SAM JOHNSON

sj110.pages.iu.edu sejohnson2020@outlook.com (812) 341-7037

EDUCATION

Indiana University - Bloomington, IN
Bachelor of Science in Data Science, May 2024 **Major GPA: 4.00/4.00**
Master of Science in Data Science, May 2025
Major: Data Science, Luddy School of Informatics

- Kelley School of Business Honors Program
- Hutton Honors College
- Founder's Scholar (cumulative GPA exceeding 3.80)

RESEARCH

Incoherence in Predictive Chess Models (in progress)
UC Berkeley Supervised Program for Alignment Research

- Exploring the influence of reward representation in various RL algorithms on the propensity for predictor-policy incoherence to arise in LLM agents trained to play chess at super-human levels
- Designing and implementing RL training loops built on REINFORCE, proximal policy optimization, and action guidance algorithms to create predictive agents from transformer architecture

Are They What They Claim: A Comprehensive Study of Ordinary Linear Regression Among the Top Machine Learning Libraries in Python
Accepted Paper at 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining

- Authored a comprehensive survey of current implementations of the Least Squares method in popular Python libraries (TensorFlow, PyTorch, scikit-learn, MXNet) to give users actionable information about state of ML
- Conducted original experiments to analyze the runtime across platforms, space requirement, performance over big data, and strength of model implementation of these popular libraries

AReS: An AutoML Regression Service for Data Analytics and Novel Data-centric Visualizations
Accepted Paper at 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining

- Designed and implemented a web service that enables users to automatically build 30+ ML models over a dataset, then provides visual assessments of the models' performances through novel data-centric visualizations
- Composed the white-paper on AReS, motivating the need for such a web service, detailing the designs of novel visualizations, and analyzing AReS's performance

TECHNICAL SKILLS

Programming Languages: Python (*adv*), R (*int*), SQL (*int*), Ruby (*int*)
Web Development: Rails (*int*), HTML & CSS (*int*), JavaScript (*beg*)
Machine Learning Software: TensorFlow, PyTorch, scikit-learn, MXNet, pandas, NumPy
Techniques: object-oriented programming, parallel programming
Miscellaneous: Microsoft Excel, Microsoft Access, Docker, Git

WORK EXPERIENCE

Groundwork – Zionsville, IN May 2023 – Present
Software and Data Engineering Intern

- Developed, tested, and maintained new features of a web application using Ruby on Rails framework to support CRM and lead qualification startup with client-base of 150+ contractors in home-improvement industry
- Established several data analysis pipelines, including an exploration aimed to increase client lead conversion by 20% and a prediction of new pricing structure expected to increase company ARR by upwards of 6%

Luddy School of Informatics, Indiana University – Bloomington, IN August 2022 – May 2023
Undergraduate Instructor

- Instructed students in Introduction to Computers and Programming honors course, covering various topics such as choice, loop, search, sorting, recursive algorithms, object-oriented programming, SQL, and Matplotlib
- Aided students in writing Python scripts to solve complex exercises in homework assignments

Bloomington Tutors – Bloomington, IN August 2021 – November 2021
Tutor for Computers in Business Course

- Employed aptitudes in database management and technical problem solving to help students master database navigation and search topics, with 7 out of 8 students achieving their target grade