

# ANKIT VAGHELA

[ankit.vaghela30@gmail.com](mailto:ankit.vaghela30@gmail.com) | [linkedin.com/in/ankit-vaghela](https://www.linkedin.com/in/ankit-vaghela) | (+1) 706-461-4428 | [github.com/ankit-vaghela30](https://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:** Java, Python, Scala, SQL, Javascript, HTML, CSS

**ML Cloud Services:** Amazon Web Services (Sagemaker), Google Cloud Platform

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

**Machine Learning Skills:** Classification - Naive Bayes, Random Forest, Logistic Regression, Decision Trees, XGBoost |

Clustering - K-Means | Regression - Linear Regression, Multiple Regression | Deep Learning - CNN

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

## WORK EXPERIENCE

### Machine learning SDE II – Amazon, New York NY

01/2022 - Present

- Ran ranking experiments on Amazon's desktop and mobile homepage to rank content which resulted in +15 MM GCCP (Growth adjusted Composite Contribution Profit) for desktop and +100 MM GCCP for mobile homepage.
- Developed hypotheses based on previous experiments to improve model performance and added feature generators and new objectives to existing featurizer flow. As a result Improved model performance for Non-Prime customers and not logged in customers.
- Collaborated with the team and published a research paper in an internal AMLC conference based on above experiment findings.
- Lead SDE from my team for Prime day 2023 where I led the efforts to scale up services, switch model objectives from GCCP to total sales and the drafting process of validating the switch and model's exploration.
- Responsible for heading efforts to migrate our service from Intel hosts to AWS's new graviton hosts which resulted in saving \$100 per month and lesser carbon footprint (more energy efficient)
- Responsible for switching our service's loadbalancers from network level Netscaler to internal application level tardigrade loadbalancers to decrease costs. Also, switched service authentication to internal Cloudauth authentication to make service more secure (HTTPS).

### Data Scientist – ADP, Alpharetta GA

07/2020 - 12/2021

- Worked on digitization of documents by extracting data using Machine Learning/Deep learning.
- Responsible for designing end to end ML infrastructure in AWS Sagemaker for our Machine learning models lifecycle.
- **ML privacy and security:** Used NER model Spacy to detect PII's in pdf documents and mask them. It reduced the risk of unauthorized access to sensitive information. Deployed the model as an API.
- Trained 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 creating end to end testing framework BDD (Behavior driven development) for Document extraction ML application.
- **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 display the offers and coupons on a javascript/HTML 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 a tremendous amount of time for the client account and I was given "Infosys Insta award" for this idea implementation.