Sayali Walke

BUSINESS INTELLIGENCE | DATA SCIENTIST

PORTFOLIO

sayaliwalke30.github.io/

CONTACT

walkesayali11@gmail.com
(206)-751-9440
Seattle, WA
sayali-walke-7712b874/

PROFILE

Data Scientist with a strong background in Business Intelligence who is proficient in delivering valuable insights through data-driven approaches. Love to learn new technologies and be curious.

EDUCATION

Aug 2018 – May 2020
Northeastern University [Seattle, WA]
MS in Information Systems
[3.84/4.0]

June 2009 - May 2013
Govt. College of Engineering [India]
BE Electronics & Telecommunication
[4.0/4.0]

TECHNICAL SKILLS

- **Python** (Pandas, NumPy, Seaborn, Scikit-learn, Matplotlib)
- **Data Analysis** (Python, MySQL)
- **Business Intelligence** (Tableau, Power BI, ETL, SSIS, SSRS, SSAS, Amazon Redshift)
- **Machine Learning** (Regression, Classification, Random Forest, Neural Networks)
- **Statistics & Probability**
- **Big Data Systems** (Hadoop Map Reduce, MongoDB, DynamoDB, AWS EMR, Hive, Pig)
- **Programming** (Python, Java)

MANAGEMENT SKILLS

- Excellent Communication
- Stakeholder Engagement
- Agile Development
- Excellent Presentation Skill

ANALYTICS & DATA SCIENCE PROJECTS

Aug 2018 - April 2020

Customer Segmentation (Python, Tableau) Github

- Performed behavioral **customer segmentation** using **k-means clustering** technique. Grouped customers based on RFM (Recency, Frequency, Monetary) attributes to determine effective & tailored marketing strategies for **500k** customers.

Judge a Book by its cover (Pytorch, Python) Github

- Created Deep Learning **Convolutional Neural Network (CNN)** to predict the genre of the book by its cover to classify 57000 images. Implemented transfer learning technique to address this **image classification** problem and improved accuracy to **72%.**

Stock Analysis & Forecast (Python, SQL, PowerBI) Github PowerBI

- Web scraped stock statistics for over **500** stocks to perform **Time Series Analysis**. Built an interactive dashboard that allows users to make data-driven decisions. Developed and evaluated different models to forecast stock prices using **ARIMA, k-NN.**

Big Data Analysis & Recommender (Python, Hadoop, EMR) Github

- Carried out data analysis on **130M+** Amazon customer reviews and derived various Numerical Summarization, Filtering and Data Organization patterns using Hadoop map-reduce. Built a **Recommender System** to give smart recommendations to users.

House Prices Prediction (Python, Tableau) Github

- Applied regression machine learning models to predict the price of houses and carried out **exploratory data analysis** on a dataset of 16000 records to understand the patterns and detect anomalies. Reduced **RMSE from 0.1257 to 0.062** using recursive feature elimination for feature selection.

EXPERIENCE

June 2015-Sept 2017

Data Analyst | Vidya Technology Solutions

- Built a **Python & Internet of Things (IoT)** based Automated Machine Monitoring & Alarming System. Collected **real-time data** from sensors attached to CNC machines. Carried out **data pre-processing** on sensor values and performed diagnostic **data analysis** to identify patterns in machine's key performance indicators. It resulted in reduction of average machine maintenance cost by **20%.**

Oct 2013-Sept 2014

Business Intelligence Engineer | Tata Consultancy Services

- Solved **Data Reconciliation** problem by creating **12** different types of business reports using customer usage and payments data which resulted in loss prevention of **$1.5 million.** Created multiple visualizations using **Tableau** and **SSRS** to present business operation **KPIs** to management which gave a unified view of data from multiple databases.
- Integrated data from multiple sources by designing **ETL** packages. Performed **data analysis** by building complex **MySQL** queries, functions, stored procedures as a part of reconciliation solution.