##### **In traditional machine learning (non-NN methods), we use feature engineering to model complex relationships between observed variables (features) and a target variable (response). When using deep learning methods, should we design and incorporate feature engineering processes? Explain why or why not.**

In general I would think that feature engineering should not be used in deep learning models, largely because it might influence or encourage the model to learn patterns that are not actually there or relevant. Traditionally with non-NN Machine Learning models features are usually chosen that the Data Scientist or subject matter expert believe might be beneficial in the model. Then through the modeling the features out of the ones started with are determined as important or not; however only the features that are tested can be chosen to be in the model. Therefore, if there are more granular details or features that are significant in predicting the target output that were not included in the modeling, the model would be missing this important explaining power. In contrast, a Neural Network model can identify the most important features within the data and what the connections are between certain features to increase the complexity of the weights being used when predicting an output.

**Explain in words or equations, why we should introduce nonlinearity in neural networks.**

Nonlinear functions in Neural Networks is important because it helps the model to learn complex patterns and be able to predict very accurate predictions on data the model has never seen before.

**You are training a deep learning model to predict sentiment of Twitter posts — the model predicts whether a post is “happy” or “sad”. Your model achieves 0.95 accuracy on the dataset you used to train the model. But when you take new posts from Twitter and use your model to predict the sentiment, the model performs much worse. What might have happened? What should you do to improve your model?**

If a model has very high accuracy on the training data and low accuracy on new test data, then the model likely has overfit the training data. This means that the model has learned the training data so well that anything that is different or not very similar to the training data the model will not know how to handle. Two ways to decrease overfitting or the chances of it occurring would be, one to add more data that are representative of the data in which you want the model to perform well on, or be sure to include various examples that capture the variation that the data may have. The second technique would be to simplify the model or ending the training for the model early, such that the accuracy on the training data might not be as high as before, but it would likely have better accuracy on data that the model has never seen.

**The MNIST dataset consists of images of dimension 28x28 pixels with one color channel (28x28x1), with each image corresponding to a label between 1 and 10. To build a classifier, we implement a multi-layer perceptron model with 3 hidden layers. The first two hidden layers have 100 perceptrons each, and the third hidden layer has 30 perceptrons. Calculate how many weights will be updated for each iteration of gradient descent. Show your work.**

For the first layer we have the images of size 28 x 28 x 1, which ends up turning in 28x28 = 784 neurons for the first layer. The second layer has 100 neurons, which ends up giving us 784 x 100 = 78,400 weights for the first layer. The third layer has 100 neurons as well, which with the first layer has 100 x 100 = 10,000 weights for the second layer. The third layer has 30 neurons, equaling 100 X 30 = 3,000 weights, and lastly 30 x 10 to get 300 weights for the output layer. Overall, the model has 78,400 + 10,000 + 3,000 + 300 = 91,700 weights that are updated with each iteration.

**Answer the following questions based on a close reading of this article and possibly additional research (remember to cite your sources).** [**https://www.nytimes.com/2023/06/28/technology/facial-recognition-shopliftersbritain.html**](https://www.nytimes.com/2023/06/28/technology/facial-recognition-shopliftersbritain.html)

1. **What is the technology being discussed in this article? How does it relate to deep learning?**

The article talks about facial recognition being used in stores to identify individuals who had previously stolen or committed a crime and have been added to a watchlist connected to a network of stores. Facial recognition is relevant to deep learning because identifying individuals’ faces is a task that deep learning models are used to solve for. Similar to how Convolutional Neural Networks are used to correctly identify objects within an image, facial recognition aims to correctly identify one person from another.

1. **What are some ethical concerns discussed in the article? Do you share these concerns? Why or why not?**

The main ethical concern as mentioned in the article is improperly identifying someone as being on the watchlist when they are not. I can see this being a possible issue, but more so depending on how the business decides to act given the response from the Facewatch facial recognition software. If the facial recognition has a very low error rate and the action taken by the store is that the individual is told to leave and police are not involved, then the software seems to be pretty safe to use, with the biggest risk being that the store has falsely accused a customer of being someone that previously shoplifted. However, if there is a more serious response like having the police involved, the risk of the facial recognition system being wrong is much higher, and therefore is of more concern.

1. **Are these concerns addressed by the company deploying this technology? How?**

Yes, it seems like the company does acknowledging the fact that their Facewatch system could have false positives (identifying someone as being on the watchlist, but actually isn’t). Also, the company is addressing concerns of local governments and the public, adding limitations on what individuals would be flagged based on the severity of the crime.

1. **What is your personal view of the way this technology is being used? Explain your reasoning.**

I think facial recognition is something that would have a net positive impact on the population, and the necessary precautions and ethics of the concept are being appropriately addressed, at least according to this article. Within the past few years, I think the anticipation and acceptance of possibly being recorded has increased with the dramatic increase in content being posted and availability of cameras. Therefore, though facial recognition software and how and where it is used might be concerning to the public, I think if properly communicated and if it has a low chance of false positives, then facial recognition software will be something that people will adapt to.