ADITYA DUTT

+1(352) 226 5295 | adityadutt1996@gmail.com | linkedin://in/adityadutt12 | github://AdityaDutt | adityadutt.github.io |
adityadutt.medium.com

EDUCATION

Ph.D., **Machine Learning**, University of Florida, Gainesville, Florida, USA   *(Aug 2019 – Present)*   ***3.68/ 4.0***
M.S., **Computer Science,** University of Florida, Gainesville, Florida, USA   *(Jan 2018 – May 2019)*   ***3.70/ 4.0***
B.Tech., **Computer Science,** Jaypee Institute of Information Technology, Noida, India   *(Aug 2013 – July 2017)*   ***6.95/ 10.0***

PUBLICATIONS

- **Dutt A.** and Gader P., "Wavelet Multiresolution Analysis Based Speech Emotion Recognition System Using 1D CNN LSTM Networks", *IEEE Transactions of Audio, Speech, and Language Processing*, 2022, **Under Review.**
- **Dutt A.**, Zare A., and Gader P., "Shared Manifold Learning Using a Triplet Network for Multiple Sensor Translation and Fusion with Missing Data", *IEEE J. Sel. Topics Appl. Earth Observ.*, Oct. 2022.
- **Dutt A.**, Dutt R., Nagaraja H., "Development of Cost-Effective Substitute of Braille Tactile Display Operated by Linear Actuators", *IEEE Xplore*, 2017.
- Kumar S., Pant M., Kumar M., and **Dutt A.**, "Color image segmentation with histogram and homogeneity histogram difference using evolutionary algorithms", *Int. J. Mach. Learn. & Cyber (Springer),* Apr. 2015.

WORK EXPERIENCE

Graduate Research Assistant for **Dr. Alina Zare and Dr. Paul Gader**   *Jan 2021 – Present*

Project Title: Adaptive Manifold Learning for Multisensor Translation and Fusion given Missing Data, Funded by NSF
ESSIE Department, University of Florida, Gainesville, FL

- The goal of the project is to build a contrastive learning based deep learning framework to translate different streams of data from individual sensors like spectrometer, LiDAR, radar, etc. into a shared manifold.
- The fused information of multiple sensors is used to perform an accurate and robust classification.
- Created a unified analysis model which is sensor independent. It will help in saving development costs for separate analysis models for every sensor.
- The embeddings of a missing sensor can also be predicted using the other sensors' embeddings. The predicted embeddings can be used to reconstruct the missing sensor's original data.

Graduate Research Assistant for **Dr. Aditya Singh**   *Jan 2020 – Dec 2020*

Project Title: Landscapes in Flux: the influence of demographic change and institutional mechanisms on land cover change, climate adaptability and food security in rural India, Funded by NASA LCLUC Program
Institute of Food and Agricultural Science, University of Florida, Gainesville, FL

- Implemented machine learning methods to estimate indicators of poverty and food security across economically depressed regions of India by analyzing regional patterns of land cover change.
- Created an SQL database to gather and analyze data from India's National Sample Survey Office. Data is comprised of housing conditions, employment, and agricultural assessment surveys of India.
- Built an autoencoder-based model and utilized the latent space to predict over 300 economy and poverty indicators from census data for regions (each district was divided into hexagonal regions) smaller than districts.

Graduate Teaching Assistant for Course **"Distributed Operated Systems"**   *Aug 2019 – Dec 2019*

Computer Science Department, University of Florida, Gainesville, FL
- Created course assignments that were approved by the supervising professor. Managed one Teaching Assistant and one Course Grader. Held weekly office hours, graded coursework, and championed student development.
- Collaborated in creating projects (and grading system) based on Twitter implementation and gossip algorithms in Elixir using the actor model to simulate the dissemination of information across a network using Honeycomb and 3D Torus.
- Received an end-of-semester evaluation score of "Outstanding" based on student feedback.

Internet of Things (IOT) Intern   *Jun 2016 – Jul 2016*

Indian Institute of Technology, Delhi, India
- Tasked with the challenge of controlling the rpm of a DC motor clock so that it synchronized with data from a GPS using Beagle Bone.
- Created a webpage using HTML, CSS, and JavaScript to monitor live data from microcontroller to detect and resolve time and clock errors.
- Gathered data from Arduino and Wi-Fi module (ESP8266) and stored it in the SQL database. Next, pushed data from database on the webpage allowing microcontroller hardware to be controlled remotely.

RESEARCH EXPERIENCE

Emotion Detection Based on Text and Speech, Advisor: Dr. Paul Gader   *Jan 2020 – Present*
University of Florida, USA

- Researching on detecting different emotions from speech by analyzing spectrograms as a visual representation speech features such as the pause between each word, pitch, loudness, and more.
- Testing an analytical blend of words with aggregate factors (pauses, pitch, etc.) to do sentiment analysis including more accurately which may alert to security breaches or health issues.

Lexicon-Based Segmentation of Offline Cursive Handwriting, Advisor: Dr. Paul Gader   *Jan 2019 – Dec 2019*
University of Florida, USA

- The objective was to develop an algorithm for cursive handwriting recognition. CEDAR dataset is used for this project.
- Implemented a combination of connected component analysis and distance transform to segment (or oversegment) word into possible characters. The small segments which are not characters can be used as a good dataset to test outlier detection algorithms.
- Created possibilistic target outputs instead of one-hot encoding to represent the ambiguity of overlapping classes effectively.
- Used a dynamic programming algorithm and **CNN** to find a match score between each string in the lexicon and the segments.
- A Siamese/ Triplet network was also used to classify handwritten words along with CNN. It improved the model accuracy. It was suitable for this problem because we have imbalanced classes and fewer samples in some of the classes. And a Siamese Network can handle imbalanced classes and work with fewer number of samples.

ACADEMIC PROJECTS

Audio Classification using Wavelet Transform *[Python, Keras, Librosa, Batch Normalization]*   *2021*
The goal of this project is to use continuous wavelet transform to perform a multi-resolution analysis and classify different speakers. Free Spoken Digit Dataset (FSDD) was used (70% training/ 30% testing per class). Developed a CNN model and used batch normalization, to reduce the problem of internal covariate shift and make model more robust. Achieved an accuracy of 97% on test dataset. Here is my article published on Medium based on this project. (Article Link: https://medium.com/mlearning-ai/audio-classification-using-wavelet-transform-and-deep-learning-f9f0978fa246)

Bird Song Classification *[Python, Keras, Siamese Networks, Few-shot Learning, 1D Dilated Convolutions]*   *2021*
Every species of bird has their unique sound. Built a deep learning model to classify 9 bird species by songs. British Birdsong Dataset was used for this project. **Spectrograms** were extracted from audio using **librosa** library and a high pass filter was applied. Since the data samples per class were few, a **Siamese network** architecture was used. Each network in the Siamese network is made of several 1-D **Dilated convolutions** along with **Batch normalization** layers. The model was trained using triplet loss. Accuracy on test set was around 97%. Because of Batch normalization layer, the model was robust and converged very quickly. Here is my article published on Medium based on this project.
(Article Link: https://towardsdatascience.com/bird-song-classification-using-siamese-networks-and-dilated-convolutions-3b38a115bc1)

COVID-10 Speech Tracker *[Python, Google Dialogflow]* *2020*

Created a speech-based tracker to answer the questions related to covid-19 heath information, deaths and confirmed cases in any county, state or country between any time period. Extracted covid data from '*csbs*' and '*jhu*' databases using a python package. Used **Dialogflow** to detect and extract intent and time period from speech and respond with appropriate answer.

Voice controlled Web Browser *[Python, Wit.ai, Selenium]* *2020*

Created a voice-controlled web browser using **wit.ai** to train the model and extract intent from voice command. Used **selenium** to access the web page and control it as per user's command. Used **Beautiful Soup** to scrape weather information and news from New York times. The various commands supported were- switching tabs/ windows, scrolling, controlling scroll speed, play video on YouTube, mute, play/pause. Used **gTTs** library to convert output to speech.

Mussel Mounds Detection with Aerial Imagery *[Python, Scikit, NumPy, OpenCV]* *2019*

Worked on a project to identify settlements of mussels. Filtered water and green region near coastal areas by converting image from RGB to HSV format and applying several color masks using **Scikit** and **OpenCV** libraries. Successfully detected greenish gray mounds through grayscale thresholding and circular blob detection. Filtered blobs within threshold size. Matched results with LIDAR ground truth data.

Music Genre Classification *[Python, Librosa, CNN, LSTM]* *2019*

Collaborated on project to accurately segment music based on different genres. Extracted multiple 30-second parts of song to compute MFCC features of each chunk using a **librosa** library and **Keras**. Applied **PCA** on features and a combination of **CNN** and **LSTM** to achieve an 85% classification accuracy. Also used CNN directly on **spectrograms** and compared the output performance with earlier model.

Implementation of Information Propagation Protocols *[Elixir/OTP, Phoenix]* *2018*

Implemented information propagation protocols [Elixir/OTP, Phoenix], Gossip and Push Sum algorithm to simulate information dissemination in a network, network topologies such as 2D, Random 2D, 3D, Sphere to compare dissemination speeds, and chord algorithm for 10,000 nodes using GenServer with fault tolerance.

Bitcoin Simulator *[Elixir, Phoenix, JavaScript]* *2018*

Created a simulator using actor model in the elixir to make a distributed system to simulate bitcoin mining and observe its behavior. Implemented bitcoin wallet and target difficulty with maximum 8 zeros. Simulated 400 bitcoin transactions for 100 nodes using GenServer and implemented pool mining.

SKILLS

Tools: Keras, TensorFlow, Tensor Board, Pandas, Geopandas, NumPy, Scikit-Learn, OpenCV, Matplotlib, Plotly, PyQt, Seaborn, Spacy, NLTK, Beautiful Soup, Jupyter Notebook, QGIS
Languages: Python, SQL, MATLAB, Elixir, Julia, C/ C++, HTML, CSS, Linux