

# Aaditya Srivastava

+91-9569802517

adityasrivastava9567@gmail.com

LinkedIn Profile

GitHub Profile

## Education

- **Bachelor of Technology in Computer Science and Engineering** 2022-26  
Vellore Institute of Technology, Bhopal CGPA: 8.37
- **City Montessori School, Rajendra Nagar-I** 2021-2022  
XII (Senior Secondary) 90.1%
- **City Montessori School, Rajendra Nagar-I** 2019-2020  
X (Secondary) 93.2%

## Projects

- **LLM Tokenizer** September 2025 - Present
  - Developed a tokenizer from scratch to efficiently convert raw text into subword token sequences for downstream NLP tasks.
  - Implemented subword tokenization algorithms such as Byte Pair Encoding (BPE) , optimized for vocabulary coverage and reduced out-of-vocabulary rate.
  - Integrated text normalization, token frequency analysis, and vocabulary generation pipelines to ensure balanced token distribution.
  - Evaluated tokenizer performance on diverse corpora, improving encoding efficiency .
  - Technologies Used: Python, NumPy, regex, tokenization algorithms (BPE)
- **License Plate Recognition System** August 2025 - September 2025
  - Built an image-processing pipeline for license plate detection using CCA followed by character segmentation and recognition
  - Trained a supervised Support Vector Classifier (SVC) on a labeled dataset of characters to achieve robust character recognition
  - Enhanced accuracy through preprocessing, vertical projection filtering, and cross-validation, improving results by 15%
  - Performed character segmentation by isolating individual symbols from license plate regions, resizing them to 20×20 pixels, and ordering them for recognition
  - Technologies Used: Python, scikit-image, scikit-learn, NumPy, Pillow, matplotlib
- **Digit Recognition using Neural Networks** July 2025 - August 2025
  - Implemented a feed-forward neural network in Keras, reaching 95.6% test accuracy on 10K unseen images
  - Preprocessed and normalized 60K training samples, reshaping pixel vectors into 784-dim inputs
  - Designed and trained a Sequential model with two hidden layers (64 neurons each, ReLU activation) and an output layer (Softmax), optimized with Adam
  - Performed model evaluation achieving robust generalization across unseen data and demonstrated predictions with real test samples
  - Technologies Used: Python, NumPy, TensorFlow, Keras, MNIST

## Achievements

- CodeChef Rating:** 1277 (aaditya8) Highest Rating: 1277  
Global Rank - 8988 in Starters 174 (Rated)
- LeetCode Rating:** 1358 (AadityaSRV) Highest Rating: 1490  
Global Rank - 1807 in LeetCode Biweekly Contest 133, out of 25,500+ participants

## Technical Skills and Interests

**Languages:** Java, Python, MySQL

**Web Dev Tools:** GitHub, Git

**Cloud/Databases:** Relational Database (MySQL)

**Relevant Coursework:** Machine Learning, Data Structures & Algorithms, Object Oriented Programming, Computer Network, Deep Learning, LLM

## Certifications

- **DeepLearning.AI** - Supervised Machine Learning: Regression and Classification March 2025