Adrien Philardeau-Planche

MAIS Assignment 4 Writeup

i) After researching the Fashion MNIST dataset, a 2-D CNN seemed to be the best approach to this problem, as the pictures have a spatial relationship. I tried to implement a multi-layer CNN, but it didn't work, so I stuck to a more simple implementation.

ii) My results were weird. As I explain in iii), my 3-layer CNN was reaching accuracies on the validation set close to 90%, but resulted in a measly 30% accuracy on Kaggle. Disappointing. A more simple CNN gave me an accuracy in the 80s on the validation set, and a 69% accuracy on Kaggle, which ended up actually being my best model in the competition. Not great at all, but not awful. Tweaking the hyperparameters would affect the accuracy on the validation set by a few percent, and each epoch would typically increase the accuracy for the first few, and then have really small gains near the end, even regressing sometimes. I tweaked the dropouts a little bit, but it didn't make a huge difference, at least on the validation set.

iii) Figuring out how to actually make and format the numpy arrays containing the images and labels was more difficult than expected. Then, understanding how CNN's and their various attributes such as MaxPooling, Dropout, Flatten, Dense, Activation work and affect the results was interesting, but took some time. At one point I implemented a multi-layer convolutional network, but the main drawbacks were that each epoch took about 3 minutes, so training the model took upwards of an hour, and then the submission resulted in an awful accuracy of about 30%, probably meaning the neural network was incredibly overfit, which was surprising considering there was a Dropout at each step, one of them being 50%. At this point it was 11:30 PM so I had to quickly simplify my model, learn about one-hot encoding, and only have time to run 5 epochs, but I submitted at about 11:58, with adrenaline coursing through my veins.

iv) I learned a lot during this assignment. From how CNN's work, to how to use Keras, to the MNIST datasets, to data processing, to working with CSV's, this assignment was a good introduction to how data science actually works. I also have a newfound appreciation for how truly impressive it is for people to come up with neural networks and machine learning algorithms in general.

v) No contributors.

https://github.com/adrienphilardeau/FashionNMIST

Adrien Philardeau-Planche