Session R Abstracts

Investigating Greenwashing in Energy Industry Communications Using Natural Language Processing
Jena Alsup
Mentors: R. Michael Alvarez and Danny Ebanks

The propagation of misinformation related to climate change through corporate rhetoric constitutes a pressing issue that demands further exploration. Greenwashing involves intentionally misleading consumers about a corporation’s own environmental practices, often to appear more eco-friendly or socially responsible than the company actually is. However, there is yet to exist a quantifiable definition of what specifically constitutes greenwashing. We leverage a diverse array of media such as SEC filings, corporate sustainability reports, and tweets from the 15 most influential energy and renewable energy companies. We analyze this cross-sectional time series data using structural topic modeling — a modern natural language processing technique based on unsupervised learning — that tests language distances between renewable energy companies and oil companies. This ultimately allows us to investigate the role of rhetoric in shaping public opinion on climate change.

Identifying and Classifying Misinformation About Cryptocurrencies on Twitter
Jamal Omosun
Mentor: R. Michael Alvarez

Misinformation on Twitter, especially around cryptocurrency, has long been a topic of conversation on the platform. There has been a long history in the crypto community on Twitter of bot accounts, spam, scam, and misinformation. Recently there have been incidents where misinformation has had an effect and actively influenced crypto trading and market prices such as in the case of the currency Ripple. The goal of our project is to train models that can accurately identify misinformation involving cryptocurrencies and classify them. We gathered a large dataset of thousands of tweets and then tested them against a preselected selection of misinformation key phrases and other such identifiers. The results of this analysis give us an idea of how likely any given tweet is to be misinformation, a scam, or from a bot account.

Exploring Misinformation in Twitter’s Musk-era
Siddhartha Shendrikar
Mentors: Michael Alvarez and Jacob Morrier

This project aims to detect the changes in climate misinformation after Elon Musk acquired Twitter. There have been previous studies specific to this goal, however we are going to look deeper into the underlying data and draw conclusions aimed at finding the reasons for the increase in misinformation. We will use hashtags with high misinformation rates to gather the data, measure the differences, and interpret them using a method called causal inference. In the early stages of the project, we gathered hundreds of millions of tweets from the past year that are related to climate. Subsequently, we cleaned up the data, and found the coherence of the topics and started topic modeling to identify trends. The upcoming stages are to create visuals and dive into the realm of misinformation.

Relationships Between Nighttime Light, Economic Variables, and Covid-19 in the United States
Katherine Marquis
Mentor: Pawel W. Janas

Covid-19 has caused over 1 million deaths in the United States by May 2022 (WHO). Nighttime light (NTL), as used as an indicator of human activity, may reflect Covid-19 transmission and severity. This study attempts to assess the relationship between NTL and Covid-19 by assuming that high NTL radiance levels indicate a large amount of human contact to spread Covid-19 and that significant drops in NTL radiance levels correspond to quarantine or deaths due to Covid-19. In order to examine this relationship, NTL monthly composite satellite images from 2018-2023 were collected from NASA, county information was collected from the U.S. Census Bureau, Covid-19 case and death statistics at the county level were provided by Johns Hopkins, Covid-19 vaccination information was collected by the CDC, and economic variables were collected by Opportunity Insights. After preprocessing this data, various machine learning models were applied to the data to analyze relationships.

GPU Optimization for Parameter Estimation in the Attentional Drift Diffusion Model
Jacob Goldman
Mentors: Antonio Rangel and Zeynep Enkavi

The Attentional Drift Diffusion Model (aDDM) is a computational model that describes the impact of visual attention on decision-making. It expands the widely used algorithm of evidence accumulation models to account for how visual fluctuations influence information intake and bias choice. Estimating free parameters that best describe empirical data poses a computational challenge because the model does not have an analytical solution for the likelihood function of choice dependent response times. One way to compute the likelihood of any observed choice
and response time given a set of parameters and fixation patterns is through a sequential updating of the probability of crossing decision boundaries. Likelihoods for each observed data point can be computed independently but together identify the parameters with the maximum likelihood. In this project, we utilized Graphics Processing Units (GPUs) to optimize model fitting performance. GPUs offer increased parallelism and higher memory bandwidth, making them suitable for this task. We developed a CUDA-based software toolbox to perform aDDM model fitting and data simulation, providing significant speedup from previous implementations. This research presents a significant step forward in increasing the broader adoption of models designed to increase understanding of the role of attention in decision-making.

**Using Webcam-based Eye Tracking Technology to Probe Atypical Gaze Behavior in Autism**

Ava Barbano  
**Mentors:** Ralph Adolphs, Na Yeon Kim, and Qianying Wu

Autism spectrum disorder (ASD) is a neurodevelopmental disorder that is commonly associated with social and communicative difficulties. The complex and heterogeneous nature of ASD symptoms has placed challenges in establishing an objective and quantifiable behavioral marker that provides a foundation for further research on the genetic and biological basis. Eye tracking has revealed atypical gaze patterns in ASD; however, conventional desktop-based eye trackers require specialized equipment and in-lab testing. This study instead utilizes Webgazer, a webcam-based eye tracking technology, to study gaze behavior in a large-scale subject population recruited from Prolific, an online study platform. The pilot implementation of this study (N = 96) featured ASD and non-ASD groups viewing naturalistic Zoom conversations featuring human actors and physical objects. All participants completed Autism Quotient and Social Responsiveness Score questionnaires, as measures of autistic traits. Preliminary analysis of the pilot data will be used to inform the final experimental design and implementation of the study, with a focus on the comparison of gaze data, self-reported engagement, demographic information, and questionnaire reports between the ASD and non-ASD groups.

**Understanding Pupil Response Patterns in Autism**

Qianhui Hong  
**Mentors:** Ralph Adolphs, Qianying Wu, and Na Yeon Kim

Autism Spectrum Disorder (ASD) features atypical social and emotional processing, which can be reflected in the pupil dilation response (PDR) to external audio-visual stimuli. This project aimed to quantify how different features in naturalistic movies affect pupil dilation using an eyetracking dataset collected in 20 healthy controls and 13 high-functioning ASD participants. In particular, we extracted changes of the pupil diameter at a high sampling rate (600 Hz) while participants were watching 18 sounded, short YouTube movies. After data preprocessing and normalization, we found that the pupil diameter is affected by various factors in the movie, including lighting, changing of scenes, sound volume, emotion arousal, etc. Compared to the control group, the ASD group in general has greater dilation when the emotion arousal is high and when a new human subject appears. These preliminary findings can thus form the basis of fine-grained future analysis that will comprehensively characterize the importance of different features in inducing atypical pupil dilation patterns in ASD.

**Using Machine Learning to Classify Autism From Oculomotor Features**

Audrey Lai  
**Mentors:** Ralph Adolphs and Na Yeon Kim

Quantifying gaze behavior through eye tracking has unveiled distinct patterns in autism, offering potential for screening and diagnosis. However, it has been difficult to identify reliable visual features that can be used to distinguish autism across different studies. This project investigates whether autism can be classified using oculomotor features obtained over a sufficiently long and diverse set of stimuli, without needing to decompose the stimuli into specific objects such as faces. Participants with autism (N = 13) and control participants (N = 19) completed a series of tasks using a Tobii eye tracker. These tasks consisted of well-validated measures to prompt specific eye movements, as well as naturalistic viewing of images and movies. We established an optimal method for characterizing oculomotor events (e.g., fixations, saccades, and smooth pursuit) through the application of a machine learning-based algorithm. We then compared properties of those events, such as duration, velocity, and acceleration, between the autism and control groups. The findings will advance our understanding of oculomotor signatures in autism, and have the potential to enhance early detection of autism.

**Probing Neural Correlates of Performance-Based Arbitration During Model-Based and Model-Free Control**

Sarita Raghunath  
**Mentors:** John P. O’Doherty and Vincent Man

Arbitration between multiple systems is the key to decision making. Previous studies in model-based (goal-driven) and model-free (habitual) reinforcement learning indicate that meta-level control mechanisms assess the reliability of each system to guide decision makers towards the strategy that predicts rewards more accurately. However, recent work from the lab suggests a novel performance-based arbitration mechanism - unlike reliability-based
models, performance-based arbitration models weigh the system that yields the highest expected reward (performance). We hypothesize that medial prefrontal (mPFC) regions are involved in performance-based strategy-switching due to their role in top-down control. We also investigate the basal ganglia, particularly the globus pallidus, given its role in strategy switching for reliability-based arbitration. To explore this, we apply this new computational model to fMRI data from participants performing a two-step decision-making task, previously used to indicate model-based/model-free decision-making. By utilizing variables from performance-based arbitration in GLM analysis, we aim to better understand the neural correlates for performance-based arbitration over multiple decision-making systems. In the future, we will use these fMRI analysis results to guide electrophysiology work which will investigate the underlying physiological signal in the globus pallidus and basal ganglia.

**Estimation of Computational Models of the Impact of Attention on Simple Choice Using Julia**

Lynn Yang  
Mentors: Antonio Rangel and A. Zeynep Enkavi

Neuroeconomics studies the neurocomputational bases of decision-making. Experiments within this field often include simplified choice scenarios, such as picking one of two snack options where experimenters record choice and response time data. Literature has shown that behavior in such tasks are well-described by a noisy evidence accumulation process that compares values associated with each item, and attention can especially bias this process. Inspired by evidence accumulation models in cognitive psychology, the attentional drift diffusion model (aDDM) formalizes how overt attention, measured by fixation patterns, biases choice and response times. Current methods for estimating parameters of the aDDM are computationally inefficient and can take hours to days. To address this issue, we developed a new toolbox in Julia. By taking advantage of the compiled language and multithreading packages we significantly sped up parameter estimation. Wider adoption of our toolbox can enable researchers to test more complex models of decision-making by significantly reducing the time between building and fitting models.