HW 6

Student Name

1/21/2024

What is the difference between gradient descent and *stochastic* gradient descent as discussed in class? (*You need not give full details of each algorithm. Instead you can describe what each does and provide the update step for each. Make sure that in providing the update step for each algorithm you emphasize what is different and why.*)

Gradient descent and stochastic gradient descent are both optimization algorithms used to minimize the cost function in machine learning models. The primary difference between them lies in how they update the model parameters based on the gradient of the cost function. Gradient descent computes the gradient of the cost function with respect to all training examples in the dataset and then updates the model parameters based on this average gradient. It's a deterministic process because it considers the entire dataset for each parameter update. On the other hand, stochastic gradient descent randomly shuffles the training examples and then updates the model parameters after computing the gradient of the cost function with respect to just one training example at a time. It's a stochastic process because it randomly selects a single training example to compute the gradient and update the parameters.

Consider the `FedAve` algorithm. In its most compact form we said the update step is $\omega_{t+1} = \omega_t - \eta \sum_{k=1}^{K} \frac{n_k}{n} \nabla F_k(\omega_t)$. However, we also emphasized a more intuitive, yet equivalent, formulation given by $\omega_{t+1}^k = \omega_t - \eta \nabla F_k(\omega_t); w_{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^k$.

Prove that these two formulations are equivalent.
(*Hint: show that if you place* ω_{t+1}^k from the first equation (of the second formulation) into the second equation (of the second formulation), this second formulation will reduce to exactly the first formulation.)

Substituting ω_{t+1}^k from the first equation into the second equation, we get:

$$\omega_{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} (\omega_t - \eta \nabla F_k(\omega_t))$$

Distributing the sum over ω_t and $\eta \nabla F_k(\omega_t)$, we get:

$$\omega_{t+1} = \omega_t \sum_{k=1}^{K} \frac{n_k}{n} - \eta \sum_{k=1}^{K} \frac{n_k}{n} \nabla F_k(\omega_t)$$

Since $\sum_{k=1}^{K} \frac{n_k}{n} = 1$, the first term simplifies to ω_t:

$$\omega_{t+1} = \omega_t - \eta \sum_{k=1}^{K} \frac{n_k}{n} \nabla F_k(\omega_t)$$

And this is exactly the first formulation of the FedAve algorithm:

$$\omega_{t+1} = \omega_t - \eta \sum_{k=1}^{K} \frac{n_k}{n} \nabla F_k(\omega_t)$$

Now give a brief explanation as to why the second formulation is more intuitive. That is, you should be able to explain broadly what this update is doing.

The second formulation of the FedAve algorithm is more intuitive because it directly illustrates the process of federated learning. In federated learning, multiple clients collaborate to train a shared global model while keeping their data decentralized and in the second formulation of the FedAve algorithm, individual groups send their model updates to a central server. The server then calculates the average of these updates, giving more importance to updates from groups with larger datasets, which ensures that the shared model reflects the contributions of all groups, with greater emphasis on those with more data.

Explain how the harm principle places a constraint on personal autonomy. Then, discuss whether the harm principle is *currently* applicable to machine learning models. (*Hint: recall our discussions in the moral philosophy primer as to what grounds agency. You should in effect be arguing whether ML models have achieved agency enough to limit the autonomy of the users of said algorithms.*)

The harm principle posits that the only justification for limiting an individual's freedom of action is to prevent harm to others which means individuals should be free to pursue their own interests and make their own choices as long as those choices do not harm others. This principle places a constraint on personal autonomy because it suggests that one's autonomy should not infringe upon the autonomy or well-being of others. When we consider the harm principle in regardsto machine learning models, we have to evaluate whether these models have achieved agency to the extent that they can potentially harm others and therefore warrant constraints on personal autonomy. Machine learning models lack the capacity for moral agency so harm that comes from these models really comes from the people using these models carelessly. Machine learning models can have significant societal impacts, such as reinforcing biases, perpetuating discrimination, and invading privacy if misused or poorly designed. Ethical considerations and regulations should focus on mitigate potential harms by guiding the actions of the individuals and organizations involved in the development and deployment of these models.