Bayesian Neural Networks

Blair, Taylor Sorgmon, Ava Conor

April 2, 2024

Abstract

Bayesian Neural Networks are...

Contents

1 Introduction

2 Machine Learning

Machine learning includes a variety of topics of which artifical intelligence (AI) and neural networks are a subset. Neural networks in particular involve an abstract imitation of the human brain using simulated neurons which is trained on data using particular algorithms.

The term supervised learning refers to a specific common method of machine learning which uses a labeled dataset to train a model to correctly predict the labels via some training algorithm. A classification problem in supervised learning involves predictions where all labels are grouped into a set of categories. The training process can be broken down into three major steps, which include a decision process, an error function, and an optimization process [**berkelyWhatIsML**]. The decision process is the set of steps the algorithm takes after receiving the data based on the goal of the model. The second step in the process is the error function, which is the method of measuring chosen to see if the algorithm gave a "good" or "bad" input. The two most common choices of the error function is a simple yes or no on if a data point was classified correctly in the case of classification,

or the difference in value between the predicted outcome and the actual observed outcome for continuous values. Finally, the third and last step is the part of the process that implements learning. This step, the updating or optimization process, requires the algorithm to review the past data and outputs of the error function in order to better correct its decision making process in the future.

Throughout this report, we'll often use the terms machine learning, deep learning, and neural networks. It is important to note that although these are fundamentally related fields, deep learning is a subfield of neural networks that in particular focuses on more complicated neural networks, while neural networks are a subfield in machine learning.

We will be training a classifier on labeled images in a supervised learning process to predict what is depicted in the image from a range of possibilities such as dog, cat, and plane.
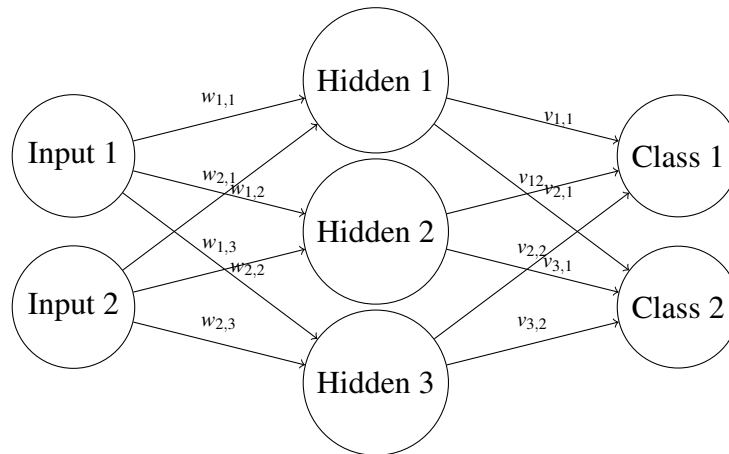
2.1   Neural Network

Figure 1: Example neural network

A neural network takes a series of inputs... In figure 1, the network takes two inputs, has one hidden layer of size
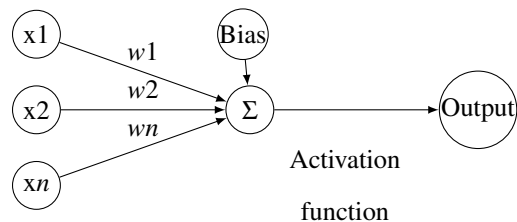

Figure 2: Neural network neuron

Neural networks are made out of a series of neurons... The neurons take a set of inputs, multiplies the inputs by the weights, sums the weighted input, adds a bias, and runs the output through an activation function...
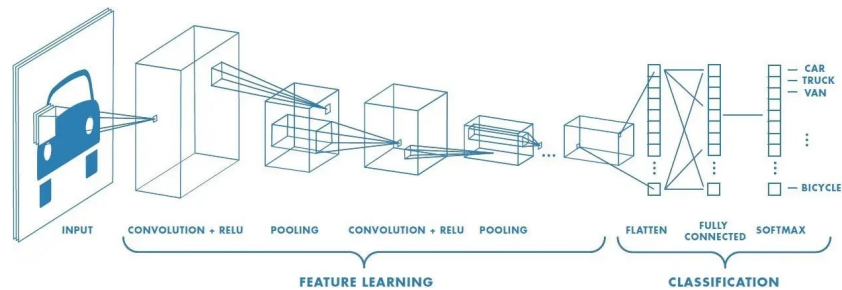
2.2   Convolutional Neural Networks

Figure 3: CNN pipeline [**eli5CNN**]

Convolutional neural networks (CNN) are a type of neural network that is better suited for image recognition. While this might sound like a seperate model structure, CNNs are largely the same. In figure 3 the difference between a traditional neural network is the convolutional layer. Instead of reading the entire image at once, a convolutional layer slides over the image...

IMAGE OF SLIDING (gif split)

...

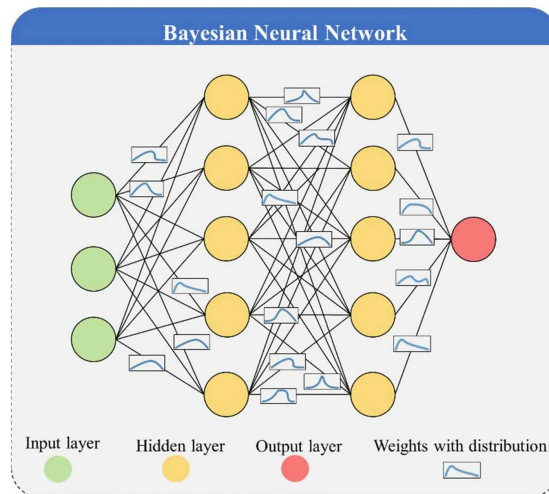# 3   Bayesian Neural Networks

Figure 4: Example BNN [**FleszarBNN**]

Bayesian neural networks take the same principle as
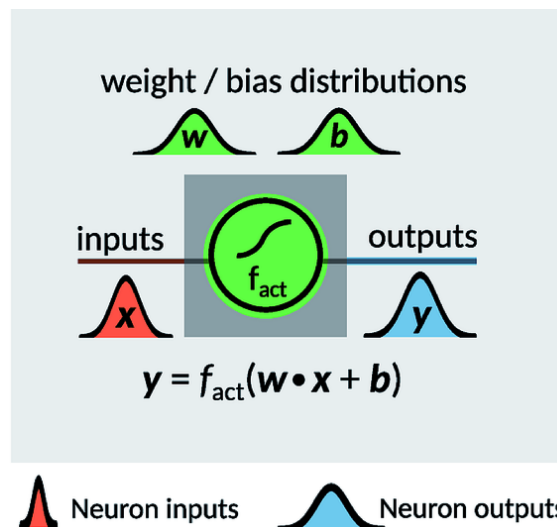
Similar to a Neural network such as...

Figure 5: Example BNN Neuron [**hase2019machine**]

3.1   Bayesian Convolutional Neural Networks

Bayesian convolutional neural network (BCNN) are similar to CNNs. The difference between is that BCNNs and a CNN is that the BCNN uses a bayesian neuron.

4   Simulation

We use a BCNN implementation from Github based on work from ... [**shridhar2019comprehensive**] [**shridhar2018uncertainty**]
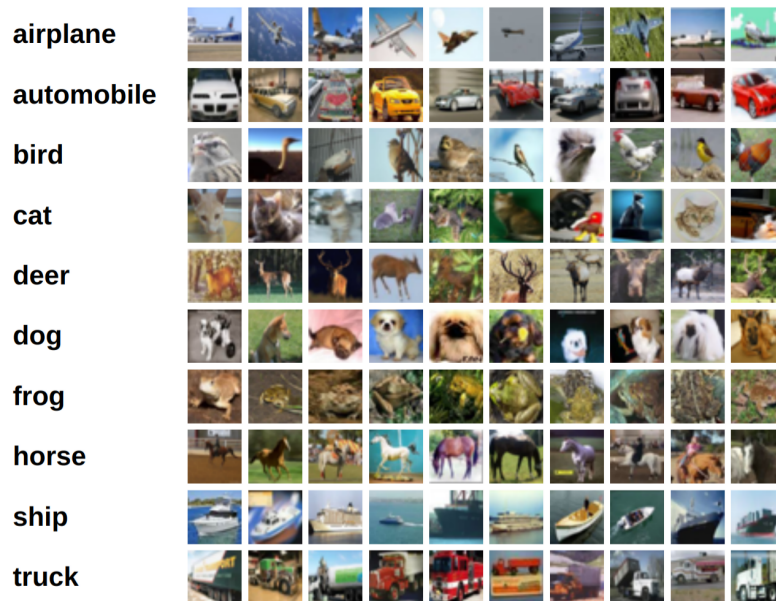
4.1   CIFAR-10

Figure 6: Example CIFAR-10 images [**cifar10**]

The CIFAR-10 ((Canadian Institute For Advanced Research 10) dataset is a machine learning benchmarking set. It contatins 60,000 16×16 RGB pictures of: airplanes, cars, birds, cats, deer, dogs, frogs, horses, ships, and trucks[**cifar10**]. It is used...

4.2   Hyperparamaters

We used the following hyperparamaters for training:

Hyperparameter	CNN	BCNN
Epochs	500	500
Learning Rate	....	May be higher (0.01 - 0.1) due to simpler structure
Regularization	L1/L2 weight decay or Dropout common to prevent overfitting	Can benefit from Dropout, but weight decay might be less crucial
Optimizer	Adamw	Adamw

In addition to the above, the two models have the same number of levels...

4.3 Results
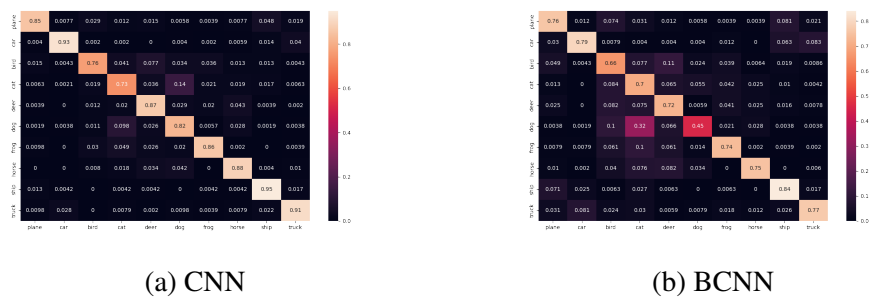
(a) CNN   (b) BCNN

Figure 7: Confusion matrices

The two have roughly similar accuracies given the model

5 Closing