

RAHUL KHANNA

Senior Big Data Engineer

Total Experience: 3.5+ Years

Relevant Experience: 3+ Year

Mobile -: +91 -7055340256/9457541957

Email -: rk.rahul22khanna@gmail.com

PROFESSIONAL SUMMARY::

PG-Diploma in Big data Analytics from **CDAC-KP(Bangalore)** with 3+ years experience in **Spark, Python, Machine Learning, Power BI**. In-depth knowledge and understanding of big data analysis and building Machine learning models in Python with ability to use them to design solutions for complex business problems. Knowledge of Data Processing(ETL Process) using PySpark, Python. Excels at providing insights from raw data, collaborating with client. Demonstrating expertise in Power BI. Brilliant analytical & mentoring skills; A cohesive team leader with ability to work with multi-disciplinary teams.

ORGANIZATIONAL SCAN::

April-2017 – Till Date (Kockpit Analytics, Gurgaon)

Business Analytics::

- Analysis of business requirements from client & translating them into efficient framework for development
- Plug N Play Visual Dashboards for **Navision(ERP)** users which populates some important Business KPI as per their Navision data
- “Sales” related KPIs such as Revenue, Gross Margins, Accounts Receivables, Payables, Orders and Inventory are covered
- Core member of Application Architecture Design team, **Apache Spark** with **Python** (Pyspark) and **Hadoop Cluster** for ETL Process and **Power BI** for visualization used with **PostgreSQL** as an intermediate stage for saving transformed data
- Designed the logic for Data transformation in Apache Spark for all Business KPI and implementing the same using Python and Spark
- Utilize **Bit Bucket** for version management and **AirFlow** for Task Scheduling
- Deployed Product on multiple clients (Retail, Insurance,Solar)

Data Analysis and Forecasting::

- Analyse and transform the data from the On premise solar site which is stored in **AZURE cloud** using Python
- Did analysis on different solar plants and visualize each plant performance and it's forecasting
- Generate the **Daily Generation Report** for all plants.Prepare **MIS report** on monthly basis
- Did deep level analysis to provide insights for different types of losses in the solar plant, Worked on the major loss factor to increase the generation and profit
- **Forecasting and Predicting Analysis** done on all the solar plants to get the maximum performance

Process Automation::

- Helped brand managers of European market in tracking post-launch activities, such as Brand and competitor Uptake, Sales with comparison to the forecast, detailing, samples and vouchers.
- Automated all the client tasks using **Python**, which earlier was done manually through **Excel**.
- **Lead the Project** and delivered it before deadline.

Project Scan::

- **Twitter Sentiment Analysis(Mini Project) in CDAC**, Analyze the sentiments of users on Twitter for a given search term whether they are Positive, negative or neutral. Use Python and Machine learning(NLP)
- **Ultrasound Nerve Segmentation (Mini Project) in CDAC**, we build a model that can identify nerve structures in a Dataset of ultrasound images of the neck. Hence, would improve catheter placement and contribute to a more pain free future. Use Python and Machine learning(CNN).

SKILLS::

- Writing PySpark(Python + Spark), Python Codes for ETL process
- Implementing PySpark Environment On top of Hadoop cluster
- Excellent knowledge of Hadoop ecosystem
- Developing Machine Learning models in Python using techniques Regression,Classification and Clustering
- Time Series Forecasting Using statsmodel, ARIMA and Prophet Library
- Design Dashboard on PowerBI for Visualization

ACADEMICS::

- **Computer Science Engineering**
SIET, Allahabad, UP
2011-2015

CERTIFICATIONS::

- **PG-Diploma in Big Data Analytics**
CDAC-KP, Bangalore
2016

IT SKILLS::

- Leading a team with successful Delivery of end to end solution
- Ability to make decisions & create workable solutions for business problems and new project requirements Architecture
- Process Improvement