Bayesian Neural Networks

Blair, Taylor Sorgmon, Ava Conor

April 2, 2024

Abstract

Bayesian Neural Networks are...

Contents

1 Introduction

1.1 History

1987 The patent a 'Method of providing digital signatures' is filed by Ralph C. Merkle[**merkle-pater**

1999 The original patent expires.

2009 Bitcoin uses Merkle Trees for 'block header commitment.'[**friedenbach_alm_2017**]

2009 BitTorent uses Merkle Trees for data integrity[**bep30**].

2 Machine Learning

Machine learning is considered to be a branch of artificial intelligence (AI) with the goal to help AI "learn" more, but specifically in a way that imitates the human process of learning through the use of data and algorithms[8]. The term supervised learning refers to what one most commonly thought of methods of employing this practice, which consists of using a labeled dataset to train, or get an algorithm to

learn, how to predict outcomes and/or classify data correctly. This process can be broken down into three major steps, which include a decision process, an error function, and an updating or optimization process [7]. In the first step, the decision process includes exactly what it sounds like; the decision process of the algorithm. This 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, or the difference in value between the predicted outcome and the actual observed outcome. Finally, the third and last step is what is considered to be the actual learning part of the process. 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, and in turn, neural networks are a subfield in machine learning. Thus when we are looking at Bayesian neural networks, we are looking at a subfield of more general neural networks, and that Bayesian neural networks are a subfield of a subfield of machine learning.

2.1 Neural Network

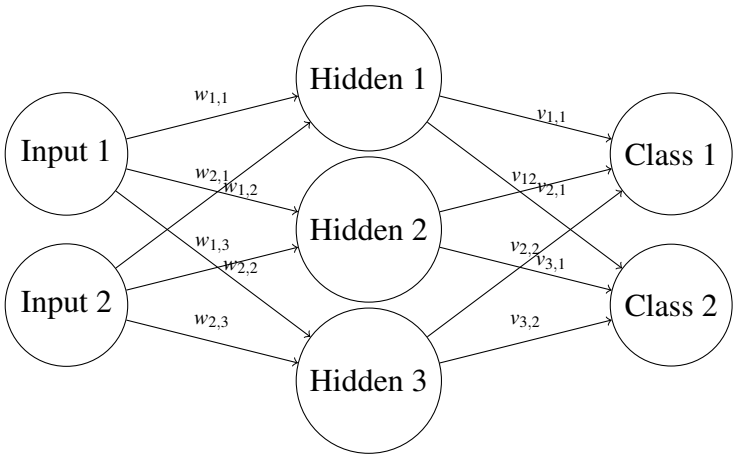


Figure 1: Example neural network

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

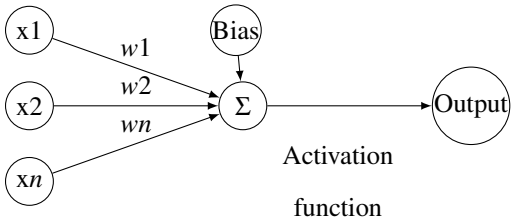


Figure 2: Neural network neuron

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

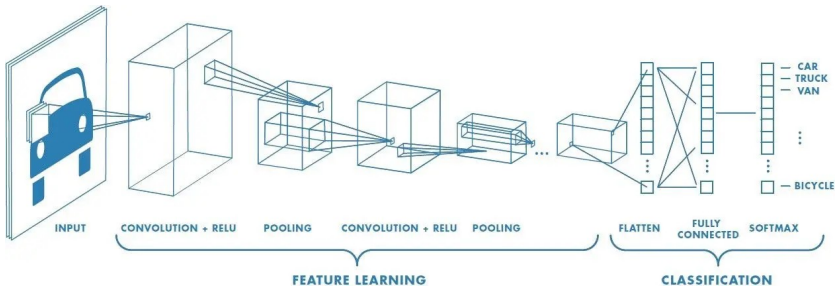


Figure 3: CNN pipeline [4]

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

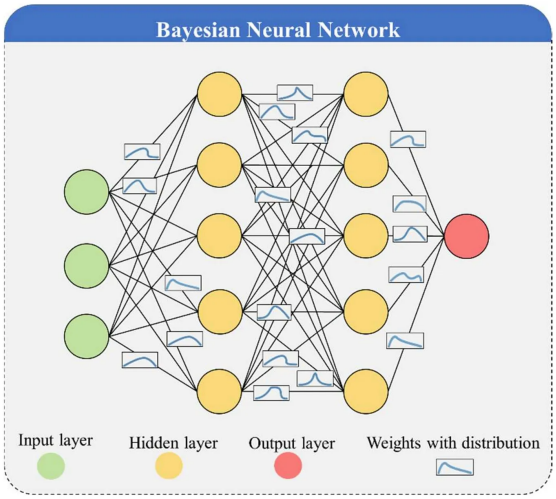


Figure 4: Example BNN [1]

Bayesian neural networks take the same principle as

Similar to a Neural network such as...

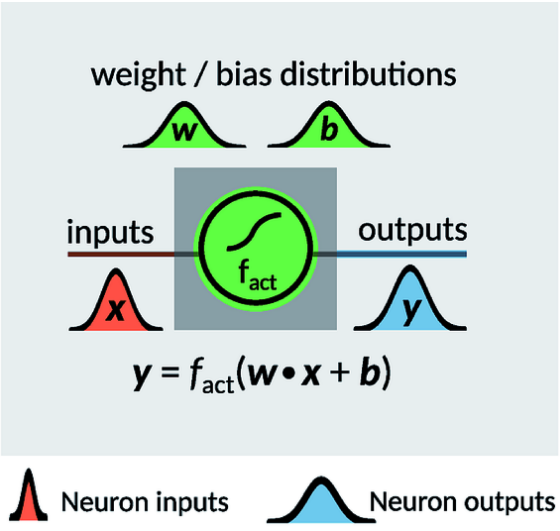


Figure 5: Example BNN Neuron [2]

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 ... [6] [5]

4.1 CIFAR-10

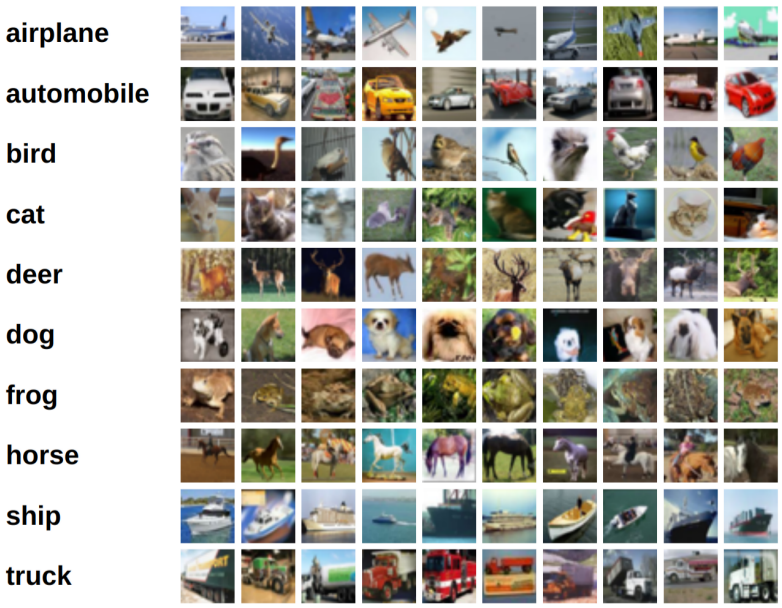


Figure 6: Example CIFAR-10 images [3]

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[3]. It is used...

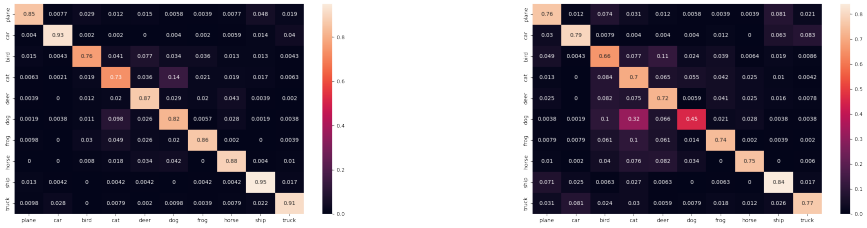
4.2 Hyperparamaters

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

References

Fleszar, J. (2023). Bayesian neural networks - capturing the uncertainty of the real world !!

Häse, F., Galván, I. F., Aspuru-Guzik, A., Lindh, R., & Vacher, M. (2019). How machine learning can assist the interpretation of ab initio molecular dynamics simulations and conceptual understanding of chemistry. *Chemical science*, *10*(8), 2298–2307.

Krizhevsky, A., Nair, V., & Hinton, G. (n.d.). Cifar-10 (canadian institute for advanced research). http://www.cs.toronto.edu/~kriz/cifar.html

Saha, S. (2018). A guide to convolutional neural networks — the eli5 way.

Shridhar, K., Laumann, F., & Liwicki, M. (2018). Uncertainty estimations by softplus normalization in bayesian convolutional neural networks with variational inference. *arXiv preprint arXiv:1806.05978*.

Shridhar, K., Laumann, F., & Liwicki, M. (2019). A comprehensive guide to bayesian convolutional neural network with variational inference. *arXiv preprint arXiv:1901.02731*.

What is machine learning (ml)? (2020).

What is machine learning (ml)? (n.d.).