

Shouvik Sharma

400 E 33rd Street, Apt 515, Chicago, IL-60616 | Phone: 3124592008 | shouvik19@gmail.com | [Linkedin](#) | [Github](#) | [Medium](#)

SUMMARY

Over 4 years of comprehensive work experience in Data Engineering, Marketing Analytics and Business Intelligence in banking and retail domains. Ability to solve complex business problems using ETL, Data Mining, Machine Learning & Data Warehousing concepts.

EDUCATION

- MS in Data Science, Illinois Institute of Technology, **GPA: 3.8** (Aug 2019 – May 2021)
- MS in Statistics, NMIMS University, **GPA: 3.35** (Jul 2016 - Apr 2018)
- Certifications: [Snowflake Pro Certification](#), SAS Certified Base Programmer for SAS 9, SAS Certified Predictive Modeler

WORK EXPERIENCE

Sr. Business Intelligence Engineer at Avant LLC., Chicago: (Aug 2021 - Present)

- Work with members of the Data Management and Data Services teams to build tooling for large-scale data validation, then leverage this tooling to validate the completeness, timeliness, and accuracy of raw and transformed data sets
- Developed dashboards to analyze marketing sales funnel using Looker that allowed business to understand customer journey experience and improve customer experience
- Enabled effective decision-making by retrieving and aggregating data from multiple source and compiling it into a digestible and actionable format.
- Identifying process and system improvement opportunities by monitoring existing metrics, analyzing data and partnering with internal teams.
- Communicating key analysis to senior stakeholders through verbal and written communication.

Business Intelligence Engineer Intern at CNH Industrial, Racine: (May 2021 – Aug 2021)

- Development of predictive analytics and statistical forecasting models utilizing forecasting related systems to improve forecast accuracy and bias reduction
- Formalizing assumptions about how demand forecasts are expected to behave, creating definitions of outliers, developing methods to systematically identify these outliers, and explaining why they are reasonable or identifying fixes for them.
- Development of tools to allow process automation, analysis & corrective action implementation by the business. - Forecast Accuracy analysis and corrective action implementation

Business Intelligence Engineer at Daten Solutions Inc., Chicago: (May 2020 – Apr 2021)

- Developed and automated **data migration pipeline** from SQL Server to Snowflake and performed **dimensional modeling** on the migrated data, further created **data dictionary** for the technical audience.
- Automated **ETL** processes using **Prefect** (Python), making it easier to wrangle data sets and reducing time by as much as 40% by performing large-scale data conversions, and transferring BAAN data into standardized formats for integration into Snowflake.
- Created **Tableau** dashboards to explain variation in success **Metrics** and **Time Series Analysis** to higher management.
- Automated reporting process using **Excel VBA (Macros)** and **MySQL** maintaining accuracy and saving ~ **75%** of time.

Business Systems Analyst – Practicum Student at Labelmaster, Chicago: (May 2020 – Dec 2020)

- Involved in designing databases, data marts, E-R model for **OLTP** and multi-dimensional model for **OLAP**.
- Optimized complex **SQL** scripts for quality checking of projects and populating output tables for deployment using **SSIS**.
- Automated hourly status report saving **10 man-hours/week**, thus decreasing response time for fixes and campaign failures.
- Achieved an accuracy of **MAPE 8%** approx. on price forecasting using Deep Learning algorithms like **LSTM** and **RNN**, further created dashboards for presenting the forecasted values to the higher management.

Business Systems Analyst at Cartesian Consulting: (Apr 2018- Jul 2019)

- Developed pipelines for **ETL** (Extract, Transform, Load) using **MySQL**, **Python**, **Airflow** and **AWS S3** for acquiring a POC project.
- Extracted data from streaming pipelines using **Flume** and **Kafka** and processed using **Spark** Structured Streaming.
- Predicted sales by **time series forecasting** in **Python** using **neural networks**, **ARIMAX** and **Prophet** for inventory management by eliminating understocking and reducing overstocking by 56%.
- Applied **K-means clustering** in **Python** for **segmentation** of customers, comparing it with **RFM** (Recency, Frequency and Monetary Value) analysis for improved campaign targeting.
- Identified the '**Most Valuable Customer**' by leveraging the customer data and deploying **Random Forest algorithm** with **True positive rate of 81%**, this led to better customer targeting and improving yearly topline revenue by 13 %
- Developed **dimensional data models** and **data warehouse** adhering to integrity and **normalization** rules to support campaign **data mart** and customer one view for marketing campaigns. Wrote **complex SQL** queries (multiple joins, CTE's, subqueries).
- Generated visualizations using **Tableau** to analyze marketing **metrics** for making recommendations and supply chain analysis.

SKILLS

- Programming:** SQL, Python, R, SAS, Pyspark, HTML, C#, Excel VBA (Macros).
- Big Data Ecosystem:** Spark, Hadoop, MapReduce, Hive, Pig, Kafka, Flume.
- Cloud Technologies:** AWS (S3, EC2, Lambda, Athena, RDS, Redshift, EMR).
- Tools:** Tableau, Power BI, Azure ML, RStudio, Jupyter Notebook, SAS E-Miner, SAS CI, IBM-Unica, SSIS, MS Office, JIRA.
- Libraries:** Numpy, Pandas, Matplotlib, Seaborn, Scikit-Learn, Keras, Nltk, Gensim, Scipy, Beautiful Soup.