

TECHNICAL SKILLS

Languages and Databases: Python, Java, R, C++, MATLAB, HTML, XML, CSS, JavaScript, MySQL, MongoDB

Libraries: Scikit-Learn, Pandas, SciPy, Flask, Selenium, Scrapy, TensorFlow, Keras, PyTorch, Matplotlib, Seaborn, Shiny, ggplot, React

Tools: Jupyter Notebook, RStudio, Pentaho, Tableau, PowerBI, Kubernetes, CATIA, ANSYS, AutoCAD, SolidWorks, ROS, Git

Big Data Technologies: Apache Spark, AWS, Pentaho, SAS, ETL, ELT

EDUCATION

Master of Science in Computer Science, Minor in Big Data Analytics | Rochester Institute of Technology 12/2022

Rochester, NY **GPA: 3.61/4**

Bachelor of Engineering in Mechatronics, Minor in Robotics | Manipal Institute of Technology 07/2019

Manipal, India **GPA: 3.6/4**

EXPERIENCE

Machine Learning Research Intern | MVSR Engineering College 05/2019 – 07/2019

Hyderabad, India

- **Analyzed and processed** the raw data from the automobile industry corpus
- Developed an Intelligent system to monitor the conditions of the vehicle's subsystem
- Enabled continuous learning from the Intelligent System using **Remaining Useful Life (RUL)** prediction, **Deviation Detection**, and **supervised classification**. Embedded the sensors and process units to study the system's robustness

CERTIFICATIONS

- **Advanced Graduate Certificate in Big Data Analytics | Rochester Institute of Technology**
- **Advanced Certification in Machine Learning | International Institute of Information Technology, Hyderabad**
- **Advanced Data Science with IBM Specialization | Coursera**
- **TensorFlow developer DeepLearning.AI Specialization | Coursera**

PROJECTS

Drug Design Using Machine Learning

- Extracted untapped bioactivity data that is publicly available and interpreted data using the **NumPy** model to perform the necessary calculations and computations
- Implemented **Proteochemometrics (PCM)** modeling to come up with new drug compositions based on how similar compounds bind to similar targets
- Performed **Support Vector Machines**, **Deep-Learning**, and **Random Forest** methods to prepare the metrics of the system built

Data Warehouse (Dimensional) Development

- Designed a Dimensional model with **Data Marts** to help the company manage financial performance amongst sister companies
- Cleaned, extracted, mapped, and populated the Data Mart using **Pentaho** and **SAS**
- Automated the process of measuring the performance of aggregated data marts using facts from the fact table

YouTube Data Analysis

- Categorized videos based on comments and identified factors that affect the popularity index
- Built **Data lake** from scratch using **Amazon S3** to organize data and used **AWS Glue** crawler and **Lambda** to understand how the data is built and ran **ETL** on top of it for transformations where required
- Used **Amazon Athena** to understand and analyze the data. Developed a dashboard of multiple charts and graphs on **Amazon QuickSight** to visualize and understand the results

Employee Salary Estimator

- **Scraped** job postings from Glassdoor using **Selenium** to extract features and performed data cleaning using **Pandas** and **NLTK**
- Performed feature engineering along with some **Exploratory Data Analysis** and combined text and numeric features using **FeatureUnion**
- Found that the **Random Forest algorithm** performed the best with a Mean Absolute Error of 11.22

Storing, Managing, and Analyzing Web APIs

- Developed a web application using **Flask**, **MongoDB compass**, and **HTML**
- Parsed and cleaned the data and loaded it to MongoDB compass
- Created the application to query the database from the server side by enabling varieties of filters using **pymongo**