

## Skills

---

- Python (Jupyter Notebook, Pandas, NumPy, SciPy, Matplotlib, Seaborn)
- Machine Learning, Deep Learning
- SciKit-Learn
- Data Visualization, Data Manipulation (Tableau)
- HTML, CSS
- MongoDB, Cassandra
- PySpark
- Web Scrapping
- Natural Language Processing (NLP)
- Amazon(AWS)
- Artificial Neural Network (ANN), Convolutional Neural Network (CNN)
- Tensorflow, Keras, PyTorch
- OpenCV
- Object Detection

## Work Experience

---

### JUNIOR DATA SCIENTIST

Oracuz Infotech PVT LTD, Trivandrum

*August 2023 – Present*

- Led data cleaning and preprocessing efforts, enhancing data quality and reliability, resulting in a **30%** reduction in model training time.
- Contributed to the development and deployment of ML models for customer segmentation, leading to a **25%** increase in targeted marketing effectiveness.
- Designed and implemented over 40 machine learning models for different programs and projects.
- Verified results of algorithms to predict future occurrences using real world programs data with **82%** precision.

### MACHINE LEARNING INTERN

iNeuron Intelligence Pvt Ltd, Bangalore

*January 2024 – March 2024*

- Improved an accuracy rate of from **78.3% to 82.23%** in predicting credit card defaults using the Random Forest algorithm using hyper parameter tuning, demonstrating strong model performance and predictive capabilities.
- Developed and deployed a web application on **Amazon EC2** instances using Flask, enabling real-time access to the credit card default prediction model for users and Utilized Cassandra database for efficient storage and retrieval of large-scale credit card transaction data
- Streamlined the data integration process and optimized model performance, resulting in a significant reduction in data latency from **8 days to 1 day**, enhancing the efficiency of the prediction system.
- Check out my GitHub repo for further information [https://github.com/alnxha7/credit\\_card\\_default.git](https://github.com/alnxha7/credit_card_default.git)

## Projects

---

### DIABETES PREDICTION WITH PYTORCH ANN

- Introduced a diabetes prediction model using PyTorch Artificial Neural Networks (ANN). Conceived a classification accuracy of **76.62%** in predicting diabetes onset based on patient data.
- Analyzed feature engineering and optimization techniques, resulting in **8%** improvement in model performance compared to baseline.
- Deployed the model locally using Flask for real-time predictions, reducing inference time by **20%** compared to traditional methods. Investigated over **768** patient records to train and validate the model, ensuring robustness and generalization.
- [https://github.com/alnxha7/Diabetes\\_prediction\\_using\\_pytorch\\_ANN](https://github.com/alnxha7/Diabetes_prediction_using_pytorch_ANN)

### PEPPER DISEASE CLASSIFICATION USING TENSORFLOW

- Established a perfect **100% accuracy** in classifying pepper diseases including classifying pepper diseases ('Pepper bell Bacterial spot', 'Pepper bell healthy') using TensorFlow CNN.
- Trained the model on a dataset comprising **2475** of pepper plant images, allocating **80%-10%-10%** split for training, validation,

and testing respectively, ensuring robust performance and generalization.

- Contributed to agricultural innovation by developing a highly accurate disease detection system, potentially aiding farmers in early diagnosis and effective management of crop diseases.
- [https://github.com/alnxha7/Pepper-disease-Classification\\_using\\_Tensorflow](https://github.com/alnxha7/Pepper-disease-Classification_using_Tensorflow)

## SENTIMENT ANALYSIS USING NLP

- Conducted sentiment analysis on over **3100** Amazon Alexa reviews using advanced Natural Language Processing (NLP) techniques.
- Achieved an impressive accuracy rate of **94%** upon deploying the model locally via Flask for efficient real-time analysis.
- Explored a range of machine learning models, including Random Forest, Decision Tree, and XGBoost, ultimately selecting XGBoost for its superior performance. Optimized XGBoost hyperparameters using Grid Search, resulting in a **10%** improvement in accuracy for sentiment analysis on Amazon Alexa reviews.
- Generated a sentiment analysis model capable of processing both text inputs and CSV files containing customer reviews on Amazon Alexa products
- [https://github.com/alnxha7/Sentiment\\_Analysis\\_using\\_NLP](https://github.com/alnxha7/Sentiment_Analysis_using_NLP)

## HOUSE PRICE PREDICTION WITH PYTORCH

- Prepared a house price prediction model achieving an RMSE of **56,874.50** using a Feedforward Neural Network resolved in PyTorch on the US Housing Index dataset.
- Conducted extensive feature engineering, improving model accuracy by **20%** through the incorporation of price per square foot and neighbourhood-based price indices.
- Utilized PyTorch for model training and optimization, resulting in a **25%** reduction in overfitting and a **15%** increment in prediction performance compared to baseline models.
- [https://github.com/alnxha7/House\\_price\\_prediction\\_using\\_PyTorch](https://github.com/alnxha7/House_price_prediction_using_PyTorch)

## Education

---

- ADVANCE DIPLOMA IN AI AND ML  
*December 2023*
- BSc COMPUTER SCIENCE  
*June 2020 – 2023*

## Certifications

---

- Certificate in Build a computer vision app with Azure Cognitive Services
  - [Microsoft](#)
- Certification in Python
  - [HackerRank](#)
- Certification in AI and ML
  - [Global India Techno Hub](#)