

## 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, NetworkX, Scrapy, Pandas, NumPy, Scikit-learn, Keras, SpaCy, NLTK, Pytest, Matplotlib, Gensim, CUDA, HuggingFace, Git, Jira, Linux, Power BI, MySQL, OpenAI, Pinecone, Redis, MongoDB, AWS OpenSearch, Jupyter Notebook, Apache Airflow, FastAPI, Visual Studio Code, Docker, Google Cloud Platform
- **Industry Knowledge:** Machine Learning, Data Science, Natural Language Processing, Large Language Models, Deep Learning, Statistics, Data Structures and Algorithms, Software Engineering, MLOps, LLMops, Statistics
- **Soft Skills:** Technical Writing, Verbal and Written Communication skills, Leadership, Collaboration, Detail-Oriented

### EXPERIENCE

---

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

- Implementing vector store database drivers for MongoDB, Redis and OpenSearch with read/write and search latencies below 1.2s, to efficiently store word embeddings from machine learning models and perform vector search similarity to aid Retrieval-augmented generation (RAG) workflows.
- Developing tools for Large Language Models (LLMs) for seamless integration with Google services such as Drive, Docs, Sheets, and Gmail. Functionalities include upload/download, listing, sharing, and, real-time editing of docs and sheets.
- Conducting comprehensive testing (unit and integration) with 100% code coverage, to ensure tool robustness and flawlessness. Additionally, curating detailed documentation for these tools in alignment with mkdocs standards.
- Building dynamic, interactive chatbots using Griptape's agent, rulesets and LLM tools specific to every industry.
- Contributed to over 25 pull requests in the Griptape open-source framework and continuing to actively contribute.

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

- Designed and led NLP research that performs exploratory data analysis (EDA), sentiment analysis and topic modeling on customer reviews for market research using multi-task convoluted neural network (CNN) that generated an accuracy above 90% and outperformed separate CNNs for each task.
- Executed fine-tuning of GPT-3 model to identify synonym phrases for key topics in fashion and technology interviews.

**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 (EDA), data preprocessing and refactoring 70 lines of technical debt.
- Leveraged Docker containers to improve portability of applications and Google Kubernetes Engine for model deployment.

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

- Developed 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.

### PROJECTS

---

**Enhanced Industrial Safety: Streamlining a Video ML Pipeline - PyTorch, TensorRT, NumPy, CUDA** Jan 2024

- Optimized and dockerized a video-based ML pipeline for manufacturing safety, involving acceleration of detection, tracking, and background blurring in videos using multi-GPU and multi-core processing, and streamlined Docker deployment.

**Long-Text Processing Using BELT - PyTorch, NumPy, HuggingFace** 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.