

## TRISHA MANDAL

### EDUCATION

University of Southern California, *Los Angeles, CA*  
Master of Science (MS) in Computer Science, Specialization: Artificial Intelligence

GPA: 3.7  
Jan 2022-May 2023

Pennsylvania State University, *State College, PA*  
Bachelor of Science (BS) in Computer Science, Minor in Mathematics

GPA: 3.5  
Aug 2017-May 2021

### SKILLS

- **Programming Languages:** Python, C/C++, Java, SQL, HTML, CSS, Javascript, MATLAB, Racket, Verilog
- **Tools and Frameworks:** TensorFlow, PyTorch, LangChain, Pandas, NumPy, Scikit-learn, Keras, SpaCy, Pytest, Matplotlib, CUDA, Scrappy, HuggingFace, Git, Jira, Linux, Power BI, MySQL, OpenAI, Pinecone, Redis, MongoDB, AWS OpenSearch, Apache Airflow, FastAPI, Docker, Google Cloud Platform
- **Industry Knowledge:** Machine Learning, Deep Learning, Natural Language Processing, Large Language Models

### EXPERIENCE

Griptape, Inc | Software Engineer | Los Angeles, CA Jul 2023-Dec 2023

- Engineered **vector store database drivers** for MongoDB, Redis, and OpenSearch achieving latencies below **1.0s** for vector storage and vector similarity search, enhancing **RAG** workflows.
- Developed integration tools for **Large Language Models** with Google Workspace, enabling file management and real-time collaboration on Drive, Docs, Sheets, and Gmail.
- Conducted thorough unit and integration testing for **100% code coverage** and created mkdocs-standard documentation.
- Designed industry-specific, dynamic chatbots leveraging Griptape's LLM tools and rulesets, and contributed to over **25 pull requests** within the Griptape open-source ecosystem.
- Technologies Used: Python, Pytest, NumPy, Google OAuth 2.0, OpenAI, Git

USC Marshall School of Business | Research Assistant | Los Angeles, CA Oct 2022-May 2023

- Spearheaded NLP research using **text-based multi-task CNN** built from scratch for customer review analysis, achieving over **90% accuracy** in sentiment analysis and outperforming separate CNNs for each task.
- Performed exploratory data analysis, topic modelling and keyword matching using Pandas and Matplotlib visualization, LDA analysis and GPT-3 respectively to distill customer sentiment and market trends from **8-million-entry** datasets.
- Technologies Used: Python, TensorFlow, Keras, NumPy, NLTK, Regex, GPT-3, Pandas, Matplotlib

Lexalytics, Inc. | Software Engineer Intern | Amherst, MA Jun 2020-Aug 2020

- Rectified anomalies in output for automated workflow built to convert PDFs to JSON output used by a machine learning model by performing exploratory data analysis, data preprocessing and refactoring **70 lines** of technical debt.
- Leveraged Docker containers to improve portability of applications and Google Kubernetes Engine for model deployment.
- Technologies Used: Python, Java, Pandas, Docker, Google Kubernetes

Lexalytics, Inc. | Software Engineer Intern | Boston, MA Jun 2019-Aug 2019

- Programmed **HTTP REST API's** for Lexalytics data analytics platform and their documentation using Java Spring Framework and Swagger UI/UX properties and improved the platform's results by **6%** through unit testing.
- Performed Named Entity Extraction using spaCy for over **800** financial documents for NLP and data analytics prospects.
- Technologies Used: Java, Python, SpaCy, SQL, Swagger UI

### PROJECTS

Optimization of Video ML Pipeline - PyTorch, TensorRT, NumPy, CUDA, Docker Jan 2024

- Optimized and dockerized a **video ML pipeline** for manufacturing safety, involving acceleration of detection, tracking, and background blurring in videos by processing frames in parallel using **multi-GPU** and **multi-core processing**.

Long-Text Processing Using BELT - PyTorch, NumPy, HuggingFace, Apache Airflow Nov 2023

- Engineered 'DataFormatter' class to standardize and structure public datasets to train **BELT (BERT for Longer Texts)** model on them for classification tasks, enabling efficient processing of long-text datasets exceeding transformer token limits. Automated this **ETL pipeline** with a DAG workflow in Apache Airflow for streamlined data handling.

Adapting Multimodal Models to Unimodal Tasks by Ensembling FLAVA with ALBERT - TensorFlow Apr 2023

- Performed a research study to adapt multimodal model: FLAVA (Foundational Language and Vision Alignment model by Facebook) to perform unimodal tasks and improve visual question answering (VQA) results by replacing baseline text encoder **BERT** with **ALBERT**, **GPT-2** and **RoBERTa**. Executed training and testing on various VQA/QA Hugging Face datasets.