HARSHAL TALELE

📞 +1 (224) 503-7866 | ✉ htalele21@gmail.com | 🌐 harshaltalele.com | 🔗 linkedin.com/in/harshaltalele

EDUCATION

Master of Science (Data Science)	**Aug 2022 – Dec 2023**
University of Rochester, USA	GPA – 3.64/4

- Courses – NLP, Statistical Machine Learning, Time Series Analysis, Computational Statistics, Data Mining

Bachelor of Engineering (Computer Science)	**Aug 2016 – Jun 2020**
Savitribai Phule Pune University, India	GPA – 8.49/10

SKILLS

Programming & Database Languages – Python, R, PySpark, SQL, MySQL, PostreSQL, Oracle SQL, NoSQL, MongoDB, Google Firebase
Software Tools & Technologies – AWS, Databricks, Tableau, Alteryx, MLFlow, IBM Cognos, Denodo, Advanced MS Excel, Git
Technical Skills – Machine Learning, Natural Language Processing, Generative Models, Time Series Forecasting, ETL, Exploratory Data Analysis, Hypothesis Testing, Statistical Analysis, A/B Testing, Quality Assurance, Quality Control

PROFESSIONAL EXPERIENCE

Data Scientist Intern	**Jun 2023 – Present**

University of Rochester **–** *Department of Business Intelligence & Data Warehousing*
- Developing classification of over **12k students** using predictive modeling to facilitate early academic support to struggling students
- Deployed **data pipelines** to integrate student academic data from multiple sources and generate data feeds for model training
- Employing **Jira**, **Scrum**, and **Sprint** methodologies to collaborate across teams for efficient project management
  Skills – Python, SQL, Feature Engineering, Scikit Learn, IBM Cognos, Denodo, Agile software development

Data Analyst	**Sep 2020 – Jun 2022**

IQVIA (159 Solutions Inc.) **-** *Healthcare Data Science company*
- Delivered sales analytics to support the pharmaceutical client's drug **sale of $5.6B** aiding an **increase of 98%** over previous year
- Designed **Tableau** dashboards and **Alteryx** ETL workflows to generate weekly reports on drug sales and disease metrics
- Collaborated with cross functional teams to migrate database from Teradata & Azure to **AWS S3** by conducting **QC and sanity checks**
  Skills – Tableau, SQL, Advanced MS Excel, Alteryx, Quality Assurance, Quality Control, Business Acumen, Data Storytelling
  Achievement – Recognized with '**Ovation Award**' for managing offshore work stream that included organizing
  daily catch ups, facilitating client engagement & stakeholder deliverables as per client needs

Data Science Intern	**Aug 2019 – Feb 2020**

Medialytics - *Data Science driven Ed-Tech developing startup*
- Developed a system to predict social media trend duration, driving a **70% boost** in **user engagement** by assisting content creation
- Utilized **web scraping** and **APIs** to efficiently store Twitter, YouTube, and Google Trends data in multi-level **PostgreSQL** database
- Performed exploratory data analysis in **R** and built a hyper-parameter tuned **XGBoost** classifier model to achieve an 84% precision
  Skills – Python, R, PostgreSQL, Machine Learning, Exploratory Data Analysis, Git, Leadership
  Achievement – Spearheaded a team of 8 in the software development phase of the project & planned the
  hand-over to a new team in limited duration of time

PROJECTS

**Forecasting Bike Inventory for Citibike	*GitHub Link***	**Apr 2023 – May 2023**

- Developed an app for **50k+ users**, to monitor bikes and empty dock availability by building a forecasting model for net bike change
- Implemented **ETL pipelines with Spark** streaming to process incoming data, creating bronze, silver, and gold data-level delta tables
  Skills – PySpark, Databricks, MLFlow, Time Series Forecasting, ETL, EDA, Version Control Tools (Git), Data Engineering

**Dynamic QA Generator for Research Papers	*GitHub Link***	**Mar 2023 – May 2023**

- Developed a QA model to aid efficient comprehension of research papers by summarizing relevant information in the form of Q&A
- Fine-tuned T5 models on OpenAI-modified QASPER dataset with **1.5k+ papers** for question generation and answer generation task
  Skills – Python, Large Language Models, Data Preparation, OpenAI API

**Analyzing Political Interest of Indian American	*Tableau Link***	**Nov 2022 – Dec 2022**

- Analyzed data from **2.8k unique** Indian Americans' tweets to identify political trends and biases ahead of 2024 presidential elections
- Designed **Tableau** dashboard illustrating the topics and sentiment within the community based on party affiliations and geographies
  Skills – Python, Topic modelling, Sentiment Analysis, Tableau, Twitter API

**100Ascent : Gamified Fitness Tracking app	*Mobile App Link***	**Aug 2021 – Dec 2021**

- Designed NoSQL database for the fitness-motivation app facilitating seamless fitness tracking and virtual travel experiences
- Achieved **50% increase in active users** by conducting user behavior analysis and leveraging insights to drive business decisions
  Skills – NoSQL, MS Excel, Data Engineering

POSITIONS OF RESPONSIBILITY

**Vice President – Rochester Indian Association	Website Link**	**June 2023 – Present**

Managing events and coordinating communications for a community of 500+ graduate students and alumni network

Graduate Teaching Assistant – Simon Business School	**Jan 2023 – Mar 2023**

Python Analytics – Supported the professor in grading assignments & labs and conducted office hours to assist students