**YASH PASAR** (315) 450 1430 | LinkedIn: [linkedin.com/in/yashpasar/](https://www.linkedin.com/in/yashpasar/)

5520 15th Ave NE, Seattle, WA 98105 Email: [yspasar@syr.edu](mailto:yspasar@syr.edu) | GitHub: [yashpasar](https://github.com/yashpasar)

Highly analytical and result-oriented college graduate with an insight for in depth knowledge of data and visualization. Confident mediator and presenter with analytics and business intelligence skills used to advance opportunity identification and corporate growth.

**EDUCATION** Data-driven decision-making exploratory data analysis Unix, Linux Salesforce

**MS in Information Management, C.A.S in Data Science , Syracuse University, GPA: 3.7**  *May 2020*

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

**Graduate Teaching Assistant**: Big Data Analytics, School of Information Studies (Spring Semester 2020)

**B.Tech. Computer Engineering , University of Mumbai**, **GPA: 3.52**  *May 2018*

**Courses:** Algorithms, Data Structures, Databases, Operating Systems, Software Engineering, Data Mining

**PROFESSIONAL EXPERIENCE**

# Data Analyst Intern | Illumina, Inc., San Diego, CA *May 2019 - Aug 2019*

* Developed a predictive model in Python to classify ServiceNow tickets into business domains with an accuracy of 82%. Worked on frequent pattern mining to proactively detect commonly occurring issues in service-now tickets for the e-commerce team
* Extracted and parsed through large volume of raw real-time logs of Splunk API to understand end-user behaviors, requirements, pain points, monitor failures and diagnose root causes
* Eliminated manual efforts worth ~1200 minutes per month by creating dashboards on Splunk for daily and weekly applications health checks, and alerts for various application and infrastructure events
* Created complex queries on SQL to perform a root cause & impact analysis on millions of impacted customers
* Spearheaded requirement analysis in agile environment; provided techno-functional project management support using Jira

# Data Analyst | Project Management Institute, Syracuse, NY *Apr 2019 – May 2020*

* Analyzed web traffic using Python and Google Analytics from client website of 5K users to identify key KPIs (such as click rate, bounce rate, session details) for member acquisition & retention
* Developed machine learning models (decision trees, neural nets and logistic regression) to predict individuals of a marketing campaign that are more likely to get converted into customers, with an accuracy of 92% and F-1 score of 93%
* Created Business dashboards to demonstrate multiple metrics to perform year over year analysis using Tableau and Excel

# Business Data Analyst | One Planet Edu. Network, Scituate, MA *Aug 2018 - Apr 2019*

* Consulted the client in developing an IoT cloud infrastructure for the Democratic Republic of Congo to collect CO2, Humidity and Temperature data from the surfaces of the active Volcano using The Things Network and LoRAWAN enabled microcontroller
* Built & deployed an auto encoder model from the collected data with 92% accuracy, to detect anomalies in the levels of CO2 and Air Temperature leading to possible volcanic eruption and mitigate the risks associated with lives using AWS SageMaker

# Teaching Assistant – Data Analytics | 2U, Inc., Lanham, MD *June 2020 – Present*

* Instruct and mentor 30+ data science students for a 24-week course as they learn multiple analytics technologies
* Areas of study include Python, Pandas, Jupyter, Matplotlib, APIs, JavaScript, Plotly, Tableau, MySQL, MongoDB, Flask, Heroku & Git

**PROJECTS**

# Social Media Engagement Analytics | Python, Tweepy, NLTK, SkLearn *Jan 2020 – Apr 2020*

* Leveraged Twitter API to scrape and filter tweets, retweets of a user along with metadata
* Extracted features such as n-grams, word embeddings and used classifiers like Naïve-Bayes, Logistic Regression and neural networks to categorize tweets and developed a dashboard on tableau to visualize tweets by day, topic and sentiment

# Walmart Store Sales forecasting | Python, Flask, AWS *Aug 2019 - Dec 2019*

* Performed exploratory data analysis for 400K records to recognize 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 41 stores and deployed it on AWS using EC2

# Loan Defaulter Prediction | Python, Spark, Databricks *Aug 2019 - Dec 2019*

* Performed ETL on 4M financial records using Spark SQL to discover patterns and trends of loan defaulters and leveraged demographic (like age, occupation) and financial (like FICO, income) to predict probability of defaulting on loan payment
* Developed ML models (SVM, Naïve-Bayes and random forest) & handled class imbalance by under sampling the data to reduce bias
* Used MLLib to compare models and tuned hyperparameters to achieve the highest F1-score of 0.89 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, and 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, stored procedures & databases and added functionalities like insert, delete, update, triggers and generating reports

# TECHNICAL SKILLS

**Languages:** Python, R, MySQL, MS SQL server, Spark SQL, NoSQL(MongoDB, Cassandra, Neo4j), BigQuery

**Tools & Frameworks:** Spark, Tableau, Power BI, Google Analytics, SAP ECC, Splunk, AWS, Azure

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

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

**Certifications:** Azure Fundamentals (AZ-900),Tableau Desktop Specialist

**Publications:** “Brain Tumor Detection and Classification – A Survey”, IEEE Conference, INDIACom – 2018