

Nishit Mistry

Data Scientist

Education

- Master of Applied Computer Science / Dalhousie University, Halifax.
January 2022 – April 2023 Grade: A+
- Bachelor of Engineering in Information Technology / Shah & Anchor
Kutchhi Engineering College, Mumbai.
August 2017 – June 2021 GPA: 3.7

Work Experience

Data Scientist / Climative, Halifax.

January 2023 – April 2023

- Performed **pre-processing, data exploration, and cleaning** of **large structured and unstructured climate** data using **R** (dplyr, tidyverse), **Python** (NumPy, Pandas, Scikit-learn), and **SQL** queries.
- Evaluated **Model Performance, Quality, and Validity** using advanced **statistical** methods such as **RMSE, Mean, Median, Accuracy Evaluation**, and various **plots**.
- Created **Data Visualizations** using **Power BI, Spreadsheets, and Matplotlib** to effectively **present information** to **non-technical** team members.
- **Deployed** models using **ETL Flow** on the **cloud** using **Azure ML Studio** and analyzed output for data-driven decision-making.
- Developed **GBM-based XGBoost, GAM-based models, and clusters** using **Gower distance** and **PAM**, resulting in a **25% Improvement in Performance and Understandability**.
- **Helped** the company **acquire clients** through **data-driven** work and **machine-learning** models.
- Worked in **Agile Methodology**, utilized **Jira** for task tracking, and participated in **SCRUM** calls for efficient **Project Management**.

Data Scientist / Bombay Leprosy Project, Mumbai.

November 2019 - February 2021

- **Led a team of 6 people** and helped **solve** the **business problem of identification of Leprosy** at its early stages by developing **Deep Learning Models** for **Object Identification** using **TensorFlow GPU**.
- Conducted **Data pre-processing** including **manipulation, transformation, and visualization** of the dataset using **SQL** queries and **Python (Seaborn)**.
- Performed **Feature Extraction** and **Outlier Detection** while performing **Data Modelling** with **training CNN models** like **InceptionV3** using **Transfer Learning** and **Ensemble Learning**, resulting in a **35% Improvement in accuracy**.
- **Validated** and **evaluated** the model results using various statistical metrics such as **Precision, Recall, F1-Score, and ROC Curve** using **scikit-learn**.

Contact

🏠 Toronto, Ontario, Canada.

✉ nmistry45@gmail.com

☎ +1-902-412-8376

🌐 <https://www.linkedin.com/in/nishitmistry24>

🔗 <https://github.com/nmistry45>

🌐 <http://nmistry45.pythonanywhere.com/>

Skills

- Python
- Pandas, NumPy, scikit-learn
- Matplotlib, Seaborn, Plotly
- TensorFlow, Keras, PyTorch
- R
- Dplyr, tidyverse, ggplot2, data.table
- SQL/NoSQL
- Power BI, Tableau
- Microsoft Azure
- ETL Pipeline
- AWS, Google Cloud Platform
- Git
- ReactJS, NodeJS
- Microsoft 365, SharePoint
- Statistics
- Agile Methodology, SDLC
- Docker
- Communication Skills
- Machine Learning, Deep Learning
- RNN, CNN, NLP
- CI / CD

Data Scientist / Shah & Anchor Kutchhi Engineering College, Mumbai.

December 2019 - March 2020

- **Designed and implemented a Machine Learning Supervised** model using **Keras** for Educational Help, including **Data pre-processing** and **Anomaly Detection** using **Python** and **SQL**.
- **Created a Random Forest Classifier Ensemble** to **predict** patterns and analyze whether a student requires assistance based on **previous data** such as scores, resulting in a **20% Improvement in Confidence**.
- Implemented **storage** functionality using **Firebase** and developed features for uploading and downloading study materials such as E-Books, PDFs, and important notes.

Projects

Stock Price Prediction using LSTM

- **Led a team** to create a **Stock Price Prediction** System using **LSTM**.
- Conducted **Data Exploration, Analysis, and Transformation** using **SQL** queries and **Feature Selection** using **Grid Search, Random Search** and handled **Data Quality Issues** while **tuning the hyperparameters** on a **time-series** dataset.
- **Performed Data Modelling** by **developing an LSTM model** and **Visualized** various **graphs** using **matplotlib** and **Seaborn** to gauge the prediction accurately with proper **evaluation metrics** and **learning curves**.

Cropifier: Crop Identification Mobile Application

- **Led a team of 6 people** to build a mobile application for **ISRO** in **Smart India Hackathon 2020 - Software Edition**.
- **Generated** dataset from various sources, **explored** data using **R**, performed **data analysis, image pre-processing, data labeling, and data augmentation** using **Numpy** and **Pandas** in **Python**.
- **Trained CNN-based models - InceptionV3 and SSD MobileNetV2** using **Transfer Learning** with **TensorFlow** and **Google Colab**.

Invoicify

- Developed a **Cloud** application on the **AWS** platform using **ReactJS, NodeJS, and MongoDB** that **helps businesses** to manage their invoices and payments while **improving user accessibility by 45%**.
- The system was created using **MERN Stack** with a proper file structure.
- The back end was entirely created using **AWS Services** like **Elastic Beanstalk, Lambda, SQS, S3, Cognito, DynamoDB, API Gateway, CloudWatch, and EC2**.

Publications

- ParaSpeak: A Speaking System for Paralyzed – GIS Science Journal <http://gisscience.net/VOLUME-7-ISSUE-6-2020/>

Awards

- Chunautee '21 Winner – IIC YMCA, Faridabad
- Certificate of Merit (Topper of Department) – Shah & Anchor Kutchhi Engineering College
- Smart India Hackathon 2020 – Software Edition Winner – Government of India
- 1st Runner Up at Intra College Robotics Competition – SAKEC Robo-Club

Certificates

- Python (Basic) – HackerRank
- Complete Python Developer in 2021 – Udemy
- Introduction to Deep Learning (with Honors) – Coursera
- Introduction to Machine Learning – NPTEL

Volunteer Experience

- O – Week Leader – Dalhousie Student Union
- Food Bank Volunteer – Dalhousie Student Union
- Outreach Volunteer – Dalhousie Student Union

Languages

- English (Proficient)
- Gujarati (Native Tongue)
- Hindi (Proficient)