

# Harshil Patel

www.harshilcpatel.site www.linkedin.com/in/harshilcpatel9944@gmail.com +1 551-804-1305

## Education

---

### Master of Science in Applied Artificial Intelligence

*Stevens Institute of Technology, New Jersey, USA*

**GPA: 3.75**

**Courses:** Deep Learning, Machine Learning, Applied Modelling and Optimization, Pattern Recognition and Classification, Data Structures and Algorithms, Probability and Statistics, Linear Algebra.

### Bachelor of Technology in Computer Engineering

*Dharmsinh Desai University*

**CGPA: 7.5/10**

**Courses:** Data Structures and Algorithms, Database Management System, Design and Analysis of Algorithm, Software Engineering Practices, Advanced Algorithms, Machine Learning, Artificial Intelligence, Image Processing, Big Data Analytics, Cloud Computing

## Skills

---

**Programming Languages and Technologies:** Python, C++, Php, JS, SQL, CUDA

**Data Analysis Tools:** Pandas, Numpy, Matplotlib, MySQL, Pillow, OpenCV

**Frameworks:** TensorFlow, Scikit-Learn, PyTorch, Keras, Langchain, Langflow

**Web-Dev:** Django, Laravel, NextJS, VueJS, React.

**Tools:** Jupyter Notebooks, Git

**Cloud Services:** GCP, AWS

**SoftSkills:** Effective communication, Critical thinking, Collaboration, Self-motivation, Attention to Detail

## Certification

---

### AWS Certified AI Practitioner

Proficient in applying AI/ML fundamentals, responsible AI practices, and leveraging AWS AI services to design data-driven solutions.

## Research Paper

---

### *A Novel Seizure Prediction Technique Using CNNs and Self-Attention.*(Yet to be Published)

A hybrid CNN and self-attention model for early prediction of epileptic seizures using EEG signals.

## Academic Projects

---

### Customer Lifetime Value Prediction

- Analyzed transactional data and engineered RFM (Recency, Frequency, Monetary) features to uncover inefficiencies in customer targeting.
- Applied K-Means clustering to segment customers into High, Mid, and Low Value tiers, enhancing retention strategy.
- Built and optimized an XGBoost model that achieved 96% prediction accuracy, enabling precise targeting of high-value customers.
- Boosted training and inference speed by implementing custom CUDA kernels and leveraging TPUs, improving compute performance and scalability.

### Advanced Hyperparameter Optimization Using Simulated Annealing

- Conceptualized the idea of using a simulated annealing approach to optimize the hyperparameter of the XgBoost classifier.
- Developed the Simulated Annealing Algorithm and used it to tune the hyperparameter of the Training model.
- Analysed and compared results of the performance of each model until the model with the best performance is found.

### Detection of Fraudulent Transaction Using Machine Learning

- Analyzed and validated data thoroughly to understand fraud patterns and identify data inaccuracies by exploring the dataset and using statistical methods to assess reliability of dataset.
- Experimented with various machine learning algorithms (i.e. Naive Bayes, Random Forest, Logistic Regression) to identify which one could best predict fraudulent activities. Through this exploration, discovered that the Random Forest algorithm stood out for its accuracy and robustness.
- Successfully applied the Random Forest model to achieve a remarkable accuracy rate of 99.5%.

## Work Experience

---

### Bhaskaracharya Institute of Space Applications and Geo-informatics

December 2022 - April 2023

**Role:** Research Intern

### Project: Noise Removal from Noisy Audio Signal Using Deep Learning Methods

- Developed a noise reduction model by sourcing diverse audio samples and converting them into spectrograms for analysis, leveraging Python and ML frameworks like TensorFlow.
- Designed and trained a U-Net convolutional neural network with dropout layers on noise-blended clean audio spectrograms, achieving a training loss of 0.002129 and a validation loss of 0.002406, while preserving audio quality.
- Utilized large-scale distributed training strategies for model optimization, ensuring efficient processing and performance across diverse datasets.