Mihir Sawant

Boston, MA | 857-415-8229 | sawant.mi@northeastern.edu | www.mihirsawant.github.io |
www.github.com/mihirsawant | www.linkedin.com/in/mihir-sawant

EDUCATION

Northeastern University, **Boston, MA, United States**   **Sept 2018-May 2020**
Master of Science in Information Systems
Relevant Courses: Data Management and Database Design, Data Architecture and Business Intelligence, Data Science Engineering Methods and Tools, Network Structures and Cloud Computing

University of Mumbai, Mumbai, MH, India   **Sept 2013-Jun 2017**
Bachelor of Engineering in Electronics Engineering

CERTIFICATIONS

- **AWS Certified Cloud Practitioner**, *Amazon Web Services Training and Certification*   **May 2020**
- **Tableau Desktop Specialist**, *Tableau*   **Feb 2020**
- **Google Analytics**, *Google Analytics*   **Feb 2020**

TECHNICAL SKILLS

Languages:	Python, SQL, Cypher, Node.js, Scala, Java, JavaScript, HTML, CSS
Databases:	MySQL, MongoDB, Microsoft SQL Server, Oracle, PostgreSQL, Cassandra, Neo4j
ML Libraries:	NumPy, Pandas, SciPy, Scikit-learn, MatPlotLib, Seaborn, Keras
Big Data:	Hive, Spark, Hadoop, Apache Airflow
ETL Tools:	Talend, Alteryx, SSIS
Tools:	Tableau, Power BI, Azure, Docker, SSRS, SAS, Big Query, Jupyter, QlikView, Microsoft Excel, Git
AWS:	EC2, RDS, RedShift, Athena, ELB, CloudWatch, Quicksight, Lambda, SNS, SES, S3, Kinesis, EMR

PROFESSIONAL WORK EXPERIENCE

Data Analyst Intern | Velero Technology | Wellesley, MA   **Jul 2020-Present**
- Modeled a **data warehouse** to contain data integrated from multiple sources in SQL Server
- Achieved **18%** efficiency by architecting **data pipelines** to integrate data in data warehouse
- Performed **exploratory data analysis** to define analytical models
- Validated and demonstrate analytical models through various relevant error metrics and calibration techniques
- Accomplished **90%** percent accuracy using **statistical machine learning** models to implement predictive analysis
- Translated data results into **ad-hoc reports** in **Microsoft Power BI** to convey information to management

SQL Developer Intern | Applab Technologies | Mumbai, India   **Jun 2017-Sep 2017**
- Devised an application for a management system to handle data of 500 + students in **MySQL** server
- Normalized the data to **3NF** to eliminate the redundancy to avoid interdependencies between entities
- Transformed **SQL** queries and stored procedures for the enhancement of query performance by **10%**
- Extracted, interpreted, and analyzed data for the evolution of KPI and other metrics for market survey
- Presented reports with **Tableau** for performance analysis and monitoring customer satisfaction

ACADEMIC PROJECTS

Data Integration for News API *(Apache Airflow, Docker, Python, Boto3)*   **Feb 2020-May 2020**
- Constructed a data pipeline to fetch data from **API**, transform and store it on **AWS S3**
- Configured **Docker** to containerized **Apache Airflow** and relevant files and dependencies
- Authored a workflow in **Apache Airflow** to schedule an array of tasks to execute once daily
- Attained **10%** optimization in python script to transform data into CSV files

Healthcare Analytics *(Python, Machine Learning, Tableau)*   **Mar 2020-Apr 2020**
- Conducted **healthcare data analysis** to identify patterns of length of stay in hospitals
- Implemented **one hot encoding** to handle categorical data
- Executed **feature engineering** and **exploratory data analysis** to identify the trends
- Mange to wangled accuracy of **43.1%** using **LGBM Classification** model

NYPD accidents and complaints *(MS SQL, ERStudio, Alteryx, Power BI, Tableau)*   **Dec 2019-Feb 2020**
- Applied concepts of **data modelling** to fashion 27 dimensions and 16 facts using **ERStudio**
- Applied data security-related transformations, which include **masking**, **anonymizing**, and **encryption**
- Achieved **17%** process efficiency by **data cleansing**, and integrated data using **ETL** processes
- Developed **visualizations** and **dashboards** for reporting and to satisfy ad hoc data analysis requests and tracking

Infrastructure and Web Application *(Terraform, AWS, PostgreSQL, Node.js, CircleCI)* **Sept 2019-Nov 2019**
- Administered the data in **RDS** instance in **PostgreSQL, DynamoDB,** and images in **S3 buckets** using REST API
- Deployed program on **AWS Lambda** to provision e-mail notification service on API calls
- Configured **AWS CloudWatch** to monitor and setting high **CPU usage alarm** at 90% and low at 60%
- Demonstrated application deployment (**CI/CD**) over an **auto-scaling group** with a standard size of 3 instances