**Title:** Investigating Model-Agnostic Meta-Learning (MAML) for Fast Adaptation of Deep Networks

**Members:** Mayank Agarwal

**Problem:** Contemporary applications of Deep Learning methods require massive amounts of data for training. If sufficient data is not available, the model runs the risk of overfitting and poor generalization. This however, is in contrast to the natural way of learning, where a human can learn though a small number of data points. This is called few-shot learning.

A recent paper by Chelsea Finn et. al. (<https://arxiv.org/pdf/1703.03400.pdf)> proposes an algorithm that trains a model to learn representations that can be quickly altered to a new task with few examples.

This domain is interesting because this provides a way to apply Deep Learning to areas with limited training data.

**Dataset:** Omniglot is a popular dataset for one-shot learning. Along with this, I’ll be experimenting with MNIST dataset, and CIFAR-10 as well, to understand the algorithm dynamics in different contexts.

**Novelty:**

The project will focus on:

1.     Implementing and replicating the results presented in the paper

2.     Investigating the effect of MAML on features learned by the network

The novelty of project comes from the second point. Specifically, I’m interested in finding answers to the following questions:

1.     What features does a network learn when trained with and without MAML?

2.     How do these learned features differ from the features learned in a traditional deep network?

3.     How do these learned features adapt themselves to new tasks with a small number of examples?

**Readings:**

·      Finn, Chelsea, Pieter Abbeel, and Sergey Levine. "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks." arXiv preprint arXiv:1703.03400 (2017).

·      Ravi, Sachin, and Hugo Larochelle. "Optimization as a model for few-shot learning." (2016).

·      Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." *European conference on computer vision*. Springer, Cham, 2014.

·      Yosinski, Jason, et al. "Understanding neural networks through deep visualization." *arXiv preprint arXiv:1506.06579* (2015).

**Evaluation:** The project will apply the algorithm to few-shot classification tasks, and therefore, would use Cross-entropy loss and test accuracy as the evaluation metrics. Different models trained with and without MAML would be compared using the accuracy on the test set.

Qualitatively, the project would employ deep visualization techniques to visualize the features learned by the network. Accompanying these would be graphs depicting the model’s performance with varying number of per class samples provided for training.