Prerak Mehta

MSDS 458 Artificial Intelligence and Deep Learning

Second Research Assignment

Abstract:

In this assignment we use the CIFAR-10 dataset (Canadian Institute For Advanced Research) which is a collection of images that are commonly used to train machine learning and computer vision algorithms. It is one of the most widely used datasets for machine learning research. The dataset contains 60,000 32x32 color images in 10 different classes such as airplanes, cars, birds, cats, deer, frogs. Horses, ships, and trucks. The images are even distributed among the classes i.e. each class has 6,000 images. This research is concerned with deep learning, a neural network with more than one hidden layer. Deep neural networks will be created to analyze how various factors affect the fitting and the test set performance of these networks. The performances of the DNNs will be compared against the performances of the Convolutional Neural Networks of differing structures in terms of pooling layers. Four experiments will be performed with differently structured CNN and DNN. Later on in the research all the experiments will be repeated by adding regularization - a process which regularizes or shrinks the coefficients towards zero. In simple words, regularization discourages learning a more complex or flexible model, to prevent overfitting. Comparisons will be made using the training, testing and validation accuracy, loss and processing time for each model.

Introduction:

This research assignment addresses a hypothetical management problem that involves developing a neural network model that basically classifies images with a high accuracy. The goal is to further this model and develop some sort of facial recognition software for a mobile device. This research assignment serves as a hands-on practical experience with not only designing, training and assessing a neural network but also understanding whether employing more hidden layers and convolutional layers positively impact the accuracy of the image classification model.

Literature Review:

There has been tons of research already done in the field of image classification using neural networks especially in the area of facial recognition. This technology has been successfully adapted by multiple mobile phone manufacturers for a few years now. Samsung was one of the first mobile device companies to have successfully implemented facial recognition for their customers to unlock their devices. This particular attribute has actually now been replaced by a retina scanner which is proved to have been far more accurate than facial recognition. However the majority of the facial recognition software/technology is used by intelligence and various government agencies for almost a couple of decades now. There have been vast improvements in this field not just in terms of accuracy but also in terms of processing speed. Intelligence agencies now possess capabilities to process and compare a face across tens of thousands of faces per minute from their database to find profiles on people. However, judging by the constant improvements in the algorithms and technology in this field for years it is safe to assume that there is still a lot of room for improvement in the future. This particular research project can be considered as a basic foundation which can lead to many opportunities in the field of facial/image recognition. The most famous social media platforms such as facebook, instagram and twitter can use this technology to tag images from the backend (excluding the tags already placed by users from the front end) and combine with other algorithms to improve content recommendation for their users.

Methods:

To begin, the CIFAR-10 data is loaded and divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly-selected images from each class. The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another. Between them, the training batches contain exactly 5000 images from each class. However the training set will be split to make our validation set and thus the final sets will be of 47,000 training, 3,000 validation and 10,000 test images. Before the preprocessing stage, a function to output random examples from the training and testing sets is built to inspect the images and their labels. The images are 28x28 NumPy arrays, with pixel values ranging from 0 to 255 - 0 is equal to black and 255 is equal to white.

Two different types of models were trained in this research assignment - dense neural networks and convolutional neural networks. Each type of model was trained with 2 layers and 3 layers without regularization, hence making 4 models in total. Then all models were trained again but this time with regularization. The goal was to compare DNN and CNN architectures. In all experiments the parameters were held constant - 30 epochs, 100 batch size, same optimizer and same loss function of cross entropy - in order for the comparisons between the models to be fair. The results of all the 8 models were put in a dataframe to have a clear idea regarding the performance of each model. Lastly outputs from 2 filters from the 2 max pooling layers were extracted and visualized in a grid as images to see whether the 'lighted' up regions correspond to some features in the original images.

Results:

The highlights of the results from this research are as follow:

- The processing time of the models increased as the layers and regularization techniques got added to the models. Convolutional Neural networks took way more time to get trained compared to the dense neural network models. Adding hidden layer(s) also increase(s) the time it takes to train a model.
- The increase in accuracy is also directly proportional to the increase in model training time and increase in hidden layers in case of this research assignment. It is important to know that increasing the hidden layers in a model does not always increase the test accuracy of the model .It may result in unhealthy overfitting at times. It really depends on the complexity of the problem that is being addressed. In case of this research of building models that perform object recognition, the problem is complex enough that increasing the hidden layers will increase the validation and testing accuracy. Hidden layers allow for the function of a neural network to be broken down into specific transformation of the data. Each hidden layer function is specialized to produce a defined output. A hidden layer that is used to identify wheels cannot solely identify a car since even airplanes, busses and trucks have wheels, however when placed in conjunction with additional layers used to identify windows, a large metallic body, and headlights, the neural network can then make predictions and identify possible car(s) within visual data.
- Another notable thing from the research was that the training loss was much lesser than the validation loss as the models started to become more complex (CNN models with more max pooling layers). Usually this is a sign of overfitting. However the test accuracy was satisfactory which means that this overfitting may be good for the model.

	Process Time (Min)	Test Loss	Test Accuracy (%)	Train Loss	Train Accuracy (%)	Validation Loss	Validation Accuracy (%)
Model 1	0.87	1.56	45.31	1.42	51.149	1.54	46.767
Model 2	1.23	1.48	48.3	1.33	54.317	1.46	48.533
Model 3	31.77	0.79	74.5	0.29	90.966	0.75	74.867
Model 4	55.76	0.67	76.78	0.27	92.483	0.63	78.7
Model 1_regularization	6.15	1.49	52.11	1.36	56.696	1.47	52.2
Model 2_regularization	13.93	1.5	52.71	1.3	60.187	1.48	54.267
Model 3_regularization	48.53	1.0	75.45	0.61	89.94	0.98	75.933
Model 4_regularization	86.09	0.71	80.13	0.26	96.177	0.67	81.533

Model 1 -DNN with 2 layers (no regularization)

Model 2 - DNN with 3 layers (no regularization)

Model 3 - CNN with 2 convolution/max pooling layers (no regularization)

Model 4 - CNN with 3 convolution/max pooling layers (no regularization)

Model 5 -DNN with 2 layers (with regularization)

Model 6 - DNN with 3 layers (with regularization)

Model 7 - CNN with 2 convolution/max pooling layers (with regularization)

Model 8 - CNN with 3 convolution/max pooling layers (with regularization)

Conclusion:

In conclusion, this research assignment was able to serve its intent which was to provide hands-on practical experience with understanding the transition from simple (single hidden layer) to deep neural (multiple hidden layers) network. This research was able to explain how hidden nodes learn to extract features from their inputs and how each successive layer extracts more generalized and abstract features. The hidden layer learns the kinds of features that are inherent in its input data and then figures out feature classes. This research assignment showed us the positive impact of the hidden layers in DNN and max pooling layers in CNN when the problem at hand is just complex enough - in this case object recognition.