PRIYANKA GOEL

DATA SCIENCE | ANALYTICS PROFESSIONAL

CONTACT

gunjan.priyanka29@gmail.com
9838960933
Pune, India
linkedin.com/in/goel-priyanka01/

PROFILE

Worked as Data scientist with 1 year of experience in executing data driven solutions to increase efficiency, accuracy, and utility of internal data processing.

Also worked as EDI Analyst in Retail Domain for Client located at United States for leading international American fashion designer and manufacturer company.

EDUCATION

2015

PRANVEER SINGH INSTITUTE OF TECHNOLOGY [KANPUR, INDIA]

Bachelor of Computer Science & Engineering

TECHNICAL SKILLS

- Programming Languages: Python
- Database: MySQL, PostgreSQL
- Machine Learning (sklearn, pandas, Numpy)
- Data Mining
- Natural Language Processing(nltk)
- Statistics and Probability (statsr, dpylr, scipy)
- Data Analysis and Visualization (Matplotlib, Seaborn, bokeh, plotly)
- Clusters: GCP, AWS, HPC, DGX-1

EXPERIENCE

2017–Present

Data Scientist | Infosys

- Developed a regression model using **sklearn** and **lightgbm** to predict the product prices to online sellers using the product description and quality and **able to achieve 80% of accuracy**.
- Analyzed input parameters to find the correlation with the prices. **Further extract the features** in terms of gender (male/female) to find the better results.
- Used **lightgbm** algorithm to avoid the long training time as clothing brands can add in future.
- Use **bokeh** to create interactive charts for deep understanding of the model.
- **Training time was 58 seconds.**

EDI Analyst | Infosys

- Worked on managing EDI data (purchase order, sales data, inventory data etc.) coming from different trading partners via different communication modes (e.g. AS2, ftp etc.)
- Convert the data into IDOC format to be processed in backend system SAP. Sending order response, advanced shipment notifications and invoices generated by SAP via EDI to customers.
- Created a platform which can support conversions like csv to idoc, idoc to positional file, idoc to csv.
- Creation/maintenance of maps, trading partners, AS2 configurations.

Self-Learning Project

- Developed a classification model to identify whether a given piece of file/software is a malware and able to achieve **the loss of 0.03**.
- Used ML algorithms – KNN, Logistic Regression, Random Forest Classifier and XGBoost Classifier to predict the malware and choose **XGBoost classifier** as the final model.
- Used **Gridsearch** to optimize the hyper parameters of the model to achieve the best result.
- Used **Multivariate analysis** on all the 9 types of malwares.