

Mehul Gupta

Data Scientist @ DBS

📍 Nagda, MP

☎ +91 8103795345

✉ mehulgupta2016154@gmail.com

in www.linkedin.com/in/mehulgupta7991

🌐 mehulgupta2016154-resume-builder-streamlit-app-ajmqjx.streamlit.app/

Summary

Experienced Data Scientist with 4+ years of expertise in developing data-driven solutions for Healthcare and Financial domains. Proficient in solving complex problems using AI, including Generative AI, Reinforcement Learning, NLP, Computer Vision, Graph Analytics, Classification, and Regression. Active AI/Data Science blogger with 750k+ all-time views.

Work experience

Associate Data Scientist

Jan'22 - Present

DBS Bank

- Trained and deployed a **multi-label text classification** model over an imbalanced dataset with ~500 labeled financial articles and 16 classes for generating geographical labels using **semi-supervised learning**—Micro F1 Score: 85%.
- Designed an **Anomaly Detection** system using Streamlit to track monthly usage patterns for BI Tools reducing dashboard failures by 50%.
- **Automated metadata creation** for production tables reducing manual efforts by 100% at <1 sec/table.

TechStack: Pandas, Tensorflow, SQL, Scikit-learn, Streamlit, Airflow, AutoML, CometML, Pyspark

Data Scientist

July'20 - Dec'21

Tata 1mg

- Formulated **Lab Labeler** for information extraction from lab reports by detecting important regions and extracting text sequences digitizing 4k reports daily & already parsed 3 Million historic reports covering 90% of total reports received.
- **Established pipeline** for Lab reports digitization using Airflow & AWS Lambda Function parsing reports at 1 report/sec.
- Ideated **patient summaries** increasing user footfall on the Health Records section of the Tata 1mg app by 25%.

Techstack: Python, Pandas, Scikit-learn, Airflow, Tableau, Cassandra, SQL, Neo4j, Metabase

Data Science Intern

May'19 - June'20

Tata 1mg

- Concocted **algorithms for medicine names extraction** Printed & Handwritten prescription images namely '**C-Cube**' & '**3 Step Filtering**' reducing the time taken for manual verification by 50%. 10% of all orders are completely automated.
- Fine-tuned **NIMA**(Neural Image Assessment) for scoring Prescription images Using TF-Lite, model size reduced to nearly 1/10 of the original size for deployment purposes.

Techstack: Python, Pandas, Keras, OpenCV, Tensorflow

Founder & Editor

May'19 - Ongoing

Data Science in your pocket

A [medium publication](#) and [YouTube](#) channel covering intermediate concepts in Data Science with 50k+ views/month

Publications

- **Author:** [Algorithms for rapid digitization of prescriptions](#), Visual Informatics (Elsevier, 2021), 4 citations
- **Co-Author:** [Attended robotic process automation of prescriptions' digitization](#), Smart Health (Elsevier, 2021), 5 citations
- 100+ [Data Science blogs](#) ranging from Reinforcement Learning, Time Series, GANs, Transformers in NLP, etc.

Relevant Projects

Data Science Portfolio using Streamlit

The portfolio [website](#) was developed using Streamlit with features like graphs, pagination, carousels, etc.

OpenAI gym environment training using reinforcement learning

Applied DQN, REINFORCE algorithm, and Actor-Critic methods to train multiple OpenAI gym environments.

Multiple GAN implementation

- Season translation (Summer-->Winter, Winter-->Summer) using CycleGANs
- Image generation for Fashion dataset using Vanilla GANs, PixelCNNs, and Conditional GANs

Vehicle Number Plate Detection

Vehicle number plate detection & extraction of Registration numbers using YOLOv3 object detection. mAP = 98%

Classification using Few-Shot Learning

Image binary classification with Siamese Network using Contrastive loss. F1 score: 95%

Education

B.Tech.

2016 - 2020

Indian Institute Of Information Technology , Design and Manufacturing Jabalpur

8.1 CPI (CSE)

Senior Secondary

2015 - 2016

Bharatiya Vidya Bhavan's K.D.K.V.M

94.2% (C.B.S.E)

High School

2013 - 2014

Bharatiya Vidya Bhavan's K.D.K.V.M

10 CGPA (C.B.S.E.)

Certificates

- Data Scientist Career Track (Data Camp)
- Data Analyst Career Track
- Statistical Thinking In Python

[Link to Certificates](#)

Achievements

- Top Writer in Artificial Intelligence, Medium.
- 1.8k+ reputation on [stack overflow](#).
- Guest Speaker, [Neo4J Live](#)
- TCS HUMAIN finalist (AI competition organized by TCS, 2019)
- Shikshan Bharati Kulapati K.M. Munshi Award in Mathematics(2014)
- Bharatiya Vidya Bhavan Shri C. Subramaniam Award for excellence in character(2008-2009,2011-2012)
- Certificate of Merit(Proficiency in Co-curricular activities) for Declamation and Extempore
- Guitarist, Drummer, and Kalimba player in free time