Capstone Projects

DLFA Program (Cohort 4)

COVID-19 detection using acoustic sounds

Domain: Audio

Dr. Sriram Ganapathy (sriramg@iisc.ac.in)

Short Description:

The COVID-19 pandemic has resulted in more than 250 million infections, and more than 5 million casualties. The pathogenesis of COVID-19 is increasingly suggesting impairments in the respiratory system. In this light, it is natural to ask - Can sound samples serve as acoustic biomarkers of COVID-19? If yes, an acoustics based COVID-19 diagnosis can provide a fast, contactless and inexpensive testing scheme, with potential to supplement the existing molecular testing methods, such as RT-PCR and RAT. The DiCOVA Challenge Series is an exploration of ideas to find answers to this question.

Considering the immediate societal relevance of a technology driven point-of-care-test (POCT) for COVID-19, the project has three aims. Explore the use of a curated dataset of sound samples (breathing, cough, and speech) drawn from individuals with and without COVID-19 during the time of recording collected in a crowdsourced manner. Developing deep learning based algorithms. Compare against the benchmark prior work and explore various modeling and optimization algorithms.

COVID-19 detection using acoustic sounds

Domain: Audio

Dr. Sriram Ganapathy

Reference: https://arxiv.org/pdf/2106.00639.pdf

Data Resource: https://github.com/iiscleap/MuDiCov

Use of NLP models for solving simple Math Word Problems

Domain: Text (NLP)

Dr. Shirish Shevade (shirish@iisc.ac.in)

Short Description:

The aim of this project is to study the performance of different Seq2seq models on simple math word problems. The relevant paper is available at https://arxiv.org/pdf/2103.07191.pdf and the dataset is available at https://github.com/arkilpatel/SVAMP/.

Reference: https://arxiv.org/pdf/2103.07191.pdf

Dataset: https://github.com/arkilpatel/SVAMP/

Study of Practical BERT models for Sequence Labeling

Domain: Text (NLP)

Dr. Shirish Shevade (shirish@iisc.ac.in)

Short Description:

In this project, you will study the use of small and practical BERT models (which can be implemented on a single CPU) for text classification and compare the performance with pretrained BERT on sentence/tweet classification. The relevant paper is available at https://arxiv.org/pdf/1909.00100.pdf. The datasets that could be used are available at https://huggingface.co/datasets/sst2 (sentence classification) and https://huggingface.co/datasets/zeroshot/twitter-financial-news-sentiment (tweet classification).

Reference: https://arxiv.org/pdf/1909.00100.pdf

Dataset(s):

1. https://huggingface.co/datasets/sst2
2. https://huggingface.co/datasets/zeroshot/twitter-financial-news-sentiment

Prompt Learning for Vision Language Models

Domain: Computer Vision

Dr. Rajiv Soundararajan (rajivs@iisc.ac.in)

Short Description:

The goal of this project is to study two variants of prompt learning of vision-language models for image classification. The experiments can be done with the ImageNet dataset.

Reference:

1. https://arxiv.org/pdf/2109.01134.pdf
2. https://openaccess.thecvf.com/content/CVPR2022/papers/Zhou_Conditional_Prompt_Learning_for_Vision-Language_Models_CVPR_2022_paper.pdf

Data Resource: ImageNet Dataset

Comparison of self-supervised features and prompt learning for image classification

Domain: Computer Vision

Dr. Rajiv Soundararajan (rajivs@iisc.ac.in)

Short Description:

The goal of this project is to understand whether multi-modal learning is better than pure feature learning with images. We will compare a self-supervised learning (SSL) framework such as Dino-v2 (https://arxiv.org/abs/2304.07193) with prompt learning (https://arxiv.org/pdf/2109.01134.pdf). The idea in SSL is that pretrained features are mapped to a distribution of classes using a single linear layer. On the other hand multi-modal models can be mapped to classes using prompt learning. The goal is to understand which of these approaches performs better. Note that in both these approaches, we wish to take pre-trained features and use them in different ways for image classification.

Reference:

1. https://arxiv.org/abs/2304.07193
2. https://arxiv.org/pdf/2109.01134.pdf

Training a speech recogniser using data from speech synthesis

Domain: Audio

Dr. Prasanta Kumar Ghosh (prasantg@iisc.ac.in)

Short Description:

Generally, training automatic speech recognition (ASRs) systems require paired data of speech and text. In this problem statement, you will be training an ASR with only text. This is done by using a pre-trained multispeaker text-to-speech (TTS) model to generate this. The generated speech, along with the corresponding text, is used to train ASR. The trained ASR will be evaluated on unseen sentences for seen and unseen speakers from the multi-speaker TTS system.

Training a speech recogniser using data from speech synthesis

Domain: Audio

Dr. Prasanta Kumar Ghosh (prasantg@iisc.ac.in)

Pretrained multispeaker TTS using coqui-ai:

https://huggingface.co/projecte-aina/tts-ca-coqui-vits-multispeaker

Training ASR using speechbrain:

https://colab.research.google.com/drive/1aFgzrUv3udM_gNJNUoLaHIm78QHtxdIz?usp=sharing

Reference(s):

1. https://arxiv.org/pdf/2306.00998.pdf

Training a Voice conversion model

Domain: Audio

Dr. Prasanta Kumar Ghosh (prasantg@iisc.ac.in)

Short Description:

Voice conversion refers to changing the speaker characteristics for audio while preserving the content. We will be trying to achieve this by learning phoneme-based content representations. For speaker embeddings, we will use a pre-trained speaker verification model. This trained model will be used on a test set consisting of unseen speakers.

Training a Voice conversion model

Domain: Audio

Dr. Prasanta Kumar Ghosh (prasantg@iisc.ac.in)

Reference(s):

Related voice conversion model: https://github.com/auspicious3000/autovc

For the speaker verification model: https://github.com/pyannote/pyannote-audio

Reinforcement Learning based Task Offloading in Mobile Edge Computing

Domain: Networking

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description:

A fundamental problem in MEC is how to efficiently offload heterogeneous tasks of mobile applications from user equipment (UE) to MEC hosts. Recently, many deep reinforcement learning (DRL)-based methods have been proposed. The project would involve a quick review of these followed by implementation and analysis of MRLCO.

Reinforcement Learning based Task Offloading in Mobile Edge Computing

Domain: Networking

Dr. Chandramani Singh (chandra@iisc.ac.in)

Reference(s):

1. J. Wang, J. Hu, G. Min, A. Y. Zomaya and N. Georgalas, "Fast Adaptive Task Offloading in Edge Computing Based on Meta Reinforcement Learning," in IEEE Transactions on Parallel and Distributed Systems, vol. 32, no. 1, pp. 242-253, 1 Jan. 2021

Client Selection for Federated Learning

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description:

In FL, the number of clients could be sufficiently large, but the bandwidth available for model distribution and re-upload is quite limited. he client selection policy is critical to an FL process in terms of training efficiency, the final model's quality as well as fairness. The project would involve a quick review of client selection algorithms followed by implementation and analysis of .RBCS-F.

Client Selection for Federated Learning

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description:

Reference(s):

1. T. Huang, W. Lin, W. Wu, L. He, K. Li and A. Y. Zomaya, "An Efficiency-Boosting Client Selection Scheme for Federated Learning With Fairness Guarantee," in IEEE Transactions on Parallel and Distributed Systems, vol. 32, no. 7, pp. 1552-1564, 1 July 2021

Federated Learning with Imbalanced Data

<u>Domain</u>: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

<u>Short Description:</u>

In FL, the data distribution of the participating agents can highly imbalanced which will increase the bias of model. Several frameworks have been proposed to counter this problem. The project would involve a quick review of client selection algorithms followed by implementation and analysis of Astraea.

Federated Learning with Imbalanced Data

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Reference(s):

1. M. Duan, D. Liu, X. Chen, R. Liu, Y. Tan and L. Liang, "Self-Balancing Federated Learning With Global Imbalanced Data in Mobile Systems," in IEEE Transactions on Parallel and Distributed Systems, vol. 32, no. 1, pp. 59-71, 1 Jan. 2021

Anomaly Detection in Industrial IoT

Domain: IoT Security

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description:

IoT networks offer fertile ground for malicious attackers to steal, manipulate, and perform nefarious activities. In typical IoT networks, data are generated continuously by IoT devices and are non-transferable to the main server. The project would involve a quick review of distributed learning based anomaly/intrusion detection algorithms followed by implementation and analysis of AMCNN-LSTM.

Anomaly Detection in Industrial IoT

Domain: IoT Security

Dr. Chandramani Singh (chandra@iisc.ac.in)

Reference(s):

1. Y. Liu et al., "Deep Anomaly Detection for Time-Series Data in Industrial IoT: A Communication-Efficient On-Device Federated Learning Approach," in IEEE Internet of Things Journal, vol. 8, no. 8, pp. 6348-6358, 15 April, 2021
2. V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha and G. Srivastava, "Federated-Learning-Based Anomaly Detection for IoT Security Attacks," in IEEE Internet of Things Journal, vol. 9, no. 4, pp. 2545-2554, 15 Feb.15, 2022

Computation Offloading Using Deep Reinforcement Learning

Domain: Mobile Edge Computing

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description:

Mobile edge computing (MEC) emerges recently as a promising solution to relieve resource-limited mobile devices from computation-intensive tasks, which enables devices to offload workloads to nearby MEC servers and improve the quality of computation experience. In this paper, an MEC enabled multi-user multi-input multi-output (MIMO) system with stochastic wireless channels and task arrivals is considered. In order to minimize long-term average computation cost in terms of power consumption and buffering delay at each user, a deep reinforcement learning (DRL)-based dynamic computation offloading strategy is investigated to build a scalable system with limited feedback. Specifically, a continuous action space-based DRL approach named deep deterministic policy gradient (DDPG) is adopted to learn decentralized computation offloading policies at all users respectively, where local execution and task offloading powers will be adaptively allocated according to each user's local observation.

Computation Offloading Using Deep Reinforcement Learning

Domain: Mobile Edge Computing

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description (continued):

Numerical results demonstrate that the proposed DDPG-based strategy can help each user learn an efficient dynamic offloading policy and also verify the superiority of its continuous power allocation capability to policies learned by conventional discrete action space-based reinforcement learning approaches like deep Q-network (DQN) as well as some other greedy strategies with reduced computation cost. Besides, power-delay tradeoff for computation offloading is also analyzed for both the DDPG-based and DQN-based strategies.

Reference(s):

Chen, Z., Wang, X. Decentralized computation offloading for multi-user mobile edge computing: a deep reinforcement learning approach. *J Wireless Com Network* **2020**, 188 (2020).

Federated Learning for Personalised Deep Neural Networks

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description:

Federated Learning (FL) is an emerging approach for collaboratively training Deep Neural Networks (DNNs) on mobile devices, without private user data leaving the devices. Previous works have shown that non-Independent and Identically Distributed (non-IID) user data harms the convergence speed of the FL algorithms. Furthermore, most existing work on FL measures global-model accuracy, but in many cases, such as user content-recommendation, improving individual User model Accuracy (UA) is the real objective. To address these issues, we propose a Multi-Task FL (MTFL) algorithm that introduces non-federated Batch-Normalization (BN) layers into the federated DNN. MTFL benefits UA and convergence speed by allowing users to train models personalised to their own data.

Federated Learning for Personalised Deep Neural Networks

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description (continued):

MTFL is compatible with popular iterative FL optimisation algorithms such as Federated Averaging (FedAvg), and we show empirically that a distributed form of Adam optimisation (FedAvg-Adam) benefits convergence speed even further when used as the optimisation strategy within MTFL. Experiments using MNIST and CIFAR10 demonstrate that MTFL is able to significantly reduce the number of rounds required to reach a target UA, by up to 5× when using existing FL optimisation strategies, and with a further 3× improvement when using FedAvg-Adam. We compare MTFL to competing personalised FL algorithms, showing that it is able to achieve the best UA for MNIST and CIFAR10 in all considered scenarios. Finally, we evaluate MTFL with FedAvg-Adam on an edge-computing testbed, showing that its convergence and UA benefits outweigh its overhead.

Federated Learning for Personalised Deep Neural Networks

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Reference(s):

J. Mills, J. Hu and G. Min, "Multi-Task Federated Learning for Personalised Deep Neural Networks in Edge Computing," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 3, pp. 630-641, 1 March 2022, doi: 10.1109/TPDS.2021.3098467.

Federated Multi-Armed Bandits based Recommender System

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description:

Local servers want to recommend the most popular items to their served customers to maximize the expected rewards. The item popularity can only be learned via interacting with customers, leading to a bandit problem (Li et al., 2010). As different local servers have potentially heterogeneous customers, their local popularities are non-IID. In addition, each local server only collects data from a small group of customers, and the central server needs to average the locally learned popularity models to have a global model, without accessing the individual recommendation for privacy protection. However, the globally most popular item may not apply to the small group of customers of a particular local server, which again leads to personalization in a federated bandit setting, i.e., a joint consideration of global and local item popularities. This work proposed a novel framework of personalized federated MAB (PF-MAB) to solve the recommendation problem.

Federated Multi-Armed Bandits based Recommender System

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Reference(s):

1. Shi, C., & Shen, C. (2021). Federated Multi-Armed Bandits. Proceedings of the AAAI Conference on Artificial Intelligence, 35(11), 9603-9611. https://doi.org/10.1609/aaai.v35i11.17156

2. Chengshuai Shi, Cong Shen, Jing Yang, “Federated Multi-armed Bandits with Personalization”, Proceedings of The 24th International Conference on Artificial Intelligence and Statistics, PMLR 130:2917-2925, 2021.

Multi-Armed Bandit-Based Client Scheduling for Federated Learning

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description:

By exploiting the computing power and local data of distributed clients, federated learning (FL) features ubiquitous properties such as reduction of communication overhead and preserving data privacy. In each communication round of FL, the clients update local models based on their own data and upload their local updates via wireless channels. However, latency caused by hundreds to thousands of communication rounds remains a bottleneck in FL. To minimize the training latency, this work provides a multi-armed bandit-based framework for online client scheduling (CS) in FL without knowing wireless channel state information and statistical characteristics of clients. Firstly, we propose a CS algorithm based on the upper confidence bound policy (CS-UCB) for ideal scenarios where local datasets of clients are independent and identically distributed (i.i.d.) and balanced.

Multi-Armed Bandit-Based Client Scheduling for Federated Learning

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Short Description (continued):

An upper bound of the expected performance regret of the proposed CS-UCB algorithm is provided, which indicates that the regret grows logarithmically over communication rounds. Then, to address non-ideal scenarios with non-i.i.d. and unbalanced properties of local datasets and varying availability of clients, we further propose a CS algorithm based on the UCB policy and virtual queue technique (CS-UCB-Q). An upper bound is also derived, which shows that the expected performance regret of the proposed CS-UCB-Q algorithm can have a sub-linear growth over communication rounds under certain conditions. Besides, the convergence performance of FL training is also analyzed. Finally, simulation results validate the efficiency of the proposed algorithms.

Multi-Armed Bandit-Based Client Scheduling for Federated Learning

Domain: Federated Learning

Dr. Chandramani Singh (chandra@iisc.ac.in)

Reference(s):

W. Xia, T. Q. S. Quek, K. Guo, W. Wen, H. H. Yang and H. Zhu, "Multi-Armed Bandit-Based Client Scheduling for Federated Learning," in IEEE Transactions on Wireless communications, vol. 19, no. 11, pp. 7108-7123, Nov. 2020, doi: 10.1109/TWC.2020.3008091.

Emotion Recognition From Speech and Beyond

Domain: Audio

Dr. Sriram Ganapathy (sriramg@iisc.ac.in)

Short Description:

With the growing demand for conversational agents and personal assistants, automatic recognition of human emotion has emerged as a key task in enabling enhanced user experience. Human emotion recognition using multi-modal data of text, speech and video has substantial impact on various applications like smartphones, wearable devices, smart speakers, driver monitoring in automotives, mood analysis and mental health. This area of developing emotional intelligence would allow machines to be more human-like in the interactions.

Reference(s):
https://www.sciencedirect.com/science/article/pii/S1746809420300501?casa_token=dPFJJddjncYAAAA:68YYkebYxgIhqcXXHYU-CVYGAGSq4vizQJHf2WOupKE9dKBQAmiVDJZFLDgBvL-WIS_2GeIqww

Data Resource: https://doi.org/10.1371/journal.pone.0196391

Q-Learning for grid world environments

Domain: Reinforcement Learning

Dr. Ambedkar Dukkipati (ambedkar@iisc.ac.in)

Short Description:

This project involves implementing a tabular Q-learning algorithm to solve simple grid-world environments in which the state and action spaces are finite and small. This algorithm can then be run and tested on the Cliff Environment, a grid world in which the agent is placed on a grid and must reach a goal position while avoiding certain other grid positions.

Reference(s):

1. Reinforcement Learning: An Introduction by Sutton and Barto, 2nd Edition, Chapter 6, for a description of the Cliff Walking environment and Q-learning.

2. The environment and example code for the algorithm: https://github.com/dennybritz/reinforcement-learning/tree/master/TD

3. Requires OpenAI gym for the environment, and Jupyter notebook along with other standard python libraries for running the example code.

Implementation of the REINFORCE Algorithm

Domain: Reinforcement Learning

Dr. Ambedkar Dukkipati (ambedkar@iisc.ac.in)

Short Description:

The tabular Q-learning algorithm cannot deal with environments with large or continuous state or action spaces. Such situations require using a function approximator such as a neural network. This project involves implementing the REINFORCE algorithm, which is a policy gradient algorithm in which a stochastic policy is implemented as a feed forward neural network whose parameters are updated after each episode using gradient descent and the policy gradient theorem. This algorithm can then be run and tested on the OpenAI gym CartPole environment.

Implementation of the REINFORCE Algorithm

Domain: Reinforcement Learning

Dr. Ambedkar Dukkipati (ambedkar@iisc.ac.in)

Short Description:

Reference(s):

1. Environment: https://gym.openai.com/envs/CartPole-v0/

2. Example code: https://github.com/udacity/deep-reinforcement-learning /tree/master/reinforce

3. Also requires the PyTorch framework for implementing the neural network and gradient descent.

4. https://towardsdatascience.com/policy-gradient-methods-104c783251e0 : A blog post describing the REINFORCE algorithm.

Interpretable GAN Controls

Domain: Computer Vision

Dr. Venkatesh Babu (venky@iisc.ac.in)

Short Description:

In this project, the latent space of StyleGAN will be explored to change the attribute of the generated image without changing the identity of the person.

Reference(s) / Project page: https://sites.google.com/view/flamelatentediting

Crafting Adversarial Samples for Deep Models

<u>Domain</u>: Vision/ML

Dr. Venkatesh Babu (venky@iisc.ac.in)

<u>Short Description:</u>

Adversarial samples are cleverly constructed inputs that are intentionally designed to deceive neural networks, leading them to make incorrect predictions or classifications. This project aims to explore and develop techniques for creating such adversarial samples, shedding light on the vulnerabilities of deep models. By understanding the methods used to craft adversarial samples, the project seeks to enhance the robustness and security of these models, which is crucial in applications ranging from computer vision to natural language processing.

<u>Reference(s)</u>: https://github.com/dangeng/Simple_Adversarial_Examples

Text only training for Image Captioning

Domain: Vision and Text

Dr. Venkatesh Babu (venky@iisc.ac.in)

Short Description:

Can we generate captions for the given image without having any Image-Text pairs? Yes, it is possible with CLIP (Contrastive Language–Image Pre-training) model, which was trained with large text-image pairs. By a clever trick you can train the CLIP model to generate the captions for the given image without using any image during training.

Reference(s):
https://github.com/DavidHuji/CapDec

Zero-shot segmentation using Stable Diffusion

Domain: Computer Vision

Dr. Prathosh A P (prathosh@iisc.ac.in)

Short Description:

In this project the goal is to use a pre trained stable diffusion model capable of segmenting anything in a zero-shot manner without any annotations.

Reference(s):

The relevant project page is here: https://sites.google.com/view/diffseg/home

Open set zero shot classification via stable diffusion

Domain: Computer Vision

Dr. Prathosh A P (prathosh@iisc.ac.in)

Short Description:

In this project, we address the question of whether a pretrained diffusion model be used for open set zero shot classification task. One possible method is described here - https://diffusion-classifier.github.io/

Reference(s):

https://diffusion-classifier.github.io/

Distillation for faster sampling in DDPMs

Domain: Model Compression / Computer Vision

Dr. Prathosh A P (prathosh@iisc.ac.in)

Short Description:

Here is the goal is to use the idea of knowledge distillation for reducing the sampling time in diffusion models.

Reference(s):

Here is the paper - https://arxiv.org/abs/2202.00512

Knowledge distillation for Speech Recognition

Domain: Model Compression / Speech recognition

Dr. Prathosh A P (prathosh@iisc.ac.in)

Short Description:

The goal of this project is to reduce the size of large models such as Wave2vec 2.0 for automatic speech recognition.

Reference(s):

Here are the relevant resources -

1. https://arxiv.org/abs/2303.09278
2. https://medium.com/georgian-impact-blog/compressing-wav2vec-2-0-f41166e82dc2

Time series forecasting via LLMs

Domain: LLMs/Time series

Dr. Prathosh A P (prathosh@iisc.ac.in)

Short Description:

Here, we would like to explore if LLMs trained on natural text be used for time series forecasting.

Reference(s):

Here is the paper - https://arxiv.org/abs/2310.07820