

SHWETALEE JITENDRA GADEKAR

Santa Clara, California 95050

+1(669)340-7910

sgadekar@scu.edu

EDUCATION

Santa Clara University, Leavey School of Business, Santa Clara, CA
Master of Science in Information Systems (Concentration - Big Data Modeling & Analytics)
GPA: 3.94 / 4.00 | Sept 2024 – June 2026 | Teaching Assistant, Winter 2025
Relevant Cousework : Data Analytics with Python & Object Oriented Programming (Java)
Club : Women in Business

SKILLS & TECHNOLOGIES

- Programming & Scripting: Python, Java, Unix
- Data Engineering Tools: Spark, Kafka, Hive, Hadoop, Sqoop, Airflow, NiFi
- Data Analytics and Visualization: Pandas, Numpy, Scikitlearn, Seaborn, Matplotlib
- Machine Learning: Classification (Logistic Regression, Decision Trees, Random Forest, SVM), Clustering (K-Means, DBSCAN), Regression (Linear, Ridge, Lasso), Naive Bayes
- Cloud Platforms: AWS (S3, EMR)
- Database & Querying: MySQL, SQL
- Version Control, Scheduling & CI/CD: Git, Crontab, Jenkins
- Data Warehousing: Hive
- Core Competencies: Data Structures, Algorithms, ETL Development, Data Modeling

ACADEMIC PROJECT

Crash Reporting Data Analysis

- Analyzed structured large-scale traffic accident data using Python, Pandas, and Seaborn to uncover patterns and correlations.
- Applied Clustering and Decision Tree algorithms to identify high-risk accident zones and factors influencing crash severity.
- Conducted extensive exploratory data analysis (EDA) to highlight peak accident times, weather impacts, and location-based trends.
- Visualized findings with Matplotlib and developed actionable insights for improved road safety policies.

EXPERIENCE

IDFC First Bank, India	Data Engineer	Feb 2024 – Aug 2024
<ul style="list-style-type: none">• Developed real-time streaming production pipelines using Kafka and Spark, enabling ingestion of 10M+ records/day from Oracle Golden Gate to HDFS.• Automated ETL processes for operational reports, ensuring 99.9% data accuracy and reducing manual intervention by 90%.• Optimized AWS S3 storage utilization, ensuring 100% compliance with data format standards.• Built data pipelines with Airflow and Nifi, enabling seamless integration with Salesforce REST API.• Collaborated with cross-functional teams to expose data pipelines through custom APIs for seamless application integration.		
Capgemini, India	Associate Consultant	March 2023 – Feb 2024
<ul style="list-style-type: none">• Automated workflows with PySpark and Shell scripts, improving efficiency by 75% and reducing errors while scheduling the jobs in Airflow, Crontab, or generating AWS EMR• Ingested and processed 5TB+ banking data/month via Sqoop and Nifi; ensured data quality.• Created Hive tables with optimized partitioning, enhancing query performance by 40%.• Delivered insights by visualizing HDFS data in Tableau and Power BI, meeting stakeholder needs.		
Crown Solutions, India	Business Associate	March 2022 – Feb 2023
<ul style="list-style-type: none">• Engineered big data solutions with Spark, Scala & SQL to transform 1TB+ client data/month for IFRS17 compliance under Tata Consultancy Services (TCS).• Collaborated with SMEs to deliver accurate data models, enabling critical financial reporting adhering to IFRS17 standards.• Deployed Spark jobs on Cloudera Distributed Platform, reducing processing time by 30% through parameterized scripts.		
Unique Mixers & Furnaces, India	Design Trainee	Feb 2020 – Sep 2021
<ul style="list-style-type: none">• Created Hive tables, with partitioning (static and dynamic), sub-partitioning, and bucketing to store the data in a meaningful manner. Creating child tables based on logical partitions.• Processed and transformed 1TB+ structured and semi-structured data using Hadoop and HiveQL.• Automated data pipelines with SERDE, achieving 95% efficiency in handling Parquet and XML files.		
Tech Mahindra, India	Business Co-ordinator	Jan 2019 – March 2019
<ul style="list-style-type: none">• Led recruitment efforts, interviewing 200+ candidates and improving hiring process efficiency by 20%.		

ACTIVITIES, ACHIEVEMENTS AND AFFILIATIONS

- **Workshops attended:**
 - 5-day workshop on Airflow, Nifi, and creating reliable data pipelines organized by Capgemini International, **2023**
 - 3-day workshop on Big Data Analytics organized by Tata Consultancy Services, Mumbai, **2022**
- **Rotaract Club of Bombay Powai, (Youth Wing Rotary International):**
 - President, **2019- 2020**
 - Secretary and President-Elect, **2018- 2019**