Aakruti Ambasana
DATA SCIENTIST

(365)-779-0078
Brampton, Ontario
krishna.aakruti@gmail.com

https://github.com/aakruti09
https://www.linkedin.com/in/aakruti-ambasana/

SUMMARY OF QUALIFICATIONS

- Applied **supervised** and **unsupervised learning** techniques for **data classification** and analysis on large data sets such as Covid-19, intrusion detection, cifar10, fashion Mnist, IMDB, stock market, Airbnb using **Machine Learning algorithms**.
- Designed goal-oriented **deep learning networks**, like Artificial Neural Networks (ANN) for big data, Convolution Neural Networks (CNN) for image classification, and Recurrent Neural Networks (RNN) for time-series forecasting.
- Deployed pre-trained **Transformers** model to build conversational AI tool using **NLP**.
- Hands on experience with ML projects using ML frameworks, which depicts solid knowledge in data science with commendable knowledge of **TensorFlow, Scikit-Learn, Pandas, NumPy**.
- Applied **exploratory data analysis (EDA)** on structured and unstructured datasets for acquisition of **insightful data analytics**.

TECHNICAL SKILLS

Data Modeling and Analysis	: Python – TensorFlow, PyTorch, Scikit Learn, Pandas, NumPy, NLTK, spaCy
Data Visualization	: Microsoft PowerBI, Seaborn, Matplotlib libraries (Python) for statistics
Programming languages	: Python, Java, C/C++, Android (using web APIs with JSON)
Software / Tools	: Jupyter Notebook, Google Colab, Webots, Omnet++, Git (version control system), Visual Studio
Relational databases	: SQL, MySQL, Oracle

EDUCATION

University of Waterloo – *Master of Engineering (MEng) in Electrical and Computer Engineering*   *Sep 2020 – May 2022*
Specialization: Artificial Intelligence and Machine Learning | *CGPA:* **83.78**/100

L.J. Institute of Engineering and Technology – *Bachelor of Engineering (BE) in Computer Engineering*   *Aug 2016 – May 2019*
CGPA: **9.26**/10

Nirma University – *Diploma in Computer Engineering*   *Aug 2013 – May 2016*
CGPA: **9.23**/10

INDUSTRY EXPERIENCE

Data Scientist   *Oct 2022 – Present*
MyBenefits AI
- Concept – **Transformers**, Natural Language Processing (**NLP**), Conversational AI, AWS
- Project – Post-care interactions for Urology patients
- Deployed pre-trained transformers model to provide answers to any questions related to discharge process from the patient document.

PROJECT EXPERIENCE

Predictive models for data classification | *covid-19 (messy real-world data), Heart disease*   *Jan 2021 – Apr 2021*
- Classification Models: **KNN, Decision Tree, Random Forest, Gradient Boosting (XGBoost), Naive Bayes, GridSearchCV**
- Applied **exploratory data analysis (EDA)** to analyze data trends for statistical modeling and data mining and **visualized data** for outlier detection and to present data-driven core insights in clear and compelling way, demonstrating my analytical skills.
- Predicting whether a person has covid or not for covid-19 data, using the above-specified classifiers.
- Trained the model and tuned hyper-parameters of classifiers for better accuracy.

Image Classification using TensorFlow | *Computer vision, Convolution Neural Network (CNN), Google Colab*   *Jan 2021 – Aug 2021*
- Built Convolution Neural Network model for predicting correct class label for image data.
- Tuned various neural network architecture, and optimized its hyper-parameters, for generalized a model by preventing overfitting.
- Achieved 90% accuracy on fashion Mnist dataset, and 94% accuracy on cifar10 dataset with test data.

Sentiment Analysis using TensorFlow | *Natural Language Processing (NLP), Google Colab*   *May 2021 – Aug 2021*
- Concept: Text Vectorization, Artificial Neural Network (ANN)
- Data processing – data is cleaned, tokenized, and converted into numerical embedding for statistical analysis using NLTK.
- Constructed ANN for predicting whether reviews are positive or negative from text data using TensorFlow and Keras libraries.

- Achieved **88% accuracy** on unknown data, after optimizing hyper parameters of the neural network.

Natural Language Processing (NLP) using Word2Vec   *Jan 2021 – Apr 2021*
- Concept: Word2Vec, GloVe, NLP, NLTK library, data source – Hugging Face
- **Data pre-processing** – data cleaning and tokenizing using the NLTK library before converting it into vectors.
- Translated words to numerical embedding using **Word2Vec** algorithm, measured similarity between words for text data analysis and compared it with the pre-trained **GloVe** model.

Stock Market Price Prediction | *Recurrent Neural Network (RNN)*   *May 2021 – Aug 2021*
- Concept: Long Short-Term Memory (LSTM), **time-series forecasting**
- Generated dataset by using features of past 3 days and analyzed the data to produce insights for better business decision making.
- Built RNN model for regression to predict the next day opening price, using **LSTM layers**.
- Analyze model performance and predicted data accuracy by reducing the loss, without overfitting.
- Gained training **loss less than 2.5**, by optimizing hyper-parameters, which enhanced my critical thinking skills.

Clustering with Unsupervised Learning | *Radial Basis Function Neural Network (RBFNN), KMeans cluster*   *May 2021 – Aug 2021*
- Produced data using mathematics and labelled data records, which resulted in data that is not linearly separable.
- Trained the model for predicting classes and learning data patterns using **RBFNN** and Gaussian kernel.
- Compared results of scenarios where centre points are all inputs or 150 random inputs or KMeans selected inputs.

Analysis of Node Attack in MANET and Intrusion Detection Using AI   *May 2021 – Aug 2021*
- Co-modelled and simulated active attacks such as Black-Hole, Hello Flooding, and DOS attacks in Omnet++ with different mobility models and node speed in ad-hoc network.
- Analysed data packet delivery ratio, end-to-end delay, network throughput, and energy consumption of nodes to collect data, and designed Artificial Neural Network (**ANN**).
- Written neural network documentation in project, and gave presentation to describe our experiments and process, implies my excellent communication skills, and also gained **teamwork** skills while working with excellent team in this project.
- Achieved **97.80%** accuracy in detecting attacker nodes, with learning data patterns.

ACHIEVEMENTS

- Rewarded for scoring 10.00 SPI in 4th Semester in bachelor's degree. I was ranked 1st in computer science department at L.J. and 3rd at Gujarat Technological University.
- Got Certificates of Appreciation for achieving 9.00 SPI consistently in all semesters in Diploma at Nirma University.
- Exhibited and presented project for innovative idea to various companies in L.J. Innovation Village 2019, implies my effective communication skills.