Suraj Anand

suraj_anand@brown.edu | (310) 987-0123 | github.com/surajk610

EDUCATION

Brown University Providence, RI

Sc.B. in Applied Math-Computer Science, GPA: 4.0/4.0 Expected Graduation May 2024

- *Relevant Coursework:* Deep Learning, Computational Probability and Statistics, Information Theory, Discrete Structures/Probability, Statistical Inference I, Honors Linear Algebra, Ordinary Differential Equations, Blockchains and Crypto, Software Engineering, Computer Systems, Compilers, Data Structures and Algorithms

EXPERIENCE

Kaiser Permanente Medical Informatics Providence, RI

Machine Learning (NLP) Intern May 2022 - August 2022

- Researched and fine-tuned various BERT/sBERT models to generate embeddings of a patient complaint to be used in logistic regression classification of the chief complaint; improved precision-recall area of production model by 10%
- Conducted ablation studies to measure the effect of patient history and demographic data on complaint classification
- Fine-tuned GPT-J, T5, and BART models to extract important kidney stone features (size, laterality, etc.) from chunks of radiology reports; Calibrated GPT-J Models with a BERT Discriminator and then used Self Learning and Active Learning to achieve 80% accuracy (writing paper about experiments)

VolKno Los Angeles, CA

Machine Learning Engineering Intern January 2022 - Present

- Deploying a Dockerized Apache Airflow DAG to compute and store scene-level attributes of trailer content including key objects, creative concepts, demographic attributes of actors, and face embeddings of actors in a MySQL database
- Trained a multi-output Resnet-50 Convolutional Neural Network to determine age, race, and gender from an aligned face image using pytorch lightning; tracked experiments and stored models in MLFlow
- Advised by Professor David Parkes of Harvard, predicted downstream Twitter sentiment about upcoming movies from emotional response data collected from volKno users viewing trailers with an AUC of 0.79 in a project for Netflix; required scoring users

Brown Computational Linguistics (CS1460) San Diego, RI

Teaching Assistant March 2022 - Present

- Instructing 100+ Brown students on Transformers, Hidden Markov Models, Dynamic Programming, Dependency Parsing, etc
- Created PyTorch coding assignment on Topic Modeling using Latent Semantic Analysis (LSA); helped develop course website

OR-Stencil Irvine, RI

Application Developer June 2020 - August 2020

- Developed an algorithm to draw relaxed facial skin tension lines using a Deep Convolutional Neural Network for facial landmark detection, quadratic interpolation, and weighted averaging
- Set up a WSGI backend server in Flask. With a team, integrated into OR-Stencil App (paper)

PROJECTS

Lucidity | *Spotify Wrapped for Messenger*

- Used Latent Dirichlet Allocation and Linear Support Vector Classifiers for topic extraction and emotion/personality classification
- Designed in Figma and developed in React TypeScript various frontend components including a login, tables compiling response time and emoji usage, and a word cloud (website, codebase)

Quantitative Testing with Concept Activation Vectors (TCAV) | *PyTorch Implementation*

- In a team of three, implemented Been Kim et al.'s Interpretability Beyond Feature Attribution paper in PyTorch
- Wrote the code to extract activations and gradients with pytorch hooks, generate the normal vector to the hyperplane separating concepts from non-concepts (CAV), and compute directional derivatives to assess how a change in input towards a concept influences model prediction (codebase)

MedBrain | *Deep Learning Tool for Interpretable Clinical Event Prediction*

- Implemented LSTMs with Attention to interpretably predict severe clinical events and BERT Language Model to highlight
- salient info in historic physician notes
- Intel International Science and Engineering Fair Finalist, LA County Science Fair Overall Winner (paper)

SKILLS & ACHIEVEMENTS

- **Languages**: Python, SQL, C, C++, MATLAB, JavaScript/TypeScript, Java, HTML/CSS, Solidity, Swift, Go
- **Technical**: Git, Docker, NumPy, Pandas, PyTorch/Tensorflow, MatplotLib, MongoDB, Heroku, AWS Server