

Gourang Patel

Boston, MA 02115 | (551) 248-3469

patel.gou@northeastern.edu | www.linkedin.com/in/gourang-patel | <https://github.com/Gourang97>

EDUCATION

Northeastern University, Boston, MA

Sept 2019 - Present

Khoury College of Computer Sciences

Expected Graduation: Dec 2021

Candidate for a Master of Science in Data Science

Teaching Assistant – Mathematics for Data Modeling [Jan 2020 – May 2020], Foundations of Artificial Intelligence [June 2020 – July 2020]

Related Courses: Linear Algebra for Data Science, Supervised Machine Learning, Data Management and Preprocessing, Unsupervised Machine Learning, Causal Inference in Machine Learning, Natural Language Processing, Deep Learning

Nirma University, Ahmedabad, India

Bachelor of Technology in Information Technology

Aug 2015 – May 2019

Related Courses: Operating Systems, Database Management Systems, Theory of Computation, Data Mining, ML, Data Structures, Big Data Analytics, Business Analysis and Optimization, Linear Algebra and Probability

TECHNICAL KNOWLEDGE

Programming Languages & Tools: Python, Pyro, R, SQL, JavaScript, C, C++, Tableau, PowerBI

Python Libraries: NumPy, Pandas, Matplotlib, Scikit-learn, Keras, SparkMLlib, OpenCV, TensorFlow, Pytorch, Transformers.

Machine Learning Algorithms: Regression, LDA, Logistic regression, KNN, SVM, K-Means, Random Forest, ANN, RNN, YOLO v3, Deep Neural Networks, Ensemble Learning, Clustering, PCA, SVD, BERT, Attention Models.

PROFESSIONAL EXPERIENCE

Axle Informatics, Rockville, Maryland, USA

Aug 2020 – Dec 2020

Data Science Intern

- Designed and developed a **Computer Vision Annotation** tool, an end-to-end application for auto and manual annotation of **CT scanned images/videos** and provided mechanism for dumping the annotations in various standard formats like **COCO, PASCAL VOC, YOLOv3**.
- Implemented **Image Recognition** and **Segmentation** techniques for the annotated images/videos dataset using state-of-the-art Deep learning models like **Faster RCNN, Mask RCNN, Yolo v3**. The dataset was stored in **AWS S3** buckets and model was trained on **AWS Sagemaker** Notebook Instance.
- Build the application using **Angular** Framework [Frontend], **Django** [Backend], **Docker** [Containerization]. Annotations exported from the application served as training datasets for the Auto-annotation models.

Crest Data Systems, Ahmedabad, Gujarat, India

Jan 2019 – July 2019

Data Analyst Intern

- Performed **Exploratory Data Analysis** using **SQL, Python, Tableau** on data collected from various endpoint devices using various API's, and **ELK stack**, also provided better visualizations by identifying key features using **feature engineering** and **Natural Language Understanding** and computed new incident id's causing threat and malicious contents into the system.
- Developed ML Security pipeline and integrated it with **IBM QRadar** and **Splunk platform** for information gathered from the Client's Rest APIs using **Python Flask, JavaScript, SQL, MongoDB, machine learning models**.

ACADEMIC PROJECTS

Image Based Product Recommendation System – An Unsupervised Way

Jan 2020 – Apr 2020

Northeastern University, Boston, MA

- Applied **feature reduction technique** and **Image Preprocessing** on images using **PCA** and, **OpenCV** performed **clustering** methods **K-Means, Gaussian Mixture Model and Hierarchical Clustering** on image vector. Hierarchical Clustering performed best to separate different categories of products.
- Recommended** similar images to the user 85% of the time using cosine similarity and spatial distance matrices.

openGAVY Package - An Open-Source Pipeline for Data Science Projects

Sept 2019 – Dec 2019

Northeastern University, Boston, MA

- Designed and Developed a package which perform Extract – Transform - Load (**ETL**), Exploratory Data Analysis (**EDA**), **Modelling** and **Visualization** functionalities on regressions and binary **classification** datasets.
- Incorporated unique functionalities like recommended plots, **binning data, information values** for feature selection.
- Simplified the process of implementing various data science projects as the lines of codes reduced by approx. 85%.

House Credit Default Risk Prediction- Supervised Machine Learning Models

Sept 2019 – Dec 2019

Northeastern University, Boston, MA

- Developed unique approach which involves feature engineering techniques to determine various parameters to provide loans and thereby removing dependencies of bank that rely on credit histories.
- Implemented Models like **Naive Bayes, Logistic Regression, Multilayer Perceptron, Decision Trees and, Random Forest** in **SparkMLlib** to **predict** and **classify** the customers loan eligibility and eliminated the use of credit histories for loans.