

JAY NARESH DHANWANT

ACADEMIC PROFILE

Degree/Certificate	Institution	Percentage/CGPA	Year
IDD	Chemistry IIT (BHU), Varanasi	8.50	2022
HSC (XII)	Hislop, Nagpur, Maharashtra	82.77	2017
CBSE (X)	Bharti Krishna Vidhya Vihar, Nagpur	91.20	2015

SKILLS

Languages & Tools: Python, C++, MySQL, MS-Excel, HTML, CSS, AWS, Flask, FastAPI

Area of Interest: Data Structures, Algorithms, Machine Learning, Computer Vision, NLP

INTERNSHIP/TRAINING

Data Science Intern | Binhex AI Technologies Pvt. Ltd, Bengaluru May 2021 - July 2021

- Ideated, experimented, & prototyped **invoice extractor** robust to key-value & template variations.
- Developed pipeline with **transfer learning** on **YOLOv5**, **OCR**, **NER (auto-annotation)** & **parsing**.
- Designed **disjoint set union** based rectifier [\[Github\]](#) for OCR & YOLOv5 increasing successful extractions from **60% to 95%+**. Developed a demonstrable prototype using **HTML**, **Flask** & **NGROK**.

Data Science Intern | Flatfolder Ltd, Helsinki, Finland May 2020 - July 2020

- Utilised scraped ad-data for (i) **Selling & Rental modelling** (ii) **Trend Analysis** (2017-2020) (iii) **Unseen apartments' rental & selling price estimation (r₂: 0.96)** for **0.4 Million+** data points.
- Automated ML pipeline for ETL(**MySQL**) clustering (**DBSCAN**) & training of tuned **Catboost**.
- Developed **FlatFolder Score** (personalized report of tenants' financial health). This mutually serves landlord & tenants for effective credit check using tenant's **transaction history (Tink API)**.
- Designed KPIs including **RTI**, **persistence**, **cash flow**, **gradients**, etc. to calculate tenant's score for a rent budget. The solution bifurcates secondary accounts & detects recurrent expenses.

PROJECTS

DoubtBuddy | Founding Team Member

- Created API to **load excel workbook** from vendors directly to **MySQL database** using **Flask**.
- Developed OCR search using **google vision**, **latex-parser**, **text preprocessing**, **SBERT**, and **Approximate Nearest Neighbours** implemented using **FAISS** on **1 Million+** questions.
- Implemented **fraud detection** based on tutor behaviour and student-tutor conversations.
- Automated Image QnA extraction using **MathPix** & **custom trained YoloV5** - saved **1000+** hours.

Pathogen Detection using spectroscopy | Dr. V. Ramanathan

- Designed and trained custom **1D-ResNet** on 60,000 spectra of 30 pathogen stains grouped into 7 treatment classes. Observed **99.1%** treatment accuracy & **82%** stains accuracy.
- **Exposure** - Data Cleaning & Preprocessing, Tensorflow, Bi-LSTMs, ResNet.

Covid-19 Risk Factor Predictor & Pathology report parser | MediPort.in

- Performed **statistical analysis** to find **significant blood parameters** affecting the prognosis under the supervision of Dr. Aman Dev (IMS-BHU). Exposed the service and **OCR-based parser** using Flask.
- Work featured in various news medium including [\[Lokmat\]](#) [\[Times of India\]](#) [\[Jagran\]](#) [\[HealthWire\]](#) etc.

Remote, Satellite office & HQ allotment - Sterlite Power | Havish M. Consulting

- Automated office allotment for **HR workflow** using form data & deployed ML classifier in **Excel**.
- **Exposure** - xlwings, chi2test, pearson correlation, f-classif & permutation based feature importance.

Customer query clustering on CRM data | IT Data Consulting, LLC

- Solved **database redundancies** by predicting **multiple queries** from the same person.
- **Exposure** - Hierarchical clustering, Caverphone, **custom distance metric** [\[Recommendation\]](#).

Relative Grading system | Dr. V. Ramanathan

- Presented drawbacks of current evaluation methods and proposed **clustering-based alternative**.
- **Exposure** - Relative Grading, **DBSCAN**, **K-Means clustering**, **Seaborn**. [\[Kaggle Kernel\]](#)

Social Contact Pattern Estimation | Dr. V. Ramanathan

- Designed susceptible-infected-recovered model with custom loss function using **SciPy** [\[Manuscript\]](#).

Beer Recommendation | AbInBev Mavericks 2.0 Hackathon

- Created time-variant beer-recommendation-system using whole-saler features & purchase history.
- Carried out **RFM** analysis for **cross-sell & upsell**. **Within top 20/2204 teams**. [\[Github\]](#)

POSITION OF RESPONSIBILITY

- **Manager, Technex '20:** Led 8 member team & executed Case Study Event with **1000+** participants.
- **Educator at Unacademy:** Online instructor with **11,400+ views** and **700+ followers**.

T: 9834123803 **E:** jaynaresh.dhanwant@gmail **Address:** 57, Popular So, Nagpur