**Federated Learning with Pretrained Text DNNs**

DATA590 Project Proposal

University of Washington Autumn 2019

Arjun Singh

Joel Stremmel

**Executive Summary / Introduction**

Federated Learning aims to train machine learning models in a distributed fashion without centralizing data but instead updating and passing model parameters from a central server to distributed entities and back to perform stochastic gradient descent. McMahan et al. propose the Federated Averaging algorithm in [Communication-efficient learning of deep networks from decentralized data.](https://arxiv.org/abs/1602.05629) This algorithm and associated experiments in the paper yield promising results but are limited to models with randomly initialized weights. We aim to address the problem of applying state-of-the-art pretrained text models like [BERT](https://arxiv.org/abs/1810.04805) (or more pruned versions, such as [ALBERT](https://arxiv.org/pdf/1909.11942.pdf)) for weight initialization in the context of the Natural Language Understanding tasks at which they excel.

**Problem Statement**

**Background / Literature Review**

**Work-to-Date / Data Review**

**Proposed Solutions**

(including specifics, toolboxes, language, hosting solutions) - final deliverables specified

**Risks & Benefits of Proposed Solution**

**Schedule**

**Team Bios**