



**Rishi Daga**  
**Economics**  
**Indian Institute of Technology Bombay**

**22B3003**  
**B.S.**  
**Gender: Male**  
**DOB: 11/11/2004**

Examination	University	Institute	Year	CPI / %
Graduation	IIT Bombay	IIT Bombay	2026	
Intermediate	CBSE	SBR Maheshwari Vidyapeeth	2022	96.80%
Matriculation	CBSE	SBR Maheshwari Vidyapeeth	2020	98.80%

Pursuing **Minor** in **Artificial Intelligence and Data Science** from **C-MInDS**, IIT Bombay

## SCHOLASTIC ACHIEVEMENTS

- Currently holding **Department Rank 2** out of 39 students in the B.S. Economics Undergraduate programme [2024]
- Attained the pinnacle of academic excellence by achieving an **SPI** of perfect **10/10** in the 4th semester [2024]
- Achieved **Perfect Grade Point (AA)** in **13** courses, demonstrating consistency in academic performance [2024]
- Secured an All-India Rank **1707** in the JEE Advanced Examination amongst over **0.15 million** candidates [2022]
- Secured an All-India Rank **923** in the JEE Mains Examination amongst over **1 million** candidates [2022]

## WORK EXPERIENCE

**Data Science Intern** | Kotak Mahindra Bank

[June 2024 - July 2024]

Project - Creation of Feature Mart and Model Development Pipeline for efficient data processing and predictive modeling

- Leveraged **hyperopt's Bayesian optimization** techniques to efficiently explore the hyperparameter space of **XGBoost** models, built a **function** to output the **iteration results** for **streamlined analysis** and comparison
- Developed function to train **XGBoost** model on customer risk data and evaluate accuracy on **test** and **OOT** datasets
- Implemented **Backward Elimination** technique to streamline **feature selection** and enhancing model efficiency

## KEY PROJECTS

**Cricket win prediction model** | The American Express Campus Challenge 2024

[May 2024 - July 2024]

Qualified for Round 2 of the hackathon and achieved a **top 20** ranking at IIT Bombay

- Engineered **over 25 features** from batsman, bowler, and match data to enhance predictive insights and accuracy
- Implemented diverse **boosting models** including **Gradient Boosting**, **XGBoost** getting **accuracy** score over 90%
- Leveraged the **hyperopt** library for precise **hyperparameter tuning**, resulting in **5%** accuracy boost on test dataset

**Image Caption Generator** | Seasons Of Code | Web and Coding Club, IIT Bombay

[May 2024 - Present]

- Implemented multimodel **encoder-decoder** & **attention** mechanism from **Show Attend and Tell** research paper
- Leveraged an advanced pretrained **VGG19** model to **encode images** into **256-dimensional** feature embeddings
- Trained **LSTM** for text decoding using the **Flickr8k Dataset**, **fine-tuned** last few layers of the **VGG19 CNN**

**A/B Testing at WorldQuant University** | Applied Data Science Lab | World Quant University [May 2024]

- Built **Choropleth Map** to depict the global distribution of **5025 ADSL students** data extracted via **MongoDB**
- Constructed Python class to implement **ETL processes** based on hypothesis and **control and treatment group**
- Designed experiment for **email campaign**, estimated **group size** of **196** required to separate signal from noise
- Analyzed the results using a **chi-square test of independence** to test statistical significance for **0.05 p-value**

**Sudoku Solver** | Deep Learning & Computer Vision | Self Project

[June 2024]

- Extracted **9x9** Sudoku board from input images using **OpenCV**, addressing lighting and potential noise in images
- Implemented a **deep learning** model to recognize and classify digits within the Sudoku grid with accuracy over **98%**
- Integrated **backtracking** algorithm with **digit recognition model** to ensure seamless and accurate puzzle solving

**Automated Fraud Detection** | Convolve | Analytics Club, IIT Bombay

[January 2024]

One in **top 100** teams who qualified for round 2 of a PAN IIT AI/ML Hackathon out of **1000+** teams

- Developed **automated transaction monitoring framework** to detect fraud accounts on basis of their transactions
- Conducted **exploratory data analysis** with **95+** features to derive insights while addressing **class imbalance**
- Applied various ML models such as KNNs, Decision Trees and XGboost classifier to obtain an accuracy of **99.04%**

Semantic Textual Similarity | Self Project

[March 2024]

- Engineered **Semantic Textual Similarity** leveraging **NLTK** and **Gensim** modules for advanced **NLP** techniques
- Training of **Word2Vec** model using **Gensim**, resulting in **document embeddings** capturing semantic relationships
- Achieved similarity score above **0.85** for semantically similar texts and deployed the model using **FASTAPI** endpoint

Volatility Forecasting in India | Applied Data Science Lab | World Quant University

[June 2024]

- Extracted Stock Market data for Ambuja Cement and Suzlon using the **AlphaVantage API** via HTTP requests
- Constructed **SQLRepository** class to **insert**, **store** and **read** structured stock market data from SQL database
- Calculated **asset volatility** for the stocks and built a **GARCH model** for forecasting the volatility of returns
- Used **correlograms** to set **3 day lags** & assessed the model using **AIC & BIC scores** & standardized residuals

Applied Data Science : Industry Experience | Course Project | DS203

[October 2023]

Guide: Professor Vinay Kulkarni | Department Of Mechanical Engineering, IIT Bombay

- Conducted an extensive **EDA**, using rigorous **Outlier detection** and **Data cleaning** on the transformer data
- Applied **one-hot encoding** strategy for efficient representation of categorical variables, enhancing model accuracy
- Employed various **Pre-processing** techniques on data from a chemical factory to carefully select features using **PCA**

Ping Pong Game | Game Development | Self Project

[December 2023]

- Developed a **2-player game** using **Turtle**, implementing **player-controlled paddles** & ball for engaging gameplay
- Guaranteed **smooth ball movement** and **collision detection**, enhancing the **virtual Ping Pong** environment
- Applied the **Python's OOP concepts** for modular design, promoting **code readability** and **maintainability**
- Implemented **dynamic ball speed adjustments** during gameplay, intensifying the challenge as the match progresses

POSITION OF RESPONSIBILITY

Seasons Of Code Project Mentor | Web and Coding Club, IIT Bombay

[May 2024 -Present]

- Mentoring a group of **8** students in developing a **Fashion Recommender System** that analyzes input images to recommend complementary clothing items leveraging advanced **computer vision** and machine learning techniques

TECHNICAL SKILLS

- Programming** Python, C, C++, Git, Github
- Libraries** Numpy, Scipy, Pandas, Matplotlib, Plotly, Seaborn, Scikit-learn, Tkinter, Selenium, BeautifulSoup, TensorFlow, Keras, Pyspark, NLTK, Spacy, Open CV, Hyperopt
- Software and Development** HTML, CSS, Flask, MYSQL, SQLite, MS Excel, MS Powerpoint

COURSES UNDERTAKEN

- Programming:** Introduction to Machine Learning, Computer Programming and Utilization, AI and Data Science, Programming for Data Science, Makerspace, Econometrics\*
- Mathematics:** Game Theory and Economic Analysis\*, Statistics-I, Statistics-II, Calculus, Differential Equations, Linear Algebra, Maths for Economics
- Economics:** International Economics\*, Intermediate Microeconomics, Intermediate Macroeconomics [\*Ongoing]

EXTRACURRICULARS

Competitions	<ul style="list-style-type: none"><li>Secured <b>3rd position</b> in <b>Pan-IIT hackathon “SARCathon”</b> in Alumination</li><li>Developed wireframes using <b>Figma</b> for Swasth Mind App in Product Analytics Hackathon</li><li>Engaged in Corefactor Case Study, in <b>optimising CRM processes</b> for revenue growth</li></ul>
E-Cell	<ul style="list-style-type: none"><li>Oversaw and directed <b>350+</b> individuals in coordination and management of competitions</li><li>Organized <b>IPL auction &amp; “Bid and Build”</b> as part of the E-summit event [2023]</li></ul>
Mentorship	<ul style="list-style-type: none"><li>Mentored <b>10+</b> UG students in <b>Machine Learning, Natural Language Processing</b></li><li>Created reading materials and project named <b>Financial News Sentiment Analysis</b></li></ul>
Social	<ul style="list-style-type: none"><li>Mentored <b>20+</b> students for <b>80+</b> hours as part of the <b>NSS</b> program at IIT Bombay [2023]</li></ul>
Online Courses	<ul style="list-style-type: none"><li><b>Applied Data Science Lab</b> 16 weeks program by <b>World Quant University</b></li><li><b>Machine Learning Specialization</b> three-course series by <b>Andrew Ng</b> on Coursera</li><li><b>Databases and SQL for Data Science with Python</b> course by <b>IBM</b> on Coursera</li><li>Learning <b>Python</b> and its different libraries through Udemy’s <b>100 days of code</b> course</li></ul>