

Skand Mani

Software Engineer - II

63D/1 Allenganj, Prayagraj - 211002, Uttar Pradesh | +918935056063 | skandmani19@gmail.com

Summary

- A competent professional with 4+ years of experience in Data Analysis, Data mining, Data Acquisition, Data Validation, Data Visualization and Web Scraping. Adept in statistical programming language like Python including Big Data technologies like Hadoop, Hive and Spark.
- Expertise in transforming business requirements into analytical models, designing algorithms, building models, developing data mining and reporting solutions that scale across a massive volume of structured and unstructured data.
- Skilled in performing data parsing, data manipulation and data preparation with methods including describe data contents, regex, pandas, BeautifulSoup, Selenium web driver.
- Working knowledge of developing CI/CD pipelines for build, test and deployment
- Good industry knowledge, analytical & problem-solving skills and ability to work well within a team as well as an individual.

Technical Skills

- Python, Hive and MS-Excel.
- Hadoop, Spark and Pyspark
- Web Scraping.
- Scrapy framework and web crawling.
- Regular expressions.
- Xpath and CSS selectors.
- Familiar with JSON, Git and Jira.
- Working experience of Jenkins, Maven and Docker.
- AWS - EC2

Work experience

Software Engineer - II

May 28, 2019 - Present

CareerBuilder India

- Analyse Hadoop clusters using big data analytic tools including Hive and Pyspark.
- Importing data from multiple sources such and converting into Spark RDD.
- Extracting features from data sets using Spark SQL
- Performed web scrapping using Python, Scrapy and Selenium to extract required informations from multiple websites to increase skills coverage for our internal database CaroteneSkills.
- Conduct in-depth research on Hive and Spark datasets to analyse partitioned and bucketed data.
- Performing QA on the data extracted, transformed and exported to excel.
- Developed Hive queries that compared new incoming data against historic data. Built tables in Hive to store large volumes of data.
- Worked on data that was a combination of unstructured and structured data from multiple sources and automated the cleaning using Python scripts.

Professional competencies:

- Communication.
- Problem solving.
- Flexibility and adaptability
- Avid learner
- People Management.

Research Analyst

January 09,2017 - May 25,2019

GlobalLogic India Pvt. Ltd.

- Member of a team which is responsible for maintaining data ingestion to Knowledge Graph either through manual or automated pipelines .
- Using Google tool to evaluate changes in Google Search and prevent bad data from reaching the search engine.
- Performed Exploratory Data Analysis, trying to find trends and clusters.Used Python to preprocess data and attempt to find insights.
- Extensively used Python's package like pandas,Beautiful Soup,regex(re).
- Write Python scripts to parse JSON documents and load the data in database.
- Mapping flow of trade cycle data from source to target and documenting the same.

Additional Information:

- Worked as a Freelancer Data Analyst for netdrivered.com.
- Knowledge of building the CI/CD process from scratch.
- Developed scripts for build, deployment, maintenance and related tasks using Jenkins, Docker, Maven, Python and Bash.
- Experienced in Automating, Configuring and deploying instances on AWS-EC2.

Education

- **B.B.D National Institute Of Technology & Management,Lucknow,U.P | B.Tech(I.T) | 2016**
- **J. T. Golden Jubilee,Prayagraj,U.P | Intermediate | 2011**
- **J. T. Golden Jubilee,Prayagraj,U.P | High School | 2009**