Rupam Jogal

Sunnyvale, CA | ☎ (857) 333-7822 | ✉ rupamjogal@gmail.com
www.linkedin.com/in/rupam-jogal-0712 | Online portfolio: https://rupam07.github.io/

SKILLS

Programming languages	R, Python (numpy, pandas, seaborn, sklearn, matplotlib,pickle), HTML, CSS, SQL, C++
Databases/ETL/Cloud Platform	SQL Server, MySQL, Spark, Cloudera, AWS (S3), SSIS, Alteryx, Hive, Databricks
Frameworks/ Platform's	PyCharm, Tableau, Microsoft PowerBI, Advanced Excel (vlookup), Jira, Confluence
IDE'S	R-Studio, Matlab, MS Office, Latex, Visual Studio, Scilab, NetBeans IDE, Eclipse, GitHub, Jupyter, Postman

EDUCATION

Northeastern University | Boston, USA | MS in Analytics – GPA: 3.75 **Apr 2020**

- **Relevant Coursework:** Probability Theory & Statistics, Data Mining Applications, Data Visualization, Predictive Analytics, Data Management and Big Data, Application of AI, Data Warehousing and SQL, Data-Driven Decision Making

Ahmedabad University | Ahmedabad, India | Bachelor of Technology in Information and Communication Technology **May 2018**

- **Relevant Coursework:** Object-Oriented Programming, Data Structures and Algorithms, Machine Learning

EXPERIENCE

Aera Technology | Mountain View, CA | Data Science Intern **Jun 2020 – Present**

- Building cognitive skills to help Pharma, CPG customers in predicting unfulfilled demand to market affiliates using lead time values based on history data, extracting features and grains values with the help of **Microsoft IDEAR** library
- Performing clustering analysis through unsupervised machine learning model, predicting lead time values, deploying it using backend **API**s in **Postman** as well as Python with model accuracy of 97%
- Implementing **Agile Scrum Software Development Lifecycle** methodology in creating test cases and monitoring the performance of the Epic and user stories using **Jira** to report and test the production environment

Northeastern University | Boston, MA | Customer Service Analyst **Dec 2018 – Mar 2020**

- Managed more than 200 students, 15 high profile faculty for providing them unrestricted access round the clock to the knowledge treasure
- Performed **ETL** using **Alteryx** and generated **dynamic dashboards** using Tableau which helped in the improvement of the inventory forecasting, designing strategies which would help in checking the available back end inventory
- Designed a Tableau frontend that displayed various inventories available which help the customers understand the shortage and availability of the inventories, performing data **quality assurance** used for improving the hardware and software technical support provided in person
- Troubleshoot an issue and raising a ticket using **ServiceNow** and help them resolve over the call

Pixometry Infosoft Pvt. Ltd. | Ahmedabad, India | Business Data Analyst **Jan 2018 - May 2018**

- Applied topic modeling with the help of **Gensim** in order to help in the risk prediction. Built a dictionary and a corpus using Python which helped in counting the words in each file, creating a **Term-Frequency – Inverse Document Frequency**
- Tokenized words further created the bigrams and trigrams from the words and performed the topic modeling with unsupervised machine learning **Latent Semantic Indexing** (LSI) algorithm which used **SVD** helped in providing the percentage contribution of each word in the files using a pre-built **Word2Vec** model and hence gave a prediction accuracy of **68%**

PROJECTS

Burial Record Image/Text Recognition **Dec 2019**

- Applied classification using **keras** and **tensorflow** in backed to the dataset with dead burial records using max-pooling and relu and classified them into 6 classes with **convolution neural network (CNN)** achieved an accuracy of **94%**
- Recognized texts with Optical Character Recognition using **opencv pytesseract** and then **AWS textract** and achieved a confidence interval with AWS and **S3** for all the images in the dataset around **95%**

Customer Churn Analysis **Aug 2019**

- Deployed a machine learning-based classification model in order to predict the churn percentage in the dataset based on the customer usage patterns and performing analysis using a classification machine learning algorithm for figuring the churn rate. Classified using **sci-kit** learn **RandomForestClassifier**, **Naïve Bayes** (GaussianNB), **kNN** (KNeighborsClassifier) and **LogisticRegression**, evaluated confusion matrix scored for every model with random forest model with the best accuracy among all other models

Sentiment Analysis **Jun 2019**

- Implemented Natural Language Processing on the dataset with textual requests for Pizza using **NLTK**, cleaning and preprocessing the data and further identifying the stop words and **stemming** and **lemmatizing** words with the help of **Parts of Speech** tags
- Vectorized the words with sci-kit learn finally applying Naïve Bayes classification and **SVM** to find out if the requester gets the pizza or not and see if their review is positive or negative

Database Design - Online Shopping Management **Oct 2018**

- Created the normalized **ER model** for the database using **Visio**
- Implemented check constraints, stored procedures, views column encryption and SQL queries that answer the question about the database in **MySQL**, creating reports based on the quality, price and geography on **PowerBI** with visualizations

Pest and Disease identification in plants **May 2018**

- Identified different stages of disease in the cotton plant, converting the RGB image into HSV to do the leaf edge detection by removing the background noise and reducing the image pixels further segmenting the leaf image with **kNN** and extracting features like texture and color for the classification of diseases with the classifier with an accuracy of **77.8%**

PROFESSIONAL DEVELOPMENT

- **Stanford University,** Machine Learning (Coursera MOOC by Andrew Ng) **Nov 2019 – Present**
- **Teaching Assistant,** Northeastern University **Dec 2019 – Mar 2020**