Rohit Chandrashekhar Bapat

Bloomington, IN | (812)606-7949 | rbapat@iu.edu | rohitcbapat | github.com/rohitbapat | Tableau Resume

EDUCATION

Indiana University, *Master of Science in Data Science*, Bloomington, IN, United States **May 2020**
Coursework: Machine Learning, Statistics, Information Retrieval, Big Data Systems, Deep Learning **GPA: 3.7/4.0**

University of Pune, *Bachelor of Engineering in Computer Engineering,* Pune, India **May 2017**
Coursework: Data Mining Techniques and Applications, Business Analytics and Intelligence **GPA: 3.6/4.0**

WORK EXPERIENCE

Myxx Inc., *Data Scientist Intern*, Cary, North Carolina, US **May 2019-August 2019**

- Developed Ruby and Python scripts for ingredient mapping and retailer integration.
- Boosted the recipe ingestion success rate from 55% to approximately 85%.
- Implemented a new cosine similarity based ElasticSearch scoring technique for ingredient mapping.
- Scraped purchase history of retailers for building user profiles with Javascript and Puppeteer packages.

TIBCO Software Inc., *Junior Consultant- Business Process Management,* Pune, India **July 2017-July 2018**

- Designed and modeled the business processes for client in telecom domain and interfaced TIBCO products with external third party resources.
- Used Spotfire- TIBCO Analytic Tool to determine time required for process execution and staged reports.

Persistent Systems Pvt. Ltd., *Project Intern*, Pune, India **June 2016-June 2017**

- Developed a business intelligence based solution to provide the retail store chain with analytics based on customer footfalls, inventory management, and sales trends by generating role specific PDF reports.

TECHNICAL SKILLS

- **Database:** SQL, MySQL, MongoDB.
- **Languages:** Python, JAVA, R, Javascript.
- **Applications:** Apache Storm, Tableau, Lucene Indexing, Apache Spark, GIT, Jenkins, ElasticSearch, Qlik.
- **Frameworks:** Pandas, NumPy, sklearn, matplotlib, seaborn, tensorflow.
- **Models:** KNN, Adaboost, Random Forests, SVM, K–Means, Regression models.
- **Deep Learning:** DNN, CNN, RNN, LSTM.

PROJECTS

Deep Learning Projects *Python | tensorflow 2.x | Numpy | Google Colab* **October 2019**

- Achieved 98.2% accuracy with MNIST digit recognition data with fully connected Deep Neural Network.
- Successfully extracted clean audio from a noisy signal using 1D and 2D Convolutional Neural Networks.

Yelp Dataset Challenge *Lucene | Python Flask | Pandas | MongoDB* **May 2019**

- Predicted the cuisine data of restaurants and evaluated results using TRECEVAL.
- Implemented collaborative filtering and cosine similarities for user input based recommendations.
- Achieved Mean Reciprocal Rank of 0.85 and Precision of 0.79 using parameter tuned LDA.

IndyCar - Performance Analysis of Anomaly Detection Application *Apache Storm | Tableau | Python* **May 2019**

- Performed latency and performance analysis on IndyCar race data.
- Successfully deployed a Storm topology using .yaml files, Apache MQTT pub-sub broker, Zookeeper.
- Used Apache Storm for streaming analysis and Tableau for visualizations and insights.

Movie Rating Prediction with Text, Emoji Ideogram Data on Twitter *Pandas | Python | tweepy* **April 2019**

- Used IMDb Data to predict Movie Rating based on the movie related Twitter Data using 'tweepy' API.
- Achieved AUC score of 0.78 for Light GBM and 0.85 F1 Score for SVM Classifier for Emoji based analysis

Image Orientation Classification *Python | Pandas | Numpy | matplotlib* **December 2018**

- Implemented Adaboost technique from scratch to identify image orientation ($0^0/90^0/180^0/270^0$)
- Achieved overall accuracy of 69.48% for test images with 4000 weak classifiers.

DonorsChoose.org Application Screening *Pandas | Python | sklearn | matplotlib* **November 2018**

- Predicted application decision of DonorsChoose.org application dataset via Kaggle competition.
- Applied Natural Language Processing (NLTK) coupled with inflect libraries and textBlob packages.
- Used scikit-learn models of Light GBM (AUC 0.766) and K-Means for prediction and essay review.

Route Optimization with Search Algorithms *Python* **September 2018**

- Worked on roadways data consisting of highways with speed limits and distances.
- Implemented algorithms like BFS, DFS, A-star, IDS, and Uniform Cost Search to find optimal paths.