**Session 3 AI and Machine Learning Abstracts**

|  |  |  |
| --- | --- | --- |
| **Faculty Presenter** | **Abstract Title** | **Abstract** |
| Katherine Weisensee | Building a Nationally Representative Forensic Taphonomy Database | Estimating an accurate time of death is critical when human remains are discovered, but currently available methods are based are small and unrepresentative sample sizes. This research utilizes a mass collaboration data collection model to build a large and representative reference data set tied to existing environmental data. The ultimate goal of the project is to use machine learning methods to develop robust and multifaceted models of decomposition to provide reliable and accurate estimates for time of death in a forensic context. |
| Apparao M. Rao | Data-driven Design of Electrolytes Enabled by Computation, Machine Learning, and Experiments | We aim to develop a multi-pronged theoretical approach combining machine learning (ML) and first-principles calculations to model the experimentally determined 60 electrolyte compositions to study their impact on the lifecycle stability in lithium sulfur batteries (LSBs). The main goal is to use the 60 compositions as a training set to predict the optimal electrolyte composition that will lead to an increased battery lifetime (gauged from the cycling stability). At first, we will directly apply ML to the experimental data for 60 compositions and identify the key descriptors of chemical properties to accelerate the electrolyte discovery and optimization process. However, if we find the current data for 60 compositions (tested for a 200-cycle stability at 7.5 C rate) to be insufficient for training our statistical model, we will expand our training data by including experimental data we have at different C rates (6 per electrolyte composition) and different scan rates (5 per electrolyte composition). The result will be at least 6x5x60 entries in our training data set. By using experiments to verify the predictions from the ML model, this study will yield a fundamental understanding of the electrolyte-dependent behaviors that impact LSBs for energy storage applications. |
| Zhen Li | Multiscale Modeling and Machine Learning for Scale-Bridging | Intrinsic multiscale features in soft matter, functional materials, biophysical processes and beyond lead to significant challenges in scientific research because the relevant physicochemical processes can span several decades or more of spatiotemporal scales. Multiscale modeling and machine learning approach that allow seamlessly coupling of heterogeneous models and data across scales is a key in creating more accurate predictive tools for analysis of these multiscale phenomena. I will present example applications of our multiscale modeling and machine learning methods for complex fluids, viscoelastic materials, and biophysics. |
| Alex Feltus | Discovering Patient-Specific Biomarker Systems Using Artificial Intelligence | In 2021, democratized research computing systems and open-source artificial intelligence techniques have matured to the point that freshman and above can access huge datasets to find solutions to grand challenges through interdisciplinary teams. As an example, the story of the GAN-based Transcriptome State Perturbation Generator (TSPG) developed for patient-centered biomarker discovery will be presented. This story revolves around faculty, students, and staff from CECAS, SCIENCE, and CCIT as well as a colleague with kidney cancer. |
| Nicole Bannister | Teaching AI-related cybersecurity through metaphors and math | This project aims to improve middle school students’ use of security “best-practices” in their day-to-day online activities, while enhancing their fundamental understanding of the underlying security principles and math concepts that drive AI and cybersecurity technologies. There is a critical need to develop such educational materials, particularly for adolescents from minority groups that are underrepresented in STEM because they unduly experience harm from AI-related cybersecurity issues and are less likely to have school-based experiences that can inform their responses. The project team is using a mixed-methods, exploratory research approach to investigate middle school students’ AI-related cybersecurity competencies relative to their mathematics knowledge and behaviors. The results of the initial grounded theory analysis will be featured in this presentation. |
| Carl Ehrett and Lucian Ghita | “Using AI To Understand Shakespeare’s Drama” | Our project is a collaboration between the Watt AI Program (Dr. Carl Ehrett and Dr. Hudson Smith) and the English Department (Prof. Lucian Ghita) that started in the Fall of 2019. Our team includes two research students, Dillon Ranwala (Computer Science) and Aarin Henning (Chemical Engineering). It started as an exploratory study aimed at using AI tools, such as Watson Natural Language Understanding, to analyze emotions, sentiments, and group dynamics in Shakespeare’s drama. The first phase of the project involved the analysis of textual and generic differences among eighteen plays (roughly half of the entire Shakespearean corpus) published both in individual quarto editions during Shakespeare’s lifetime and in the Folio Complete Works of 1623. Using Natural Language Understanding tools and a range of other computing tools (HTML web scraping), we also examined emotions and sentiments associated with various groups of characters (gender, race, ethnicity, class) in order to study the relationship between language, group identity, character dynamics, and the dialectic between self and other. We have currently moved into a new stage of our project in which we are using Labelbox to train our own customized model for classifying compassion and aggression as categories for analyzing patterns and structures of Shakespeare’s plot and character construction while we continue to figure out the implications of these result. |