Jack C. Yeung

734-929-8513 | jackyeung99@gmail.com | linkedin.com/in/jackcyeung | jackyeung99.github.io

EDUCATION

Indiana University - Bloomington **May 2026**
Master of Science in Data Science, Applied Data Science in Economic Analysis

Indiana University - Bloomington **July 2024**
Bachelor of Science in Informatics, Minor & Cognate in Psychology

Relevant Coursework: Machine Learning, Applied Machine Learning, Applied Algorithms, Engineering Cloud Computing, Performance Analytics, Network Science, Information Infrastructure, Information Representation

TECHNICAL SKILLS

- **Programming Languages:** Python, SQL, Shell, HTML5/CSS, Git
- **Frameworks:** SciKit Learn, Pytorch, Pandas, Network X, NumPy, Flask, SciPy, Scrapy, Selenium, BeautifulSoup, GeoPandas, Pytest, Matplotlib, Seaborn

EXPERIENCE

Center for Complex Networks and Systems Research(CNetS) **June 2023 - February 2025**
Research Assistant *Bloomington, IN*

- Extended an efficient probabilistic ranking method to evaluate competitive player strength in multi-entity competitions such as Formula One races and elections, improving its ability to handle complex group dynamics
- Experimented with the model's predictive performance and convergence rate on real and synthetic datasets demonstrating an increased likelihood of ratings and an average of 14x faster convergence
- Collaborated with a team in weekly meetings to design and refine experiments, Co-authoring a research paper detailing the methodology and findings, which is currently under journal review

PROJECTS

End-to-End Machine Learning On Real World Apartment Data

- Automated a data retrieval and storage pipeline using Scrapy for advanced web scraping to gather, validate, and store real estate data for over 300,00 units into a SQLite relational database
- Conducted exploratory data analysis (EDA) with Geo Pandas for visualizations and spatial mappings of different neighborhoods, providing insights into market analysis and investment decision-making
- Trained and Tuned an XGBoost(Extreme Gradient Boosted Trees) predictive pricing model, incorporating natural language processioning for feature engineering of amenity text data, reducing model error by 40%, for accurate identification of under-priced apartments

Cloud Based Applications

- Implemented a key-value store deployable on Google Cloud Platform (GCP) virtual machines, ensuring efficient storage and retrieval of data on the cloud
- Recreated and optimized a research-backed matrix multiplication schema using Map-Reduce, leveraging GCP serverless functions for parallel computation, achieving an 8x speed improvement to the standard matrix multiplication
- Applied heuristics for efficient virtual machine allocation, minimizing resource wastage and optimizing cloud infrastructure utilization

Portfolio Optimization

- Designed an algorithm to optimize portfolio contributions from different sub-companies and product categories, aligning resource allocation with business objectives and financial constraints
- Projected future revenue using stochastic processes, modeling revenue growth, risk exposure, and profitability under multiple macroeconomic scenarios to support data-driven decision-making
- Simulated and analyzed various optimization strategies to identify the most effective approaches for maximizing revenue and profit margins under different constraints, enhancing financial planning efficiency

Equipment Rental Interface

- Built a full-stack equipment rental management system using Flask, HTML, and MySQL to streamline inventory tracking, allowing employees to easily keep track of rentals and business processes
- Designed a user-friendly front-end interface with Bootstrap, enhanching the UX experience for customers and employees
- Integrated a Python-Flask backend with a MySQL database to enable real-time updates, efficient data storage, and seamless retrieval of rental information