All Hands On Deck Assignment #1 - James R. Herbick
"I GIVE PERMISSION TO SHARE THIS WITH OTHERS"

Deep Learning Neural Networks (DLNN)

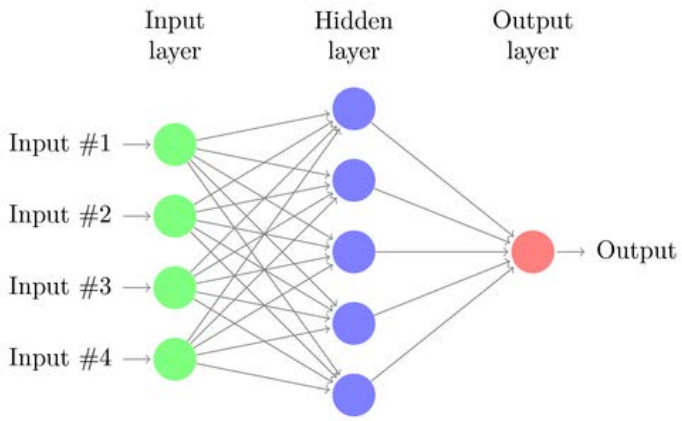
Overview

Deep learning neural networks (DLNN) are a machine-learning modeling technique that extends artificial neural networks (ANN). Therefore, it is helpful to have an understanding of artificial neural networks when learning about DLNN. Neural-network machine-learning techniques seek to create mathematical models for prediction that simulate the thought patterns of the human brain. ANN models can be used to predict numeric outcomes (e.g. the price of a product, the number of items sold) as well as categorical outcomes (e.g. does something fall into category 1, 2, or 3). DLNN models are still evolving as are their applications to particular problems. One area where DLNN are being introduced is with image recognition. DLNN models process vast amounts of data, such as images, learn and extract particular characteristics, and then look for these characteristics in new data. I will explain the applications of DLNN in more detail shortly, but first let's briefly discuss the important components of neural network models.

Components of Neural Network Models

Figure 1 shows the main components, or layers, of a neural network model; these basic concepts apply to ANN and DLNN. Let's briefly examine each layer of neural network models.

Figure 1: Basic Neural Network Diagram



Input Layer

This layer represents the attributes of the data upon which a model will be trained and built. For example, if we are predicting house prices, the various inputs would then be characteristics of each

house (i.e. square footage, number of rooms). The inputs to a DLNN would be more complex, a series of full images for example. The DLNN would then attempt to generalize or abstract characteristics of these images.

Hidden Layer

Think of this layer as a series of on/off switches (nodes) that are triggered in specific combinations based upon the data coming into the model. The combinations of on and off values for these switches then produce a prediction, or output. It is here in the hidden layer that the main distinction is made between ANN and DLNN models. The hidden layer can contain any number of nodes, and the neural network model itself can have more than one hidden layer. As the number of nodes within a hidden layer increases and as the number of hidden layers within a model increases, the model becomes more complex. It is really the level of model complexity that differentiates DLNN from ANN models. DLNN models typically include multiple hidden layers, each made up of many on/off switches.

Output Layer

The output layer represents the model's final predictions. If the model is predicting a numerical value such as the price of a product, then there will be only one node in the output layer. However, if the model is predicting whether or not something falls into one of many categories, then there will be many nodes in the output layer. Each node will represent one of the categories.

Applications of DLNN – Image Recognition

A major area of focus for DLNN approaches is that of image recognition. Whereas an ANN model may be applied to a character recognition problem (i.e. is this hand-written letter the letter 'a' or 'b'), DLNN models go a step further by abstracting characteristics and learning patterns from entire images and then recognizing those characteristics in new images. For example, the medical imaging community is exploring DLNN approaches for disease diagnosis. Diagnosing a disease requires capturing images and evaluating those images for particular characteristics. Limitations currently exist in the image evaluation process. For example, a doctor may introduce subjectivity into an evaluation or differing conclusions may be reached by different doctors. Through the application of DLNN to medical image evaluation, the medical community hopes to reduce the subjectivity and variance of outcomes in the process. The medical community is examining DLNN in areas such as colon cancer and lung disease detection as well as mammogram risk scoring.

Other areas where DLNN are being applied to image recognition include facial recognition and self-driving automobiles. DLNN models can be used in facial recognition situations within social media applications and to help identify terrorists. The auto industry is exploring this technology to help the self-driving cars identify objects on the road such as other cars, open road, and stop signs.

Additional Resources

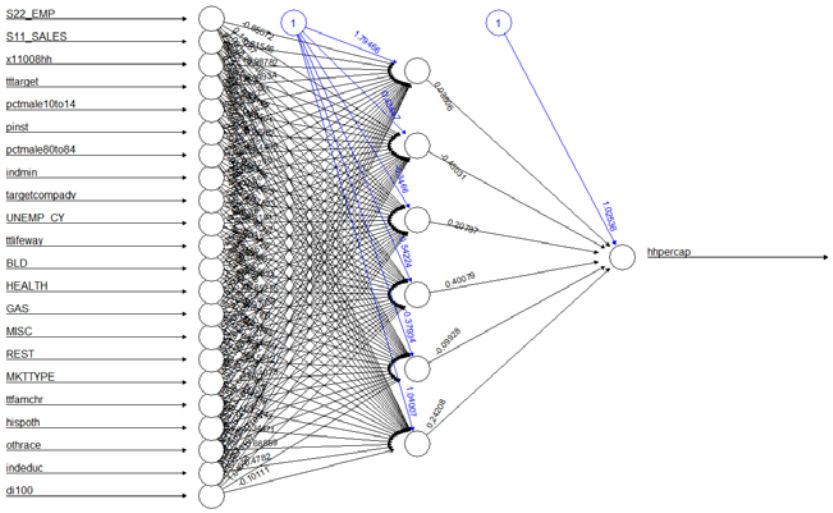
http://venturebeat.com/2016/04/13/google-launches-tensorflow-0-8-with-support-for-distributed-model-training/

https://cloud.google.com/products/machine-learning/

DLNN and the JAM Analytics Project

The analysis that JAM Analytics is performing for Mardel stores includes building a model to predict store sales per household. As a part of this process, JAM Analytics has been very thorough in the modeling techniques that we've investigated. Our philosophy is that of analytical street brawlers, we let the predictive accuracy of the models determine which to select. One of the modeling techniques that we explored was ANN. As mentioned previously, the ANN technique can be viewed as a less complex version of DLNN. In fact, our ANN model is among the top 3 best-performing models from over 15 models built. Figure 2 shows a sample representation of our ANN model for Mardel stores.

Figure 2: JAM Analytics ANN Model Example



Because the primary objective in predicting store sales per household is accuracy, the use of an ANN model is appropriate. Since the ANN model's accuracy is in the top 3 of all of our models, JAM Analytics will consider this approach as a part of its analysis.

However, the use of DLNN approaches for the Mardel analysis is not appropriate at this time. Although DLNN approaches have promise and are being applied to interesting areas, the techniques are still in their infancy. In addition, the application of DLNN approaches are more focused on image and audio analysis, as opposed to pure numeric prediction. Therefore, JAM Analytics will not explore DLNN techniques as a part of its current analysis.

As Mardel stores begins to better understand their primary customers with the help of the JAM Analytics project, there may be future opportunities to introduce DLNN approaches. One example might be analyzing site images for a new Mardel store to determine if it would be an ideal location based upon identified store characteristics (e.g. distance to competitors, other retail stores).