# Manal Shah

manal.shah@stonybrook.edu | manal3112.github.io | github.com/manal3112 | linkedin.com/in/manalshah31

## Education

**Stony Brook University** — Stony Brook, NY
Master of Science in Computer Science | GPA: 3.86/4 — May 2023 (Expected)

**Courses:** Machine Learning, Data Science Fundamentals, Computer Vision, Big Data Analytics, Probability and Statistics

**Malaviya National Institute of Technology** — Jaipur, India
Bachelor of Technology in Computer Science and Engineering | GPA: 8.31/10 — May 2018

## Technical Skills

**Languages**: Python, Java, JavaScript, TypeScript, Enaml, PHP, MySQL, HTML/CSS
**Libraries and Tools:** PyTorch , PySpark, React, Dash, Keras, Tensorflow, OpenCV, Hugging Face, Pandas, NumPy, Hive, HDFS, AWS SageMaker, Domino Data Lab, GCP, Git, JIRA, Apache Ambari

## Experience

- **New York Life Insurance Co.** — New York, NY
  **ML Engineer Intern** — May 2022 - Present
  - Built a web application on Domino Data Lab using Web-HDFS API and Dash framework for Mortality Risk Classification model project to automate the model validation and generate reason code analytics.
  - Worked on Mortality Risk Classification model wrapper and corner case testing.
- **Stony Brook University** — Stony Brook, NY
  **Graduate Research Assistant at HLAB** — Feb 2022 –Present
  - Working on fine-tuning masked language modeling (MLM) with RoBERTa, BERTweet as seed to capture the language use in social media for token and sequence classification tasks.
- **JP Morgan Chase & Co.** — Mumbai, India (3 Years)
  **Associate Software Engineer** — Feb 2021 – Aug 2021
  - Worked with Quant Research team to build dynamic schemas to store calibration outputs of pricing models in Hydra. Re-designed workflows to incorporate LIBOR to RFR migration of derivative products.
  - Led a team of developers in a global team hackathon and developed a critical business web application prototype using React.JS and TypeScript, deploying it on the private cloud - GAIA.

  **Software Engineer** — Jul 2018 – Jan 2021
  - Onboarded Price Valuation automated jobs workflows for Fixed Income Finance business in JPMC's Athena framework saving annually 800+ person-hours.
  - Implemented external vendor data snapping APIs using encrypted connections in a client-server application to snap data from Bloomberg, Reuters and other broker firms.
  - Re-engineered legacy processes by significant design changes and efficient use of computing resources.
  - Provided production support for the live dashboard of workflow processes of various business teams.

  **Summer Intern** — May 2017 – Jul 2017
  - Developed an application to query trade data and plot informational views using the Enaml framework.
  - Adapted to the Agile model of software development and CI/CD pipelines for releases.

## Publications

- **Journal Publication:** AVDNet: A Small- Sized Vehicle Detection Network for Aerial Visual Data. [ IEEE Geoscience and Remote Sensing Letters ( Volume: 17, Issue: 3, March 2020) ]
- **Conference Publication:** SSSDET: Simple Short and Shallow Network for Resource Efficient Vehicle Detection in Aerial Scenes. [ IEEE International Conference on Image Processing (ICIP) 2019 ]

## Projects

**Situational Crime Prevention** — April 2022 – May 2022
- Performed similarity search using MinHashing and Locality Sensitive Hashing (LSH) algorithm to determine spatial (county-level) stability of crime patterns in the US over 20 years using PySpark and HDFS on GCP.
- Conducted hypothesis testing to determine correlations of crime types with unemployment and poverty.

**Retrieval Patterns of Physical Objects from Retail Sales** — Oct 2021 – Dec 2021
- Implemented Prod2Vec algorithm which uses the local product co-occurrence information established by the product sequences to create a distributed representation of products and recommends similar products.
- Performed comparative analysis of GloVe, FastText and Word2Vec models on retail product embeddings.

**First-Person Activity Recognition in Videos** — Nov 2021 – Dec 2021
- Finetuned ImageNet pre-trained Resnet50 CNN model to serve as a feature extractor and using it performed classification task on JPL interaction video frames to identify human activities in PyTorch and OpenCV.
- Temporal pooling was used to extract spatial features by using fusion of ResNet50 and VGG16.