

Venkata Krishna Anirudh Nuti

Phone: 857-919-1884 | nuti.krish@gmail.com | [linkedin.com/in/nvkanirudh/](https://www.linkedin.com/in/nvkanirudh/)

ABOUT

Graduated from Boston University with a master's degree in Applied Data Analytics and with almost **2 years** of experience as a Machine Learning Engineer. Looking forward to applying my skills in the industry and learning from its experience.

EXPERIENCE

Research and Development Engineer, Machine Learning, Intain Technologies (FinTech) June 2019 - Aug 2020

- Leveraged Python libraries, including Pandas, Regex, Object Detection using Mask-RCNN, TensorFlow framework, and OCR to develop and deploy predictive models in sales and marketing applications.
- Digitized and extracted relevant fields from Indian government documents such as passports, driver's licenses, voter IDs, and more as part of product development.
- Deployed Docker image of the product, used for automation, built on REST APIs in Google Cloud Platform (GCP).
- Built a prototype with appropriate documentation to automate the process of extracting fields, reducing manual effort of extraction **by over 80%** through the development of KYC product, which generated **\$500k** in revenue for the company.
- Improved **accuracy by 7% (97%)** using Mask-RCNN trained on around **500** images per class (medium-sized datasets) compared to **90%** accuracy achieved using RegEx.
- Managed source control and collaborated with team members using Git (GitHub).
- Collaborated with non-technical stakeholders to effectively communicate complex ideas through written, verbal, and data visualization presentations.

Machine Learning Engineer Intern, Intain Technologies (FinTech) Feb 2019 - June 2019

- Collaborated with the company's software development team to implement best practices in a team-oriented approach.
- Utilized Google OCR to extract relevant fields with **90%** accuracy from over four scanned documents and financial statements.

Teaching Assistant, Boston University Jan 2022 - May 2022

- Assisted students in the course CS 677 Data Science with Python, which covered topics about supervised and unsupervised learning algorithms, under the guidance of Prof. Eugene Pinsky.
- Evaluated assignments and final projects of **50** students and interacted with them twice a week to answer all their queries throughout the semester.

PROJECTS

Ad Fraud Detection: Python, Supervised Learning, Data Science

- Utilized Python and supervised learning to detect ad fraud.
- Performed Exploratory Data Analysis (EDA) on existing features and created new features using feature engineering along with visualizations using python libraries, including Pandas and Matplotlib.
- Developed machine learning models (*classification*) using algorithms such as Logistic Regression, Decision Tree, Random Forests, and LightGBM. Achieved highest accuracy of 93% with LightGBM.

Campus Recruitment Analysis: Statistical Analysis using R

- Performed Multiple Linear Regression to estimate the employment test percentage for given high school, secondary education, and degree percentages.
- Implemented a one-way Analysis of Variance (ANOVA) to compare means of employment test percentages across two groups of gender – male and female.

Extract, Transform, and Load (ETL): Apache Airflow and Apache Kafka

- Performed ETL on National Highway Toll data by building data pipelines using Apache Airflow's DAGs (*Directed Acyclic Graphs*) and loaded the data into PostgreSQL.
- Implemented streaming data pipelines to load the data into SQL Server database using Apache Kafka.

Covid Analysis: MySQL and Tableau

- Analyzed COVID dataset of different countries to understand rate of deaths and vaccinations using SQL queries with MySQL and visualized the results on the Tableau dashboard.

Image Generation of butterflies using Diffusion Models: Python, Hugging Face, Transformers, Deep Learning, PyTorch

- Generated images of beautiful butterflies by training a diffusion model using Hugging Face's diffusers library and PyTorch framework.
- Added noise to 1000 clean butterfly images using noise scheduler and trained a UNet architecture to reverse the process of generating clean images from noisy images passed to it as input.
- Used AdamW optimizer and Mean Squared Error loss function to update the parameters in backward denoising process.

SKILLS

Languages: Python (NumPy, Pandas, Scikit-learn, SciPy), R, SQL (SQL Server, PostgreSQL, MySQL, MongoDB)

Tools and Frameworks: Data Science, Machine Learning, Neural Networks, Deep Learning, TensorFlow, PyTorch, OpenCV, Computer Vision, Natural Language Processing (NLP), Diffusion Models, Apache Airflow and Kafka, Statistics, Tableau

EDUCATION

Master of Science, Boston University Jan 2023

Applied Data Analytics, GPA: 3.56

Bachelor of Technology, GITAM University April 2019

Computer Science Engineering, GPA: 8.46

Relevant Coursework: Data Structures and Algorithms, Data Science with Python, Machine Learning, Advanced DBMS