

# Rajeev Joshi

✉ [rj1234@nyu.edu](mailto:rj1234@nyu.edu) ☎ +1-347-216-4621 🌐 [www.github.com/rj-1234](https://www.github.com/rj-1234)  
🌐 [www.rj-1234.github.io](https://www.rj-1234.github.io) in [www.linkedin.com/in/rj1234](https://www.linkedin.com/in/rj1234)

## EDUCATION

<b>New York University, New York USA</b>	May 2019
Master of Science in Computer Science	GPA: 3.77/4.0
<b>Academic Achievement Award</b>	Nov 2018
• <b>Teaching Assistant</b> for Introduction to Operating Systems CS-GY 6233.	Fall 2018, Spring 2019
• <b>Teaching Assistant</b> for Information Security and Privacy CS-GY 6813.	Fall 2018
<b>Maharaja Agrasen Institute of Technology, Delhi India</b>	Jul 2017
Bachelor of Technology in Computer Science & Engineering	GPA: 69.37/100

## TECHNICAL SKILLS

*Programming Languages:* Python, JavaScript, SQL, HTML, CSS, PHP, Shell Scripting.  
*Softwares/Tools:* Flask, CodeIgniter, Web2py, OpenCV, Git, Docker, PostgreSQL, Oracle.  
*Big Data:* Spark, YARN, Map-Reduce, HDFS, Hive.

## EXPERIENCE

<b>Machine Learning Engineer</b> , The Vanguard Group., Malvern, PA	Aug 2019 – Present
<ul style="list-style-type: none"><li>Engineered a scalable multiprocessing pipeline within the vendor's cloud platform to pre-process and run inference on roughly 85+ million rows of structured data (70% reduction in execution time) with testing and notification capability solely using python and shell scripting.</li><li>Designed and implemented an event based database architecture(long term non volatile and short term volatile memory) for processing large amounts of requests at the backend of an in house audio transcription engine.</li><li>Extensive experience building scalable NLP pipelines to Extract-Load-Transform-Clean utilizing multi-processing.</li><li>Developed an in house NLP data labelling platform to fast track label generation and model development.</li><li>Deployed an enterprise wide API (&lt;0.5 sec latency) serving fund similarity backed by periodic model refresh.</li><li>Developed an in house package to accelerate Exploratory Data Analysis phase which automatically scales based on the underlying data to reduce model development cycle.</li><li>Built an extensive suite of web-based applications to aid and assist leaders in integration of ML-Ops and Model governance practices.</li><li><i>Technologies:</i> Python, BitBucket, Spark, YARN, S3, Domino Data Science Lab, Gensim, sklearn, Keras, Tensorflow, Jira, Confluence.</li></ul>	
<b>Software Intern (R&amp;D)</b> , BotFactory Inc, New York City, USA	Jun – Aug 2018
<ul style="list-style-type: none"><li>Developed computer vision algorithms using OpenCV to improve the accuracy of the Pick-n-Place Head.</li><li>Designed (UX/UI) a critical webpage for a new feature implementation in the next version of the product.</li><li>Migrated Instruction set of hardware abstraction layer for the previous model to the latest software architecture.</li><li><i>Technologies:</i> Web2py, OpenCV, Git, Linux, JavaScript, Jira.</li></ul>	

## ACADEMIC PROJECTS

<b>CodeJudger</b>	Feb 2019
<ul style="list-style-type: none"><li>An In-house tool for professors to allow real-time in class coding test for students. 🌐</li><li>Developed using the DevOps/Agile methodology of Continuous Integration/ Continuous Deployment. 🌐</li><li><i>Technologies:</i> Docker-compose, Travis CI, CodeIgniter PHP framework, Github.</li></ul>	
<b>Protein Residue Distance Matrix Prediction</b> 🌐	Sep 2018
<ul style="list-style-type: none"><li>Given the primary and secondary structure of a protein sequence, we created a deep neural network model to predict the distance matrix containing the distances between each residue within the given protein sequence.</li><li><i>Technologies:</i> Keras, pandas, numpy, Jupyter Notebook.</li></ul>	
<b>Computer Vision Projects</b>	Aug 2018
<ul style="list-style-type: none"><li>Human Detector - A 2-layer Perceptron that uses an HOG vector to find Humans in a 2D color picture. 🌐</li><li>Canny Edge Detector - Implementation of a Canny Edge Detector including all four steps i.e. Gaussian Smoothing, Gradient operation, Non-Maxima suppression and Thresholding. 🌐</li><li><i>Technologies:</i> Python.</li></ul>	
<b>Rating Movies and Predicting Movie Ratings</b> 🌐	May 2018
<ul style="list-style-type: none"><li>Designed a model to rate movies based on IMDB ratings and user tweets by doing Twitter Sentiment Analysis.</li><li>Built a model to accurately predict movie ratings based on several key features using IMDB dataset.</li><li><i>Technologies:</i> Spark, Twitter API, Google Cloud, Text Blob, Zeppelin, Jupyter, Python, HTML, CSS, JavaScript.</li></ul>	

## EXTRA CURRICULAR

- Hackathons - Social Good Hackathon, HackNY Spring 2018, HackNY Fall 2018 and HackNYU 2018.
- NYU 2018 Commencement Volunteer.