

Ashwini Khedkar

Boston, MA(Open to Relocate)| (857)-544-9921 | khedkar.as@northeastern.edu | [Linkedin](#) | [Github](#)

EDUCATION

Northeastern University

Master of Science in Information Systems

Boston, MA

May 2024

TECHNICAL SKILLS

Programming Languages: Python, C/C++, RDBMS/ SQL, C#, Java ,JavaScript,R, NoSQL, WebAPIs

Frameworks: Flask, Django, Nifi, Apache Spark, Streamlit, Dash, Kubernetes, Docker, dbt, Pyspark, Hadoop, FastAPI

Other Tools: Git, GCP, AWS, Azure, Jupyter, Excel, RStudio, Apache Airflow, Databricks, Tableau, Snowflake, PowerBI, Alteryx, Talend, ER Studio, SaaS, Looker, Bigquery

Libraries: Pandas, NumPy, sci-kit-learn, Matplotlib, Seaborn, Yolo, Pytorch, Plotly, GGplot

PROFESSIONAL EXPERIENCE

Data Analyst

Aug. 2019 – Jul. 2022

Virtusa Corporation

Bangalore, India

- Performed **profiling, EDA, data profiling and statistical analysis like regression analysis** of raw P&C internal and external data sets for the enhancement of client's insurance analytic engine in MongoDB and Python
- Streamlined ETL processes with Nifi for data ingestion by implementing Root Cause Analysis and **reduce data processing time by 10%** through custom python transformation scripts
- Developed **70+ Python Flask and Jinja-based** programs to improve MongoDB data modeling, mining and mapping, improving risk analysis engine's accuracy and got awarded with Excellence Work Award
- Collaborated with internal and external stakeholders to gather, validate, and implement data requirements, ensuring **data quality and accuracy** through the development and monitoring of data validation tools

Data Science Intern (Autonomy)

May 2023 – Aug. 2023

Rivian Automotive

Palo Alto, CA

- Built a Streamlit dashboard with Kubernetes deployment for autonomous vehicles(self driving) KPI metrics, **surpassing Tableau Dashboard by 11%**, emphasizing data visualization principles, & enabling data-driven adhoc analysis for cross-functional teams
- Engineered ETL pipelines to profile PCAP sensors data with Airflow and Databricks(Pyspark), resulting in a **30% reduction** in data processing duration and improved data consistency
- Conducted in-depth data analysis and developed actionable insights using SQL and Python, **supporting strategic recommendations** for autonomous vehicle performance improvements and data-driven decision-making

Graduate Research Assistant/Data Analyst

Sep. 2022 – Present

Northeastern University

Boston, MA

- Develop a data pipeline using python and unity scripts to collect and synchronize eye-gaze, joystick keypress and screen recording data to analyze player's performance using **advanced predictive analytics**
- Performed **statistical tests such as t-tests, one-way and two-way ANOVA**, and group analysis to derive actionable insights and support data-driven decision-making
- Train and Fine-tune **YOLO(computer vision model MAP=0.9)** to detect gameobjects annotated using CVAT to study spacio-temporal awareness. Published a paper at **IEEE-COG 2023 - [Measuring gaze behavior to characterize spatial awareness skill in Rocket League](#)**

ACADEMIC PROJECTS

IMDb Data Analytics and Integration | SQL, Talend, Alteryx, PowerBI, Tableau

Sep. 2023 – Dec. 2023

- Led a collaborative team to integrate multiple data sources(CSV,XML) for a data-driven IMDb warehouse (SQL, Azure SQL) by orchestrating the process by Alteryx & utilizing PowerBI and Tableau for data visualization
- Executed efficient & scalable ETL workflows with Talend to **optimize data flow by 35%** by emphasizing data modeling, data quality % integration principles,dimensional modeling and Business Intelligence Tools

Consumer-Business Marketing Analytics | Snowflake, Tableau, Machine Learning

Sept. 2023 – Dec. 2023

- Developed interactive Streamlit app for forecasting & anomaly detection using Snowflake's Forecast ML models
- Integrated **Customer Lifetime Value ,Customer Spend and ROI prediction** into a multipage Streamlit app, leveraging Snowflake, Snowpark, and ML models like XGBoost and RandomForest