Shivani Sampat Raut

IN 46202 | shivanisraut@gmail.com | 463-256-5469 | https://github.com/shivaniRaut
https://www.linkedin.com/in/shivani-raut-2233b4219

EDUCATION

- ***Master of Science in Applied Data Science -*** Indiana University, USA ***(Dec 2023)***
- ***Bachelor of Engineering in Computer Science -*** Pune University, India ***(May 2018)***

TECHNICAL SKILLS

- **Tool**: Visual Studio code, Jupyter, Anaconda, Dataiku, WinSCP, Putty, Oracle SQL Developer, Tableau, PowerBI, Sterling, Postman, Eclipse IDE, Azure Devops, Azure Databricks, Github
- **Query Language**: SQL, MySQL, MongoDB, Oracle SQL, JSON file, Teradata SQL, Hive QL, Spark SQL
- **Programming Languages**: Python, R, Unix Shell scripting, C++, Java, C, C#
- **Data science**: Statistics, SciKit-Learn, TensorFlow, PyTorch, SciPy, Seaborn, BioPython, Pandas, PySpark, Numpy, matplotlib, Spacy, NLTK, NLP, Deep Learning, Time Series, ISLR, ARMA, ARIMA, Apache Airflow, Facebook Prophet
- **Big Data/Cloud**: Hadoop, HDFS, Hive, Spark, Azure Databricks, Azure Dev-ops, Azure ML-ops
- **Certification**: MS Azure Associate Data Scientist, MS Azure Associate Data Engineer, MS Azure Fundamentals, SAFe Agile practitioner

WORK EXPERIENCE

Geodis Logistics LLC., *(Full time) Tennessee, USA* ***(May 2023 – present)***
IT Intern *(Integration Team)-* ***Data Scientist***

- New **database modeling** and **data extraction** methods for **Raptor UI**. Technologies and skills include **cloud computing**, **Data Analytics**, Machine Learning, Cloud Computing, **Data Science**, and Data Engineering. **Predicting** number of transactions for next hours using **Time Series ARIMA/SARIMA** and **Prophet Model**.

Indiana University, *Indianapolis, USA*

1. ***Research Assistant*** *to Dr. Khomtchouk Bohdan (Part time)* ***(May 2023 – present)***
   - Creating **Question Answering Large Language Model** for **clinical guidelines** taking data collected from AHA sources.
   - Developing **hinge loss function** in **Generative adversarial networks** model for GRAPE: Genomic Regional Assessment of Peak Elements. Used **Neural network**, **Deep Learning**, **Machine Learning**, **Natural Language Processing (NLP)**
2. ***Research Assistant*** *to Dr. Jiang Ming (Part time)* ***(Nov 2022 – May 2023)***
   - Research on improving the quality of **Knowledge graph** using various techniques. Used **BERT**, **GPT-3**, **ALBERTA**, **ROBERTA**.
3. ***Adjunct Professor*** *(Part time)* ***(Aug 2022 – Nov 2022)***
   - INFO-B 211 Information Infrastructure II course covering Data Analysis and Bioinformatics libraries of python.

Atos Syntel pvt. Ltd., : Data Scientist *(Full time) India* ***(Jul 2018 – Dec 2021)***
Client: FedEx pvt. ltd. *(Remote)*

- Designed and developed complex data pipelines using ***Numpy, Pandas, PySpark, NLTK, Spacy etc. python libraries in the Dataiku tool.***
- Maintained their data quality using several ***QA test cases***, to support the rapidly growing business needs of client FedEx. Also, Analyzed impact on accuracy by scoring the data sources.
- ***Optimized*** up to ***2 hours*** of execution time and ***scaled up*** the Data flows and ***machine learning pipelines*** using ***Spark*** and ***HIVE***. For that, used ***incremental data*** loading/ingestion.
- ***Productionized*** and ***Automated*** Machine Learning model flows using ***Dataiku, Apache Airflow, Abinitio tools*** and ***Azure MLops cloud*** platform.

PROJECTS UNDERTAKEN

FedEx ***(Sep 2018 – Dec 2021)***

- Used NLP and Various Data flows to predict HS(Harmonic System) code as part of HS-Search Application.
- Using Classification and clustering Machine learning algorithms, predicted caging of the shipments at customs and True/False hits of ***Restricted Party Screening (RPS)*** system.
- Used Hadoop, Spark framework to ***extract*** and store the data through Dataiku tool. Used SQL(Teradata, Oracle), Hive and Python for ***transforming*** the data. ***Loaded*** the data into scaled up ML model flows after ***feature engineering*** using python. ***Validated*** the predictions using QA test cases.
- Achieved 2 hours of execution time optimization. For that, implemented ***incremental*** data ingestion based on the date and ***partitioned*** the data for ***parallelism***.

INDUSTRIAL TRAININGS UNDERTAKEN

Atos Syntel pvt. ltd.

- Training covered the following technologies: Azure Fundamentals, Azure Synapse, Azure Cosmos-DB and Big Data Fundamentals such as Hadoop ecosystem, Spark framework.

ACADEMIC PROJECTS UNDERTAKEN

Microcontroller Based Intelligent Traffic Signal Light Control System - *Aug 2017 - Feb 2018*

- Successfully developed IOT(Internet of Things) based system for detecting traffic density in particular lane and assigning traffic lights accordingly.
- Using the microcontroller's computing power we were able to calculate correct traffic signal light by accumulating IR sensor's signals indicating traffic density on vehicular lanes.
- Integrated the hardware system of Intelligent Traffic signal lights to the front end UI of Java JSP pages and created a website for end users to monitor the traffic.
- This Intelligent Traffic signal lights calculating system generated real time results on real time data and could be monitored from UI present on web pages.

RESEARCH PAPERS Published

- Microcontroller Based Intelligent Traffic Signal Light Control System
  IJRASET (International Journal for Research in Applied Science & Engineering Technology)
  e-ISSN: 2321-9653, Volume 5, Issue 12 (Dec. 2017), Paper ID: IJRASET12195

EXTRACURRICULAR

Throughout the years, I have engaged in various activities and competitions. In 2010, I won a consolation prize in a drawing competition, and in 2009, I participated in a science day speech competition. In 2016, I took part in a coding competition and volunteered for the organization of the college festival 'Cynosure'. During my college years, I visited an orphanage school, organized activities for the children, and participated in a project competition. I also volunteered in community service events, such as cleaning historical forts and assisting rice farmers. In 2019, I received recognition and awards from Atos Syntel organization for my dedicated work.