

(+91) 7017070256
Noida, UP
shanugargsg@gmail.com

Shanu Garg

Data Scientist

[linkedin.com/in/shanu-garg](https://www.linkedin.com/in/shanu-garg)

Data Science Associate with 4+ years of rich experience in Natural Language Processing and in Machine Learning. Skilled in modelling, statistical, mathematical, and other analytical techniques. Highly organized, motivated, and diligent with significant background in technical skills.

SKILLS

Languages	Python, C/C++, Javascript
Databases	SQL, PostgreSQL, MongoDB, CouchDB
Libraries	Scikit Learn, NLTK, Spacy, PySpark
Models	BERT, BIO-BERT, Word2Vec, Embedding, NER
WebFrameworks	Flask, FASTAPI, NodeJS
Utilities	Anaconda, Git, Jira, VS Code, Jupyter Notebook, Databricks, Docker, Apache Airflow, PTC Thingworx

TECHNICAL EXPERIENCE

Data Science Associate / AMGEN **AUG 2021 — Present**
Tata Consultancy Services *NOIDA, UP*

- Extracted Data from multiple data sources like Pubmed with the help of automated scripts implemented in Airflow.
- Worked on machine learning models (like Naive Bayes, Logistic Regression, Decision Tree and so on) which enable to assist speciality of Doctor.
- Implemented a Doctor's name matching with Pubmed author's using fuzzy matching algorithms.
- Worked on Geography prediction with the help of multiple data sets.
- Developed pipeline on Databricks with PySpark for data information saved in s3 storage.

Data Science Associate / TCS Safety Literature Surveillance (SLS) **May 2020 – JUL 2021**
Tata Consultancy Services *NOIDA, UP*

- Designed NER models in NLP by using hybrid approach of pre-trained BIOBERT and custom NER's on the abstract part of article (like protein, drugs etc.). Also, Implemented a Disease Prediction of Patient by using different ML algorithms (Logistics, Random Forest, Naive Bayes and so on).
- Designed NER pipeline steps with the help of AIRFLOW (like article extraction, author details etc.)
- Analyzed and designed the Database schema in POSTGRESQL and in MONGODB to save the information of data.
- Also, Developed REST APIs using python flask for multiple usage.

Data Science Associate / TCS PromoRev MLR **APR 2019 — MAR 2020**
Tata Consultancy Services *NOIDA, UP*

- Worked on multiple algorithms according to business rules on text (like reference, authors)
- Developed PDF Layout Analysis, Text Extraction programs and scripts.
- Developed REST APIs using python flask and designed a swagger documentation.
- Also, Analyzed and Designed the machine learning architecture for the project.

IOT Developer / American Multinational Consumer Goods Corporation **AUG 2018 — MAR 2019**
Tata Consultancy Services *NOIDA, UP*

- Developed an IOT Application by using Thingworx.
- Developed an IOT fluctuation Analysis at multiple levels by using Thingworx.
- Also, Developed REST APIs for application using Nodejs.

EDUCATION

Bachelor of Technology, Computer Science and Engineering, GLA University Mathura **GPA – 8.17/10, 2014–2018**
HSC (PCM), Delhi Public School Mathura **CGPA – 8.72/10, 2014**
SSC, Delhi Public School Mathura **CGPA – 8.4/10, 2012**

TECHNICAL CERTIFICATIONS

- Microsoft Certified : Azure Data Scientist (DP 100)
- Machine Learning A-Z™: Hands-On Python & R In Data Science.
- Machine Learning by Stanford University.
- Master SQL for Data Science.
- The Complete Hands-On Introduction to Apache Airflow.
- REST APIs with Flask and Python.

ACTIVITIES

- Secured 94 percentile in GATE 2018.
- Group Representative in TCS Initial Learning Programme (ILP) and the part of TCS Innovations.
- Selected for Technical Grade in Indian Air force (IAF).