Pendela Hemanth Kumar

Data Engineer (Python, PySpark, SQL)

⌂ S.N.S Luxury Homes, Whitefield, Bangalore - 560066

✆ +91 8951528138 ✉ hemanthkumarpendela@gmail.com

Professional Summary

- Technology driven professional with **2.5 years** of experience in **Python** Scripting**.**
- Presently working as **Big Data Engineer** in **Tata Consultancy Services** at Bangalore, Karnataka.
- Experience in processing and handling huge data sets using **PySpark**.
- Good Understanding about **HDFS** and **YARN** architectures.
- Hands on experience with importing and exporting data from Relational databases to **HDFS** and **Hive** using Spark.
- Implemented Spark using Python and utilized **Data frames** and **Spark SQL** API for faster processing of data.
- Data visualizing and publishing reports using **Tableau**.
- Experience in writing database queries using **SQL.**
- Experience in using **Sqoop** for transferring data between HDFS and RDBMS.
- Experience in using **Unix** Shell Scripting.
- Experience in **Agile/SCRUM** enterprise-scale software development and good exposure in JIRA.

Achievements

- Received 'Young Achiever - **Star of The Month**' award of CFT at Tata Consultancy Services.
- Acquired **On the Spot** award.

Certification

- Completed the online course **Spark and Python for Big Data with PySpark** from Udemy

Technical Skills

Tools	PyCharm, Tableau and Databricks.
Programming languages	Python, SQL
Methodology	Agile Scrum
Operating Systems	Unix, Linux
Big Data Technologies	Spark, HDFS, Sqoop, and Hive
Repositories	GitHub

Education

Qualification	Name of the Institution	Year	Percentage/CGPA
B.Tech	Amrita VishwaVidyapeetham	2018	8.28 CGPA
Intermediate	Narayana Co-Spark	2014	95%
S.S.C	Sri Chaitanya Techno School	2012	9.7 GPA

Work experience and Key Projects

Taxone WH Renovation April 2019 - Present

Taxone is an application to validate the transactions of Morgan Stanley's old and new systems which are legacy and distributed respectively for various tax events like Interest Type, Dividend Type, Gross Proceeds and Accured Interest, besides generating reports of their discrepancies present in both systems and publishing them in Tableau.

Client: Morgan Stanley

Role & Responsibilities
- Responsible for loading data from database to HDFS using Sqoop.
- Preprocessing and handling huge data sets using PySpark processing engine.
- Written extensive Spark programming using Data Frames & RDD's for transforming transnational database data and load it into HDFS and csv files based on the requirement.
- Publishing reports to clients in Tableau for daily, weekly and monthly cadences as per the client's requirement.
- Worked on Spark SQL in querying for a better performance.
- Scheduling AutoSys jobs.
- Interacting with client's for requirement gathering.
- Developing application components and unit testing as per functional requirements.

Technologies Used
- **Data Ingestion**- Sqoop
- **Data Storage-** HDFS
- **Data Processing** - Spark
- **Data Visualization**- Tableau

Achievement
- Processed and did a reconciliation of **50 million records** for Withholding Interest tax type in 15 mins and nominated for Morgan Stanley's Tax Fleet Day.

Corporate Actions FPL & PNS Oct 2018 - March 2019

The project involves to distribute the CA data from data warehouse to FPL & PNS teams based on Document Type by performing modifications and analytics to the data based on Client Type.

Client: Morgan Stanley

Role & Responsibilities
- Preprocessing and handling huge data sets using PySpark processing engine.
- Performing Big Data Analytics on the data.
- Handling Data Skew.
- Making the entire process as 100% Automated.
- Executing database queries.

Technologies Used
- **Data Storage-** HDFS, **Data Processing** – Spark, **Data Base** - MySQL