

ANKIT VAGHELA

ankit.vaghela30@gmail.com | [linkedin.com/in/ankit-vaghela](https://www.linkedin.com/in/ankit-vaghela) | (+1) 706-461-4428 | github.com/ankit-vaghela30
Website: <https://ankit-vaghela30.github.io/>

EDUCATION

University of Georgia, Athens

Master of Science in Computer Science

Relevant Coursework: Data Science Practicum, Advanced Data Analytics, Algorithms, Distributed Computing Systems

Expected Graduation: May 2019

Current GPA: 3.86/4.00

Institute of Technology, Nirma University

Bachelor of Technology in Instrumentation and Control

Graduated: May 2013

TECHNICAL SKILLS

Programming Languages: Python, R, MATLAB, Java, Javascript, Node.JS, HTML, CSS

Big Data Technologies: Apache Spark, Hadoop, Hive, Pig

Machine Learning Skills: Classification - Naive Bayes, Random Forest, Logistic Regression, Decision Trees | Clustering - K-Means | Regression - Linear Regression, Logistic Regression | Deep Learning - CNN, UNET, Tiramisu, VGG16

Data Visualization/Reporting: matplotlib, HBase, Excel

Data Science/Machine Learning Libraries: Weka, Tensorflow, Keras, Theano, NumPy, Scikit learn, Pandas, NLTK

Cloud Services: Google Cloud Platform, Amazon Web Services

ACADEMIC PROJECTS

- **American sign language detection (September/2018 - November/2018):** Created a real-time sign language detection platform which can be used by disabled people to communicate on video based chat platform. I learned how difficult it is to choose best dataset for given problem and how to deploy Deep learning model on user-facing web application.
- **Google Landmark Recognition Challenge (Kaggle Competition - April/2018 - May/2018):** Developed a CNN model that can predict landmark labels directly from image pixels and help people better understand and organize their photo collection. I learned how to analyze and visualize big dataset, perform data cleanup and handle infrastructure limitations.
- **Cilia Segmentation (April/2018):** Developed a model which can segment moving cilia from time-series images of human body cells. I worked on Tiramisu model and gained knowledge about segmentation problem in biomedical images.
- **Neuron Finding (March/2018):** Created a model to find coordinates of region of interests (neurons) from time-series images of calcium activated neurons. I worked on UNET model and learned how time-series data can be complex to work upon and tricks like masking can help.
- **Microsoft Malware Classification challenge (February/2018):** Created Naive Bayes and Random forest classifiers to classify malware data. I learned to work on distributed dataset using RDDs of Spark. I also realized Natural language processing model building complexity especially while working with clusters on cloud services infrastructure.
- **Documents Classification using Naive Bayes (January/2018):** Classified corpus of documents using Naive Bayes. I learned basic NLP preprocessing like stop-words removal, TFIDF feature extraction and stemming.

WORK EXPERIENCE

Software Engineer Intern – Verifacto Inc., Atlanta GA

June/2018 - August/2018

- Integrated insurance data from diverse sources of vehicle insurance providers by creating a pipeline in python which resulted in saving 30% manual labour.

Software Engineer – Infosys Ltd. & Cognizant, India

October/2013 - July/2017

- Created a pipeline in Hadoop using MapReduce and Hive to filter customer data and visualize reports in HBase.
- Developed Java based web services, web applications and designed databases for multiple clients.

ACHIEVEMENTS

- Won 3rd place in MLH (Major League Hacking) Hackathon organized at Georgia State University (March 2018).
- “American sign language detection” project was voted as best project in Advanced Data Analytics course.