

# Pragya Jatav

 [pragya42](#) |  [pragyajatav11@gmail.com](mailto:pragyajatav11@gmail.com) |  +91-7522092221

## WORK EXPERIENCE

---

### Blend 360-Data Scientist

Jul 2022 - present

- **Discount optimization (Python, Genetic algorithm)**

- Created a discount optimization framework to create discount strategies for an e-commerce firm.
- Created a demand forecasting model using ridge regression to predict demand and Genetic algorithm to calculate optimum discounts to generate maximum revenue
- Increased the average weekly revenue by 7% and increased stock depletion rate by 5%

- **Personalized email generator (Python, GPT 3.5)**

- Created a tool that develops hyper-personalized marketing emails for users
- Used clustering algorithms to develop hyper-personalized marketing emails to users
- Generated product recommendations and created personalized emails by Retrieval-Augmented Generation using GPT 3.5 turbo
- Recommended products with 95% accuracy and increased email response by 10%

- **Recommendation Engine (SQL, Tensorflow)**

- Led a team of 3 to develop a repeat purchase recommender system for an E-commerce firm
- Implemented session-based recommender system utilizing GRU session encoding, attention mechanisms, and mode prediction for enhanced item recommendations. Deployed the model using GCP.
- Helped developing 3 probabilistic models based on Poisson, and Binomial distributions.
- The deep learning-based recommender system resulted in 15% boost in conversion rates.

- **Customer Segmentation Analysis (DBSCAN Clustering, SQL, Tableau)**

- Developed a customer segmentation model and dashboard to analyse and acquire customers.
- Used RFM, behavioral and demographic data to generate insights. Created customer segments using DBSCAN clustering. Created a Tableau dashboard to showcase insights.

### ICICI Lombard General Insurance Company- Data Scientist

Jul 2020 - Nov 2021

- **Google Analytics Lead prioritization (XGBoost, Azure Data Factory)**

- Created datamart from Google analytics data, conducted EDA to generate insights.
- Evaluated previous model using CSI. Developed an XGBoost classification model for potential customers
- Increased average monthly lead generation by 9 percent

- **Insurance Claim Report Analysis (Tensorflow, Tesseract)**

- Develop a tool that checks all resources have been filed for health insurance claims.
- Developed an object detection model using transfer learning to check if all forms have been filed.
- Developed an OCR model using Tesseract OCR to extract and cross-check details from the forms

## SKILLS

---

Languages Python, R, SQL, Spark

Frameworks Pyspark, StreamLit, Advanced Excel, Google Cloud, Tensorflow, GIT, Tableau, MLFlow

Libraries OpenCV, pandas, numpy, sklearn, nltk

## COURSES AND CERTIFICATIONS

---

GCP Associate Cloud Engineer Certification, Probability and Statistics, Deep Learning by Open AI, Natural language processing, Introduction to data engineering, Data Structures & Algorithms

## EDUCATION

---

2016 - 2020 B.Tech (Materials Science and Engineering) at **Indian Institute of Technology, Kanpur**