Tanveer Mittal

Machine Learning Enthusiast & Data Science Student

(630) 640-5246 | tamittal@ucsd.edu | tanveermittal.github.io | TanveerMittal | Tanveer Mittal

Education

University of California: San Diego — *La Jolla, CA*

Bachelor of Science in Data Science, Provost Honors, GPA: 3.5/4.0, — *Sep. 2018 - May 2022*

- **Technical Coursework:** Natural Language Processing, Database Management, Data Science, Algorithms, Data Structures
- **Theoretical Coursework:** Statistics, Graph Theory, Probability Theory, Multivariable Calculus, Linear Algebra
- Volunteered as data science student panelist to speak to incoming data science students and answer questions for Triton Day

Coursera

- **Deep Learning Specialization:** Deep Learning, Optimization, Structuring ML Projects, Convolutional Neural Networks, and Sequence Models
- **Tensorflow in Practice Specialization:** Intro to Tensorflow, Time Series Analysis, Natural Language Processing, and Computer Vision

Skills

Languages Python, SQL, Java, Javascript

Frameworks Keras, Tensorflow, Pytorch, ReactJS, NodeJS, Kubernetes, Docker, SpaCy, OpenCV, Spring, Matplotlib, Seaborn

Experience

American Express — *Phoenix, AZ*

Software Engineering Intern — *July 2020 - Aug. 2020*

- Created a 5 node Kubernetes cluster to deploy a containerized Couchbase database and CI/CD pipeline with Jenkins
- Developed a data migration tool to migrate data from a MongoDB collection to PostgreSQL using NodeJS and transform streams

Cardiac Mechanics Research Group — *La Jolla, CA*

Software Engineer — *Nov. 2019 - Present*

- Developing full stack with ReactJS, Java, and SQL for a database application that stores and presents clinical data and MRI images
- Performing programming, scripting, troubleshooting, and computer repair tasks as requested by researchers and lab management

Data Science Student Society @ UCSD — *La Jolla, CA*

Machine Learning Workshops Lead — *May 2019 - Present*

- Collaborating with companies and other student organizations to organize machine learning workshops for the UCSD data science community
- Planning and facilitating logistics for data-science focused educational, professional, and networking events throughout the school year

The InfoPost — *La Jolla, CA*

Machine Learning Engineer — *Feb. 2020 - May 2020*

- Quantified news articles' political bias using named entity recognition, sentiment analysis, and coreference resolution models
- Designed and created a chat bot that verifies COVID-19 claims utilizing SciBERT as a search engine and ROBERTA for textual entailment

iDTech Camps — *Lisle, IL*

Machine Learning Instructor — *June 2019 - Aug. 2019*

- Taught deep learning and python programming to 50+ high school students using Keras, Tensorflow, and OpenAI Gym
- Managed student behavior, administered safety procedures, and planned camp logistics

UC San Diego Health Department of Medicine — *La Jolla, CA*

Data Programmer — *March 2019 - June 2019*

- Migrated data from central data warehouse to department's personal SQL server by writing automated T-SQL procedures
- Performed scripting projects for fuzzy logic algorithms, data processing, and python/SQL integration

Swarm Robotix — *Naperville, IL*

Software Engineering Intern — *May 2018 - July 2018*

- Developed computer vision and deep learning solutions for image segmentation and deployed final model as a ROS node
- Collaborated to create a motion planning interface for precise autonomous control and odometery of a swerve drive mechanism

iManage — *Chicago, IL*

R&D Intern — *June 2017 - Aug. 2017*

Projects

Time Series Modeling of AirBnB Prices in San Diego (Data Science Class Final Project)

- Forecasted the average prices of AirBnB's in San Diego using SARIMAX and FB Prophet time series models
- Created an interface for users to input search criteria for an AirBnB and receive price predictions from the time series models

Textual Analysis of Chronic Illness Surveys (1st Place Entry for UCSD DataHacks 2020)

- Applied T-Stochastic Neighbor Embeddings, GloVe, and TF-IDF to visualize survey questions in a 2d space using Tableau dashboards
- Analyzed the effectiveness of Tobacco laws in the US for restricting underage tobacco sales by performing permutation testing

Sequence to Sequence Question Answering System (NLP Class Final Project)

- Utilized the Stanford Question Answering Dataset and pretrained GloVe embeddings to train models capable of answering questions
- Created RNN, LSTM, bidirectional LSTM, and multi-headed self attention deep learning models using Pytorch for automatic differentiation