**Pratibha Patro, MS**

(669) 255 - 8486 | pratibhapatro007@gmail.com

LinkedIn: <https://www.linkedin.com/in/pratibhapatro/>

Data Engineer/ Python Developer

*Key player in Mission-Critical Data Strategies*

Hands-on 8 years of experience and results-driven senior customer-facing product management and deployment professional with the highly- coveted blend of business savvy and technical proficiency to best address business specific needs in driving growth and differentiations in fiercely competitive markets.

**Summary of Accomplishment**

Proven capacity to deliver complex Data engineering task on-time and within budget despite shifting priorities, demanding expectations, ambiguity, and stringent deadlines

* + **Druid Database and Data pipeline with ETL**
  + **Realtime Ingestion with Kafka from AWS**
  + **Python for Data Science (Pandas, NumPy)**
  + **Python for REST, Django and Flask**
  + **Advanced SQL for Data analytics and data science**
  + **MySQL, PostgreSQL, Oracle and SQL server**

**Summary of Experiences**

* Over 8 years of hands on Python coding experience in web application development.
* Having 4 years of experience in writing SQL queries to get the data and analyze the data.
* Worked on Data Pipeline and full stack developer with python.
* Strong experience using Web Services, REST API’s using FLASK in python.
* Experience building data pipelines from ingestion to delivery in data applications.
* Developed the ETL data pipelines from Hadoop deep storage to Druid developed with Python code.
* Design and develop ETL UI to get the process and environment to ingest data daily basis into druid and meet the customer’s business processes. Develop customer-centric API integration of python.
* Design and Develop data pipelines for real-time data processing using Kafka,Pyspark
* Analyzing the data and the performance using SQL analytical queries. Ran the hive commands in Hadoop application server to explore about the data.
* Experience in creating initial website prototype from Django skeleton and building out Views, Templates using CSS, JavaScript for the whole site following Django MVC architecture.
* Familiar with UNIX / Linux internals, basic cryptography & security. Experience in using Design Patterns such as MVC and frameworks such as Django.
* Experience in working with AWS (Amazon Web Services) cloud platform.
* Experience in working with AWS Services like Glue, Lambda,S3.
* Experience in Data Science and Machine learning tools: Python (Pandas, Numpy, Scikit-learn, SciPy), Tableau, R, SAS, MS Excel (Macros, VLOOKUP, Pivot table, Solver, Descriptive Statistics).
* Experience in Machine learning algorithms- Customer Churn Prediction, Predictive Analytics, Sentiment Analysis, A/B testing, Principal Component Analysis, Logistic Classifiers, Null Hypothesis, Simulations and Optimizations, Linear &logistic Predictions, Clustering, Time Series Analysis, Statistical Modeling, Forecasting.
* Excellent Programming, debugging, problem-solving, optimization and testing skills.

**Professional Experience**

**Apple, Sunnyvale, US**

Python Developer Data Engineer March 2019 – Till Date

* Worked as a data engineer in druid platform team of Apple.
* Designing and developing code, scripts, and data pipelines that leverage structured and unstructured data integrated from multiple sources.
* Conducting feasibility of the requirements, creating logical and physical detail level design of web service and interfaces. This will include attending design reviews meetings ensuring that business requirements are being met and all design documents are developed as per requirements.
* Created API’s to consume data from Flask REST endpoints and cleansing the data and store it in druid database from Hadoop deep storage. Worked on configuration from Hadoop deep storage to druid in all properties in Unix environment and integration to all downstream applications.
* Developed the frontend dashboard using Angular Js and HTML and CSS for business querying for further usage and analysis.
* Involve end to end integration Testing in all stages like Unit Testing, Integration Testing, and User Acceptance Testing
* Developing a fully automated continuous integration system using Git and deploy by CI/CD pipeline across different environments like dev, test, Production environment.
* Design and build scalable data pipelines to ingest, translate, and analyze large sets of data. Design, build, and debug Hadoop Hive based data pipelines. Developed the ETL data pipelines using Python code.
* Automate the ETL framework to get the metadata by enqueueing the batch from upstream. Set up all the required configuration for running the map reduce job into druid and deep storage for dashboard users.
* Created and automate the real-time data pipelines from our storage (AWS S3) to application(druid). Using AWS s3 storage to get the real data into a Kafka platform using Kafka topic. Using Kafka producer to send events to druid to consume data for real time ingestion. Kafka Consumer consumes the message and running the ingestion to ingest the segment into druid.
* Developed the database schemas database using MySQL, SQL Alchemy and written several queries and python API’s to extract data from database
* Analyzing the data loading into system and creating the model. Visualize the data and query performance of dashboard using tableau, clarity.
* Creating application access protocol and training documentation for the developed application, track any defects in the production environment ensure defect free implementation and deliver L2 support for production issues related to infrastructure, development, and interfaces

**California College of Arts, San Francisco, US**

Python Developer November 2018- February 2019

* Involved in the requirement study, design and development. Provided the reports to clients based on their needs.
* Designed the Web Application Using Python on Django Web Framework pattern to make it extensible and flexible for California College of Arts web portal and accounts
* Used the Model View Controller (MVC) framework to build modular and maintainable applications. Build script and templates using python and Django.
* Build all database mapping classes using Django models with the help of ORM like SQL Alchemy.
* Used data structures like directories, tuples, object-oriented class-based inheritance features for making.
* Wrote and executed various MYSQL database queries from python using Python-MySQL connector and MySQL dB package. Created database using MySQL, wrote several queries and Django API’s to extract data from the database.
* Deployment and Build of various environments including Linux and UNIX
* Worked with JSON based REST Web services.
* Worked on automation, setup, and administration of build and deployment tools
* Developed a fully automated continuous integration system using Git, SQL Server and custom tools developed in Python and Bash.
* Debug the issues and provide the solution to customers and business.
* Created database using MySQL, wrote several queries and Django API’s to extract data from the database.
* Generated various report using Python Report lab and sent to Business users to improve their decision  
  making.

**Cisco, Bangalore, India**

Python Developer Apr 2014 – Oct 2016

* Provided input into developing and modifying systems to meet client needs and develop business specifications to support these modifications.
* Deployment and Build of various environments including Linux and UNIX
* Built python module, to extract data from FTP offloads in JSON format.
* Developed ETL scripts in Python for SQL query execution and Writing SQL scripts and validating the data. Write unit testing for testing purposes.
* Created database using MySQL, wrote several queries and Django API’s to extract data from the database.
* Exploratory data analysis and profiling of data using SQL, Python.
* Assist the team in data transformation and analytical data preparation process for developing a statistical model.
* Utilized PyUnit, the Python unit test framework for testing the functionality of the application.
* Involved in front-end designing using HTML, CSS, JavaScript, and Bootstrap.
* Profiled python code for optimization and memory management.
* Interpreted raw data using a variety of tools (Python, Excel Data Analysis), algorithms, and statistical/econometric models (including regression techniques, decision trees, etc.) to capture the bigger picture of the business.
* Developed SQL queries for visualization and reporting systems.

**Aviva Inc, Bangalore, India**

Software Engineer May-2011 –Feb-2014

* Experience in writing Sub Queries, Stored Procedures, Triggers, Cursors, and Functions on
* MySQL and PostgreSQL database
* Provided input into developing and modifying systems to meet client needs and developed business specifications to support these modifications.
* Prepared all required and user manuals, conducting CRP, UAT and user training for implementations.
* Used HTML, CSS, jQuery, JSON and JavaScript for front-end applications.
* Implemented locking mechanisms using multi-threading functionality.
* Designed business system test cases based on the the business process of the Client and tested all business cases in development and test instances.
* Worked on SQL and Oracle OBIEE reporting and analytics
* Identified and reconciled errors in client data to ensure accurate business requirements.

**Education**

**Saint Mary’s College, Moraga CA**

**Master of Science in Business Analytics**, **GPA: 3.9** **Sep 2018**

**Business Analyst Projects:**

* Built a predictive analytics model for e-commerce by analyzing 2.8 millian records on women's apparels. Measured the likelihood of a customer to recommend an item to others based on the costumer's age group. Technologies: Excel, SAS, Tableau.
* Designed an MYSQL database to determine the performance and behavior of homeless clients. Recommended the funding cutoff value and suggested switching to consultation services after a certain period. Technologies: Excel, SQL
* Built predictive model to predict customer’s next move to optimize e-commerce conversion rate of an e-commerce company. Used python for developed a machine learning model.

Technologies: Machine Learning Algorithm using Python, Simulation, Multivariate Testing, Tableau, SQL, MYSQL Database.

**Biju Patnaik University of technology, Odisha, India.**

**Bachelor of Technology, Computer Science and Engineering, GPA 3.58** **Jul 2010**