Yash Khaitan

Ashoka University, Haryana, India
yash.khaitan_ug2023@ashoka.edu.in — +91 74340 88328

EDUCATION

Ashoka University, Haryana, India — Aug 2023 — Aug 2027
BA. Economics with Research (Minor in Mathematics) — Major & Minor GPA: 4.00/4.00
Deans list in all semesters — Cumulative GPA: 3.97/4.00

SKILLS

Advanced: Python, R, Stata, LaTeX, Git, QGIS
Intermediate: Google Earth Engine, MATLAB, Machine Learning, Deep Learning, SQL, NLP

RESEARCH

In-Progress
The Fiscal Multiplier of Disaster Aid and Banking Implications (with Parush Arora and Derek Tran)

COURSEWORK (*A: received highest grade, O: ongoing*)

Econometrics[A], Machine Learning[A], Linear Algebra[A], Game Theory[A], Growth Economics[A], Real Analysis[O], Spatial Economics[O], Empirical IO[O], Development Economics[O], Statistics[A], Calculus[A]

ACADEMIC EXPERIENCE

Krea University, Prof Rohan Gudibande & Prof Parush Arora — Tamil-Nadu, India
Research Assistant — May 2025 — Present

- Collected and analyzed detailed event-level data on 1,000+ unique land conflict cases across India by web scraping secondary sources and integrating geospatial identifiers.
- Built and deployed a custom NLP pipeline to classify conflicts into predefined categories based on textual features.
- Currently compiling a structured database for descriptive and inferential analysis of land conflicts to facilitate future research into conflict resolution, policy response, and land rights.

Nature Conservation Foundation, Prof Meghna Agarwala and Prof Aalok Thakkar — Karnataka, India
Research Assistant — May 2025 — Present

- Developing high-resolution land cover classification products for Northeast India by integrating multi-sensor (Sentinel, Landsat) remote sensing data with supervised machine learning algorithms, and producing geospatial raster datasets to support quantitative assessments of forest loss and resource change.
- Processing and analyzing multi-temporal satellite imagery to detect and quantify spatial–temporal patterns of deforestation and land-use change.
- Investigating socio-environmental drivers of deforestation using econometric modeling, combining remote sensing–derived indicators with field-based observations and secondary socio-economic data.

Ashoka University, Prof Parush Arora — Haryana, India
Research Assistant — March 2025 — May 2025

- Conducted MATLAB-based simulation experiments to assess Bayesian opinion pooling approaches for inflation density forecasting, contributing to the working paper "Regularized Opinion Pools for Density Forecasts Under a Bayesian-Inspired Framework."
- Implemented Markov Chain Monte Carlo (MCMC) methods to estimate model parameters and perform comparative evaluation across multiple pooling specifications.
- Developed and tested over 50 model configurations by systematically varying regularization techniques (Ridge, Lasso, entropy penalties), prior distributions (Laplace, Dirichlet, truncated normal), and scoring rules (logarithmic, quadratic, spherical) to identify optimal ensemble forecasting strategies.

Ashoka University, Prof Meghna Agarwala — Haryana, India
Research Assistant — June 2024 — December 2024

- Compiled a comprehensive dataset combining district- and grid-level variables such as temperature, precipitation, land cover, and crop patterns using remote sensing to identify drivers of human-elephant conflict.
- Processed and analyzed geospatial data in raster and vector formats using Google Earth Engine, QGIS, Python, and R.

PROJECTS

Forecasting NIFTY Movements Using Economic Variables
Github Link
Developed and compared deep learning models (RNN, LSTM, GRU, CNN) to forecast NIFTY and Bank NIFTY movements, incorporating the impact of domestic and global GDP indicators to enhance prediction accuracy.

Portfolio Optimization Dashboard
Github Link
Developed a Stock Portfolio Optimization Dashboard in Python using Markowitz Modern Portfolio Theory, integrating metrics like Sharpe Ratio, Sortino Ratio, CVaR, and Volatility to create optimal, risk-adjusted portfolios. Built with Streamlit and Plotly, the tool enables real-time analysis, efficient frontier visualization from 10,000 simulations, and risk assessment through VaR/CVaR with benchmark comparisons.

Black-Scholes Pricing Model Visualization Dashboard
Github Link
Developed a Black-Scholes options pricing model to compute theoretical prices and Greeks (Delta, Gamma, Theta, Vega, Rho) for European options using key market inputs. Built an interactive Streamlit-Plotly dashboard to dynamically analyze parameter impacts and visualize pricing sensitivities under varying market conditions.

Heart Disease Risk Prediction
Github Link
Developed a heart disease risk prediction model using BRFSS data, addressing severe class imbalance with SMOTE and selecting 15 key predictors through chi-squared analysis and recursive feature elimination. Optimized Logistic Regression for high recall via Grid Search, and deployed the model through a Streamlit app that provides users with predictive insights based on their health parameters.

CERTIFICATIONS

- **Machine Learning Specialization:** 3-course series covering supervised/unsupervised learning, model evaluation, and feature engineering.
- **Deep Learning Specialization:** 5-course series on neural networks, CNNs, sequence models, and optimization techniques.
- **Harvard CS50W:** covered HTML5, CSS3, Git, SQL, JavaScript, Python/Django, UI/UX, testing, CI/CD, scalability, and security.
- **Forecasting: Principles and Practice:** applied time series analysis using `R` and statistical forecasting models.

INTERNSHIPS EXPERIENCE

Centre for Effective Governance of Indian States (CEGIS) Delhi, India
Public Policy Data Science Intern May 2025 — July 2025

- Led a predictive modeling study (MoU between CEGIS and MoWCD) to forecast height and weight outcomes for children under 5, integrating anthropometric data with remote-sensing-based geospatial variables. Designed a Random Forest–based pipeline, engineered domain-specific features, and authored a whitepaper on methodological approach and policy implications for malnutrition risk assessment.
- Performed statistical and spatial analysis using NFHS-5 and other socio-economic datasets to identify regional disparities, applying Stata, R, and Python for data cleaning, modeling, and visualization.
- Analyzed operational workflows of the Support for Statistical Strengthening (SSS) scheme (MoSPI) through document review and process mapping, developing structured consultation tools and briefs to inform ministry-level decision-making.

Government of Punjab Haryana, India
Strategy Intern January 2024 — April 2025

- Conducted a comprehensive labour market and industry landscape analysis for Punjab, combining qualitative field research with secondary economic datasets to capture regional employment patterns and growth sectors.
- Mapped career pathways, higher education trajectories, and entrepreneurship opportunities to identify skill gaps and sector-specific training needs.
- Informed the design of vocational training curricula aligned with regional economic structures and emerging sectoral demands, ensuring alignment with local employment ecosystems.

Centre for Data Science and Analytics, Ashoka University Haryana, India
Data Engineering Intern December 2024 — January 2025

- Designed and deployed end-to-end data pipelines for Ashoka University's DataLake project, enabling seamless integration of structured (Excel) and unstructured (image) data into a PostgreSQL database with dynamic schema handling.
- Developed a Django-based frontend interface to facilitate the upload and management of Excel and image data, streamlining researcher workflows and improving data submission efficiency.

- Implemented data validation, preprocessing, and visualization workflows to ensure metadata consistency, enhance analytical readiness, and improve data quality, accessibility, and usability.

Jainam Broking Gujarat, India
Financial Analytics Intern May 2024 — July 2024

- Developed algorithmic trading strategies and technical indicators using statistical modeling and time-series analysis in Pine Script on TradingView, automating entry decision-making and improving accuracy.
- Designed and implemented options trading strategies incorporating volatility modeling during the Union Budget period, yielding a 35% return on investment.
- Constructed and refined multiple financial models integrating fundamental and technical analysis with predictive analytics for price forecasting and target estimation.
- Developed an interactive Excel-based dashboard with automated data ingestion, processing, and visualization, improving analytical efficiency by over threefold and reducing manual workload by 70%.

Kashi Weaves Gujarat, India
Analytics Consultant April 2023 — Present

- Applied univariate and multivariate time-series forecasting models (ARIMA, SARIMA) to predict demand patterns, enabling data-driven inventory planning and reducing stock discrepancies in a retail apparel supply chain.
- Implemented machine learning–based clustering and classification techniques on customer anthropometric data to standardize shirt sizing, enhancing fit consistency across product lines and minimizing sizing uncertainty.
- Employed geospatial data analytics and spatial statistical methods to identify high-performing sales regions, informing targeted marketing strategies and optimizing promotional campaign allocation.