ADITI DUTT

Phone: 919.344.8380 | Email: aditidutt.nmims@gmail.com | GitHub | Linkedin

Professional Summary

Data Scientist with **4+ years** of experience delivering **machine learning solutions, anomaly detection** systems, and **interactive dashboards**. Skilled in **Python, Snowflake, Power BI,** and **Explainable AI** methodologies. Adept at driving improvements in **data integrity**, reducing **manual processes**, and implementing robust **data governance practices** to provide actionable business insights. **Proven success** in enhancing data accuracy by 20% and cutting operational costs by 30%.

Skills Summary

Languages & Frameworks: Python (including scikit-learn, pandas, NumPy, Matplotlib, plotly), NLP, R, PySpark, Snowpark, Java
Technologies & Concepts: Machine Learning, Predictive Modeling, Explainable AI, Big Data, Azure, Hadoop, Object Oriented Programming
Databases: Snowflake, MySQL, MS SQL, Oracle
Tools: Streamlit, Power BI, Spark, RStudio, Microsoft Excel
Methodologies: Data Governance, Anomaly Detection, Agile/Scrum
Certifications: Microsoft Azure Fundamentals (AZ-900)

Education

The University of Texas at Dallas (August 2021) -*Master of Science, Information Technology and Management*
NMIMS Mumbai, India (August 2014) - *Bachelor of Technology, Computer Science*

Experience

Invesco - Data Scientist (January 2023-Present)

• Engineered and deployed **predictive models and anomaly detection systems** using machine learning and statistical techniques, enhancing data quality and risk mitigation.
• Integrated **explainable AI techniques** with interactive dashboards to provide real-time, actionable insights for executive-level decision-making for various machine learning models.
• Developed a **comprehensive data quality framework** in Snowflake (leveraging Streamlit), enabling accurate detection and resolution of missing values and duplicates. Resulted in a **20% increase in data integrity**.
• Engineered a **FactSet Identifier Lookup App** using Streamlit, Snowflake, and fuzzy matching algorithms, **improving data retrieval accuracy by 20%** and **reducing manual lookup efforts by 30%.** This innovation enhanced data standardization and accessibility, key components of enterprise data governance.
• Developed a **comprehensive risk management report** for Separately Managed Accounts (SMA) that automated monthly reporting through SQL, Python, and Power BI, **reducing manual processing time by 30%** and ensuring compliance with enterprise data governance principles for accurate and reliable insights.
• Spearheaded the development of the **Agile team performance dashboard** in Power BI, enhancing data quality, **reducing scope creep by 15%,** and **improving sprint planning accuracy by 10%,** directly supporting the organization's ability to ensure alignment with strategic goals.

Predactica - Data Scientist (July 2021-December 2022) & Data Science Intern (January 2021-June 2021)

• **Designed and implemented** advanced data visualizations and interpretation tools (using Python and Flask) to **demystify machine learning model predictions**, boosting user comprehension by 40%.
• Conducted data cleaning and exploratory data analysis on large datasets to **enhance model accuracy**.
• **Developed scalable Spark jobs** for big data processing, contributing to successfully explaining **complex machine learning models**.
• **Collaborated on open-source projects** to refine model interpretability, directly contributing to the core logic of a **user-friendly analytics dashboard**.

LTI, Pune, India – *Software Engineer* (September 2014 -August 2018)

• **Developed and tested** software components focusing on robust design and defect reduction, leading to a **20% decrease** in system downtime.

• **Engineered API proxies** and built RESTful SOAP services using APIGEE Edge, **streamlining application deployment** and maintenance.

• **Participated in full-cycle software development**, including tech design documentation, unit testing, and integration support, ensuring high-quality releases.

PROJECTS

Suicide Rate Predictions (September 2020-December 2020)

- Utilized R and WHO global data to analyze trends in suicide rates, applying statistical techniques to identify key factors and inform public health strategies.

Operational Control Platform (January 2020-April 2020)

- Analyzed and proposed process improvements to automate compliance testing, **reducing human hours by 25% and cutting error rates by 75%.**
- Modeled system improvements using BPMN and object-oriented analysis/design, delivering a detailed proposal for **implementation**.

Research Publication- "A Comparative Study of Grid Computing and Cloud Computing" (March 2014)

- Authored and published a research paper comparing grid and cloud computing, leveraging Core Java and a web application to illustrate **key differences and performance metrics**.

Food Ordering System (October 2012-January 2013)

- Developed an end-to-end food ordering website using Java and SQL Server, **enhancing user experience** and **streamlining backend operations**.

Awards and Acknowledgements

• **Employee of the Month** (3x) at LTI for exceptional performance and project success.

• **Acknowledged by Senior Managing Directors** (2x) at Invesco for my efforts in going above and beyond on a project that significantly enhanced the internal team experience.