Ayan Gaur

+1 (858) 281 - 1254 || ayan.gaur@gmail.com || ayangaur26.github.io || github.com/ayangaur26

Education

University of California San Diego Sep. 2022 – Mar. 2026
B.S. Mathematics and Computer Science

Cornell University May 2024 – Aug. 2024
Machine Learning Foundations Certificate

Experience

American Express Aug. 2024 – Dec. 2024
ML Intern

- Fine - tuned DeBERTa and DistilBERT models to develop a sensitive data discovery pipeline , achieving 79% accuracy and applying custom regex solutions to accurately identify incorrectly classified data by 30%
- Optimized BIO - label alignment using tokenization strategies reducing over - redaction and false positives by 25% through cross - validation and hyperparameter tuning , using iterative model optimization
- Integrated contextual NER techniques, achieving 76% precision , 72% recall across diverse financial text datasets

Image Infosystems July 2024 – Sep. 2024
Computer Vision Intern

- Developed high - accuracy OCR models with 94.4% accuracy and ¡4.8% loss while using a custom ONNX function converting TensorFlow models to PyTorch ensuring complete retention of model accuracy
- Improved cheque verification accuracy by 18% through algorithm enhancements for image preprocessing and feature extraction in high - security bank applications

Projects

User Rating Prediction Recommender System Paper | Sep. 2024 – Dec. 2024

- Conducted EDA on a Food.com dataset (200k+ datapoints), identifying patterns for predictive modeling
- Enhanced scalability by filtering data and applied regularization techniques to reduce model overfitting
- Proposed future enhancements, including hybrid collaborative filtering and deep learning - based embeddings
- Implemented Word2Vec similarity and SVD latent factor models , outperforming baseline models by 18.3%

Computer Vision Personal Project | Jan. 2025 – Mar. 2025

- Implemented Bag of Words for image classification using SIFT feature extraction and KMeans clustering
- Applied Bayesian Estimation for email spam classification, leveraging Max Likelihood and Max A Posteriori
- Developed image processing pipelines integrating Edge & Corner Detection and feature - based matching
- Designed and rendered Photometric Stereo models, exploring light - source variations for 3D surface reconstruction
- Trained a FashionMNIST CNN model with backpropagation , achieving 92% classification accuracy and applying transfer learning on STL - 10 for feature extraction

F1 Qualifying Performance Analysis – EDA & Statistical Modeling | Jan. 2025 – Mar. 2025

- Applied statistical modeling and correlation analysis to assess whether high - curvature sectors had a stronger influence on qualifying performance than straight sectors using race data from FastF1 API
- Preprocessed and standardized data by filtering lap times, handling null values, and normalizing sector times
- Visualized sector importance through performance graphs and impact tables , providing data - driven insights

AI Coffee Dial-in Recommender | *Gemini API*

- Used the Gemini API to recommend optimal coffee brewing parameters based on bean type, grinder, and brew method.

Technical Skills

Languages: Python, Java, C/C++, Dart, Bash, R, ARM, HTML, CSS, JavaScript, TypeScript
Frameworks: TensorFlow, PyTorch, Keras, Flask, Flutter, Next.js, Tailwind CSS
Developer Tools: Git, ONNX, Power BI, Firebase, MATLAB, Microsoft Office (Excel, Word, Outlook)
Libraries: Pandas, NumPy, Seaborn, Matplotlib, Scikit - Learn, React