

# Nikhil Naik

Jersey City, NJ 07307 | [nnaik2@stevens.edu](mailto:nnaik2@stevens.edu) | [LinkedIn](#) | [GitHub](#) | +1 (917)282-6968

## EDUCATION

**Stevens Institute of Technology**, Hoboken, New Jersey

May 2024 (Expected)

Masters in Machine Learning | Current GPA: 3.965

Courses: Deep Learning, Data Acquisition for DL, Big Data Tech, Knowledge Discovery, 3D & 2D CV, ML Fundamentals & Application

**Datta Meghe College of Engineering**, Mumbai, India

May 2018

Bachelor of Engineering in Electronics

## TECHNICAL SKILLS

Programming Technologies:	Python, C, C++, JavaScript, Shell Script, HTML, CSS, Bootstrap, jQuery, Ajax, ExtJS, REST API, JSON	
Databases:	MySQL, SQLite, SQL, Distributed Systems	
Frameworks:	Apache Spark, Hadoop, TensorFlow, PyTorch, OpenCV, Web2py, Django, Flask, Selenium	
Other Tools & Tech:	Data Visualization, Statistical Analysis, Cloud Computing, Computer Vision, NLP, SageMaker Linux (CentOS), Docker, Apache Web Servers, scikit-learn, NumPy, Pandas, Full Stack, SDLC	
Certifications:	AWS: Solutions Architect Associate - <a href="#">Cert</a>	IIT Roorkee's Applied Data Science - <a href="#">Cert</a>
	Fundamentals of Data Mining by Pirple - <a href="#">Cert</a>	UCSD's Data Structures, Coursera - <a href="#">Cert</a>
	UCSD's Algorithmic Toolbox, Coursera - <a href="#">Cert</a>	C programming Seed Infotech

## PROFESSIONAL EXPERIENCE

**Gajshield Infotech**, Mumbai, India | **Software Developer**

Sept 2018 – Mar 2021

- Improved efficiency of Gajshield Firewall's email security system by single-handedly researching and developing outlier detection machine learning model which filtered malicious emails with an accuracy of 97%
- Architected and managed implementation of Centralized VPN control project for Gajshield Firewall using Python, Web2py framework and open-source VPN in Linux adding robust management capability, increasing reliability of Gajshield VPN services
- Led a team of 4 to implement SDWAN capability in all 30 models of Gajshield firewall by segmenting and benchmarking network traffic to automatically switch primary network path in accordance with user defined policies using Python, C++ and Linux modules
- Developed an NLP based Contextually Intelligent Troubleshoot Assistant using open-source RASA chatbot and Flask web server, which guided L1 support team to solve basic customer issues, reducing load on L2 engineers and developers by a factor of 50%
- Acted as Gajshield's Technical representative and collaborated with Bharat Petroleum Corporation, architected a Zero Touch Provisioning System to deploy over 1000 Firewalls with centralized Cloud Management System for remote management
- Enhanced and transformed frontend of Gajshield Firewall's UI, utilizing Web2py, Bootstrap, ExtJS, Pygal and Chartjs
- Delivered a technical presentation demonstrating major upgrades to Gajshield's new Firewall during its international launch event
- Studied and optimized development team's SDLC process to improve productivity by incorporating Kanban Agile techniques

## RESEARCH & TECHNICAL PROJECTS

**WMT 2024 – Machine Translation Research Competition** (*Hugging Face, FairSeq, Transfer Learning, Generative AI*)

Present

- Representing Stevens in WMT 2024 conference, showcasing expertise in machine translation, contributing to cutting-edge research
- Actively researching previous winners, developing an advanced LLM pipeline, poised for fine-tuning based on competition's task

**Foreign Exchange Rate Prediction – Deep Learning** (*Colab, Keras, Keras\_Tuner, LSTM, RNN, Python v3.7*) - [GitHub](#)

Mar 2023

- Developed RNN and LSTM neural networks and Hyperparameter tuned the best model for 0.0046 MSE and 0.45 R2 score
- Architected a baseline model with In-Domain generalized data and further fine-tuned model using transfer learning techniques

**CitiBike User Usage Forecast – Big Data** (*GCP, Spark, MLlib, SQL, Zeppelin notebook, Python v3.7*) - [GitHub](#)

Apr 2023

- Leveraged PySpark to perform ETL on 40GB of CitiBike & weather data, creating big dataset to train a distributed regression model
- Successfully executed on Google Cloud Platform using Dataproc, demonstrating scalability & performance in handling big data

**University Recommendation System** (*SkLearn - Cluster, Recommendation System, Jupyter notebook, Python v3.7*) - [GitHub](#)

Nov 2022

- Used Elbo and Silhouette Cluster Analysis to create University Tiers and used historic data to create a ML model for prediction
- Compared various models using ROC curves with Random Forest having best performance of AUC – 80.2 %

**Telemarketing Campaign Prediction** (*SkLearn, ImbLearn module, GridSearchCV Jupyter notebook, Python v3.7*) - [GitHub](#)

Sept 2021

- Performed Exploratory Data Analysis to gain data insights, selected and HyperParameter tuned best Machine Learning algorithm
- Awarded completion, internship certificate by UniCoverage Technologies for achieving accuracy of 95% employing KNN Classifier

**Sign Language Interpreter Glove** (*Arduino, Bluetooth, MPU-6050 Gyroscope & Accelerometer, Flex & contact sensors*)

Apr 2018

- Built a Glove to translate ASL to English utilizing Arduino Uno, with User Interface on Android/windows application
- Paper published on proprietary gesture detection logic and participated in a national project competition organized by IETE-DMCE

## PUBLICATION

Title: Sign Language Interpreter for Hearing and Speech Impaired Community

Apr 2018

Details: IEEE Bombay Section's ICATE-2K18 Event | ISBN: 978-93-5267-422-0