Trishal Muthan
Period 2
Fast.ai 1 Responses

Chapter 1:

- **Why is it hard to use a traditional computer program to recognize images in a photo?**
    - It's hard to use a traditional computer program to recognize images in a photo because these traditional computer programs function very differently compared to actual brains. The brain's overall architecture works in a very distinct way allowing it to do things like analyze pictures or understand language, things that traditional computers have a very hard time with. By constructing an approach more similar to how the brain works, computer programs might better be able to handle these types of tasks.
- **What did Samuel mean by "weight assignment"?**
    - By "weight assignment" he essentially meant a particular set of values for set variables or "weights" that define the specific operation of the program. Your specific weight assignment can have a drastic impact on the capabilities of your model.
- **What is overfitting? Provide an example.**
    - Overfitting is when you train for too long or too frequently on the same set of data, causing the model to learn the specific tendencies and details of that particular dataset. This causes the model to have a lower accuracy when classifying data points it has not encountered before. For example, if you are creating a handwritten-digit classifier, if you only trained on digits written by a single person, the model might overfit to that person's handwriting and cause future predictions on other data points to be incorrect.
- **What kinds of features do the early layers of a CNN find? How about the later layers?**
    - The main distinction between early and later layers of a CNN is that the early layers mainly detect low-level features such as different colors or edges, while the later layers of a CNN are able to distinguish high-level features such as more complex patterns, objects, and shapes.

Chapter 2:

- **What are possible negative societal implications of text generation models?**
    - There are numerous possible negative societal implications of text generation models. The primary one is that of inaccuracy. It is very easy for a model to create text that is contextually compelling but is actually inaccurate. Some negative consequences of this might be the spread of disinformation and unrest

or the harming of people (something like this might occur if NLP is used for medical advice, where the advice is actually incorrect).

- **What's a key downside of directly using a deep learning model for recommendation systems?**
    - The key downside is that these models will only tell you what products a particular user might like, rather than what recommendations would be helpful for a user. For example, the model might show products that the user is already familiar with or different versions of products they already purchased.
- **What is data augmentation? Why is it needed?**
    - Data augmentation refers to the creation of random variations of input data, such that they appear different, but do not actually change the meaning of the data. It allows us to create more, artifical data to train on, allowing us to improve the accuracy and capabilities of our models.
- **What are 3 examples of problems that could occur when rolling out a bear warning system in practice?**
    - Out of domain data is one issue where the model might see some data in production that it never saw anything like in training. Another is domain shift where the types of data that the model sees changes over time, causing the original training data to become irrelevant. The last example is that the model might change the behabior of the system that it is part of (for example a model that predicts crime in certain neighborhoods might have a large impact on things like policing and the justice system).

Chapter 4:

- **What is the difference between tensor rank and shape? How do you get the rank from the shape?**
    - A rank is the the number of axes or dimensions in a tensor while the shape is the size of each axis of a tensor. The length of a tensor's shape is equal to the rank.
- **What is broadcasting?**
    - Broadcasting is a technique used to perform subtraction between two tensors of different rank. It will automatically expand the tensor with the smaller rank to have the same size as the one with the higher rank.
- **Why can't we always use a high learning rate?**
    - If we use a high learning rate then there might be a chance of your loss function reaching diverging. The gradient might keep jumping around the actual absolute minimum and ends up unable to reach the minimum because of the high learning rate.
- **Why do we have to "zero" the gradients?**

- In PyTorch, gradients are accumulated, so it becomes necessary to zero the gradients before doing back propagation or every time you update the weights.