

Capstone Project Proposal

- ADVANCEMENTS IN CNN ARCHITECTURES -

By: Tanvi Gandhi

April 21, 2020

- MOTIVATION

In recent years, we have witnessed the explosion of Deep Neural Networks, especially Convolutional Neural Networks (CNNs) due to their ability in recognizing visual patterns directly from pixel images with minimal preprocessing. CNNs have been a part of the machine learning family since 1989 and were considered breakthrough due to their exemplary performance. This has inspired several ideas to bring advancements in CNNs have been explored, such as the use of different activation and loss functions, parameter optimization, regularization, and architectural innovations. Thus, we want to further explore the current state of the art practices in Deep Learning, relating to image recognition. These architectures include LeNet, Alexnet, ResNet and VGG-16. It would be interesting to tweak and play with the architectural components to try and come up with an improved model for the handwritten digits.

- WHO MIGHT CARE?

The domain of deep learning has still a lot to explore and it would be wise to understand the intrinsic details of the different dimensions of a Deep Neural Network so that we could come up with an improved version. Even the mere understanding of the basic principles would be fruitful and pave a way to dive into it. Further, it would be very beneficial for the organizations who are working in the image recognition field to have the comprehended architectures. The goal of this study is not only to have a deep understanding of the working of the NNs but also demonstrate myself and my knowledge and skills in this vast domain.

- DATA

The data comprises the MNIST dataset which is kind of a benchmark dataset for trivial image processing systems. Due to the limited GPU availability, we would be using the

mentioned dataset although better datasets are available like ImageNet, CIFAR10 etc. The MNIST dataset contains 60,000 training images and 10,000 testing images.

- APPROACH

As of now, we can't clearly decide on the approach to be followed and can keep on changing as we learn new concepts and methods. However, as an initial step, it seems apt to decide the two or three Neural Network architecture that we will move forward with. We will further tweak and play with the parameters and compare their accuracy on the MNIST dataset.

- DELIVERABLES

- **Jupyter Notebook:** A python notebook which will contain all the code involved in the process.
- **Final Report:** A document to highlight the entire process followed and key takeaways.
- **Presentation and/or Paper:** A presentation for the audience and/or a research paper to highlight the significant insights.