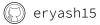
# Yash Gupta

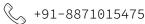
### Data Scientist











#### **EXPERIENCE**

**ADIDAS** | DATA SCIENTIST

May 2023 - Current | Gurgaon, Haryana, India

- → Developed a neural sequence model for in-season demand forecasting, gradient boosting model for pre-season demand forecasting and productionized them using AWS SageMaker.
- → Exploring advancements and methods to enhance forecasts aligned with specific business needs through experimentation.
- → Supported horizontal team building new product that estimates true consumer demand when products are partially or completely out-of-stock.
- → Created a Generative AI PoC that generates product description using a fine-tuned large language model.

#### TATA CONSULTANCY SERVICES | DATA SCIENTIST

May 2021 - May 2023 | Pune, Maharashtra, India

- → Delivered 2 digital transformation projects geared towards high-stake business challenges in SAS Migration, Data Engineering & Machine Learning ensuring a smoother experience for clients & exceeding business targets. (TCS@Nielsen)
- → Successfully rebuilt the sales forecasting model facilitating market demand analysis, cross-sell, up-sell, and offer coupons by reducing the RMSE score beyond 10% compared to existing model.

PAY1 | DATA SCIENCE INTERN

May 2020 - Nov 2020 | Mumbai, Maharashtra

- → Developed face recognition system for Aadhaar & Pan Card based on Convolution Neural Network attaining an accuracy of 92.02% reducing the cost from 3rd party API and human efforts.
- → Real-time data extraction from multiple sports websites to develop a fantasy cricket application.

#### **PROJECTS**

Predict Demand for an Online Classified Advertisement - End to End 7

BIG DATA | DEEP LEARNING | AVITO DEMAND PREDICTION CHALLENGE | KAGGLE

→ Build a complex neural network for regression that handles diverse large data types (numerical, categorical, text, images) attaining RMSE of 0.222.

Predictive Algorithms for Different Subjective Aspects of Questions-Answers ↗

NATURAL LANGUAGE PROCESSING | MACHINE LEARNING | GOOGLE QUEST Q&A LABELING

→ Build combined system of regression models scoring 30 aspects of Stack Exchange Q&A, reaching Spearman's correlation coefficient of 0.8205. Top performers: Random Forest and XGBoost.

#### **PUBLICATION**

Handwritten Signature Verification Using Transfer Learning and Data Augmentation 7 COMPUTER VISION | DEEP LEARNING | B. TECH RESEARCH PROJECT | 2021

Outbreak Trends of Fatality Rate into Coronavirus Disease-2019 Deep Learning 7

TIME SERIES PREDICTION | LSTM | DEEP LEARNING | DATA ANALYSIS | RESEARCH | 2022

## SKILLS

DATA SCIENCE:

- Machine Learning Deep Learning
- Transformers Time-Series Data DEVOPS / MLOPS:
- Jenkins Amazon SageMaker CI/CD LANGUAGES:
- Python SAS C++ SQL FAMILIAR FRAMEWORKS:
- NumPy Pandas PySpark Pytorch
- Scikit-Learn TensorFlow Keras

#### **EDUCATION**

**Indian Institute of Information** Technology - Nagpur, India

**Computer Science & Engineering** 2017 - 2021 | Bachelor of Technology

#### CERTIFICATIONS

**Applied AI Course** Applied Roots (now Scaler)

MLOps with AWS Manifold Al Learning

#### **ACHIEVEMENT**

- 3x Kaggle Expert @eryash15
- Stunning Debut, Trading Sciences, Adidas, 2023
- Gems Award, TCS, 2022
- · Co-founder, Coding Club, IIITN, 2019
- · Co-Head, Technical Feast, Gaming Event, IIITN, 2019
- Mess Representative, Hostel, IIITN, 2018-2019
- 4 star, CodeChef, 2018-2020
- Top 3% IIT-JEE Rank, 2017

Last updated on 01-04-2024.