JHANCY AMARSINGH

Austin, TX 78735 | (309)-684-0049
jhancya1@umbc.edu | LinkedIn | GitHub

SUMMARY

Detail-oriented **Data Scientist** with **8** years of IT experience building multiple data solutions and embedded systems in an Agile model. Expertise in **ETL automation**, big data **warehousing**, data **visualization**, Business Intelligence, building **Machine learning** models, and **statistical** and **time series** analysis.

PROFESSIONAL EXPERIENCE

Snowflake Data Engineer, *Caterpillar, Illinois, USA*   Feb 2023 – Present

- Spearheaded the automation of ETL processes, enhancing data transformation and loading efficiency across various databases (Postgres, SAP, DB2) using Python and integration tools like Apache Airflow and Prefect.
- Formed a new **Snowflake (SaaS)** Database focusing on US Drive Train with Data Lake, Data Warehouse, and Data Mart layers and created ~ 300 tables/Views under multiple schemas in Snowflake using DBeaver.
- Implemented **clustering**, partitioning, **hashing**, archiving, and materialized views; complex queries enhanced query performance by 70% and reduced costs by 20%.
- Developed reports to enhance supplier quality management, collaborating with supply chain teams and achieving 20% cost savings.
- Designed and deployed 22 custom dashboards for supply chain analytics, facilitating critical business insights and strategic process improvements.
- Performed Data Validation and Data Management and Maintained data integrity and protection; Used **GIT**, **GITHUB CoPilot,** and **Azure Repos** for version control and made over 40 Pull Requests.
- Used **Azure DevOps Boards** for Workflow Management to monitor and track Epics, Stories, and Features.

Data Scientist, *HCMI, California, USA*   Jun 2022 - Dec 2022

- Developed BI dashboard to enable reporting for ~50 Human Capital Financial (Key Performance Indicator) KPIs.
- Formed new data architecture using **Azure Synapse Analytics, Azure Data storage, Pipelines, Python**, and **PySpark** for data collection, cleaning, and wrangling to create **Azure SQL Database**.
- Created data analytics pipeline, data-driven insights, data modeling (star schema) to evaluate industry trends, and **scenario-based** models to forecast KPIs with 70% accuracy.

Research Assistant, *UMBC, Maryland, USA*   Aug 2021 – Jan 2022

- Anonymized network data and achieved 90% privacy, retaining data utility, and published a **Research paper** on ACM Digital. Measured the accuracy and privacy rate of 6 algorithms using Python scripts, ARX, and Amnesia.

Senior Data Scientist, *Centurylink (Lumen Technologies), India*   Mar 2019 - Aug 2021

- Executed advanced predictive analytics on Toyota's injection molding processes, achieving an 80% recall rate by applying sophisticated EDA and machine learning techniques (**random forest**) and performance tuning by **XG Boost**.
- **Statistical Analysis** (Quantitative analysis) to find the visitor count of 270 TD Ameritrade offices using CISCO router data (Azure Cosmos DB data) based on the Signal strength distribution using **Time Series Algorithm EWMA** and **KNN classification** and achieved 85% accuracy.
- Annotated and labeled training dataset of web element images and trained **RCNN Inception** neural network to convert Wireframe to webpage form using **Tensor Flow** object detection (AI), achieving 80% accuracy.
- Showcased leadership, collaboration skills, project management, providing feedback & mentoring skills while leading a team of 7.
- Built a methodical innovative solution to read any PDF 90% accurately by image preprocessing and text extraction using **Open CV** and Open-Source **OCR libraries** like Camelot and Tabula and hosted as a web service.
- Built conversation flow (stock **chatbot**) on **Dialog Flow** specifying 100s of entities and training phrases; Used Flask, python, and Google Cloud services to build API for real-time data and integrated with four platforms.

Data Analyst, *Accenture, India*   Jun 2017 – Mar 2019

- Performed data analysis and developed a **Predictive analytics Machine Learning model** to improve business operations and customer experience by 20%.
- Provided analytical support and change management processes using mathematics, statistics, hypothesis, and **A/B testing**, leading to a 15% improvement in user engagement.
- Educated co-workers through dozens of Knowledge Transfer sessions and **Technical Documents** with clear interpersonal communication skills.
- Used **Excel** scripting and advanced **SQL queries** with CTEs and **Windows** functions for reporting device availability daily; proposed and implemented fixes to increase device availability by 60% and for cost reduction.

Application developer, *Accenture, India*   Jun 2014 – Jun 2017

- Built **the Embedded WINCE application and deployed it** in 5000 Metro devices to validate payment use cases.
- Coded ~10,000 lines for NFC reader to read and write into the card, Web services, CRE **XML** integration, APIs Integration, and Business Logic as per the software systems, technical and business requirements.
- Unit testing with **code coverage** of 90 % and **System Integration testing** delivered code with no critical defects.
- Worked in a fast-paced, **Agile/Scrum** environment, able to work independently and flexibly.
- Worked on **Migration and production** deployment project and migrated software for 4000 devices using the SOTI simulation tool.
- Showcased accountability and excellent **organizational skills** while customer-facing using innovative **POC**s.
- Implemented **object-oriented programming and design patterns** per technical discipline to increase usability.
- Confident and Collaborative with cross-functional teams to understand workflows and problem-solving ability.
- Written ~500-unit **testing** cases and debugging, Quality Checks, System Integration, UI, and regression testing.
- Experience using JIRA Software with **Jenkins** for **CICD**, reducing deployment time by 20%.

TECHNICAL SKILLS

- Computer Programming Languages: Python, C#, VB Scripting, DAX
- Frameworks/Libraries: Pandas, NumPy, sci-kit-learn, Matplotlib, Spark, TensorFlow, Beautiful Soup, Dialog Flow
- Software/Tools: Jupyter Notebook, VS Code, Microsoft Office, Excel, Microsoft PowerBI, Tableau Jira
- Database Management Systems: SQL Server, MySQL, Azure Cosmos (nosql), Spark, Snowflake (OLAP), dbt
- Big data / Cloud: Databricks, Snowflake, Azure Synapse Analytics, Prefect Dataflow, Airflow, Microsoft Azure
- SDLC Methodologies: Agile, Waterfall

CERTIFICATIONS

- Microsoft certification (DP-900) Azure Data Fundamentals
- Microsoft certification (DP-203) Azure Data Engineer Associate
- EdX certification on Analyzing and Visualizing Data with Power BI
- Microsoft certification (70-532) Developing Microsoft Azure solutions.

EDUCATION

University of Maryland, Baltimore County, MD, USA
Master of Science in Data Science   *Aug 2021 to Dec 2022*   CGPA: 3.5
Welingkar Institute of Management, India
PGDM MBA in IT Project Management   *Jan 2020 to Dec 2021*   CGPA: 3.4
PSG College of Technology, India
Bachelor of Engineering in Computer Science   *Aug 2010 to Jun 2014*   CGPA: 3.4