Sanket Sud

Dublin, Ireland | +353 (0) 892558980 | sanketsud21@gmail.com

Social: https://linkedin.com/in/sanket-sud

Profile:

3+ years of experience in developing python applications and tools as data engineer. Expertise in handling API services and building microservices. Implementing machine learning model and data mining techniques. Automation using python scripting.
Languages: English (Fluent – verbal & written), Marathi (Native), Hindi.

Technology Skills:

• **Programming**: Python, C, JavaScript, SQL.
• **Framework**: Django, NodeJS, ExpressJs.
• **Operating Systems**: Red Hat Enterprise Linux 8(RHEL), Ubuntu 18.04 LTS, Windows 10.
• **Database:** MySQL, Postgres, MongoDB.
• **Cloud:** Amazon web services (AWS)-EC2 instance, OpenStack.
• **Python Libraries:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-Learn, Beautiful Soup, Requests, TensorFlow, Keras, JSON.
• **Big Data:** HDFS, MapReduce, Apache Spark.
• **Other tools/technologies**: Git, Anaconda, Jupyter Notebook, Ansible, Visual Studio Code.

Work Experience:

Infosys, India | Senior System Engineer   **November 2018 –December 2021**

• Handled CSV file which contains vendor data and classification records also communicated with client for data understanding.
• Experienced with JavaScript to handle web applications also collaborated with the team for business model.
• Used python for implementing supervised learning models to classify the data according to the vendor.
• Created rule-based classification for 10 million records based on MySQL.
• Implemented Linear classification and Naive Bayes classification for data of size 2 TB.
• Stored and handled metadata of the project in PostgreSQL and performed ETL operations on the data.
• Stored and accessed data into Hadoop creating Spark Connection.
• Experienced with MYSQL and Postgres for storing and data transformation.
• Experienced in Django and NodeJS frameworks and microservices.
• Performed Data Mining on vendor's Data, created charts using Tableau, and contributing to Business decisions.
• Experienced with Docker on the AWS instance for deployment and working with installation and delivery.
• Scripted various installation processes and working with CI/CD pipeline for deployment using python.

Academic Experience:

Query Cost Forecasting using Machine Learning
• Research on Finding Time Required and Total bytes used for executing SQL query statement using the Machine learning
• Collected BigQuery metadata using Python Script and storing it into MySQL for execution.
• Performed Data Mining on the metadata using various matplotlib, and seaborn visualization libraries.
• Implemented the K-Nearest Neighbour algorithm and model tunning technique.
• Creating a Machine learning model for predicting the cost of the query giving 78 % accuracy for KNN model.

Certifications:

Python For Data Science and Machine Learning (Udemy).
MongoDB-The Complete Developer Guide (Udemy).
The Complete NodeJS Developer (Udemy).

Other Skills:

• Data Structures, OOP Concepts and Data Engineer concept
• Basics of XML, Golang, C# and visual basic .NET
• Brief idea on Software testing tools like Selenium
• Leadership experience and Teamwork skills
• Networking concepts
• Using tool Jenkins.
• Research oriented work
• Documentation of work
• Object Oriented Programming
• Basic knowledge of terraform, Kafka, ETL, sns, shell scripting, cloud computing
• Jira, GitHub, hive, NoSQL, AI, Scala, Elasticsearch, airflow, bash, relational databases.
• AGILE, SDLC and DevOps
• Presentation and speech
• Swift action on errors and issues
• Inquisitive, hardworking and problem-solving perspective

Education:

Masters in Data Analytics| National College of Ireland, Dublin
                                                    January 2022 - January 2023
• Modules: Database and Analytics Programming, Data Mining and Machine Learning, Data Intensive Architecture, Scalable System programming, Statistics for Data Analysis, Modelling Simulation and Optimization.

Bachelor of Engineering in Computer Science | University of Mumbai
                                                    June 2014 – June 2018
• Modules: C Programming, Algorithm and Data Structure, Parallel and Distributed Systems, Software Engineering, Microprocessors, Operating Systems, Computer Networks, Mobile Computing, Distributed Databases, Object Oriented Programming.