**SAMARTH HADAWALE**

Boston, Massachusetts **|** (857)-214-9039 **|** [hadawale.samarth4@gmail.com|](mailto:hadawale.samarth4@gmail.com|)

[https://samarthhadawale.github.io/ |](https://samarthhadawale.github.io/%20%20|) [www.linkedin.com/in/samarthhadawale](http://www.linkedin.com/in/samarthhadawale%20)

# EDUCATION

## Northeastern University, Boston, MA Expected May 21

Master of Science in Information Systems **(GPA: 3.76/4.00)**

## Savitribai Phule Pune University, Pune, India Jul 14 – Jul 18

Bachelor of Engineering in Computer Engineering

**TECHNICAL SKILLS**

**Business Intelligence Tools** Tableau, MS Power BI, Qlik Sense, Einstein Analytics, Microsoft Excel

Data Integration & **DB Management** Talend Studio,MySQL, Oracle 18c, MSSQL Server

**Programming Languages** Python, SQL, R, Java

**ML frameworks** NumPy, Pandas, Scikit-learn, Matplotlib, Plotly, TensorFlow, Conx, Keras

**Data Wrangling & Modeling tools** Tableau Prep, Trifacta, XSV, Toad Data Modeler, Visual Studio

**Cloud Technologies** AWS (S3, Redshift, Athena, EC2), Snowflake, Docker

# ACADEMIC PROJECTS

## Northeastern University, Boston, MA.

## Book Store Marketing and Analytics (Python, R, Flask, Tableau, AWS S3, EC2) Jul 20 - Aug 20

## Performed & analyzed Customer Segmentation using RFM modeling, CLV, Cohort Analysis & Sales forecasting

## Developed ‘Books Recommendation System’ for customers to get similar books based on their choices & popularity

## Implemented ‘Predictive Analysis System’ for suggesting similar books on the basis of confidence score using R

## Explored the dataset staged in AWS S3 & Visualized it for getting valuable insights using Tableau

## Integrated all above systems into the Web Application using Flask, HTML, CSS, Bootstrap & deployed it on AWS EC2

## Recommendation System (FastAPI, Streamlit, Python, EC2) Jun 20 – Jul 20

## Developed Algorithmic Analytics Application which uses ML Models to build Recommendation system

## Implemented LightGBM & LightFM Models to recommend the products of the dataset containing 7M+ Records

## Stored the models in pickle file which were called using FastAPI as a backend & Streamlit as front end to relay requests

## Deployed application on AWS EC2 & tested load using Apache JMeter

## 

## Customer behavior analysis using Data Pipeline (XSV, Pandas, Trifacta, Snowflake, Einstein Analytics) May 20 – Jun 20

## Joined & explored datasets having 34M+ records using XSV, Pandas

## Sampled and performed some key operations on the dataset using Trifacta Data Wrangler

## Staged that data into Snowflake data warehouse and forwarded it through live connection to Einstein Analytics

## Visualized the data & predicted customer behavior through some exciting dashboards using Einstein Analytics

## Health Effects of Increasing Toxic Dumping & Poor AQI (Python, ANN using Conx & Keras) Jan 20 – May 20

## Cleaned and integrated AQI, TRI, Cancer and Life expectancy data for 65 US counties from publicly available data sources

## Performed EDA on a merged dataset to find correlation between AQI, TRI parameters and cancer rates

## Employed Random Forest Regression and PCA to determine most significant features influencing cancer rates

## Predicted future cancer rate with an accuracy of 85% using Multi Output Regressor and CNN for model training

## United States Health Insurance Program (Talend, Tableau, MS Power BI, AWS). Jan 20 – May 20

## Integrated 2M+ records of data from multiple sources to devise a multi-dimensional data warehouse using Talend

## Designed ETL workflows using Talend to integrate data into facts and dimension tables of Insurance Data Warehouse

## Built 20+ dashboards to analyze main KPI’s using Tableau and Power BI

## Used Amazon Web Services (AWS) Redshift to store the warehouse and using AWS Athena to query the warehouse

# PROFESSIONAL EXPERIENCE

## Jawed Habib Group| India Jan 19 – Jul 19

## Head of Digital Marketing

## Focused on Email, content marketing, social media & Paid as well as organic search while leading the team of 8.

## Analyzed digital marketing performance against metrics through Data Visualization reports built using Tableau.

## Managed client relationships, digital marketing agencies and handled firm's social media handles effectively.

## 

## Opulent, Pune | India Aug 17 – Jan 18

**Data Engineer Intern**

## Worked on data cleaning & data profiling on the various datasets of clients having around 650 tables and 7M+ rows

## Reduced loading time of data by approx.70% by streamlining & capturing historical records by designing Type-2 SCD jobs

## Performed ETL workflows by staging that dataset into Talend open studio and further pushed that dataset into Oracle DB

## Visualized critical KPI’s in Power BI to analyze the market with respect to prospective and existing clients