






Jiyanshu Dhaka

Final Year Undergraduate
Major: Statistics and Data Science
Minors: Computer Science - Machine Learning , Cognitive Sciences

 jiyanshudhaka2003@gmail.com
 Jiyanshu Dhaka |  jiyanshud22@iitk.ac.in
 jiyanshud22 |  +91-9413186425

ACADEMIC QUALIFICATIONS

Year	Degree/Certificate	Institute	CPI/%
2022 - present	Bachelor's	Indian Institute of Technology Kanpur	8.1/10
2021	RBSE (XII)	Disha Delphi Senior Secondary School	97.8%
2018	CBSE (X)	Gurukul International School	94.8%

ACHIEVEMENTS

- **KVPY(SB)** Scholar 2021 with **AIR 28** and **KVPY(SA)** Scholar 2019 with **AIR 832** in CRL conducted by IISc, Bangalore
- Secured **AIR 3846** in **JEE (Advanced) 2022** (0.16M shortlisted) and **AIR 4025** in **JEE (Main) 2022** (1M+ candidates)
- Awarded the **Reliance Foundation UG Scholarship**, a merit-cum-means scholarship granted to 5,000 students nationwide
- Achieved **A*** in **Human Centered Computing & Computational Cognitive Science**, and secured **A/A*** in **14** courses
- Achieved **Rating of 1600+** (**Codeforces Expert**) with **Global Rank 95** in Round 1035 | Profile : **Imvengeance3846**

WORK EXPERIENCE

CHAMPHUNT INC | Machine Learning Internship (Sep'24-Jan'25)

- Built **Hybrid Post Recommender** (content + collaborative filtering), boosting relevance by **28%** & engagement by **3 min**
- Built **Q-Learning friend recommender** using Q-table updates from user interactions, achieving **78%** user acceptance rate
- Built **Cricket Ball Detection Algorithm** using **Contour Detection + YOLOv8, HSV masking & temporal checks**
- Applied **Gaussian blur & morphological operation** to reduce noise; Used **Optical Flow & Kalman Filter** for smoothing
- Automated extraction of **ball-pitch contact frames; classifying deliveries** as yorker, bouncer, etc. to improve analysis

SALTMINE USA INC | Workspace Design Automation Intern | (Feb'25-Jul'25)

- Built **Workspace Stacking algorithm** using **greedy allocation, proportional distribution, & adjacency modeling**
- Generated **85%+ valid** stack plans, with **~ 90% match** to manual outputs, reducing manual workload by (**~3-4 hrs/day**)
- Built grid-based zoning engine using **ILP optimization** to satisfy adjacency, periphery, diagonal, and block-pattern constraints
- Implemented **MCMC + combinatorial optimization** based sampling to optimize **objective functions** across constraints
- Developed zoning web app with **NLP-driven interface** translating custom rules into constraints and visualized grid layouts

RESEARCH EXPERIENCE

Passive Image Forgery Detection | Prof. Nisheeth Srivastava | (Dec'23-Feb'24)

- Implemented **Error Level Analysis** with **HSV contour analysis** to expose tampered regions via pixel-level inconsistencies
- Detected **fake medical scans, X-rays, and morphed reports**, aligning with ongoing **cybersecurity forensics research**
- Classified **507 bonafide** and **210 morph** images as **original or forged**, achieving **80.1%** and **78.7%** accuracy respectively

PROJECTS

Cells | Course Project MTH312 | Prof. Dootika Vats | k | (Feb'25-Mar'25)

- Clustered 10,000 cells into 8 distinct types via **spectral clustering**, outperforming **PCA+CCA** by 79.6% in **ARI** metrics
- Achieved the highest **ARI score of 0.87981** in class by integrating multi-omics data using a **Dual-branch Autoencoder**
- Visualized cell latent space with **UMAP**, revealing separation of **cell type** and overlapping modalities across 2 omics layers

Web Surfing Behavior Analysis | Course Project CGS616 | Prof. Nisheeth Srivastava | (Jan'24-Feb'24)

- Collected **web log data** to model user browsing behavior and applied **LDA** with **dwell time** to quantify topic transitions
- Built **within-day predictive model** estimating likely next-hour topics and visit counts, improving **topic revisit** prediction
- Proposed **stochastic framework** to find **recurrent** and **transitional** websites, predicting future topic switching and revisits

Bayesian Causal Inference in Time Series | Course Project MTH422 | Prof. Arnab Hazra | (Feb'25-Apr'25)

- Built Bayesian **time series** model to estimate advertising causal effects on sales with G-Wishart prior & stationarity constraint
- Designed a **2-stage estimation: EMVS for sparse variable selection** followed by **MCMC sampling** for latent states
- Proposed **Kolmogorov-Smirnov** distance causal estimand, validated via simulations, improving detection of causal impacts

Algo Findev | Finance and Analytics Club, IITK | (May'23-Jul'23)

- Applied financial indicators (**MA, FRAMA, RSI**) to analyze equity trends; Backtested **KC, SO** strategies on 79 NSE days
- Conducted **pair trading** on **AAPL-MSFT (corr = 0.25)**, achieving **1.20 returns** vs **0.85 net** after sensitivity analysis
- Benchmarked **Sharpe ratio**, alpha, beta; showing volatility outperformed momentum by **7.8%**(large-cap), **11.4%**(small-cap)

TECHNICAL SKILLS

Python | R | C | C++ | Tidyverse | ggplot2 | tseries | R Shiny | Rvest | RMarkdown | Quarto | pracma | Profvis | NumPy | Pandas | statsmodels | scikit-learn | TensorFlow | PyTorch | Keras | yfinance | Matplotlib | Seaborn | SQL | MongoDB | RStudio

RELEVANT COURSES

Data Structure & Algorithm	Introduction to Machine Learning*	AI Techniques in Data Mining*
Data Science Lab (I),(II) & (III)*	Statistical Computing & Time Series	Elementary Stochastic Processes(I) & (II)
Bayesian Models & Data Analysis	Introduction to Probability Theory	Matrix Algebra & Linear Estimation*

POSITIONS OF RESPONSIBILITY

Academic Mentor, ICS, IITK	Web Secretary, CWC, IITK	Executive, Stamatics Club, IITK
----------------------------	--------------------------	---------------------------------