williams, DHSc

Testosterone is a significant hormone and plays a major role in the growth, development, and repair of muscles in the human body. Because of the major role testosterone plays in the human body, various methods have been used to enhance testosterone production, including resistance training, endurance training, nutritional supplementation, and pharmaceutical interventions. This literature review evaluates current research on how these approaches influence testosterone levels and related physiological outcomes. Resistance training shows a significant increase in testosterone temporarily, whereas endurance training may initially increase testosterone but has been associated with long-term declines in production. Nutritional supplements such as vitamin D and zinc have been shown to support testosterone; however, many over-the-counter supplements have little to no effect on testosterone production and may function as placebos. Anabolic-androgenic steroids and testosterone replacement therapy significantly elevate testosterone levels but are associated with adverse effects, including suppression of endogenous hormone production and reduced fertility.

Physiological Responses to Extreme Temperatures and Their Impact on Endurance Performance

Audrey Franco

Advisor: Joshua Williams, DHSc

This literature review examines how weather conditions impact endurance athlete performance. When participating in an outdoor endurance sport, athletes cannot control how the weather will be. Weather conditions could be high heat, extreme cold, humidity, rain, and wind. These conditions influence how athletes train and, more importantly, how they perform. This literature review explores how extreme weather conditions affect athletes physiologically and whether other parts of the body are negatively impacted. Weather conditions can range from extreme heat over 90 degrees to extreme cold below 32 degrees. The main factor that stands out tends to be humidity. Humidity mixes in with extreme heat, but the sweat cannot evaporate from the skin, leading to the body not being able to cool down. When combining endurance athletics with the weather conditions, it can lead to a negative effect on the athlete, such as dehydration, decreased performance, inefficient muscular function, and challenges with blood flow. The purpose of this literature review is to increase awareness among endurance athletes and coaches of how weather conditions can affect overall performance both internally and externally.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Why Some People Build Muscle Faster Than Others: Genetic Influences on Muscle Mass and Bone Density of Microplastics/Bisphenols on Blood

Ciara Brown

Advisor: Joshua Williams, DHSc

Every individual has a different genetic structure, which plays a crucial role in why muscle growth and bone density vary among people. Some individuals are able to gain muscle mass more rapidly than others, and much of this can be explained by genetic differences. However, genetics is not the only factor that can influence muscle development and bone health. This literature review explores how genetic factors, along with other influences such as ethnic background, family history of osteoporosis, childhood development, and nutritional intake, affect muscle growth and bone density. Research suggests that genetics plays a significant role in determining both muscle mass and bone mineral density, while also influencing the risk of conditions such as osteoporosis. In addition, factors such as early-life development and adequate intake of nutrients like protein, vitamin D, and calcium can further impact skeletomuscular health. Overall, this review highlights how both genetic and environmental factors contribute to differences in muscle growth and bone density among individuals.

Central Nervous System Regulation of Running: Neural Signaling, Motor Control, and Fatigue

Luisana Flores

Advisor: Joshua Williams, DHSc

This literature review examines how the Central Nervous System (CNS) regulates running performance through neural control, signaling, and fatigue. While running is popularly known as an activity associated with cardiovascular health and muscular endurance, studies have found that the CNS works as the main regulator of coordination and movement. This review aims to analyze how the brain, the spinal cord, and the muscles of the human body work as one so that, together, they may achieve coordination while maintaining rhythm and sending real-time sensory feedback. Research shows that fatigue is not only a result of physical exhaustion but is also influenced by the brain's internal mechanisms. Central fatigue occurs when the brain reduces its efficiency and decreases muscular performance even before the muscles have reached physical exhaustion, acting as a protective mechanism against physiological stress. Neural signaling from the motor cortex and spinal cord plays a massive role in organizing movement and improving muscular efficiency, with oxygen availability further influencing the CNS's ability to sustain muscle activation. These findings confirm that running is a highly coordinated activity driven by continuous interaction between the brain, spinal cord, and muscles, offering insight into how training strategies can improve endurance and delay fatigue.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

The Effectiveness of High-Intensity Exercise in Cardiac Rehabilitation

Kayla McKinnon

Advisor: Joshua Williams, DHS

This literature review examines the effectiveness of high-intensity interval training (HIIT) in cardiac rehabilitation programs. Cardiovascular disease remains a leading cause of mortality worldwide, highlighting the need for effective rehabilitation strategies following cardiac events. Traditional programs have relied on moderate-intensity continuous training (MICT) due to it being safe and accessible. However, emerging research suggests that HIIT may provide greater improvements in cardiovascular function and aerobic capacity. This review will analyze the structure and physiological effects of HIIT, the impact on maximal oxygen uptake, cardiac efficiency, and metabolic health. It will also compare high-intensity interval training to moderate-intensity continuous training, with most evidence indicating that HIIT may be able to produce similar or superior outcomes in less time. Additionally, safety considerations such as supervision and individualized programming are discussed. Overall, the findings suggest that HIIT is a time-efficient and effective alternative to traditional rehabilitation methods when appropriately implemented. Although it is promising, further research is needed to establish standardized guidelines and determine its suitability across patient populations.

Comparing Intrinsic & Extrinsic Motivation in Resistance Training & Goal Achievement

Lamine Diop

Advisor: Joshua Williams, DHS

This literature review examines different types of motivation and their effects on resistance training and goal achievement.

Motivation plays a significant role in maintaining exercise consistency and achieving long-term fitness goals. Intrinsic motivation refers to internal drivers such as enjoyment, satisfaction, and personal interest, while extrinsic motivation refers to external rewards such as weight loss, changes in appearance, or social factors. This review explores how these types of motivation influence adherence and performance in resistance training. Evidence from multiple studies, including mixed-methods research, suggests that individuals with higher intrinsic motivation demonstrate greater progress, stronger adherence to training programs, and better ability to handle setbacks. In contrast, individuals driven primarily by extrinsic motivation may show high initial engagement but are more likely to discontinue training when external rewards diminish. Overall, the findings indicate that intrinsic motivation may be more effective for long-term goal achievement, while extrinsic motivation may play a role in short-term engagement.

How Social Media Data Shapes Public Opinion

Aidan Forrest, Emanuel Almanzar

Advisor: Glenda Guerrero, Ed.D.

This project focuses on how data and artificial intelligence influence public opinion through social media platform. We plan to use a social media dataset, such as Twitter/X sentiment data or another public dataset that tracks opinions and trends. This will allow us to look at how people react to certain topics over time. Our goal is to show how public opinion changes during major events and how data can reveal patterns in peoples reactions online.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Do the Benefits of Modern Major Technological Advancement Outweigh the Negative Effects on America's Small Cities?

Olwethu Mona, Lyna Baal
Advisor: Glenda Guerrero, Ed.D.

The role of artificial intelligence plays in society, with a specific focus on how the rapid expansion of AI data centers affects small communities and neighborhoods. While AI is often seen as a major technological advancement, it simultaneously has detrimental consequences, especially on small communities, such as environmental strain, economic shifts, infrastructure pressure, and growing social inequality. We will weigh out the positive benefits of AI development and the negative effects, with a focus on small and/or communities living below the poverty line.

How Data Shapes Business Decisions in the Age of AI

Godtime Onokpasa, Justin Rodriguez
Advisors: Glenda Guerrero, Ed.D.

This project explores how artificial intelligence (AI) is transforming the way decisions are made in today's world. From business strategies to everyday consumer choices, AI uses data to predict behavior and improve outcomes. It also examines important ethical issues such as data privacy, algorithmic bias, and the broader impact of AI on society. By combining real-world examples with data analysis, this project highlights both the power and responsibility that come with using AI.

Data Spill: When Information Goes Where It Shouldn't

Valeria Rodriguez, Victoria Johnson
Advisor: Glenda Guerrero, Ed.D.

The research focuses on privacy/security breach concerns in data. User information is in constant transmission all over the internet especially on social media platforms. Social media has various doorways that many users are highly influenced by. By looking at charts and research experiments we would be able to visually see the impact of social media amongst us human beings whether it be good or bad. We're doing a widespread view in research, seeking informative data to showcase the total count in reports on privacy laws in the U.S and other countries, implementation and compliance, surveillance and government access, mobile privacy and app tracking, and more interesting facts about how our phones/devices keep files and create our personality digitally.

Predicting Protein Drug Complexes for Crohn's Disease-Associated Targets

Emily Hernandez
Advisor: Hugo Guterres, Ph.D.

HGFAC (hepatocyte growth factor activator) is a serine protease that activates hepatocyte growth factor (HGF), a protein with cell growth and tissue repair functions, making HGFAC a strong target for drug development. In Crohn's disease, an HGFAC R509H mutation leads to a slower healing response in the gut. Through AutoDock Vina, HGFAC's PDB id and 1,500 FDA-approved drugs' SMILES codes were individually tested to compare binding affinity. The top 3 drugs with the highest binding affinity were: Apomorphine, Doxorubicin, and Daunorubicin. Apomorphine is used to treat hypomobility associated with advanced Parkinson's disease.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Predicting Protein Drug Complexes for Crohn's Disease-Associated Targets (cont.)

Doxorubicin is used to treat various cancers, but has cardiotoxicity as a major side effect. Daunorubicin is used as a treatment to induce remission of certain types of leukemia. All three drug molecules show hydrogen bonding interactions with GLU 571, VAL 572, SER 569, LYS 566, CYS 583, and PRO 570, as well as ring stacking. With further research, we hope that these drugs can be repurposed to be used as potential therapeutics for patients with Crohn's disease.

Do moral and pragmatic judgments differ in their physiological correlates?

Keily Barrera, Marcus Bell, Jaylen Hargrove
Advisor: Daniel Wisneski, Ph.D., & Brittany Hanson, Ph.D.

The current study explores possible ways that moral judgments differ from non-moral judgments. Previous research has found that evaluations of everyday behaviors (e.g., "studying", "wearing a seatbelt") framed in moral terms tend to be more extreme (i.e., more right/wrong) and applied more universally compared to non-moral evaluations (Van Bavel et al., 2012). Furthermore, morally framed evaluations were also found to be made more quickly – a finding consistent with theories that moral judgments are based on brief flashes of affect laden "moral intuition" (Haidt, 2001). Building off this past research, the current study sought to also test the hypothesis that moral judgments are more "intuitive" by testing whether they are associated with greater physiological arousal. Participants (N = 87) evaluated a series of 60 different behaviors while we collected data on their skin conductance levels. Furthermore, we also measured several individual difference variables related to affective/intuitive responding to help validate any skin conductance results we found.

Our results showed mixed evidence for the intuition hypothesis. Although moral (compared to non-moral) judgments were made more quickly and associated with higher skin conductance levels, analyses including relevant individual difference measures showed a pattern of results opposite what the intuition hypothesis would predict.

The Effects of High- vs. Low-Intensity Exercise and Its Effects on Brain Chemistry and Focus in Individuals with ADD/ADHD

Mia Dominguez

Advisor: Joshua Williams, DHS

Attention Deficit Hyperactivity Disorder (ADD/ADHD) is associated with challenges in impulse control, attention, and general cognitive functioning, which can affect performance in school, jobs, and social settings. Growing research has examined exercise as an alternative or complementary treatment, even though standard approaches such as medication and therapy remain the most common. The effects of different types of exercise, specifically high-versus low-intensity exercise, on brain chemistry and focus in people with ADD/ADHD are discussed in this review. Exercise has been associated with increased levels of key neurotransmitters such as dopamine, norepinephrine, and serotonin, which are important for motivation, attention, and emotion regulation. Improvements in inhibition control and prolonged focus have all been associated with regular exercise. Moderate- to high-intensity exercise often produces greater short-term cognitive benefits, while lower-intensity activity may contribute to longer-term mental health improvements, although results vary based on factors such as age, consistency, and exercise intensity.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

The Effects of High- vs. Low-Intensity Exercise and Its Effects on Brain Chemistry and Focus in Individuals with ADD/ADHD (cont.)

Overall, research suggests that exercise can be an effective method for improving concentration and cognitive function in people with ADD/ADHD. These results suggest that incorporating physical activity into daily routines may boost daily task performance as well as neurological functioning.

Predicting the “protein-drug” complex structure for Crohn’s Disease-associated targets

Misheel Myagmarjav

Advisor: Hugo Guterres, Ph.D.

MST1 (Macrophage Stimulating Protein-1) plays a crucial role as a protein kinase and a key component of the Hippo signaling pathway, which governs vital cellular processes such as proliferation, apoptosis, and organ size regulation. This study employs computational biochemistry tools to predict the complete three-dimensional structure of the MST1 protein, consisting of 480 amino acid residues. Includes a library of 3,200 drug molecules that have been approved by the FDA. Our ongoing research has identified the top three drugs exhibiting stable binding interactions with MST1. The first, carbinozamine, is commonly used to alleviate allergic symptoms, irritant cough, and motion sickness. The second compound, azlocillin, is an acylampicillin with a broad spectrum of activity and superior in vitro potency compared to traditional carboxypenicillins. The third candidate, zinc gluconate, is an essential mineral present in nearly all human cells. All three drug molecules demonstrate stable noncovalent interactions with key MST1 residues, including ILE 121, GLU 124, PHE 125, and PRO 112, highlighting their potential

as therapeutic agents targeting MST1. With further research, we expect that these existing FDA drugs can be repurposed as potential therapeutics for patients with Crohn’s disease.

The Bizarreness Effect is Inconsistent in Free Recall

Maria Medrano

Advisor: Maryellen Hamilton, Ph.D.

Morita and Kambara (2021) demonstrated the inconsistency of the color bizarreness effect in object recognition. We extended these inconsistent findings to free recall. Participants saw objects in either their typical color (pink pig) or a bizarre color (blue tomato). Orientation tasks were manipulated across three experiments. Similar to the findings for object recognition, we found that a typical bizarreness effect (greater memory for bizarre versus typical colored objects) was not consistently found in free recall

Resistance and Climbing-Specific Training for Injury Prevention in Youth Climbers

Noah Santos

Advisor: Joshua Williams, DHS

This literature review focuses on resistance training and climbing-specific training for injury prevention in youth climbers. It proposes that combining strength training with skill development will enhance performance while also limiting the potential for injury. By developing essential motor skills alongside strength, such as footwork and balance, youth climbers can improve their overall proficiency. This approach also leads to a healthier relationship with climbing, allowing individuals to maintain their physical well-being and continue the activity as a lifelong hobby. This review highlights the importance of using appropriate training methods and following proper injury prevention strategies to promote safety and long-term sustainability in youth climbing..

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Putative Protective Effects of Ginkgolide B and Folic Acid Individually, and in Combination, Against Polystyrene-Induced Developmental Dysregulation in *Xenopus laevis* (African Clawed Frog)

Darwin Breton

Advisor: Laura Twersky, Ph.D.

The objective of this experiment is to see if there is a correlation and causation relationship between ginkgolide B, individually and in combination with folic acid, and any protective effects against polystyrene with respect to development in *Xenopus laevis*. This is appraised by investigating whether ginkgolide B counteracts the effects that polystyrene might pose to the developmental stages of *Xenopus laevis*. Ginkgolide B is documented to possess protective effects, such as neuroprotective, anti-inflammatory, and antioxidant effects. Microplastics exert teratogenic effects on amphibian development, such as delayed metamorphosis and morphological malformations, with polystyrene documented to induce adverse developmental effects on *Xenopus laevis* development, such as abnormal swimming behavior, bradycardia, and reduced motility. Folic acid is known to be a protective agent in development and can act as an enhancer of any protective effects that ginkgolide B can possess. There are nine experimental groups: control, 0.1% ethanol, 5µg/mL ginkgolide B, 5µg/mL folic acid, 5µg/mL polystyrene, 5µg/mL of ginkgolide B and folic acid mixture, 5µg/mL ginkgolide B and polystyrene mixture, 5µg/mL folic acid and polystyrene mixture, and 5µg/mL mixture of ginkgolide B, folic acid, and polystyrene. Each group contained 15 *X. laevis* tadpoles that were exposed to these concentrations starting at Nieuwkoop and Faber stage 39.

Throughout development any changes that differ from that of the control group were recorded, such as heart rate, developmental rate, motility, mortality, and any malformations. The hypothesis is that the tadpoles exposed to ginkgolide B, and even more so those exposed to the combination of ginkgolide B and folic acid, will have fewer or no developmental defects, or malformations, and a higher chance of undergoing normal development as opposed to those only exposed to polystyrene. Due to the promising results acquired in the folic acid, polystyrene, and combination of folic acid and polystyrene groups, a second trial was conducted with four experimental groups: control, 5µg/mL folic acid, 5µg/mL polystyrene, and 5µg/mL of polystyrene and folic acid mixture. Each experimental group contained 10 tadpoles and exposure to these concentrations began at Nieuwkoop and Faber stage 37. The purpose is to see whether the same results of folic acid being protective against polystyrene effects will be observed again, thus enhancing credibility of results.

The Great Backyard Bird Count 2026: Joining Together to Count Winter Birds

Alexandra Ramirez, Suleimy Gomez Colon, Phoenix Campbell

Advisor: Katherine Wydner, Ph.D.

The Great Backyard Bird Count (GBBC) is an event that joins people together from all over the world to count birds! It encourages citizen scientists from every corner of our planet, wherever they are, to appreciate their local avians and connect with nature. The count is set in mid-February, so in the northern hemisphere, the GBBC is essentially a mid-winter census of birds. GBBC participants take at least 15 minutes to notice the birds around them, identify them, count them, and submit the counts to help scientists better understand and protect birds around the world.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

The Great Backyard Bird Count 2026: Joining Together to Count Winter Birds (cont.)

The data collected are submitted through eBird, a website managed by the Cornell Lab of Ornithology. Observations of birds can be stationary or traveling, in your backyard or in a park, or anywhere. In our poster, we describe checklists that we made during the GBBC, provide photos of some of the species, and report about our experiences and highlights. Birds are one of our best connections to nature, have imserve as an indicator of the health of our environment, and are an indispensable and important part of human culture.

The Twelfth Season of FeederWatch: Fostering a Diverse Avian Community Amidst a Historic Winter!

Darwin Breton, Alexandra Ramirez, Phoenix Campbell, Genesis Serpas, Kaidy Abreu, and Suleimy Gomez Colon

Advisor: Katherine Wydner, Ph.D.

In the twelfth season of Project FeederWatch on the Saint Peter's University campus, 18 species have been reported. The highlight of this season includes an unusually cold and snowy winter. This winter season was the coldest in a decade (the coldest since 2014-2015) and featured "bomb cyclones" that brought historic cold and snowfall to many regions of the Northeast, such as New Jersey. However, despite this winter being colder than the previous winter there was an increase in total unique species observed this season compared to last season, in which 17 species were reported as opposed to 18. This further emphasizes the need to provide winter birds with a safe space and food as colder and harsher winters make survivability harder and increases food scarcity.

Our PFW data collection began in 2014, a native plant garden was established and feeders expanded in 2018, and food offerings at feeders were further diversified in 2023. A total of 40 species have been reported on campus throughout the years, accompanied with a positive trend in terms of the numbers of species reported seasonally. PFW runs from November to April annually. For 39 years and counting, the Cornell Lab of Ornithology has been tracking the status of North American winter bird populations, which is a task made possible with the data reported on our very own campus.

Predicting Comets' Trajectories: Integrating Data Science with Physics

Agha Sibte E Ali Qazalbash

Advisor: Gulhan Bizel, Ph.D.

This honors thesis explores the application of data-driven methods to predict comet trajectories using modern computational tools. Accurate prediction of comet paths is essential for understanding celestial mechanics, assessing potential Earth impact risks, and advancing space research. The study utilizes datasets from the Comet Observation Database (COBS) in combination with the Astropy library to model orbital parameters and simulate future trajectories.

The research focuses on processing observational data, cleaning inconsistencies, and applying mathematical models based on Keplerian motion and gravitational dynamics. By integrating statistical analysis with computational astronomy techniques, the study evaluates the accuracy of predicted trajectories against known comet paths. Additionally, the project examines sources of error, including observational limitations and perturbations caused by planetary interactions.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Predicting Comets' Trajectories: Integrating Data Science with Physics (cont.)

The findings demonstrate that combining real-time observational data with computational modeling can significantly improve trajectory predictions. The results highlight the importance of continuous data collection and refinement of algorithms in achieving higher precision. This work contributes to the growing field of computational astrophysics by providing a framework for analyzing and predicting comet motion. Future research may expand on this model by incorporating machine learning techniques and larger datasets to further enhance predictive accuracy.

Interest Rates and Government Bond Yields across Countries: Comparative and Predictive Analysis

Hadeed Khan

Advisor: Gulhan Bizel, Ph.D.

Government bond yields dictate borrowing costs for the government and they shape how global investment decisions are made. However, the way bonds respond to interest rate policy differs across countries. While there is existing research and strong links between monetary policy and bond yields, there aren't that many that package a side by side cross country comparison using recent data (post 2015 for now) and a clear predictive base. The project will test how consistent the relationship between interest rate and bond yield is across selected countries and if it can be modeled to improve forecasting.

The Sonic Evolution of Hip-Hop: A Data-Driven Analysis of Audio Features Over Time

David Dandridge

Advisor: Gulhan Bizel, Ph.D.

Hip-hop has undergone significant stylistic changes over time, yet these shifts are often described qualitatively rather than measured quantitatively. This study presents a data-driven analysis of hip-hop's sonic evolution using track-level audio features and machine learning techniques. A Spotify dataset of over 100,000 songs was filtered to approximately 1,000 hip-hop tracks based on genre labels. Additional lyrical data was incorporated to support interpretation. Data preparation and feature engineering were conducted using Altair AI Studio, including the creation of variables such as duration, explicit content indicators, and a binary popularity-based target variable ("hit_flag"). Exploratory analysis reveals clear trends over time, including increases in energy, loudness, and danceability, along with decreases in speechiness, indicating a shift toward more melodic and production-driven styles in the streaming era. To evaluate whether these features relate to song success, Logistic Regression and Random Forest models were developed to predict hit songs. Model performance was assessed using accuracy, precision, recall, and F1-score, with Random Forest showing stronger overall performance. These findings suggest that hip-hop has evolved toward a more energetic and commercially optimized sound, highlighting the role of data science in analyzing cultural change.

ABSTRACTS-UNDERGRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Vitamin D Deficiency in Athletes: A Literature Review of Performance and Health Consequences

Justin Stotz

Advisor: Joshua Williams, DHS

Vitamin D deficiency is a common health problem worldwide. Every year, the number of people becoming vitamin D-deficient rises. The dominant explanation for the upsurge in vitamin D deficiency is inadequate sunlight exposure and a diet with minimal vitamin D intake. This literature review draws on data from 14 studies and 1 systematic review to examine the association between vitamin D, athletic performance, and the health consequences of inadequate levels. These findings suggest that low vitamin D levels have a strong relationship with decreased muscular strength, power output, and endurance, as well as an increased risk of sustaining injuries. Vitamin D plays a crucial role in the health of the human skeletal system and in how the body defends itself against illnesses and diseases. Overall, it is imperative to maintain adequate vitamin D levels, especially for athletes, to optimize athletic performance and support long-term health.

Visualizing the Impact of COVID-19.

Alberth Rojas, Dev Patel & Sraboni Haque
Advisor: Glenda Guerrero, Ed.D.

The COVID-19 pandemic has reshaped global health, economies, and communities. This project offers an overview and timeline, using different data visualizations to analyze key trends and patterns. Examining multiple data sources, it shows how COVID-19 affected public health systems, economic stability, and social structures. This analysis stresses the importance of data literacy in informing responses to global crises and reflects on lessons for future preparedness.

Cold Case Map

Ashley Ramcharran

Advisor: Gulhan Bizel, Ph.D.

This cold case map project gives an overall visual look at unsolved cases by showing them on a map and helping reveal patterns that may not stand out right away in a chart or spreadsheet. The project is meant to make the information easier to understand by organizing the cases in a visual way and showing how location can play a role in identifying trends or possible connections. Instead of only looking at rows of data, the map makes it easier to see where cases are concentrated and how different cases may relate to one another geographically.

This project also shows how data visualization can be useful in criminal justice and investigative work. By turning case information into a map, it becomes easier to notice details, compare areas, and think more critically about what the data might suggest. It helps create a clearer picture of the bigger story behind the cases rather than just focusing on one case at a time.

This project highlights the value of using visual tools to better understand unsolved crimes. It brings the data together in a way that is more engaging, easier to analyze, and helpful for discussing patterns, questions, and possible next steps in an investigation.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Credit Card Fraud Detection Using Data Mining Techniques

Student(s): Sneha Gurampally Malkan & Roystin Ronaldo Felix

Advisor: Vijay Kumar Reddy Voddi, MSDS

Credit card fraud detection is a critical application of data mining that helps financial institutions prevent economic losses and protect consumers. This project develops a data mining framework to identify fraudulent credit card transactions using the publicly available Kaggle Credit Card Fraud dataset, which contains over 20 anonymized transaction features. The study applies exploratory data analysis and preprocessing techniques, including normalization and class imbalance handling. Feature selection methods are used to identify key predictors, while clustering techniques explore behavioral transaction patterns. Because fraudulent transactions often appear as anomalies, outlier detection methods such as Isolation Forest and Local Outlier Factor are implemented alongside classification models including Logistic Regression and Random Forest. Model performance is evaluated using precision, recall, F1-score, and ROC-AUC metrics. Beyond technical modeling, the project examines societal concerns including data privacy, fairness in automated decision-making, de-anonymization risks, and the impact of false positives on legitimate customers.

Network Intrusion Detection Using Machine Learning

Student(s): Gurpreet Kaur

Advisors: Vijay Kumar Reddy Voddi, MSDS

With the rapid growth of internet usage, cybersecurity threats such as network intrusions and cyber attacks have increased significantly. This project presents a machine learning-based intrusion detection system that identifies

malicious network traffic using classification algorithms. The system uses the CICIDS-2017 dataset developed by the Canadian Institute for Cybersecurity to detect attack patterns. Various machine learning models, including Isolation Forest, Random Forest, and Logistic Regression, are implemented and evaluated. The results demonstrate the effectiveness of machine learning techniques in detecting network intrusions. Data preprocessing, feature analysis, and performance visualization were also conducted to improve detection accuracy and interpretability.

Detection of Bank Account Fraud: Securing Digital Onboarding through Machine Learning and Resampling Techniques

Student(s): Mohammed Aroui

Advisors: Vijay Kumar Reddy Voddi, MSDS

Digital onboarding, while convenient, has introduced new vulnerabilities for bank account fraud, including the use of stolen identities and synthetic profiles. This project addresses the critical need for robust fraud detection systems by developing a machine learning pipeline that incorporates advanced data preprocessing, feature engineering, and sophisticated resampling techniques. We evaluate the performance of LightGBM model, demonstrating how this method significantly enhance the detection of fraudulent activities, thereby protecting users, improving bank credibility, and strengthening data security in the evolving financial landscape.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Global Warming Patterns: A Data Mining Study

Student(s): Biyue Wang & Chander Rekha Sehgal

Advisor: Vijay Kumar Reddy Voddi, MSDS

This study applies data mining techniques to identify heterogeneous climate regimes and examine their long-term warming behavior using the Berkeley Earth Surface Temperature dataset (experimental version). Twenty-six engineered features capturing temporal, geospatial, and statistical characteristics were constructed. K-Means clustering ($k=4$) revealed four distinct urban climate regimes differentiated primarily by baseline temperature and seasonal variability. Subsequent regression analysis demonstrates unequal warming acceleration across these regimes, with higher-latitude continental urban regions exhibiting stronger long-term warming trends ($\sim+0.10^{\circ}\text{C}$ per decade). Results highlight latitude-dependent warming gradients while acknowledging that findings are constrained by urban-only sampling and dataset structure. The study emphasizes that climate data mining reveals patterns bounded by representation and methodological design.

Airline Delay Prediction using Data Mining

Student(s): Tejaswini Konakanchi & Vinesh Nangunoori

Advisors: Vijay Kumar Reddy Voddi, MSDS

Flight delays cause inconvenience to passengers and financial losses to airlines. This project focuses on analyzing airline delay data using data mining techniques to identify key factors responsible for delays and to build predictive models. The dataset was preprocessed, cleaned, and analyzed using statistical and machine learning methods. Logistic Regression and Random Forest models were applied for prediction.

The results show that the Random Forest model performs better with higher accuracy. The study demonstrates how data mining can help airlines improve decision-making and operational planning.

Doctor's Handwritten Medical Prescription Recognition Using Deep Learning

Student(s): Ganesh Tilukuri

Advisors: Vijay Kumar Reddy Voddi, MSDS

Handwritten medical prescriptions are commonly used in healthcare systems, yet they are often difficult to read due to variations in handwriting styles, abbreviations, and poor image quality. Misinterpretation of such prescriptions can lead to medication errors and serious health consequences. This project presents a deep learning-based approach for recognizing handwritten medical prescriptions and converting them into accurate digital text. The proposed system accepts an image of a handwritten prescription as input and processes it using Convolutional Neural Networks (CNNs) to extract relevant features and recognize textual content. The model is trained on a dataset of handwritten prescription images written in the English language and optimized using techniques such as the Adam optimizer and batch normalization to improve performance and stability. Experimental results show that the system achieves an accuracy of 71.25%, along with strong precision and recall values, demonstrating its effectiveness in identifying medicine names. This solution aims to support pharmacists and healthcare professionals by reducing ambiguity in prescription interpretation and minimizing the risk of medication errors. Furthermore, the digitization of prescriptions enables efficient storage, retrieval, and integration with electronic healthcare systems. Future work includes training the model with larger and more diverse datasets, improving recognition accuracy, supporting multiple languages, and deploying the system through mobile and web-based healthcare applications for real-world use.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Heart Disease Risk Prediction using Machine Learning

Student(s): Irfan shaik, Abhi Sekhar Rao Talluri

Advisors: Vijay Kumar Reddy Voddi, MSDS

Heart disease remains one of the leading causes of mortality worldwide. Early prediction of risk factors can significantly improve preventive healthcare. This project develops a machine learning model to predict heart disease using clinical attributes such as age, cholesterol, blood pressure, and chest pain type. The dataset is preprocessed and analyzed using Python, and classification models are trained and evaluated using standard performance metrics. The results demonstrate that machine learning can effectively assist healthcare professionals in early diagnosis and risk assessment.

Detecting Depression Signals from Social Media Using Data Mining Techniques

Student(s): Nikhila Boppana

Advisors: Vijay Kumar Reddy Voddi, MSDS

Depression is a major public health concern affecting millions of individuals worldwide. With the increasing use of social media platforms, users often express emotions, thoughts, and mental states through online posts. This project explores the application of data mining techniques to analyze social media text data and identify patterns associated with depressive behavior. Using a publicly available mental health social media dataset, this study performs data preprocessing, feature engineering, exploratory data mining, clustering, association rule mining, and classification. The goal of the project is not to diagnose individuals, but to uncover linguistic, emotional, and behavioral indicators correlated with depression while also addressing ethical, privacy, and societal

implications of mining sensitive mental health data.

Cardiovascular Integrated Risk Analytics System

Student(s): Manish Lakshmpuram, Shahidullah, Md Alamgir

Advisor: Vijay Kumar Reddy Voddi, MSDS

This project introduces the Cardiovascular Integrated Risk Analytics System, a cutting-edge machine learning setup to boost the precision of heart disease predictions. The study tackles the issue of scattered medical information by combining two separate datasets. One contains 13,000 records of demographic and lifestyle data, while the other has 1,190 records of clinical diagnostic information. The project uses SQLite to manage data and create new features.

The system checks out several prediction methods, such as Random Forest, Naive Bayes, and SVM. The Decision Tree model comes out on top getting 92.9% of predictions right and scoring 0.937 on the AUC scale. To make it useful in clinics, the project includes a Power BI dashboard you can interact with and a GUI app that assesses risk in real-time. By connecting social and economic factors with clinical signs, this all-in-one approach gives doctors a complete tool to step in for heart problems and make choices based on data in healthcare.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

S&P 500 Stock Prediction & Clustering

Student(s): Praveen Kumar Inti

Advisor: Vijay Kumar Reddy Voddi, MSDS

This project applies data mining techniques to S&P 500 stock data to test whether technical indicators can predict next-day price direction and to discover natural company groupings. Daily OHLCV stock prices were cleaned and merged with company metadata, then transformed into engineered features including daily returns, moving averages, volatility, price range, and volume change. A time-based split was used to prevent leakage, training models on 2010–2019 and testing on 2020–2024. Logistic Regression and Random Forest models were trained for binary classification (Up vs Down) and evaluated using Accuracy, Balanced Accuracy, and ROC-AUC. Results showed only small improvements over baseline, confirming that short-term stock direction is difficult to predict using technical indicators alone. A simple equal-weight trading backtest further demonstrated that modest predictive gains do not guarantee higher trading returns. Finally, K-Means clustering was applied to company-level return and volatility features, revealing two distinct volatility-based groups. Overall, the project highlights the limitations of technical-only prediction models and shows how clustering can add value by summarizing market structure.

Hospital Inpatient Length of Stay

Student(s): Bethlehem Seid, Srijana Lama

Advisor: Vijay Kumar Reddy Voddi, MSDS

Hospital length of stay (LoS) is an important metric influencing the availability of beds, patient outcome, cost management and staffing. The primary objective is to apply various machine learning algorithms to the SPARCS (Statewide Planning and Research Cooperative system) dataset to accurately categorize the length of stay and identify factors influencing the length of stay.

Clinical Pattern Discovery & Mortality Risk Modeling

Student(s): Deepali Poudel

Advisors: Vijay Kumar Reddy Voddi, MSDS

Hospitals generate large volumes of electronic health record (EHR) data, yet early identification of high-risk patients remains challenging. This project explores whether early clinical features can reveal patient phenotypes and support in-hospital mortality risk stratification using the SyMile-MIMIC dataset, which combines laboratory results, imaging labels, and demographic information from 11,622 hospital admissions.

I performed exploratory data analysis to examine relationships between key laboratory markers and mortality, followed by dimensionality reduction using principal component analysis and unsupervised(KMeans) clustering to identify patient subgroups.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Clinical Pattern Discovery & Mortality Risk Modeling (cont.)

Predictive modeling was conducted using logistic regression and random forest classifiers, with performance evaluated via ROC-AUC. Fairness was assessed by comparing model discrimination across gender and racial subgroups.

Clustering revealed clinically meaningful phenotypes, including a high-acuity group characterized by metabolic acidosis, renal dysfunction, and elevated inflammatory markers, with substantially higher mortality. Logistic regression achieved an AUC of 0.77 and random forest an AUC of 0.78, indicating moderate but clinically meaningful discrimination. Key predictors reflected known physiologic processes in critical illness, including shock, coagulopathy, organ dysfunction, and frailty. Subgroup analysis showed broadly similar performance across demographic groups, though modest variation highlights the importance of ongoing fairness monitoring.

These findings demonstrate that early structured EHR data contain meaningful signals for risk stratification. However, ethical deployment requires attention to privacy, representation, and clinician oversight. Future work will incorporate longitudinal trends, clinical notes, and social determinants of health to provide a more complete understanding of patient risk.

Predicting and Interpreting Residential Rent Patterns : A Modeling and Interactive Decision-Support Approach

Student(s): Yuqiu Gan, Yujie Liu

Advisors: Vijay Kumar Reddy Voddi, MSDS

This project analyzes and predicts residential rent patterns at the census tract level in Seattle by integrating predictive modeling, attribution analysis, and an interactive decision-support prototype.

Using socio-economic and housing-related variables, we developed machine learning models to estimate inflation-adjusted rents while capturing non-linear relationships among income, employment, housing supply, and neighborhood characteristics. Model performance was evaluated using standard regression metrics to ensure robust predictive accuracy.

Beyond prediction, the project emphasizes interpretability. Feature importance analysis and model explanation techniques were applied to identify the most influential drivers of rent variation and to understand how these factors shape neighborhood-level price differences.

To translate analytical findings into practical insights, we built an interactive dashboard that allows users to explore spatial rent patterns, examine key drivers, and test how changes in selected features affect predicted rent levels.

The final product functions not only as a predictive model but also as a business-oriented analytical tool that supports real estate stakeholders and urban planners in data-driven decision making.

AI-Driven Predictive Modeling and Time Series Forecasting of Employee Burnout Using Behavioral, Productivity, and Lifestyle Analytics

Student(s): Rishabh Perewar, Dakshayani Uppu

Advisors: Vijay Kumar Reddy Voddi, MSDS

Employee burnout has become a major organizational and public health challenge, resulting in reduced productivity, increased employee turnover, and significant financial losses. Early prediction and forecasting of burnout risk can enable organizations to implement preventive measures and improve workforce well-being. This study proposes an AI-driven predictive analytics framework to identify and forecast employee burnout using behavioral, productivity, and lifestyle data. Multiple machine learning models, including Linear Regression, Random Forest,

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

AI-Driven Predictive Modeling and Time Series Forecasting of Employee Burnout Using Behavioral, Productivity, and Lifestyle Analytics (cont.)

Support Vector Machine, and Gradient Boosting, are applied to predict burnout scores and classify burnout risk levels.

Additionally, time series forecasting methods such as ARIMA and Prophet are used to predict future burnout trends, allowing proactive intervention and resource planning. Feature engineering and dimensionality reduction techniques, including Principal Component Analysis, are employed to enhance model performance and interpretability. Model performance is evaluated using standard metrics such as accuracy, precision, recall, F1-score, RMSE, and MAE.

The results demonstrate that predictive and forecasting models can accurately identify burnout risk and forecast future burnout levels. This framework provides actionable insights for organizational decision-making and supports effective workforce health monitoring and burnout prevention strategies.

Delhi AQI Prediction

Student(s): Sagar Maru, Pankaj Sharma, Garima Chaudhary

Advisors: Vijay Kumar Reddy Voddi, MSDS

Delhi, home to over 20 million people, faces one of the world's most severe air quality crises. This study develops and evaluates machine learning models to forecast daily PM_{2.5} concentrations in Delhi by integrating three complementary public data sources: ground-level pollutant measurements from OpenAQ (16 monitoring stations), daily weather observations from NOAA's National Centers for Environmental Information (NCEI), and satellite-based fire detections from NASA's Fire Information for Resource Management System (FIRMS).

Analyzing 2,463 days of measurements from 2019 through 2025, we document chronic year-round pollution with a median PM_{2.5} concentration of 67.8 $\mu\text{g}/\text{m}^3$ (4.5 \times the WHO 24-hour guideline of 15 $\mu\text{g}/\text{m}^3$), with severe seasonal exacerbation during post-monsoon months reaching 200–350 $\mu\text{g}/\text{m}^3$. Through systematic feature engineering—including lag variables, rolling statistics, and temporal encodings—and evaluation of 12 models spanning four model families (baseline, linear, tree-based, and gradient boosting), we find that a Random Forest model achieves the best predictive performance ($R^2 = 0.327$, RMSE = 100.91 $\mu\text{g}/\text{m}^3$, Cross-Validation RMSE = $42.42 \pm 5.5 \mu\text{g}/\text{m}^3$), outperforming all baselines including the Seasonal Mean ($R^2 = 0.316$). Cross-correlation analysis reveals that agricultural fires in Punjab-Haryana precede Delhi's PM_{2.5} spikes by approximately 2–10 days (peak $r = 0.263$ at lag-2), confirming the atmospheric transport pathway. Feature importance analysis identifies yesterday's PM_{2.5} (lag-1) as the dominant predictor (78.6% importance), with rolling statistics (8.7%) and fire-related features (3.0%) providing additional predictive signal. Correlation matrix analysis identifies two distinct pollution regimes: a combustion cluster (PM_{2.5}, PM₁₀, CO, NO₂; $r = 0.71$ – 0.92) and independent industrial (SO₂) and photochemical (O₃) components. An interactive map-based prototype demonstrates the practical application of these findings for public health decision-making. The study establishes a reproducible, API-first methodology for environmental data integration that can be adapted to other pollution-affected cities globally.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Recidivism Prediction & Algorithmic Fairness

Selin Ozguler, Hasan Mert Sungur, Umut Baris Basol

Advisors: Vijay Kumar Reddy Voddi, MSDS

This study evaluates recidivism prediction and algorithmic fairness using the COMPAS dataset. Logistic Regression and Random Forest models were developed under different configurations, including models with race, without race, and with the COMPAS decile score. Predictive performance was assessed using Accuracy and ROC-AUC, while fairness was examined through group-based error rates.

Results show that removing race had minimal impact on overall performance, and including the COMPAS score slightly improved accuracy. However, racial disparities in false positive rates persisted across models. These findings highlight the tradeoff between predictive accuracy and fairness, emphasizing the need for fairness-aware modeling in criminal justice applications.

GateKeep: An Agentic AI Platform for Personal Financial Intelligence

Samir Sanjeev Sengupta

Advisors: Gulhan Bizel, Ph.D.

GateKeep is an agentic AI-powered platform designed to provide users with intelligent, automated insights into their personal finances. Inspired by platforms like Rocket Money, GateKeep enables users to securely upload financial documents such as official bank statements and authenticate via Google login for seamless data integration and storage.

The system leverages advanced AI agents to analyze spending patterns, categorize expenses, and generate actionable insights.

It identifies unnecessary expenditures, highlights opportunities for savings, and recommends personalized financial rules to help users build disciplined financial habits. By combining automation, contextual understanding, and user-centric design, GateKeep aims to transform raw financial data into meaningful, decision-ready intelligence, empowering users to take control of their financial future.

Modeling Fatal Outcomes of Motor Vehicle Collisions Using Random Forest

Student(s): Sonali Gupta

Advisors: Vijay Kumar Reddy Voddi, MSDS

Motor vehicle collisions remain a critical public safety issue in densely populated urban areas. This study will utilize the Motor Vehicle Collisions – Crashes dataset from NYC Open Data to develop a machine learning model for predicting fatal outcomes of traffic accidents in New York City.

The dataset includes temporal and geographic attributes such as crash date, time, and borough, along with recorded fatality counts. Data pre-processing will involve cleaning data, removing missing values, encoding categorical variables, and scaling numerical features. A Random Forest classification model will be implemented to estimate the likelihood of fatalities occurring in a given collision.

Feature importance analysis will be conducted to identify factors most strongly associated with fatal outcomes. Model performance will be evaluated using standard classification accuracy metrics. This approach highlights the potential of machine learning to support traffic safety initiatives and inform strategies aimed at reducing collision-related fatalities.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Credit Card Fraud Detection Using Data Mining

Afreen Tabasum begum Fnu

Advisors: Vijay Kumar Reddy Voddi, MSDS

Credit card fraud is a huge financial crime that costs billions of dollars each year. As digital payments and online transactions grow rapidly, they have become a hot spot for financial fraud, which has advanced through the use of special technical means that make detection with classical rule-based systems more complex. To enhance the accuracy of fraud detection, various financial institutions are increasingly using data mining and machine learning techniques. The aim of this study is to build and compare supervised machine learning algorithms against a highly imbalanced real-life dataset for credit card fraud detection. The analysis was done using the publicly available Kaggle Credit Card Fraud Detection dataset of 284,807 transactions made over two days by European cardholders in September 2013, from which only 492 resulted in fraud. We implemented Logistic Regression, Decision Tree, and Random Forest models. Among all the models, the Random Forest model achieved the better balance between fraud detection based on class distribution.

Besides technical performance, the research also examines critical societal implications such as privacy, fairness, and the impact on customers and financial institutions due to automated decision-making systems.

The Global Heartbeat: A Computational Analysis of Rhythmic Universals and Cross-Cultural Neural Resonances

Gaurab Shrestha, Bishal Gupta

Advisors: Vijay Kumar Reddy Voddi, MSDS

This research investigates the existence of "Rhythmic Universals"—mathematical patterns in percussion and tempo that transcend geographic and cultural boundaries. While musical genres vary significantly across the globe, this study hypothesizes that certain "core beats" are hard-coded into human biology.

By leveraging Machine Learning (ML) algorithms and Digital Signal Processing (DSP), the project analyzes a massive cross-cultural dataset ranging from traditional West African drumming and Indian Classical Talas to modern Western Electronic Dance Music (EDM). The goal is to isolate the specific BPM (Beats Per Minute) and syncopation ratios that most effectively trigger Neural Entrainment—the process where brainwaves synchronize with external rhythms.

Using a Word2Vec-inspired sequence learning approach, the study maps how these "Universal Beats" have evolved and adapted into different modern subcultures. The final result is a Neuro-Acoustic Framework that defines the "Ideal Human Rhythm." This research moves beyond simple musicology into the realm of Evolutionary Neuroscience, providing a data-driven explanation for why certain rhythms have dominated human history and continue to drive the global music industry today.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Predicting Colorectal Cancer Survival with Machine Learning

Michael Uhler, Suman Lamsal
Advisors: Gulhan Bizel, Ph.D.

Colorectal cancer is one of the leading causes of cancer-related mortality worldwide, and survival outcomes vary based on a combination of clinical, demographic, and lifestyle factors. The purpose of this study was to develop a machine learning model to predict colorectal cancer survival outcomes and identify the most important predictors associated with poor prognosis. A retrospective dataset of 89,945 colorectal cancer patients was analyzed using a Random Forest classifier. Data preprocessing included encoding categorical variables, standardization, and principal component analysis. The dataset was split into training and testing sets using an 80/20 split, and model performance was evaluated using accuracy, precision, recall, and F1-score. The model achieved an accuracy of 75%. Feature importance analysis revealed that BMI, age, stage at diagnosis, physical activity level, alcohol consumption, and tumor aggressiveness were the most important predictors of survival. The results suggest that both clinical and lifestyle factors significantly influence colorectal cancer survival outcomes. Machine learning models may be useful tools for risk prediction and may support early intervention and prevention strategies.

Language Development & Evolution in Children

Angelica Amegashie
Advisors: Dong Ryeol Lee, Ph.D.

This study analyzes the CHILDES dataset which contains real transcriptions of conversation between adults and children.

This study aims to answer questions such as the following:

- 1) does speech get longer with age, complexity development;
- 2) does sentence structure become more complex, child v. adult speech;
- 3) how does children's speech differ from caregivers, input effects;
- 4) does parent/caregiver speech influence child language?

AI-Driven Market Intelligence for Identifying High-Growth Startup Opportunities: Insights from the Pet Nutrition Industry

Laura Becerra, Dari Dambaeva; Jaykumar Panchal
Advisors: Gulhan Bizel, Ph.D.

This study examines how an AI-inspired market intelligence framework can be used to identify high-growth startup opportunities within the pet nutrition industry. Using an exploratory and descriptive mixed-methods approach, the research analyzes secondary data from 2021 to 2026, focusing on growth-stage firms and selected industry leaders. Data were collected from financial databases, corporate reports, and market intelligence sources, and subsequently standardized to ensure comparability.

The analysis integrates four key indicators: market trends, investment activity, product market alignment, and financial performance. Rather than relying on predictive modeling, the study adopts a signal-based approach, where opportunity identification emerges from the convergence of multiple indicators.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

AI-Driven Market Intelligence for Identifying High-Growth Startup Opportunities: Insights from the Pet Nutrition Industry (cont.)

Laura Becerra, Dari Dambaeva; Jaykumar Panchal
Advisors: Gulhan Bizel, Ph.D.

Qualitative analysis was used to evaluate strategic positioning and product innovation, while quantitative analysis assessed funding patterns and financial trajectories.

Smart Closet

Nihaar Sirikonda
Advisors: Gulhan Bizel, Ph.D.

Many people own more clothes than they can keep track of, yet still struggle to put together outfits every day. Smart Closet is a wardrobe management and outfit planning app designed to solve this problem. Users photograph their garments, and the system automatically identifies each item's category, type, and visual attributes using a deep learning model built on ResNet50 with transfer learning. All clothing data is stored digitally so nothing gets forgotten. When a user wants outfit ideas, they select simple filters like occasion, season, weather, and color preference. The system then searches their wardrobe and builds matching outfit suggestions by comparing clothing features across items. It also alerts users when a newly uploaded item looks nearly identical to something they already own, helping reduce unnecessary purchases and wardrobe clutter. The system is powered by a convolutional neural network trained on large-scale fashion image data, with a React-based interface that displays the digital closet, outfit suggestions, and duplicate alerts in a clean, easy-to-use layout.

Smart Closet aims to make getting dressed faster, reduce clutter, and encourage more mindful and sustainable use of clothing.

Clinical Intelligence System

Fnu Maliha, Sree Udhai Raj
Advisors: Gulhan Bizel, Ph.D.

NLP-Based Clinical Intelligence System
Healthcare systems produce vast amounts of information, much of which is locked in unstructured clinical notes. This project proposes an NLP-based clinical intelligence system to transform these raw narratives into structured data to assist in clinical decision-making. The methodology utilizes text preprocessing, TF-IDF feature extraction, and Named Entity Recognition to identify symptoms, diagnoses, and medications. A central goal of the system is to save practitioner time by consolidating complex patient data into an intuitive interactive dashboard. This dashboard allows practitioners to efficiently view patient history, including the number of visits and medication records. To enhance diagnostic speed, the system employs transformer-based deep learning to predict and display severity levels of a patient's health. Furthermore, it integrates RxNorm for drug standardization and the OpenFDA API for real-time safety and adverse reaction analysis. By utilizing datasets like MIMIC-III, this pipeline converts unstructured text into organized clinical intelligence, directly reducing the manual burden on healthcare providers.

M.A.R.K AI: Agentic AI for Job Hunt

Arsalan Hassan, Samir Sengupta, Afreen Tabasum Begum NIn
Advisors: Gulhan Bizel, Ph.D.

M.A.R.K is an agentic AI system that candidates use to land interview calls for jobs they have applied for. This uses a YAML format structure that collects all the data of the user used to update and create resume / cover letters based on job descriptions. The system will use structured data, natural language processing (NLP), and template-based rendering to create pdf documents that are ready to present and ATS friendly.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

M.A.R.K AI: Agentic AI for Job Hunt (cont.)

This system will collect information and structure it in a YAML format, process it through an AI layer, and use TeX templates to generate pdf documents. This pipeline should create resume and cover letters that align with the job descriptions they select through a click of a button.

Credit Card Fraud Detection

Haritha Renuguntla, Jahnvi Somu
Advisors: Vijay Kumar Reddy Voddi, MSDS

Credit card fraud is a pervasive issue that causes significant financial losses and compromises consumer trust in financial systems. This project focuses on utilizing data mining techniques to develop a predictive model capable of identifying fraudulent credit card transactions. The dataset for this project includes transaction details such as amount, location, time, and user behavior, which will be analyzed to detect patterns indicative of fraud. Machine learning algorithms like decision trees, random forests, and support vector machines (SVM) will be applied to classify transactions as legitimate or fraudulent, with an emphasis on optimizing the model's accuracy.

In addition to the technical aspects of fraud detection, this project will also explore the societal impact of data mining in the context of credit card fraud. The ethical implications surrounding privacy and the potential for data breaches will be addressed, considering the risks associated with anonymized data and the possibility of de-anonymization. Furthermore, the fairness of both the dataset and the model will be assessed to ensure unbiased outcomes. By integrating societal considerations into the project, the goal is to develop a robust, fair, and responsible credit card fraud detection system that can help reduce fraud while protecting privacy and promoting fairness.

Credit Card Fraud Detection using Data Mining Techniques with Societal Impact Analysis

Venkata Kasi Praveen Yarru, James Davis
Advisors: Vijay Kumar Reddy Voddi, MSDS

Credit card fraud is a major financial issue affecting both customers and financial institutions. Fraudulent transactions lead to significant monetary losses and reduce trust in digital payment systems. The goal of this project is to develop a data mining model that can accurately detect fraudulent transactions while minimizing false positives.

Predicting Breast Cancer Diagnosis Using Machine Learning: A Data Driven Approach Using the Wisconsin Diagnostic Dataset

Joseph Liggio, Udumaga Onyeukwu
Advisors: Vijay Kumar Reddy Voddi, MSDS

Breast cancer remains one of the most significant public health challenges in the United States and globally. As emphasized in the National Cancer Plan, "Although there have been tremendous advances in treatment, far too many cancers are diagnosed late, when treatment is often less effective. Therefore, it is important that we develop more effective screening and early detection approaches to find cancer in its earliest and more treatable stages, particularly for cancers that are most deadly."(Detect cancers Early - National Cancer Plan. (2025, April 7). This project aims to contribute to that mission by developing a machine learning model capable of distinguishing between malignant and benign breast tumors using the Breast Cancer Wisconsin (Diagnostic) dataset. The objective of this project is to evaluate multiple classification algorithms and identify the most effective model for supporting early diagnostic decision making.

ABSTRACTS-GRADS

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

AEGIS: A Three-Layer Security Proxy for Healthcare AI Agents Using Confidence-Gated Classification

Andrew Espira, Umut Baris

Advisors: Gulhan Bizel, Ph.D.

As healthcare organizations deploy AI agents for autonomous clinical billing tasks such as prior authorization, claims adjudication, and denial resolution, these agents inherit direct access to protected health information (PHI) and are vulnerable to adversarial manipulation. Benchmarking with SENTINEL, a mid-reasoning audit framework evaluated on 150 healthcare agent scenarios, revealed that 52.3% of security probes reached red tier and agents leaked patient data without detection. AEGIS addresses this gap as a three-layer inline security proxy implemented in Go with Python-trained ML components. Layer 1 applies an ONNX-optimized classifier trained on Tensor Trust and HackAPrompt datasets for sub-50ms input classification. Layer 2 uses a hardened LLM auditor for semantic policy enforcement against HIPAA-aligned rules. Layer 3 performs NER and regex-based output scanning to detect and redact all 18 HIPAA-defined PHI identifiers across six exfiltration channels. The system enforces confidence-gated autonomy: decisions above 0.85 confidence auto-proceed, those between 0.60–0.85 are held for 30-second human review, and those below 0.60 are blocked with escalation, with a 1.5× multiplier tightening thresholds for PHI-touching operations. Evaluated with five-seed cross-validation and bootstrap 95% confidence intervals across 1,000 resamples, AEGIS achieves a 0.25% attack success rate, 0.30% false positive rate on benign clinical queries, 0.00% PHI leak rate across 65 scenarios, and 45.4ms p50 latency overhead — demonstrating that production-grade AI agent security is achievable within the real-time latency constraints of healthcare workflows.

ABSTRACTS- DOCTORAL

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Using Large Language Models to Predict Analyst Questions from Earnings Call Content

Natasha Kapoor

Advisors: Dong Ryeol Lee, PH.D.

Earnings calls are critical moments where companies present their quarterly financial performance and face questions from sell-side and buy-side analysts. These questions often uncover information gaps, highlight concerns, or probe controversial topics, shaping perceptions of the company's health and strategy. While existing research typically analyzes Q&A sessions after the fact—focusing on metrics like response tone, question similarity, or evasive answers—few studies explore predicting analyst questions before they are asked. Such predictions could leverage prepared remarks, historical earnings call patterns, or broader sector events, enabling better preparation for corporate leaders. Moreover, prior studies often isolate the Q&A portion without considering the prepared earnings script, which sets the context for analyst inquiries. This study proposes a novel framework using large language models (LLMs) to anticipate the categories of questions analysts are likely to pose and generate example question texts based on a company's prepared earnings script and historical transcripts. We will construct a comprehensive dataset linking quarterly earnings scripts to their subsequent Q&A sessions across multiple firms and quarters. This dataset will train a hybrid model combining classification (to predict question topics) and generation (to produce candidate questions). The model's ability to accurately forecast question topics could significantly enhance senior management's preparation for earnings calls, helping them address potential concerns proactively. However, low prediction accuracy may leave companies unprepared for unexpected questions, risking negative outcomes in earnings call performance.

This research aims to bridge the gap between prepared disclosures and analyst curiosity, offering a practical tool for corporate communication strategies.

Investigating Microplastic Interactions with Biological Molecules

Andres Zuniga

Advisors: Hugo Guterres, Ph.D.

Polyethylene terephthalate (PET) and bis(2-hydroxyethyl) terephthalate (BHET) are types of microplastics that significantly contribute to environmental pollution. These particles can pass through cell membranes and interact with essential biological molecules, which may have harmful effects on human health. This research investigates the detailed 3D structures of how microplastics (PET and PE) interact with key biological molecules, including the cell membrane and several important proteins like fibrinogen, haptoglobin, and fibronectin. Previous studies have shown that these proteins can bind to microplastics using laboratory techniques such as centrifugation and SDS gel electrophoresis, but their atomic-level interactions remain unclear. Here, we modelled the structures of PET & BHET and performed molecular docking to obtain multiple complex structures of proteins with microplastics. We observed the binding affinities and 3D pairings of ligand and receptor of our structures. Our analysis can help elucidate the atomic-level interactions between microplastics and biological molecules.

ABSTRACTS- DOCTORAL

NOTE: These abstracts have only lightly been edited to preserve students' intellectual property and authenticity.

Measuring ESG Commitment and Credibility of Tech Companies Through NLP

Moustapha Gebai

Advisors: Dong Ryeol Lee, Ph.D.

This study examines the relationship between ESG disclosure quality and corporate performance through natural language processing (NLP) analysis of sustainability reports from three major technology companies: Cisco, Intel, and Microsoft. Covering a 20-year period, this paper constructs two disclosure-based ESG indices at the firm-year level. The first index is the ESG Commitment Index, which captures the extent to which firms articulate targets, implementation plans, governance mechanisms, and progress toward ESG goals. The second index is the ESG Credibility Index, which measures the specificity, verifiability, and reliability of disclosures through quantified targets, timelines, standards, assurance language, governance detail, and operational follow-through. The indices are derived using paragraph-level text analysis of corporate ESG and sustainability reports, applying either rule-based NLP methods or machine-learning-based NLP classification. After constructing the indices, the study compares them with financial performance indicators and relevant ESG-related KPIs to assess whether more credible ESG disclosure is associated with better corporate outcomes. The study contributes a structured and replicable approach to measuring ESG disclosure quality from text rather than relying solely on external ESG ratings. It aims to provide insight into whether corporate ESG language reflects meaningful strategic commitment and whether disclosure quality has measurable links to firm performance.

Comparing Zero-Shot Prompting, Few-Shot Prompting, and Traditional ML Approaches for Email Intent Classification

Chen Zhou

Advisors: Dong Ryeol Lee, Ph.D.

Categorizing high volume of emails for a professional as a day-to-day task is time consuming, especially some are given long contexts or noise but require immediate responses.

Inspired by the powerful capabilities of Large Language Model (LLM), this research plans to compare the AI solutions for the email intent classification task: Zero-shot prompting, Few-Shot prompting, and traditional Machine Learning baseline: TF-IDF with logistic regression, where the first two are LLM involved in.

"Lose Yourself" in the Data

Nnamdi Okoli

Advisors: Dong Ryeol Lee, Ph.D.

Hip-hop lyrics offer a unique window into artistic evolution and cultural narratives, yet traditional analysis methods fail to capture their semantic complexity. This study presents a comprehensive computational framework analyzing 25 years of Eminem's discography through three complementary lenses: semantic similarity using Sentence-BERT embeddings, vocabulary anomaly detection via Isolation Forest, and emotional topography mapping with the NRC Emotion Lexicon. Analyzing 521 songs by Eminem spanning 1996-2020, we identify distinct thematic clusters corresponding to career phases, detect vocabulary outliers marking artistic innovation, and reveal a possible temporal shift. The interactive visualization system makes these insights accessible to researchers and general audiences alike. This work demonstrates how modern NLP techniques can systematically analyze artistic evolution, with applications extending beyond music to any text corpus.

