

# YASH PASAR

Business Data Analyst Intern @ Illumina, Inc. | Graduate Teaching Assistant

Syracuse, NY | Email: [yspasar@syr.edu](mailto:yspasar@syr.edu)

LinkedIn: [yashpasar](#) | GitHub: [yashpasar](#)

## EDUCATION

**MS in Information Management | Syracuse University**, Syracuse, NY

Aug 2018 – May 2020

(Certificate of Advanced Study in Data Science)

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

**B.Tech. Computer Engineering | University of Mumbai**, India

Aug 2014 – May 2018

## PROFESSIONAL EXPERIENCE

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

Jan 2020 - Present

- Assisted students with solving problems in Hadoop (MapReduce) & Apache Spark for on-premise and cloud (Databricks/AWS)
- Conducted extensive research in the Big Data technologies like \* to design a new curriculum for future cohort
- Designed assignments to perform analytics and exercise machine learning for large-scale datasets using Spark SQL & MLlib

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

May 2019 - Aug 2019

- Analyzed large datasets using complex SQL queries to solve issues faced on the eCommerce website for seamless user experience
- Automated manual processes on Splunk & Tableau by creating 12 varied interactive dashboards to reduce process time by 60%
- Delivered business reviews to senior executives about progress and roadblocks to provide a unified solution for multiple teams
- Developed business requirement documents for orders and quote pricing on SAP from various stakeholders

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

Apr 2019 - Present

- Analyzed web traffic from client website of 5K users to identify KPIs for member acquisition & retention using Google Analytics
- Designed SEO strategies, website implementation plan and web platform comparison reports to increase members by 10%
- Conducted stakeholder interviews for requirement gathering on existing system infrastructure and marketing strategy
- Developing an improved web platform using WordPress for better data management and provide support for maintenance

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

Aug 2018 - Apr 2019

- Gathered stakeholder input and feedback through interviews and surveys to develop DFD, ERD & UML diagrams
- Identified process boundaries and determined opportunities to automate processes and functions
- Analyzed historical data to generate projections, developed key business insights and addressed data discrepancies

## PROJECTS

**Brain Tumor Detection and Classification | MATLAB**

- Built pipeline to detect & classify tumors containing Image Pre-Processing, Segmentation, Feature extraction & Classification
- Extracted features to build a fine-tuned CNN model to isolate the tumor which bumped up the prediction accuracy by 22%
- 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](#)]

**Walmart Store Sales forecasting | Python, Flask, Heroku**

- Performed EDA & data wrangling for 400K records to understand the key factors that boost the weekly sales using Python
- Built and fine-tuned different ML models to achieve an RMSE of 5%
- Developed a web app to perform online prediction of weekly sales for stores with Dash & deployed it on Heroku [[webapp](#)]

**Big Data, Big Bank | Python, Spark, Databricks**

- 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 & built multiple ML models using MLlib
- Tuned hyper parameters to achieve the highest F1-score of 0.76 with Random Forest

**Airline Customer Satisfaction Analysis | R**

- Performed exploratory data analysis and cleaning to understand 1.3M data records of airline customers
- Investigated factors influencing customer satisfaction & developed predictive models with SVM at 96% accuracy
- Delivered data-driven actionable insights for Airline companies to work on to positively impact customer satisfaction

## TECHNICAL SKILLS

**Languages:** Python, R, SQL(T-SQL, SQL server, PostgreSQL), NoSQL(MongoDB), BigQuery, Git

**Tools & Frameworks:** Spark, Tableau, SAP, Dash, Splunk, Heroku, Azure Databricks, Hadoop, Amazon Redshift, Google Analytics

**Libraries & Packages:** Scikit-Learn, MLlib, NLTK, Word2Vec, Keras, TensorFlow, Matplotlib, Seaborn,

**ML & Analytics:** Classification, Regression, Clustering, Neural Network, Deep Learning, Bayesian Inferences, Time Series Analysis

## LEADERSHIP EXPERIENCE

**Financial Secretary | Computer Society of India, KJSCE Chapter**

Jul 2016 – Jul 2017

- Maintained the accounts for the organization and implemented various measures to generate additional funds
- Led a team of 18 members and managed council affairs like conducting workshops, seminars, tech fests and technical events
- Spearheaded the PR team leading to the induction of 220 new members and acquiring sponsorship up to Rs. 55,000