SAMBHAV JAIN | 22GG10038

B.S.(Hons.) in APPLIED GEOLOGY

EDUCATION

Year	Degree/Exam	Institute	CGPA/Marks
2026	4YRS B.S	IIT Kharagpur	7.61 / 10
2020	Class XII CBSE	Vidyagram International School	92.4%
2018	Class X CBSE	Vinayak International School	81.8%

INTERNSHIPS

Data Science Internship | Inside The Box | New Delhi — [May'25 - Jul'25]

- Engineered ML models **(Random Forest, XGBoost)** for outlet demand forecasting across 45+ brands, improving accuracy by **2%**
- Applied NLP pipelines on **50k+** reviews using BERT ebemddings to extract sentiment trends, and insights for marketing optimization
- Automated feature engineering workflows (Python, Pandas, Spark) for demand prediction, reducing manual preprocessing time by **~50%**
- Built predictive dashboards integrating ML outputs with Power BI, enabling managers to monitor KPIs and take descisions

Research Internship | University of Oulu, Finland | Supervisor: Dr. Shuai Li — [Mar'25 - Aug'25]

- Designed spiking neural network models with temporal self-attention for financial time-series forecasting under volatile regimes
- Engineered feature pipelines for large-scale stock datasets, improving robustness of SNN models against noise and market drift
- Benchmarked models against **RNN** and **LSTM** baselines using **RMSE**, Precision/Recall, and Sharpe Ratio, demonstrating consistent gains
- **Drafted research manuscript** titled *"SA-SNN: Spiking Attention-based Neural Network for Optimized Algorithmic Trading"*

PROJECTS

Translator Using Self Built Transformers | NLP | LLM — [May'25 - Jun'25]

- Built Transformer seq2seq model in **PyTorch**, with **multi-head attention** & **positional encoding**, improving BLEU score by **28%** vs RNN
- Applied **BPE tokenization** via SentencePiece on 50k+ corpus, cutting OOV rate by **35%** and improving semantic coherence in translations
- Evaluated with **BLEU, perplexity & cross-entropy loss**, achieving **BLEU 32.4** and **25% lower perplexity** than LSTM-based baselines
- Tuned with **AdamW, LR warmup, dropout & gradient clipping**, reducing training instability by **40%** and accelerating convergence speed

Credit Card Customer Churn Prediction | Self Project — [Mar'25 - Apr'25]

- Performed **EDA** on **10k** records using heatmaps & plots to identify key churn dirvers and customer behaviour patterns
- Engineered **20+ features** (spending ratio, balance utilization, transaction frequency, income ratio), boosting model complexity & signal
- Implemented ML classifiers including **Logistic Regression**, **Random Forest**, & **XGBoost**, **tuning hyperparameters** with **GridSearchCV**
- Achieved **86.2% accuracy**, **0.85 F1-score**, and **0.90 ROC-AUC** with XGBoost, demonstrating superior performance over baseline models

Rubiks Cube Solver | C++ Project — [Dec'24 - Jan'25]

- Designed a 3×3 Rubik's cube solver in C++, modeling cube states & transformations with **OOPs** principles like polymorphism & inheritance
- Implemented **BFS, DFS**, and **IDDFS** algorithms, achieving consistent sub-3s solves on cubes scrambled with 8 random moves
- Optimized solver with **Korf's IDA** using pattern databases & admissible heuristics, delivering **<10s** optimal solves for 13-move scrambles

COMPETITION/CONFERENCE

General Championship Data Analytics | IIT Kharagpur (2024)

- Collaborated in a team to integrate **5+** datasets (orders, media spend, weather, holidays, NPS) into a unified analytics pipeline
- Applied EDA, adstock transformations, Random Forest, Bayesian MMM, RL to optimize marketing budget allocation across **9+** channels
- Achieved **22–62% uplift** in GMV predictions, with SHAP analysis ranking Digital **(38%)** and SEM **(28%)** as the most influential ROI drivers

COURSEWORK INFORMATION

Academic: Programming and Data Structures | Probability and Statistics | Linear Algebra Numerical and Complex Analysis | Advanced Calculus | Partial Differential Equations | Pattern Recognition in Geophysics | Economics | Data Structures and Algorithms - Algozenith | Coursera - Machine Learning Specialisation | Database Management System | Operating Systems | Computer Networks

SKILLS AND EXPERTISE

Programming & Tools: Python, C++, SQL, MySQL, PostgreSQL, MongoDB, Git/GitHub, Jupyter, Colab, Anaconda, VS Code, MS Excel
Frameworks & Libraries: PyTorch, TensorFlow, Keras, Scikit-learn, XGBoost, Pandas, NumPy, Power BI, Tableau, Matplotlib, Seaborn

AWARDS AND ACHIEVEMENTS

- Selected for **Amazon ML Summer School 2025**, out of 20,000+ students across India for advanced ML/DL coursework and training
- Achieved a maximum rating of **1672 (Expert)** on Codeforces under the handle *The_Sambhav*, ranking in top global contests

POSITIONS OF RESPONSIBILITY

Finance Head | Rajasthan Cultural Association | IIT Kharagpur — [May'24 - Apr'25]

- Curated **Traditional Food Fest (2000+ students)** & a **Cultural Night (1000+ audience)**, showcasing Rajasthani cuisine and Indian culture
- Coordinated with **PR, Design &Web portfolios**, and managed a **3L+ budget** for 5 events, ensuring branding and optimal resource use

General Secretary Students' Welfare | Gokhale Hall of Residence — [Nov'22 - Jun'23]

- Mentored **150 freshmen**, supporting seamless integration into institute life and driving high engagement through Orientation 2023
- Managed a **2 Lakh budget** with Counselling Service, ensuring efficient resource allocation and delivering a well-received program

EXTRA CURRICULAR ACTIVITIES

- Secured Global Ranks #275 in Codeforces Round 1033 (Div. 2) and #641 in Round 1022 (Div. 2) among thousands of participants
- Mentored 3 first-year students as a SWG Student Mentor**,** guiding their academic transition and smooth integration into campus life