End-to-End Deep Learning Regression for Measurements with the CMS Experiment

Abstract:

This project is about to train and deploy an End-to-End Deep Learning Regression **model for estimating particle properties.**

Different channels can be used to the study **different patterns in the images using convolutional and pooling layers and** extracting only the portion which conveys the maximum information **using dimensional reduction techniques such as PCA etc.** for the training of the deep learning regression model.

Different feature extraction techniques can be used to study the **interaction between the channels of the images.** This can be treated as a supervised machine learning problem i.e., taking the images as independent variable and labels as the particle name as the dependent variable.

Personal Information:

Name: Saksham Dhawan

University: Guru Gobind Singh Indraprastha University, New Delhi

Institute: Maharaja Agrasen Institute of Technology

Email: sakshamdhawan179@gmail.com

Github: **https://github.com/saksham2172000**

Phone No: +91-9625899537

Location: Delhi

Time zone: Indian Standard Time (UTC +5:30)

➢ Community Bonding Period (May 17, 2021 – June 7, 2021)

- Discuss with mentor, a plan about implementing the project.
- Interacting with the community to know more about the use cases.
- Finalizing the plan for implementing the features.

➢ Timeline (June 7, 2021- July 16, 2021)

- **Week 1 (June 7, 2021 – June 13, 2021)**
  - Making the pull requests to the repository and writing the necessary documentation.
  - Writing the python script for gathering the dataset ready.

- **Week 2 (June 14, 2021 – June 20, 2021)**
  - Data cleaning process.
  - Checking for the outliers in the given image dataset.

- **Week 3 (June 21, 2021 – June 27, 2021)**
  - Data exploration of different channels of images.
  - Formation of different visualization for the images to get hold of some patterns in images.

- **Week 4 (June 28, 2021 – July 4, 2021)**
  - Performing feature selection and feature extraction technique.
  - Studying the interaction between the different channels of the images.
  - Getting the correlation of channel at each pixel.

- **Week 5 (July 5, 2021 – July 11, 2021)**
  - Getting the data ready for training a model i.e., normalizing the images etc.

- Searching for appropriate regression technique.

- **Week 6 (July 12, 2021 – July 18, 2021)**
  - Trying out different deep learning regression techniques, transfer learning etc.
  - Finalizing the regression technique.
  - Building the Model Architecture and training the model.

- **Week 7 (July 19, 2021 – July 25, 2021)**
  - Completing the process of training the model.
  - Checking for overfitting and underfitting on the training data.

- **Week 8 (July 26, 2021 – August 1, 2021)**
  - Using methods to overcome overfitting such as use of affining techniques on images using dropout layers etc.
  - Applying different metrics such as recall score, precision score, accuracy, f-score and confusion matrix to evaluate how good is the model.

- **Week 8 (August 2, 2021 onwards)**
  - Work on debugging and solving the issues according to the situation.
  - Creating the report for the results obtained throughout the project.

About Me:

I am currently pursuing Bachelor of Technology in Electronics and Communication Engineering. I am in third year; I have great knowledge of Machine Learning which I started learning last year.

I have used machine learning algorithms for creating projects such as caption bot, music generation, mountain car climb using reinforcement learning (Q-Learning), Covid Detection using X-ray images which involves deep learning etc. I have solid knowledge of python.

I have done one internship with one teammate which involved a project of Covid Twitter Data Analysis and we were declared as the Topper for the project for giving most accurate results and the project was organized by Spotle.ai.

I love to code in my free time, I will be full time available during the summer. Apart from coding I like to watch movies and play cricket.

Certificate of Internship – (Link)

Ranked 1st in the Internship Project

Evaluation Test Link – (Link)

Other Commitments:

- I would be able to commit to full time for GSoC during my summer holidays.
- I am applying to only this organization; I am ready to work if selected.
- I will be able to give 6-7 hours daily.

Post GSoC And Future Work:

I will continue to contribute for this project. I will assist the new members who are willing to contribute for this project. I will actively take part in the discussions and I will contribute by creating, solving issues and adding improvements.

Why me?

I believe that I am well suited for this project. I already know the vision of this project. I have clear understanding of different deep learning regression techniques. I have good knowledge of machine learning. I will put my full efforts to contribute to this project.