JITESH PABLA

+1-480-930-2008 • jpabla1@asu.edu • jiteshpabla.github.io • linkedin.com/in/jiteshpabla

EDUCATION

Master of Science - Computer Science **Expected May 2021**
Arizona State University, Tempe, AZ GPA: 3.89/4.0
Courses: NLP Methods for Biomedical Text Mining, Statistical Machine Learning, Introduction to Artificial Intelligence, Data-Intensive Systems for Machine Learning, Perception in Robotics, Introduction to Human-Computer Interaction

Bachelor of Technology (with honors) - Computer Science and Engineering **May 2019**
Jaypee Institute of Information Technology (JIIT), Noida, India CGPA: 8.1/10
Courses: Data and web mining, Introduction to Deep Learning, Artificial Intelligence, Statistics, Quantitative methods for social sciences, Data structures, Algorithms, and Problem Solving

TECHNICAL SKILLS

Languages: Proficient: Python, C++, SQL; Competent: C, Lua, PHP; Some knowledge: JavaScript, Java
Misc: Tools: Git, GitHub, Jupyter Notebook, Anaconda, Docker, Android Studio; OS: Linux, Windows; Hardware: Arduino, Raspberry Pi; Machine Learning: NumPy, Pandas, Scikit-learn, Matplotlib, Keras, PyTorch, TensorFlow; Web: CSS, Bulma, Vue.js
Certifications: Deep learning specialization - deeplearning.ai (Coursera)

WORK EXPERIENCE

Graduate Service Assistant, Arizona State University, USA **January – May 2020**
- Worked with Dr. M. Devarakonda on classifying 50k COVID19 articles related to vaccines and therapeutics using several techniques:
    - scraping Google search results to obtain noisy data and training a scientific-text based Bidirectional Encoder Representations from Transformers (BERT) model called SciBERT.
    - utilizing BERT as an embedding generator and finding each article's Cosine similarity with keywords related to vaccines etc. as a ranking criterion. Used Elasticsearch to speed up the process.
    - implementing a BM25 + RM3 approach to rank the articles according to relevance.
- Identified Randomized Controlled Trials (RCTs) from over 50k PubMed articles by modifying the BERT architecture and manipulating its inputs along with various NLP techniques using PyTorch and transformers.
- Implemented a visualization to understand the attention architecture in BERT and the effects of modifications made to it and its inputs.

Participant with LuaRocks -The Lua package manager, Google Summer of Code 2018 **June – August 2018**
- Refactored the core functionalities of LuaRocks commands for - listing, uninstalling, and showing details of packages, searching and installing rocks from the web, opening documentation, linting the rockspec, selecting a rock-tree, etc., to modularize them.
- Programmed a complete Application Programming Interface (API) to provide access to the LuaRocks functionality using Object-Oriented design patterns and used Git extensively for contributing to the main code-base.
- Designed a responsive and interactive web-based GUI using HTML, CSS, Bulma, and Vue.js to give access to the LuaRocks functionality. Interfaced the GUI with the LuaRocks-API in the backend using CGILua.

Intern, Python development, Internity Foundation and Rannlab Technologies Pvt. Ltd., India **June – August 2017**
- Applied machine learning models like - K Nearest Neighbours (KNN), Support vector machines (SVMs), logistic regression, etc. for classification on various datasets utilizing NumPy, Pandas, and Scikit-learn.
- Built a proof-of-concept chatbot based on Stanford's CS20 chatbot by implementing a seq2seq model using TensorFlow, trained on Cornell's movie dialogue corpus.

Intern, Data analysis, Team Computers Pvt. Ltd., India **June – July 2017**
- Applied data preprocessing techniques, statistical and machine learning methods such as moving averages, linear regression, spectral clustering, etc. on dummy datasets using "Alteryx" (a data science tool).
- Predicted prospective car customers using car sales and inquiry data (with millions of data points spanning across 1 year) using time series analysis as an individual project.

ACADEMIC PROJECTS

Tweet Sentiment extraction (Kaggle competition) **May – June 2020**
- Fine-tuned several BERT models and other NLP transformers like XLNet, RoBERTa in a question-answer format to extract the substring of a given tweet which represents the given emotion of the tweet.
- Achieved a Jaccard similarity score of 0.687, better than around 600 other submissions.
- Programmed in Python and utilized PyTorch, transformers, Pandas, and Jupyter notebooks.

Model Integration with Mixture-of-Experts **January – May 2020**
- Researched and studied various methods for integrating machine learning models together based on their similarity.
- Programmed a proof-of-concept machine learning model which could integrate three separate Convolutional Neural Networks (CNNs)

(the experts) trained to classify different classes within the MNIST dataset, to showcase the feasibility of a dynamic machine learning system which can be expanded or contracted based on the use case.

Clinical Semantic Textual Similarity (STS) **August – December 2019**

- Preprocessed the clinical text to remove stop words, punctuation, etc. and utilized various word2vec pre-trained models to extract token embeddings in order to create a single vector representation for each sentence.
- Fine-tuned multiple BERT models on the given STS dataset and extracted vector representation for each sentence.
- Engineered several similarity features based on the extracted sentence vectors and applied gradient boosting regression to achieve a Pearson correlation greater than 0.84 between the ground truth and the model's predictions.

Text-to-face generation **August 2018 – May 2019**

- Investigated and summarized various methods for facial image generation using a text description of a face.
- Collected a dataset of text descriptions of hundreds of images from the Labelled Faces in the Wild (LFW) dataset and utilized word2vec to create text embeddings.
- Programmed a Keras implementation of StackGAN (a variation of Generative Adversarial Networks) and trained it to generate facial images using the collected dataset.

Crop yield prediction based on temperature and rainfall for India **September – November 2018**

- Predicted the temperature and rainfall for a set of Indian districts using Recurrent Neural Network (RNN) and it's variation Long short-term memory (LSTM) and selected the method with the least mean absolute error.
- Utilized the rainfall and temperature prediction to further predict the yield of various crops in Indian districts using different methods - Linear regression, Random Forests, K- nearest neighbors (KNN), and a Feed-Forward Network; performed a comparative analysis for all the methods with Random Forests giving the least error.
- Used Pandas, NumPy, Scikit-learn, Keras, and Matplotlib on Jupyter notebook for implementation.

Developing a Secure Soldier Monitoring System using Internet of Things and Blockchain **January – May 2018**

- Built a compact health and location monitoring system for soldiers on a battlefield using Raspberry Pi, Arduino and sensors to capture body temperature, heart rate, and GPS coordinates, along with a panic button and LCD to display messages.
- Re-engineered a blockchain prototype in Python to store AES encrypted data being transmitted from the monitoring system via GSM in an immutable and trustworthy fashion.

Anomaly detection on Intel lab data **September – November 2017**

- Applied simple moving average (SMA), Density-based spatial clustering of applications with noise (DBSCAN) and LSTM to detect anomalous readings from various sensors in the dataset.
- Used Pandas, NumPy, Scikit-learn, Keras, and Matplotlib for implementation.

PUBLICATIONS

- Pabla, Jitesh, Vaibhav Sharma, and Rajalakshmi Krishnamurthi. "Developing a Secure Soldier Monitoring System using Internet of Things and Blockchain." 2019 International Conference on Signal Processing and Communication (ICSC). IEEE, 2019.

ACTIVITIES

Vice President, GRADient, ASU, USA **May 2020 – present**

- Organizing online social events for LGBTQ+ graduate students and allies to foster inter-department communication and collaboration.

Student mentor, 'Algorithms and Problem Solving lab' and 'Artificial intelligence lab' at JIIT, Noida, India **July 2018 – May 2019**

- Assisted professors with setting assignments, proctoring exams, and solving student doubts for a class of sixty students.

Workshop teacher, Computational Methods for Medical Image Analysis, India **April 2019**

- Taught Image segmentation using Python to faculty and graduate students.

Student Coordinator, Open Source Developers Club at JIIT, Noida, India **July 2017 – May 2018**

- Coordinated, organized, and taught at various workshops, meetups, and hackathons with hundreds of participants.

Student Coordinator, Graficas - Graphics and animation Club at JIIT, Noida, India **July 2017 – May 2018**

- Coordinated, organized, and taught at various workshops and competitions.

Director of digital and technical department, Jaypee Model United Nations at JIIT, Noida, India **September 2017 – January 2018**

- Managed and led a team of eight students to create the website, social media posts, physical posters, banners, and booklets for the JMUN 2018 event - attended by over five hundred participants.

Student Volunteer, IEEE student chapter at JIIT, Noida, India **July 2016 – May 2017**

- Taught the basics of C++ and web development to students across multiple workshops.

Hackaton 2nd place, Build-a-thon by Hackerearth.com **November 2016**

- Build an Android application for women's safety, with multiple ways to record and alert about any dangerous situation.

Intern - Marketing and Recruitment, Umeed - a drop of Hope (NGO), Delhi, India **December 2015 - January 2016**

- Marketed about the Umeed NGO to people ranging from school students to senior citizens.
- Assisted more than 20 people to register as a volunteer for the NGO.
- Went to very low-income regions to help and support young children with their studies.