Yahan Yang

yahany@andrew.cmu.edu | (+1) 412-954-8993 | https://www.linkedin.com/in/yahan-yang0221/
https://github.com/yahanyang0221?tab=repositories

EDUCATION

Carnegie Mellon University — Pittsburgh, PA
 Master of Science in Applied Data Science — Aug 2024 - May 2025
Highlighted Coursework: Data Visualization, Data Engineering, Statistical Computing, Time Series, Machine Learning, Applied Linear Models(Statistical Methods), Natural Language Processing, Statistical Methods in Epidemiology, Research Experience (Capstone - Spring 2025)

Peking University — Beijing, China
 Bachelor of Arts in Economics — Sept 2020 - Jun 2024
Peking University Scholarship Recipient: Third Prize Scholarship (2021-2022, 2022-2023), **3.6/4.0** GPA
Highlighted Coursework: Advanced Mathematics, Computer Science, Data Solutions, Finance (Financial Markets, Financial Concepts)

SKILLS

Coding (programming languages): Python, R, SQL, SPSS, STATA.
Technical & Analytical Skills: Data Analytics (Big Data), Machine Learning, NLP, A/B testing, Predictive Models, Causal Inference, Linear Regression, Statistical Modeling, Hypothesis Testing, Risk Management, Project Management, Data Reporting, Testing Cycles (unit testing, system integration testing, stress testing), Business Insights, MS Excel, Version Control (Git, Github), AWS, Azure, PostgreSQL, PySpark, PyTorch.
Data Visualization Software: Streamlit, R (ggplot, shiny), Seaborn, Matplotlib, Tableau, ArcGIS, Microsoft Office, Power BI.
Soft Skills: Stakeholder Management, Verbal Communication Skills, Written Communication, Organizational Skills, Business Intelligence, Technical Communication, Leadership, Technical Documentation, Product Innovation, Detail Oriented, Self-motivated, Time Management (handling multiple projects), Collaboration, Product Development, Client Service, Corporate Planning, Resilience.

PROFESSIONAL EXPERIENCE

Apple, Channel Data Application Team — Beijing, China
Data Science Intern — Oct 2023 - Jun 2024
- Drove $2M revenue growth by leveraging deep-dive data analysis to optimize product offerings, processing over 300,000 entries in SQL databases and improving data accuracy for sales strategy decisions.
- Designed and monitored A/B testing experiments, improving experiment efficiency by 50%. Insights from these experiments directly influenced sales initiatives, improving purchase conversion rates by 15%.
- Developed and automated Tableau dashboards to provide real-time tracking of ongoing experiments, enabling product and sales teams to take data-driven actions that enhanced user engagement and conversion funnel performance.

Apple, Monobrand Team — Beijing, China
Business Data Analysis Intern — Mar - Jun 2023
- Optimized sales channel management through regression analysis and cleaned data from 500+ sales channels, contributing to a 10% increase in forecast accuracy and driving $1.5M in additional revenue.
- Performed in-depth funnel analysis on 3 million+ records using Python and SQL, uncovering actionable insights that improved customer conversion rates by 12% and enhanced user engagement.
- Collaborated with product and engineering managers to identify actionable insights through exploratory analysis of user behavior data, driving improvements to sales strategy execution and contributing to a 5% improvement in customer lifetime value (CLV).

ByteDance, Pico — Beijing, China
Consumer Research Intern — Feb - Apr 2022
- Led extensive market research across 5 countries, compiling user profiles and uncovering key consumer trends that contributed to the launch of two new VR products, boosting product adoption rates by 20%.
- Designed and administered data collection methods (surveys, interviews) to gather actionable insights, which influenced key business development decisions and improved targeted marketing strategies.
- Consolidated research findings into data-driven reports and presented 30+ pages of insights, leading to strategic adjustments in product design and marketing, driving a 15% increase in customer engagement.

Yahan Yang

yahany@andrew.cmu.edu | (+1) 412-954-8993 | https://www.linkedin.com/in/yahan-yang0221/
https://github.com/yahanyang0221?tab=repositories

RELEVANT DATA SCIENCE PROJECTS

Growth Signals   Pittsburgh, PA
Capstone Project   Jan 2025 – Present
- Conducting research on leveraging AI to enhance R&D efficiency by revealing and communicating commercial value using semantic search, retrieval-augmented generation (RAG), and specializing prompt engineering.
- Reimplementing sparse autoencoders (SAE) in PyTorch to analyze business and technology abstracts, focusing on semantic clustering and multimodal data integration.
- Investigating the effectiveness of sparse autoencoders in surfacing core concepts, re-ranking search results, and integrating various embeddings (OpenAI, NV-Embed, Jina CLIP).

Carnegie Mellon University   Pittsburgh, PA
Data Science Club – Datathon Competition   Nov 2024
- Led a team of four data scientists on a project to analyze the environmental impact of artificial intelligence tool adoption within the Financial Services Sector.
- Led Exploratory Data Analysis and developed Principal Component Analysis (PCA) models to visualize and interpret adoption of AI tools by organization within the Financial Service Sector.
- Used machine learning methods including XGBoost to predict carbon emissions related to AI adoption. Documented and presented a proposal to industry sponsors with recommendations aligned with sustainability goals.

College Data Engineering Pipeline   Pittsburgh, PA
Data Engineering Course Project   Nov 2024
- Designed and implemented a PostgreSQL database to store and manage over 30k rows of unstructured college data, including developing a robust schema and performing efficient data ingestion.
- Automated the creation of interactive dashboard reports using Python and Streamlit, enabling dynamic data visualization and streamlined analysis.

LEADERSHIP & ACTIVITIES

Carnegie Mellon University   Pittsburgh, PA
Teaching Assistant for Probability Theory for Computer Scientists   Aug 2024 – Present
- Assisted students with complex computer science concepts and graded assignments for an undergraduate-level course.

Nanyang Technological University HeForShe Club   Singapore
Marketing Head   Aug - Dec 2022
- Organized and promoted workshops aimed at empowering LGBTQ+ communities and fostering leadership among male-identifying students during a study-abroad semester.