

ANKIT VAGHELA

ankit.vaghela30@gmail.com | [linkedin.com/in/ankit-vaghela](https://www.linkedin.com/in/ankit-vaghela) | github.com/ankit-vaghela30

Website & Blog: <https://ankit-vaghela30.github.io/>

EDUCATION

University of Georgia, Athens

Master of Science in Computer Science

Graduated: May 2019

GPA: 3.90/4.00

Institute of Technology, Nirma University

Bachelor of Technology in Instrumentation and Control

Graduated: May 2013

TECHNICAL SKILLS

Programming Languages: Python, SQL, Java, Javascript, HTML, CSS

Big Data Technologies: Hadoop, Apache Spark, 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

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

Cloud Services: Google Cloud Platform, Amazon Web Services

WORK EXPERIENCE

Data Scientist – ADP, Alpharetta GA

07/2020 - Present

- Responsible for designing end to end MLOps infrastructure in AWS for our Machine learning models lifecycle.
- ML privacy and security lead: Worked on automatically detecting sensitive PII's from pdfs and mask them. Used NER model Spacy for identifying PII's. Deployed the model as an API.
- Used Deep learning models on EMNIST dataset for handwritten text/number recognition in wage forms.
- AI for QA: Worked on a model to select which tests to run based on git changelog which reduced build time by 30%.
- Worked with fellow Data scientists to help them integrate and deploy ML models in the pipeline.
- Responsible for designing a Model drift and data shift application which can automatically trigger model retraining.
- Responsible for creating end to end testing framework BDD for the document extraction ML pipeline.
- Code quality lead: Responsible for code reviews and making sure coding standards are maintained in the repo.

Software Engineer – NCR Corporation, Atlanta GA

08/2019 - 07/2020

- Worked with a cross-functional team to come up with a marketing strategy which can interest customers filling gas at the NCR's fuel kiosk - Optic to visit the inside store.
- Based on the marketing strategy, developed a module in C++ to fetch Customer loyalty information using web services and showed the offers and coupons on a javascript based UI of Optic device.

Graduate Research Assistant - Schmitz Lab, Athens GA

09/2018 - 05/2019

- Created a web application using Javascript and MongoDB based stack which teaches UGA's undergraduate biology students about genetics and inheritance through simulated experiments: <https://heredity.genetics.uga.edu/>

Software Engineer Intern – Verifacto Inc., Atlanta GA

06/2018 - 08/2018

- Integrated data from vehicle insurance providers by creating a pipeline in python which reduced manual labour.

Associate – Cognizant Ltd., India

10/2016 - 07/2017

- Provided consulting for a Salesforce based telecom retail application to a Netherlands based telecom giant.
- Developed features for order orchestration flow using apex code and lightning UI in Agile SDLC environment.

Senior Software Engineer – Infosys Ltd., India

10/2013 - 10/2016

- Created a pipeline in Hadoop to filter customer data based on features like location, age etc. to show different type of reports in HBase. Used Pig, Hive and MapReduce for processing and warehousing.
- Worked on Java based reporting application to mask sensitive banking information such as TIN from non-privileged employees and provided functionality to unmask it for employees having the privilege.

- Designed a Java tool to convert excel based records to SQL Insert/Update queries which saved tremendous amount of time for the client account and I was given “Infosys Insta award” for this idea implementation

ACADEMIC PROJECTS

- **Video Chat Platform for Sign Language (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 platforms.
- **Google Landmark Recognition Challenge (Kaggle Competition - April/2018 - May/2018):** Developed a CNN model that can predict landmark labels directly from image pixels which can help organize photo collection.
- **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 stemmin

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.