

Srashti Soni

Tampa, FL (Open to relocation) | +1(202) 937-6415 | sonisrashti3007@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

EDUCATION

M.S. in Data Science, University of Maryland, Baltimore County (GPA 3.97) (August 2021 – May 2023)

- Coursework: Machine Learning, Natural Language Processing, Big Data & Analytics, Statistics, Data Visualization

B.E. in Computer Science, Rajiv Gandhi Technical University, India (July 2012 – June 2016)

- Coursework: Data Structures & Algorithms, Data Mining, Cloud Computing, Mathematics, Database Management
-

WORK EXPERIENCES

Senior Credit Risk Analyst (CCFI Companies, LLC USA) (October 2023 – Present)

- Designed, built, and deployed several end-to-end **credit risk and response models** for large-scale Prescreen Campaigns using **Python, SQL, and Databricks**, improving targeting precision and driving a 40% increase in campaign effectiveness.
- Led enterprise-wide Prescreen Campaign execution across all three major credit bureaus, collaborating directly with bureau partners to ingest, validate, and operationalize bureau-specific datasets for monthly campaign launches.
- Redefined Prescreen strategy in July 2024 by segmenting offers based on credit scores and credit profiles, significantly improving acquisition quality and driving a 30% increase in new-customer response rates.
- Conducted in-depth analysis of charged-off customer portfolios, identifying high-risk patterns and actionable risk drivers; implemented mitigation strategies that resulted in a 15% reduction in credit-related losses.
- Performed high-impact ad hoc analyses of campaign performance, customer behavior, and risk trends, translating insights into data-driven recommendations that directly improved Prescreen targeting strategies.
- Developed and delivered a Prescreen Health Check dashboard using **Tableau**, enabling year-over-year visibility into credit loan performance, campaign health, and portfolio risk for leadership and stakeholders.
- Led cross-functional strategy meetings with credit bureaus and internal stakeholders, presenting analytical findings and forward-looking Prescreen strategies to drive alignment and inform future campaign planning.

Senior Revenue Operations Analyst (Calyx Containers, USA) (August 2023 – October 2023)

- Conducted in-depth data analysis to identify trends, guiding sales and marketing teams in targeting high-value customers.
- Generated regular reports and dashboards to provide senior leadership with insights into revenue performance, KPIs, and key trends.
- Collaborated with sales and marketing teams to align revenue strategies, ensuring consistent messaging and goal attainment.

Software Engineer Intern (Data Science) (Clinisy, USA) (June 2022 – August 2022)

- Predicted patient churn in a healthcare setting by implementing **machine learning models** (Logistic Regression and Random Forest Classifier) on large datasets, resulting in a 20% improvement in customer engagement through targeted interventions.
- Created over 10 data-driven Tableau **dashboards** that provide insights into patient falls, enabling stakeholders to perform effective risk management.
- Conducted data analysis and **reporting**, streamlining **pipelines** and workflows, leading to a 20% increase in efficiency.

Data Engineer (Infosys Limited, India) (April 2019 – August 2021)

- Established a centralized **Data Lake on Azure**, creating a streamlined pipeline for data ingestion, processing, storage, and analysis.
- Integrated **Spark** with **HDFS** and relational databases for seamless data ingestion and extraction.
- Boosted product demand by 33% by creating data visualizations and interactive Power BI dashboards.
- Improved efficiency by 20% and reduced user attrition by leveraging **ETL** processes to extract insights from raw data.
- Optimized over 500 **SQL** queries and managed a **Python** migration script, achieving a 38% reduction in query processing time.

Data Analyst (DXC Technology, India) (December 2016 – April 2019)

- Built a Python script to analyze complex insurance data, generating insights to forecast premium trends and aiding the detection and prevention of fraudulent claims by 15%.
 - Leveraged an automated **ETL process** to integrate machine learning algorithms, generating alerts for suspicious claims and enhancing fraud detection mechanisms, significantly increasing the efficiency of the data pipeline by 30% while maintaining data quality.
 - Pioneered an automated ETL process using **Azure Data Factory**, resulting in a 40% reduction in processing time.
-

PROJECTS

News Summarization Application | Python, Scikit-learn, Matplotlib, Machine Learning, Transformers, Streamlit

- Conducted data analysis on the CNN dataset (85000 rows), addressing anomalies and missing values.
- Fine-tuned T5 transformer model for news summarization, achieving a high accuracy of 91% and 94% ROUGE precision.
- Developed a user-friendly interface using Streamlit, providing live news summaries and article links.

Cyberbullying Tweet Recognition Application | Python, Scikit-learn, Matplotlib, Machine Learning, NLP, Streamlit

- Developed a web application to detect cyberbullying instances in tweets and articles, considering parameters like Age, Ethnicity, Gender, and Religion, achieving 86% data accuracy through fine-tuned Support Vector Classifier and AdaBoost Classifier models.

TECHNICAL SKILLS

- Programming Languages: Python, R, SQL, Java
- Databases: MySQL, PostgreSQL, T-SQL, MongoDB, Snowflake, MSSQL, Oracle, SQLite, DynamoDB
- Machine Learning: Pandas, Numpy, Scikit-Learn, TensorFlow, Keras, PyTorch
- Tools: Tableau, Power BI, MS Excel/Pivot, Jira, Git, GitHub, AWS Quicksight, Visual Studio, Docker
- Big Data and Cloud: Hadoop, Apache Spark, Amazon Web Services, Microsoft Azure
- Others: Natural Language Processing, Statistical Analysis, Data Mining, Predictive Analysis, Data Modeling, Data Warehousing, ETL (Extract, Transform, Load), Flask, Swagger, Postman, RESTful services.