Implementation of Convolutional Neural Networks

Lab Assignment 4

Sumedha Janani Siriyapuraju
Dept. of Electronics and Communication Engineering
Visvesvaraya National Institute of Technology
Nagpur, India
sumedhasjs@gmail.com

***Abstract*—Linear regression is the most commonly used method of predictive analysis. It uses linear relationships between a dependent variable (target) and one or more independent variables (predictors) to predict the future of the target. The prediction is based on the assumption that the relationship between the target and the predictors is dependent or causal.**

Here, we worked on finding an analytical solution to fit a curve for Matlab_accident Dataset.

***Index Terms*—Liner Regression, Pseudo Inverse, Gradient Descent, Learning Rate, Optimum Weights**

I. Introduction

Convolutional neural networks are distinguished from other neural networks by their superior performance with image, speech, or audio signal inputs. They have three main types of layers, which are:

- Convolutional layer
- Pooling layer
- Fully-connected (FC) layer

The convolutional layer is the first layer of a convolutional network. While convolutional layers can be followed by additional convolutional layers or pooling layers, the fully-connected layer is the final layer. With each layer, the CNN increases in its complexity, identifying greater portions of the image. Earlier layers focus on simple features, such as colors and edges. As the image data progresses through the layers of the CNN, it starts to recognize larger elements or shapes of the object until it finally identifies the intended object.

GoogLeNet is a 22-layer deep convolutional neural network that's a variant of the Inception Network, a Deep Convolutional Neural Network developed by researchers at Google.

II. Method

A. GoogleNet

- The input layer of the GoogLeNet architecture takes in an image of the dimension 224 x 224.
- Type: This refers to the name of the current layer of the component within the architecture
- Patch Size: Refers to the size of the sweeping window utilised across conv and pooling layers. Sweeping windows have equal height and width.
- Stride: Defines the amount of shift the filter/sliding window takes over the input image.
- Output Size: The resulting output dimensions(height, width, number of feature maps) of the current architecture component after the input is passed through the layer.
- Depth: Refer to the number of levels/layers within an architecture component.
- 1x1 3x3 5x5: Refers to the various convolutions filters used within the inception module.
- 3X3 reduce 5x5 reduce: Refers to the numbers of 1x1 filters used before the convolutions.
- Pool Proj: This is the number of 1x1 filters used after pooling within an inception module.
- Params: Refers to the number of weights within the current architecture component.
- Ops: Refers to the number of mathematical operations carried out within the component.

The output of a softmax activation function is a vector in which its set of values represents the probability of a class or event occurrence. The values within the vector all add up to 1.

III. Discussion

- The GoogLeNet architecture consists of 22 layers (27 layers including pooling layers), and part of these layers are a total of 9 inception modules
- The first conv layer in figure 2 uses a filter(patch) size of 7x7, which is relatively large compared to other patch sizes within the network. This layer's primary purpose is to immediately reduce the input image, but not lose spatial information by utilising large filter sizes.
- The second conv layer has a depth of two and leverages the 1x1 conv block, which as the effect of dimensionality reduction. Dimensionality reduction through 1x1 conv block allows the decrease of computational load by lessening the layers' number of operations.

IV. Conclusion

A function for performing the operations in the convolution layer of a Convolutional Neural Network has been successfully implemented.

Appendix (Code)

the code is available here