

# Suraj Prakash

## Data Analyst

+91 7767985884   suraj0prakash@gmail.com   <https://www.linkedin.com/in/surpraka/>   Gurgaon, IN

### SUMMARY

6+ years experienced data analytics professional possessing a prolific track record of driving user growth and boosting user retention. Proficient in data cleaning and data visualization to deliver compelling business value to clients & successfully execute projects. Suraj has actively contributed to three comprehensive projects, resulting in a notable 20% enhancement in operational efficiency in his most recent undertaking. He has adeptly managed datasets comprising lacs of records, particularly excelling in data cleaning tasks.

### KEY SKILLS

- Machine Learning Methodologies • Optimization Techniques • Data Visualization
- Predictive & Statistical Modelling • Sentiment Analysis • Data Analysis
- Project Delivery • Predictive Modelling & Analytics • Programming
- Team Coordination & Leadership • Debugging

### TECHNICAL SKILLS

**Data Visualization:** Tableau, Matplotlib and seaborn

**Languages:** Python, Java

**Cloud Computing:** Azure

**Machine/Deep Learning:** Supervised/Unsupervised Algorithms , ANN, CNN, LSTM/GRU

**Database:** MySQL, Oracle using SQL Developer, MYSQL workbench

**Statistics/ML:** Linear/Logistic Regression, Ensemble Trees, Gradient Boosted trees , Sckit-learn

**SCM:** GitHub

**Scheduler:** Jenkins

**ETL:** IICS Informatica

### EDUCATION

**Post Graduation Certification in Data Science**

Jan '24 - Dec '22

**IIIT Bangalore & upGrad**

Bengaluru, IN

Secured: 3.4/4 CGPA

**B.E. in Electronics and Telecommunications**

Apr '13 - May '17

**Pune University**

Pune, IN

- Secured ~72%

### KEY DATA SCIENCE PROJECTS

- **Eye for Blind | Tech Stack: Python**
  - Pre-Processed FLICKR dataset and evaluated model performance using greedy search, predicting titles based on **BLEU score**.
  - Developed and implement this deep learning project utilizing Python to develop custom model incorporating **CNN,RNN and attention models**.
  - **Data Source:** Kaggle
- **Gesture Recognition | Tech Stack: Python**
  - Created **NN (Conv3D)** model to categorize a series of hand gesture images from video into one of five predefined classes representing different gestures.
  - It uses custom generator to read images into batches which put into conv3D OR CNN+RNN architecture to fit and to evaluate on the **basis of loss**.
  - **Data Source:** Kaggle
- **Robo-Tester | Tech Stack: Python/JAVA/Selenium**
  - Tool to create web automation suites using the manual test cases using multi-class classification based on **logistic regression**.
  - It employees Natural Language processing to identify the entity and intent from English sentence and these elements are classified on trained model which has been evaluated using **Correlation matrix and ROC curve**. All is employed in creation of JAVA code.
  - Accuracy : 70 Percent
  - **Data Source:** Kaggle

- **Credit Fraud Detection | Tech Stack: Python**

- Constructed binary class classification model on credit data (Individuals with credit payment history) which was pre-processed to handle **missing** data, **imbalance** data and **feature selection**.
- Processed data is fit on various models like **Logistic Regression, Random Forest and SVM**. Then, used correlation matrix, F1 score and ROC curve to finalize the model to test.

---

## PROFESSIONAL EXPERIENCE

### Specialist Development ( Band 5.1 )

Present - Mar '21

#### Sunlife Financials

Gurgaon, IN

- Assisted in **migrating schemes** to SLR scheme and gained a thorough understanding of legacy data extraction and transformation for Bancs Oracle database loading.
- Managed the processes of **extracting** data, **patching** data, and applying data **masking** to prepare the data for staging in ETL pipelines.
- Explored **on-shore** to offer production support, enhanced both the product delivery time and client relationships.
- Forged **SQL** scripts and utility tool using **JAVA** to support data cut cleansing to BA team for analysis of output data files.
- Oversaw to secure **backups** and **purging** of the database before go-live and automate the patching of records.
- ETL **mappings** and **MCT** creation to manipulate and transform the schemes date for HK Mandatory Provident Fund pensions. And, mappings creation for data exclusion from staging 2.
- Developed **Java Based Utility** to compare large amount of data between two set of data taken at different timeline. It helped to reduce time taken for data cleansing by 80 percent.

### Associate L2

Mar '21 - Dec '17

#### Publicis Sapient

Gurgaon, IN

- Experienced in managing **Agile projects** and implementing Agile methodologies
- Participated in different **ceremonies** like Sprint Planning, Scrum, Backlog Refinement, and Sprint Retrospective
- Automated Web Services based on **REST architecture** using Rest Assured and Java to switch from SOAP tool. (It reduced time effort from 6 hours to 2.5).
- Creation and execution of Jenkins Pipeline/Job (s) to handle production jobs using **bitbucket SCM** to get code access.
- Mastered Software Development Life Cycle (**SDLC**), Bug Life Cycle(**BLC**) and Agile methodology.
- Constructed **Robo-Tester utility** - using Logistic Regression with NLP via Python which helped to reduce manual effort by 30 percent.

---

## CERTIFICATIONS

- Microsoft Azure Data Scientist | 2022
- Certified Scrum Master | 2021
- Azure Cloud fundamentals Certified (AZ-900/DP-900) | 2021
- Oracle Java Certified | 2019

---

## Languages

- English | 4/5
  - Hindi | 5/5
-