Rupam Jogal

Boston, MA | ☎ (857) 333-7822 | ✉ jogal.r@husky.neu.edu
www.linkedin.com/in/rupam-jogal-0712 | Online portfolio: https://rupam07.github.io/

SKILLS

Programming languages	R, Python (numpy, pandas, seaborn, sklearn, matplotlib), HTML, CSS, C, C++, SQL, MATLAB
Databases/ETL/Cloud Platform	SQL Server, MySQL, Spark, Cloudera, AWS, Microsoft Azure
Machine learning algorithms	Statistical Analysis, Deep Learning, Classification, Decision Tree, Forecasting, Support Vector Machines, KNN
Frameworks/ Platform's	Tensorflow, NLTK, PySpark, PyCharm, Tableau, Microsoft PowerBI, Advanced Excel (vlookup), Alteryx
IDE'S	R-Studio, Matlab, MS Office, Latex, Visual Studio, Scilab, NetBeans IDE, Eclipse, GitHub, Jupyter, Hadoop

EDUCATION

Northeastern University | Boston, USA **Apr 2020**

*MS in Analytics – **GPA: 3.83***

- **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 **May 2018**

Bachelor of Technology in Information and Communication Technology

- **Relevant Coursework:** Object-Oriented Programming, Data Structures, and Algorithms, Linear Algebra, Database Management Systems, Operating Systems, Machine Learning

PROJECTS

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

- Applying 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%
- Recognizing texts with Octical 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**

- Deploying a machine learning-based classification model in order to predict the churn percentage in the dataset based on the customer usage patterns and churn data whether the customer has churned or not
- 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 best accuracy among all other models

Sentiment Analysis **Jun 2019**

- Implementing Natural Language Processing on the dataset with textual requests for Pizza using **NLTK**, tokenized and lowercased the sentences and words in the data and applying the statistical model to gain information from data
- Identifying the stop words as well as punctuations and removing them to only have a gist of the document. **Stemming** and **Lemmatizing** words with the help of **Parts of Speech** tags and applying tangential note on the respective lemmatization. Vectorizing the words with sci-kit learn
- Applying Naïve Bayes classification and **SVM** to find out if the requester gets the pizza or not and see is their review is positive or negative

Online Shopping Management **Oct 2018**

- Creating the normalized **ER model** for the online shopping management database using Visio
- Implementing check constraints, stored procedures, views column encryption and SQL queries that answers 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**

- Identifying different stages of disease in the cotton plant, converting the RGB image into grayscale to do the leaf edge detection by removing the background noise and reducing the image pixels
- Segmenting the leaf image with **k-means clustering** and extracting features like texture for the classification of diseases with the neural network classifier with an accuracy of **77.8%**

EXPERIENCE

Northeastern University (Information Technology Services)| Boston, MA **Dec 2018 – Present**

Customer Service Analyst

- Performing **ETL** using Alteryx and generating **dynamic dashboards** using Tableau which helped in improvement of the inventory forecasting, designing strategies which would help in checking the available back end inventory
- Designing 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
- Submitting a Request for Change (RFC) to request service not offered at the ITS, or change a current product or service
- Troubleshooting an issue and raising a ticket using **ServiceNow** and help them resolve over the call

Pixometry Infosoft Pvt. Ltd. | Ahmedabad, India **Jan 2017 - May 2017**

Business Data Analyst

- Applying 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**
- Tokenizing 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%**

PROFESSIONAL DEVELOPMENT

- **Stanford University,** Machine Learning (Coursera MOOC by Andrew Ng) **Nov 2019 – Present**