
EDUCATION

Syracuse University, School of Information Studies, Syracuse, NY

Aug 2018 - May 2020

MS in Information Management, C.A.S in Data Science

CGPA: 3.67

Courses: Data Science, Big Data, Data Analytics, Natural Language Processing, Business Analytics, Database Admin Concepts & Management**University of Mumbai**, Mumbai, India

Aug 2014 - May 2018

B.Tech. Computer EngineeringCGPA: 3.52

TECHNICAL SKILLS

Languages: Python, R, Spark SQL, SQL(T-SQL, SQL server, PostgreSQL), NoSQL(MongoDB), BigQuery, Git**Tools & Frameworks:** Spark, Tableau, SAP HANA, Dash, Splunk, Heroku, Azure Databricks, Hadoop, Amazon Redshift, Google Analytics**Libraries & Packages:** Scikit-Learn, MLib, NLTK, Word2Vec, Keras, TensorFlow, Matplotlib, Seaborn,**ML & Analytics:** Classification, Regression, Clustering, Neural Network, Deep Learning, Bayesian Inferences, Time Series Analysis

PROFESSIONAL EXPERIENCE

Graduate Teaching Assistant – Big Data Analytics | Syracuse University, Syracuse, NY

Jan 2020 - Present

- Assisted up to 6 students weekly with solving problems in MapReduce & Apache Spark for on-premise and cloud (Databricks/AWS)
- Conducted research in Big Data technologies like Apache Spark & NoSQL to design a new curriculum for future cohort
- Designed assignments to test ML skills for large-scale datasets using Spark SQL & MLib for a batch of 100 students

Business Data Analyst | iConsult Collaborative – Project Management Institute, Syracuse, NY

Apr 2019 - Present

- Analyzed web traffic using Google Analytics from client website of 5K users to identify KPIs for member acquisition & retention
- Designed SEO strategies, website implementation plan and web platform comparison reports to increase members by 10%
- Migrated the site from StarChapter to WordPress to improve its location, platform, SEO, content, design, UX & streamline structure
- Conducted stakeholder interviews for requirement gathering on existing system infrastructure and marketing strategy

Business Data Analyst Intern | Illumina, Inc., San Diego, CA

May 2019 - Aug 2019

- Automated processes on Splunk & Tableau by creating 12 varied interactive dashboards to reduce overall process time by 60%
- Retrieved large datasets using complex SQL queries to analyze issues faced on the eCommerce site for seamless user experience
- Liaised with software developers, offshore team and senior business analysts to solve bugs & issues reported in the backlog
- Developed business requirement documents for orders and quote pricing on SAP from various stakeholders

Business Technology Analyst | iConsult Collaborative – One Planet Edu. Network, Syracuse, NY

Aug 2018 - Apr 2019

- Optimized internal processes by planning and analysis of complex, strategic business problems solved using automated systems
- Interfaced with business stakeholders and IT teams to elicit, analyze, validate and document business requirements (BRD)
- Developed user stories to glean user requirements using data analysis/feedback for gap analysis and business process mapping

PROJECTS

Walmart Store Sales forecasting | Python, Flask, Heroku

Aug 2019 - Dec 2019

- Performed exploratory data analysis for 400K records to understand the key factors that boost the weekly sales using Python
- Compared accuracy of 4 different ML models to train, tune & deploy the best model, achieving an RMSE of 5%
- Developed a web app to perform online prediction of weekly sales for stores with Dash & deployed it on Heroku [[webapp](#)]

Loan Defaulter Prediction | Python, Spark, Databricks

Aug 2019 - Dec 2019

- Analyzed & performed ETL using Spark SQL on 4M financial records to discover patterns and trends of loan defaulters
- Handled class imbalance by under sampling the data to reduce bias & compared multiple ML models using MLib
- Tuned hyper parameters to achieve the highest F1-score of 0.76 with Random Forest

NYSE Stock Trading System | MySQL, MS Access, MS Visio

Jan 2019 - Apr 2019

- Designed and developed a database management system to store, manipulate & search stocks for an investor's portfolio
- Translated staff designed ER model into a relational schema and implemented application using MS Access and SQL
- Created tables and databases in MySQL and added functionalities like insert, delete, update, triggers and generating reports

Brain Tumor Detection and Classification | MATLAB

May 2017 - May 2018

- Built pipeline to detect & classify tumors, comprising image Pre-processing, segmentation, feature extraction & classification
- Extracted features to build a hyper-parameter tuned CNN model to isolate the tumor with an accuracy of 92%
- Packaged the classification pipeline in an application built with MATLAB employed by KJ Somaiya Hospital & Research center
- Published & Presented research paper in INDIACOM'18 [[research paper](#)]

CERTIFICATIONS

- Machine Learning by Stanford University, [Coursera](#) (Mar 2020)
- Advanced Google Analytics, [Google](#) (Mar 2020 - Mar 2023)
- Python for Data Science Essential Training, [Lynda](#) (Feb 2019)