ANITTA SAJU THERATTIL (아니타)

Data Engineer

About Me

I have been consistently appreciated for my accountability, technical acumen, and out-of-box thinking by my peers and leaders throughout my 4.5 years of professional experience in the field of Big Data and Data Engineering. I am a quick learner, a proactive problem solver, and open to learning new platforms. I look forward to being a part of an organization that would help me grow and in turn help me use my skills to contribute to the company's vision.

Contact

010-96407835

anittasaju1996@naver.com

경기도 성남시 분당구

in Anitta Therattil

VISA F-1

Area of Expertise

- Big Data Platforms
- Data Visualization
- Data Analytics
- Reporting infrastructure
- Data Cleaning
- Exploratory Data Analytics
- Automating processes

Education

Bachelor of Technology

Electronics & Communication Engineering

National Institute of Technology, Surat, India

Graduation Year - 2017

Bronze Medalist (동메달리스트)

CGPA - 9.41/10

Professional Experience

American Express, Gurgaon, India

Senior Analyst - Product Development

(Dec 2019 - Feb 2022)

Credit and Fraud Risk Department

- Involved gathering metric requirements from our strategic partners as well as designing, implementing, testing and deployment of an end-to-end automatic process for cleaning, analyzing, processing, and visualization of these KPIs for the senior leadership.

Technical Skills

- **Programming Language**

  Python, Shell Scripting

- **Big Data Platforms**

  Hadoop ecosystems like : Hive, PySpark, Sqoop, Oozie, HDFS, Spark SQL

- **RDBMS**

  Oracle, MySQL

- **Data Visualization**

  Tableau, Python, Excel Pivot

- **BI Tools**

  Microsoft SSIS and SSAS

- **Machine Learning**

  Basic algorithms and scripting using python

- **Other Beginner level skills:**

  Teradata, R, Java, Jupyter Notebook, Jethro

Soft Skills

- Hardworking
- Motivated
- Accountability
- Out-of-the-box thinking
- Quick learner
- Interpersonal communication
- Team Player
- Curious
- Eager to learn
- Patience

- Maintaining the quality of these metrics after their deployment as well as ensuring their timely updates.
- Undertook several ad hoc projects involving optimization and automation in our infrastructure and processes.
- Proactively involved in identifying areas of improvement and implementing alternate solutions for existing infrastructure.

  Some of my major contributions are :

  a) Smart design and implementation for our infrastructure to tackle the existing problem of rigidity and redundancy in our processes. The design was well appreciated for its robustness.

  b) Design and implementation of auto-matic tracking/alerts of our processes and data to decrease the manual efforts of the developer.

- Actively involved in planning assessment/ analysis during activities like patching, migration, and presentation of our products to senior leadership.
- Self-initiatives to quickly learn and adapt to several technologies like Microsoft SSIS, SSAS, Jethro, and Tableau to provide our partners with visualization suited to their needs.
- Mentoring incoming group of interns to my team every 6 months.

Infosys, Bangalore, India

Systems Engineer (Dec 2017 - Dec 2019)

*Production Support for **Apple***

- Ensuring error-free processing of data using platforms from Hadoop ecosystem (like HDFS, Hive and Spark) and smooth functioning of infrastructure for timely reporting to vendors.

Certifications

- Big Data Hadoop Certification by Edureka
- Data Analyst Nanodegree Program by Udacity
- Microsoft Python Certification for beginners

Awards

- Leadership in Action

    Awarded by Amex

- Top Performer in training

    Awarded by Infosys

- Bronze Medalist

    Awarded by Department of Electronics and Communication, NIT, Surat

- District topper in High School

Interests

- Reading
- Recycling
- Foodie
- Foreign cultures
- K-Dramas

Languages

- English (Professional)
- Hindi (Native)
- Korean (Interested to learn)

- Troubleshoot bugs in case of infrastructure failures and if required coordinate with necessary teams to expedite the fix.
- Extensive client communication was an integral part for the success of this project.

Projects

- **Fraud Dashboard** *(American Express)*

    Why : To report key fraud metrics to senior leadership and strategic partners for daily insights into fraud trends

    Highlight : This is a daily interactive dashboard with monthly and YoY components which is also capable of interactively visualizing trends in different categories like Markets and types of fraud across years.

    Technology : Hive, SSIS, SQL, Tableau, Shell scripting

- **Event Engine** *(American Express)*

    Why : A workflow scheduler was created to tackle the problems of rigidity and redundancy in the existing infrastructure.

    Highlight : Is customizable to any process with minor alterations and all the dynamic parameters can be passed via the GUI.

    Technology : Shell Scripting and existing internal tool for scheduling

- **Regression Testing** *(American Express)*

    Why : An automation utility that can be used for an extensive comparison of metrics between two hive tables or between baseline data in excel to a hive table. Useful during UAT or migrations.

Other Activities

♦ I have held the position of Senior Under Officer in the National Cadets Corps

♦ Awarded medals for debates and group discussions

♦ Member of Colleagues Engagement team in Amex

♦ Have organized fun events for my team as well

♦ Coordinator for various exhibitions and workshops during the college fest

Highlight : This utility is able to compare data for all the selected metrics across all the selected dimensions and as it is built using Pyspark and python, it has good performance on large datasets as well.

Technology : Pyspark, Shell scripting, Python, Hive

• **Kaggle dataset of top 500 dramas from MydramaList** *(Personal Project)*

Why: To generate a comprehensive dataset of the top 500 dramas that contains general details along with review and reviewer details. Has been uploaded on Kaggle for public use. (click on Kaggle for link to the dataset)

Highlight: This data has been scrapped from MyDramaList website and sentimental analysis has been performed using the Hugging Face "cardiffnlp/twitter-roberta-base-sentiment" model. Multithreading and Multiprocessing approaches were used to decrease the time of execution and this script can also run for fetching 5000 dramas.

Technology: BeautifulSoup, transformers, Python, Jupyter Notebook, multithreading and processing modules. (script can be found by clicking here)

I hereby confirm that all the details furnished above are authentic and accurate to the best of my belief.

ANITTA SAJU THERATTIL