# Bayesian Neural Networks

Blair, Taylor Sorgmon, Ava Conor

April 15, 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 [7]. 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

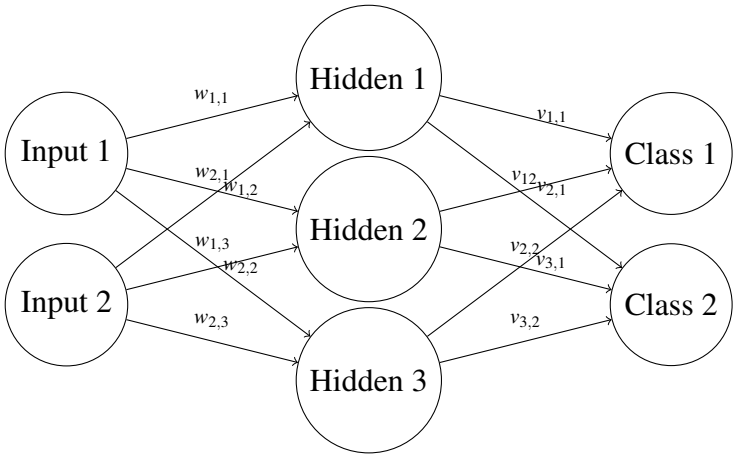



Figure 1: Example neural network

A neural network takes a series of inputs. An example of a neural network with two inputs, one hidden layer of size three, and two outputs appears in figure 1

## 2.2 Neural Network Neurons

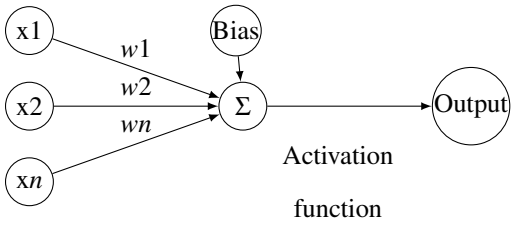



Figure 2: Neural network neuron

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...

## 2.3 Convolutional Neural Networks

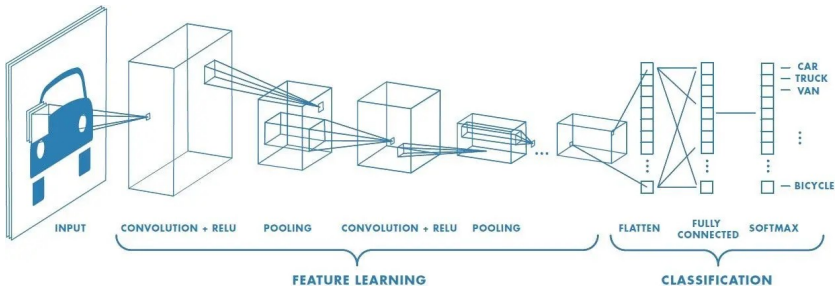



Figure 3: CNN pipeline [4]

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

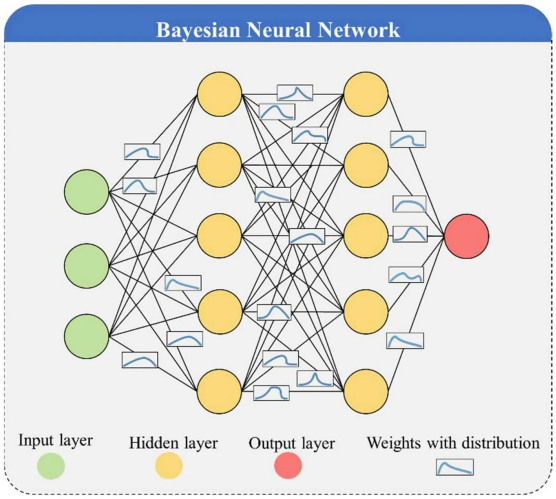



Figure 4: Example BNN [1]

Bayesian neural networks operate the same way as a normal

## 3.1 Bayesian Neural Network Neuron

Similar to a neural network such as...

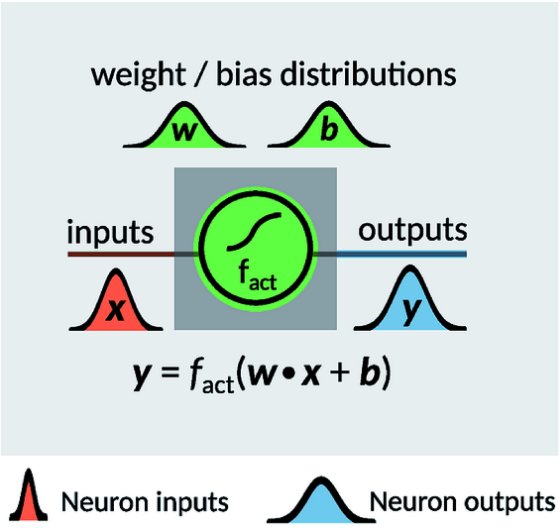



Figure 5: Example BNN Neuron [2]

### 3.1.1 Priors

This section was requested by Micheal.

## 3.2 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

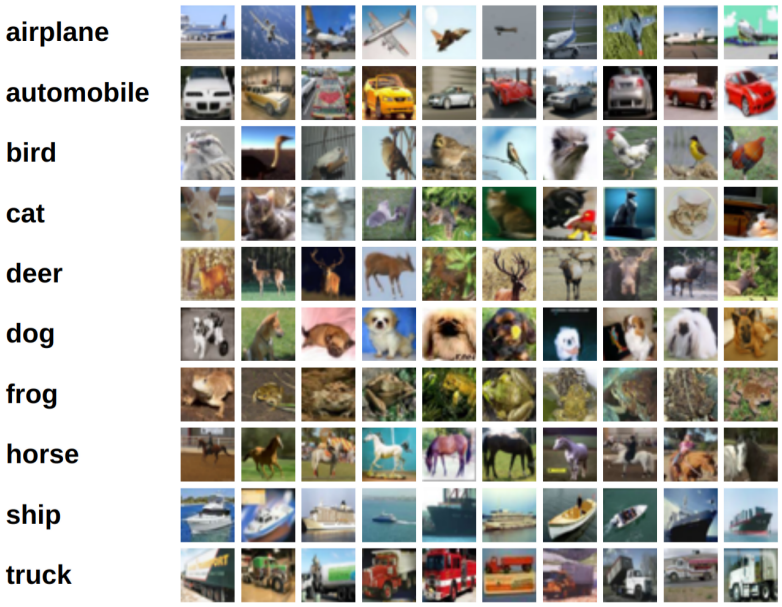



Figure 6: Example CIFAR-10 images [3]

The CIFAR-10 ((Canadian Institute For Advanced Research 10) dataset is a popular machine learning dataset set. It contatins 60,000 $16 \times 16$ color pictures of: airplanes, cars, birds, cats, deer, dogs, frogs, horses, ships, and trucks[3]. Because of its simple task and variety of images, it is often used as a baseline task for new model structures.

## 4.2 Hyperparamaters

We used the following hyperparamaters for training:

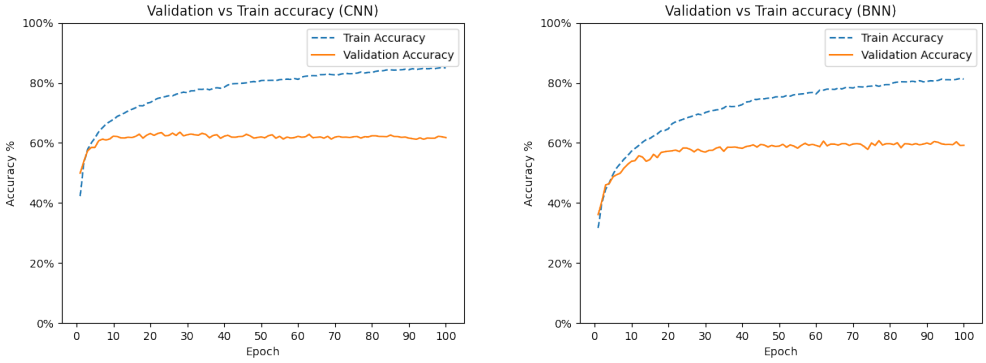
| Hyperparameter | CNN | BCNN |
|---|---|---|
| Epochs | 100 | 100 |
| Learning Rate | 0.001 | 0.003 |
| Regularization Rate | 0.001 | 0.001 |
| Optimizer | Adamw | Adamw |

In addition to the hyperparamaters above, the two models have the same number of layers and levels. We did not adjust the priors, as those were set by the BCNN layer code from [5].

## 4.3 Results

| Metric | CNN | BCNN |
|---|---|---|
| Train Accuracy | 84.96% | 81.27% |
| Validation Accuracy | 61.76% | 59.21% |
| Time to Train | 16 min 11 sec | 22 min 11 sec |

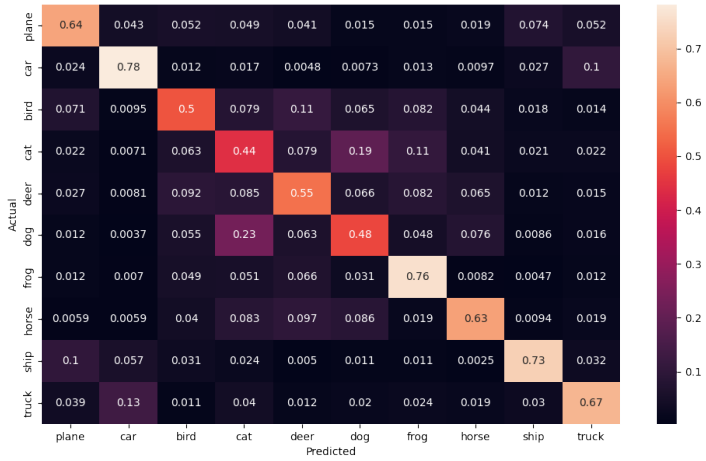
The results we achieved with training were about the same between the BCNN and the CNN. However, the bayesian model took longer to train.



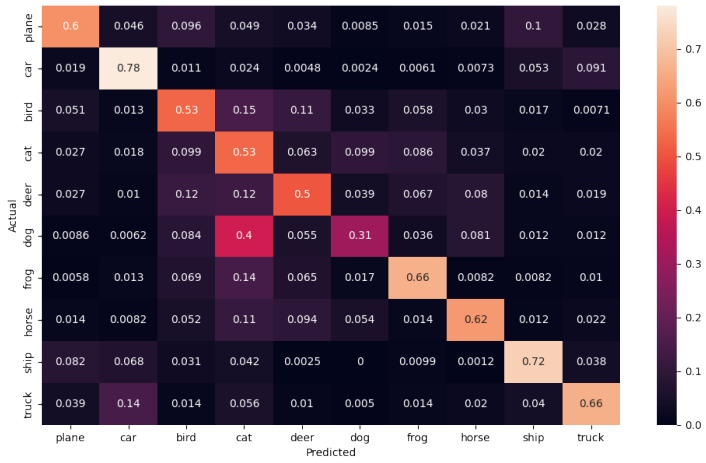

(a) CNN (b) BCNN

Figure 7: Train vs validation accuracy over time

Additionally, neither model converged with the train and validation accuracy, this is an indication that we might need to increase the regularization rate. Another thing to note from figure 7 is that the BCNN took longer to divirge from the train accuracy.




(a) CNN




(b) BCNN

Figure 8: Confusion matrices

The two models have the same set of confusions

# 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.).