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
- Integrated **location match, follower overlap**, & **noise factors** with weight tuning, boosting **personalization** for **5k** users
- 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 **stacking algorithm** using **greedy allocation**, **proportional distribution**, and adjacency modeling for constraints
- 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)*

- Reviewed **IEEE** research on passive forgery detection, studying **demosaicing artifacts**, **JPEG traces**, and **CFA** methods
- 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* | (*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

EMG Classification | *Course Project MTH209* | *Prof. Subhajit Datta* | (*Feb'24–Apr'24*)

- Developed gesture recognition model using **6 Time-Domain features** per channel, classifying **12 gestures** with **66k** records
- Captured **dual-channel sEMG** signals at **100Hz**, normalized with **Min-Max** scaling, and segmented into **1000ms** windows
- Modeled **SVM (RBF kernel)**, **LR (softmax)**, & **kNN (Euclidean)**; optimized hyperparameters via **gradient descent**
- Applied **PCA** to identify top features and reduce dimensionality; Achieved **94.1%** accuracy with **SVM** & **92.7%** with **kNN**

Recommendations | *Course Project CGS616* | *Prof. Nisheeth Srivastava* | (*Feb'24–Mar'24*)

- Processed **MovieLens-100k** dataset (943 users×1682 movies), incorporating user demographics, genres, timestamped ratings
- Implemented **item-based**, **user-based**, & **SVD** models; designed **hybrid recommender** integrating similarity+factorization
- Predicted ratings for unwatched movies using **cosine similarity**–based **collaborative filtering** and **SVD**, boosting relevance

World Demographics | *Course Project MTH208* | *Prof. Dootika Vats* | (*Aug'23–Nov'23*)

- **Scraped demographic indicators** (population, fertility, mortality, GDP) for **115+** countries over **50** years using **R rvest**
- Normalized variables for cross-sectional analysis and built **R Shiny** App, enabling **Data Visualization** in structured manner
- **Data interpretation** using **heatmaps and statistical measures**, identifying **outliers, patterns, trends** and **exceptions**

TECHNICAL SKILLS

C | C++ | Python | R | SQL | HTML | CSS | LaTeX| NumPy | Pandas | Matplotlib | scikit-learn | TensorFlow | PyTorch | statsmodels | tseries | pracma | Tidyverse | ggplot2 | R Shiny | Rcpp | Rvest | RMarkdown | Quarto | Profvis | Plotly | RStudio

RELEVANT COURSES

*$\rightarrow A$, **$\rightarrow A^*$

Data Structure & Algorithm	Introduction to Machine Learning*	AI Techniques in Data Mining*
Data Science Lab (I),(II) & (III)*	Human Centered Computing**	Computational Cognitive Science**
Linear Algebra*	Time Series Analysis	Matrix Algebra & Linear Estimation*