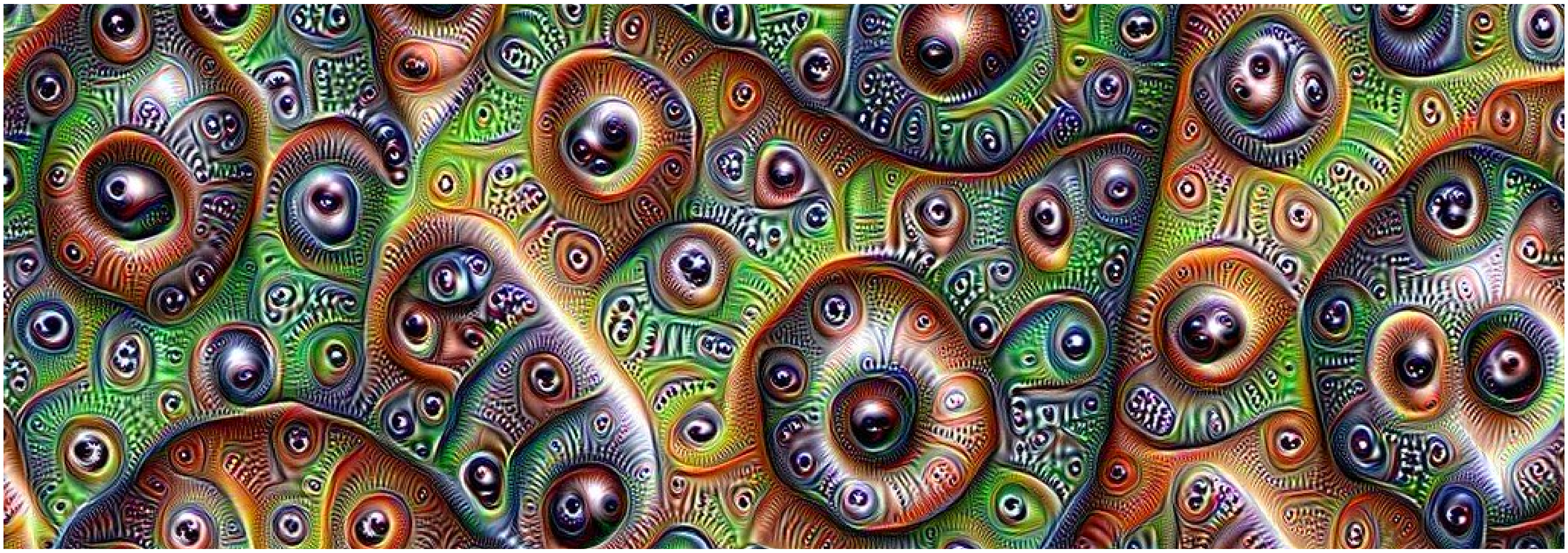
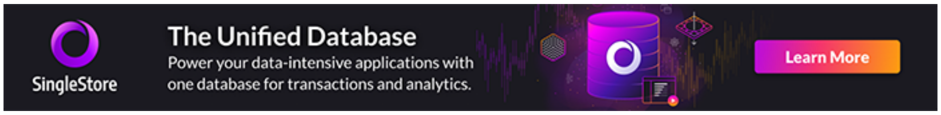
Hands-on Guide to Bayesian Neural Network in Classification

08/08/2020





Deep neural networks learn to form relationships with the given data without having prior exposure to the dataset. They perform very well on non-linear data and hence require large amounts of data for training. Although more information is better for the network, it leads to problems like overfitting (https://analyticsindiamag.com/tackling-underfitting-and-overfitting-problems-in-data-science/).

The neural network will work really well with training data but underperforms when it is fed unseen data. This makes the network blind to the uncertainties in the training data and tends to be overly confident in its wrong predictions. Consider an example where you are trying to classify a car and a bike. If an image of a truck is shown to the network, it ideally should not predict anything. But, because of the softmax function, it assigns a high probability to one of the classes and the network wrongly, though confidently predicts it to be a car. In order to avoid this, we use the Bayesian Neural Network (BNN).

What will we discuss in this article?

- Why do we need to use BNN?
- How a Bayesian network works
- Implementing a BNN

Scroll To Top

Why do we need to use the Bayesian Neural Network (BNN)?

As discussed above, we need to make sure our model does not overfit. You may be wondering that there are other methods to do this like Batch Normalisation, Dropout etc. But all these methods do not solve the problem of identifying uncertainty in the model. Hence the need for BNNs. Bayesian Networks work well on small datasets and are robust for avoiding overfitting. They also come with additional features like uncertainty estimation, probability distributions etc.

How Does a Bayesian Neural Network work?

The motto behind a BNN is pretty simple — every entity is associated with a probability distribution, including weights and biases. There are values called 'random variables' in the bayesian world which provides a different value every time it is accessed. Let X be a random variable which represents the normal distribution, every time you access X, it'll have a different value. This process of obtaining a different value on each access is called Sampling. What value comes out of each sample depends on the probability distribution. The wider the probability distribution, the more the uncertainty.

In a traditional neural network, each layer has fixed weights and biases that determine the output. But, a Bayesian neural network will have a probability distribution attached to each layer as shown below.

For a classification problem, you perform multiple forward passes each time with new samples of weights and biases. There is one output provided for each forward pass. The uncertainty will be high if the input image is something the network has never seen for all output classes.

Implementation

We will now implement the BNN on a small dataset from sklearn called iris dataset. This dataset has 4 attributes and around 150 data points.

Loading the dataset and importing essential packages

```
import numpy as np
from sklearn import datasets
import torch
import torch.nn as nn
import torch.optim as optim
import torchbnn as bnn
import matplotlib.pyplot as plt

dataset = datasets.load_iris()
```

Splitting the dataset into data and target and converting them to tensors

```
data = dataset.data
target = dataset.target
data_tensor=torch.from_numpy(data).float()
target_tensor=torch.from_numpy(target).long()
```

Defining a simple Bayesian model

Scroll To Top

```
model = nn.Sequential(
    bnn.BayesLinear(prior_mu=0, prior_sigma=0.1, in_features=4, out_features=100),
    nn.ReLU(),
    bnn.BayesLinear(prior_mu=0, prior_sigma=0.1, in_features=100, out_features=3),
)

prior_mu (Float) is the mean of prior normal distribution.
prior_sigma (Float) is the sigma of prior normal distribution.
```

Defining loss function

The two-loss functions used here are cross-entropy loss and the BKL loss which is used to compute the KL (Kullback–Leibler) divergence of the network.

SEE ALSO

(https://analyticsindiamag.com/comprehensive-guide-to-image-datasets/)

DEVELOPERS CORNER (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/DEVELOPERS_CORNER/)

Comprehensive Guide To 9 Most Important Image Datasets For Data Scientists (https://analyticsindiamag.com/comprehensive-guide-to-image-datasets/)

```
cross_entropy_loss = nn.CrossEntropyLoss()
klloss = bnn.BKLLoss(reduction='mean', last_layer_only=False)
klweight = 0.01
optimizer = optim.Adam(model.parameters(), lr=0.01)
```

Training the model

The model is trained for 3000 steps(this would have lead to overfitting for a traditional network)

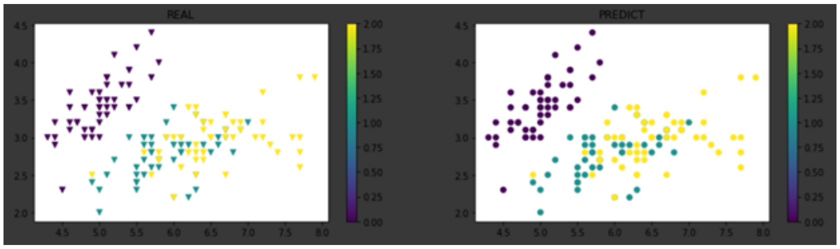
```
for step in range(3000):
    models = model(data_tensor)
    cross_entropy = cross_entropy_loss(models, target)
    kl = klloss(model)
    total_cost = cross_entropy + klweight*kl

    optimizer.zero_grad()
    total_cost.backward()
    optimizer.step()

_, predicted = torch.max(models.data, 1)
final = target_tensor.size(0)
correct = (predicted == target_tensor).sum()
print('- Accuracy: %f %%' % (100 * float(correct) / final))
print('- CE : %2.2f, KL : %2.2f' % (cross_entropy.item(), kl.item()))

Output : - Accuracy: 98.000000 %
- CE : 0.05, KL : 2.81
```

Scroll To Top

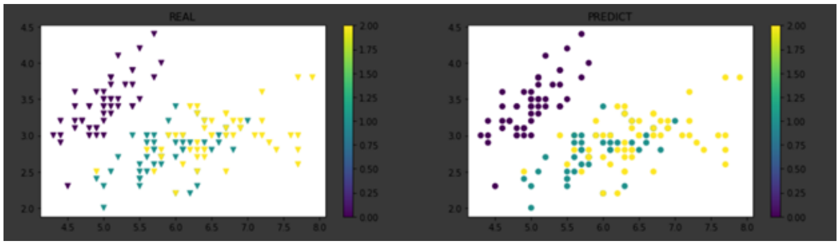
Visualisation

Let us now visualise the model and see how it has performed. To understand how sampling works, run the model multiple times and plot the graphs. You will notice minor changes with each iteration.

```
def draw_graph(predicted) :
    fig = plt.figure(figsize = (16, 8))

    fig_1 = fig.add_subplot(1, 2, 1)
    fig_2 = fig.add_subplot(1, 2, 2)

    z1_plot = fig_1.scatter(data[:, 0], data[:, 1], c = target,marker=’v’)
    z2_plot = fig_2.scatter(data[:, 0], data[:, 1], c = predicted)

    plt.colorbar(z1_plot,ax=fig_1)
    plt.colorbar(z2_plot,ax=fig_2)

    fig_1.set_title("REAL")
    fig_2.set_title("PREDICT")

    plt.show()

Run 1 :
models = model(data)
_, predicted = torch.max(models.data, 1)
draw_graph(predicted)
```



```
Run 2:
models = model(data)
_, predicted = torch.max(models.data, 1)
draw_graph(predicted)
```



The above graph indicates the difference between the scattering of data points in the actual dataset versus the scattering of points in the predicted bayesian network. Since there are 3 output classes, each of them are indicated in a different color. Each time the data is sampled, the network assigns a

Scroll To Top

probability distribution to the entire input data. On close observation near coordinates (2.3,6.2) of the predicted graph, there is a different prediction made because of the change in probability.

Conclusion

The above implementation makes it clear that not only the problem of overfitting is much more robustly (http://machinehack.com) solved with the help of Bayesian Networks but also determines the uncertainty involved in the input.

What Do You Think?

0 Comments

Sort by Oldest

Add a comment...

Facebook Comments Plugin

Join Our Discord Server. Be part of an engaging online community. Join Here (https://discord.gg/Sbtj3jdeAz).

Subscribe to our Newsletter

Get the latest updates and relevant offers by sharing your email.

ENTER YOUR EMAIL

SUBSCRIBE NOW

BHOOMIKA MADHUKAR (HTTPS://ANALYTICSINDIAMAG.COM/AUTHOR/BHOOMIKA-MADHUKARANALYTICSINDIAMAG-COM/)

I am an aspiring data scientist with a passion for teaching. I am a computer science graduate from Dayananda Sagar Institute. I have experience in building models in deep learning and reinforcement learning. My goal is to use AI in the field of education to make learning meaningful for everyone.

SHARE (https://www.facebook.com/sharer.php?u=https://analyticsindiamag.com/hands-on-guide-to-bayesian-neural-network-in-

TWEET (https://twitter.com/intent/tweet?text=Hands-on%20Guide%20to%20Bayesian%20Neural%20Network%20in%20Classification&via=Analyticsindiam&url=https://analyticsindiam
on-guide-to-bayesian-neural-network-in-classification/)

in (https://www.linkedin.com/cws/share?url=https://analyticsindiamag.com/hands-on-guide-to-bayesian-neural-network-in-classificatior

(https://wa.me/?text=Hands-on%20Guide%20to%20Bayesian%20Neural%20Network%20in%20Classification%20https://analyticsindiamag.com/hands-on-guide-to-
neural-network-in-classification/)

(mailto:?subject=Hands-on%20Guide%20to%20Bayesian%20Neural%20Network%20in%20Classification&body=Hands-on%20Guide%20to%20Bayesian%20Neural%20Network%20in%20Classification%20https://analyticsindiamag.com/hands-on-guide-to-
neural-network-in-classification/)

(https://share.flipboard.com/bookmarklet/popout?v=2&title=Hands-on%20Guide%20to%20Bayesian%20Neural%20Network%20in%20Classification&url=https://analyticsindiamag.com/hands-on-guide-tc
neural-network-in-classification/)

Scroll To Top

OUR UPCOMING EVENTS

Masterclass

Simplifying Data Architectures With A Unified Database

24th Aug 2021

Register>> (https://register.gotowebinar.com/register/3614906576975157260)

Hands-On Workshop

Mastering Exploratory Data Analysis

28th Aug 2021

Register>> (https://www.adasci.org/continuous-learning/mastering-exploratory-data-analysis-16910993)

Virtual Conference

oneAPI DevSummit, Asia-Pacific & Japan

15th Sep 2021

Register>> (https://software.seek.intel.com/devsummitAPJ2021?cid=oth&source=aim&campid=apj_q3_aim&content=live-event_cross-seg&registration_source=aim)

Virtual Conference

Deep Learning DevCon 2021

23-24th Sep 2021

Register>> (https://dldc.adasci.org/)

Scroll To Top

RELATED POSTS

DEVELOPERS CORNER (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/DEVELOPERS_CORNER/)

Guide to DORO: Distributional and Outlier Robust Optimization (https://analyticsindiamag.com/guide-to-doro-distributional-and-outlier-robust-optimization/)

(https://analytics to-doro-distributional-and-outlier-robust-optimization/)

19/08/2021 · 11 MINS READ

DEVELOPERS CORNER (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/DEVELOPERS_CORNER/)

Hands-On Guide to Multi-Class Classification Using Mobilenet_v2 (https://analyticsindiamag.com/hands-on-guide-to-multi-class-classification-using-mobilenet_v2/)

(https://analytics on-guide-to-multi-class-classification-using-mobilenet_v2/)

18/08/2021 · 6 MINS READ

DEVELOPERS CORNER (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/DEVELOPERS_CORNER/)

Hands-On Guide To ROCK Clustering Algorithm (https://analyticsindiamag.com/hands-on-guide-to-rock-clustering-algorithm/)

(https://analytics on-guide-to-rock-clustering-algorithm/)

17/08/2021 · 5 MINS READ

EDUCATION (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/EDUCATION/)

Scroll To Top

[Top 15 YouTube Channels To Follow For Deep Learning Enthusiasts (https://analyticsindiamag.com/top-15-youtube-channels-to-follow-for-deep-learning-enthusiasts/)

(https://analytics 15-youtube-channels-to-follow-for-deep-learning-enthusiasts/)

16/08/2021 · 4 MINS READ

DEVELOPERS CORNER (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/DEVELOPERS_CORNER/)

Beginner's Guide To OpenNRE Toolkit (https://analyticsindiamag.com/beginners-guide-to-opennre-toolkit/)

(https://analytics guide-to-opennre-toolkit/)

16/08/2021 · 6 MINS READ

DEVELOPERS CORNER (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/DEVELOPERS_CORNER/)

A Beginner's Guide To TensorFlow (https://analyticsindiamag.com/a-beginners-guide-to-tensorflow/)

(https://analytics beginners-guide-to-tensorflow/)

15/08/2021 · 5 MINS READ

CONNECT

About Us (https://analyticsindiamag.com/about/)

Advertise (https://analyticsindiamag.com/advertise-with-us/)

Weekly Newsletter (https://analyticsindiamaga

Write for us (https://analyticsindiamag.c for-us/)

Careers (https://www.linkedin.com/jobs/search/?

OUR BRANDS

MachineHack – ML Hackathons (https://www.machinehack.com/)

AIM Research (https://aimresearch.ai/)

AIM Recruits (https://recruits.analyticsindiamag.com/)

Workshops (https://www.machinehack.com/bootcamp)

Practice (https://www.machinehack.com/practices)

OUR CONFERENCES

Cypher (https://www.analyticsindiasummit.com/)

The MachineCon (https://themachinecon.com/)

Machine Learning Developers Summit (http://mlds.analyticsindiasummit.com/)

The Rising (https://rising.analyticsindiasummit.com/)

plugin (https://plugin.analyticsindiasummit.com/)

OUR VIDEOS

Documentary – The Transition Cost (https://www.youtube.com/watch?v=7pvGJbzTTWk&list=PL9Kc1zSa46OzMfx0I1SJZOGpN_p371ybd)

Web Series – The Dating Scientists (https://www.youtube.com/watch?v=WQbKblRKQsk&list=PL9Kc1zSa46QxzJqQEJa-qI55CtLZxFl2v)

Podcasts – Simulated Reality (https://www.youtube.com/watch?v=gvZfaeVVbGE&list=PL9Kc1zSa46OwqKv91i8W6vZ-

BRAND PAGES

Intel AI Hub (https://analyticsindiamag.com/on-intel/)

ASSOCIATION OF DATA SCIENTISTS

Chartered Data Scientist(TM) (https://www.adasci.org/cds program)

Lattice – Machine

LISTS

Academic Rankings (https://analyticsindiama rankings/)

Best Firms To Work For (https://analyticsindiama best-firms-in-india-for-data-scientists-to-work-for-2021/)

Top Leaders (https://analyticsindiama s=most+influential&categ

Data Scientists

Scroll To Top

currentJobId=2595617641&f_C=10283931&f_TPR=r604800&geoId=92000000)

Reuse our content (https://analyticsindiamag.com/reuse-our-content/)

Contact Us (https://analyticsindiamag.com/contact-us/)

MENTORSHIP

AIM Mentorship Circle (https://analyticsindiamag.com/mentorship-circle/)

Assisted Mentoring (https://analyticsindiamag.com/mentorship-circle/assisted-mentoring/)

EVENTS

AIM Custom Events (https://analyticsindiamag.com/our-events/aim-custom-events/)

AIM Virtual (https://analyticsindiamag.com/our-events/aim-virtual/)

(https://analyticsindiamag.com/)

AWARDS

Analytics100 (https://themachinecon.com/awards/)

40 under 40 Data Scientists (https://mlds.analyticsindiasummit.com/awards/)

Data Science Excellence (https://www.analyticsindiasummit.com/about/awards/)

Women in AI Leadership (https://rising.analyticsindiasummit.com/awards/)

V5DucIU4Y)

Analytics India Guru (https://www.youtube.com/watch?v=oQDZMdeyzgw&list=PL9Kc1zSa46Oxc1zxdtKwCO2SU

The Pretentious Geek (https://www.youtube.com/watch?v=Q7d1UR_PRGg&list=PL9Kc1zSa46OxcBA2Y2QFCEwjOgk

Deeper Insights with Leaders (https://www.youtube.com/watch?v=Z3Z7Riw1G9o&list=PL9Kc1zSa46Ow8fAFzC23cuLYNRUZyXAG)

Curiosum - AI Storytelling (https://www.youtube.com/watch?v=AfsqH5EzjIg&list=PL9Kc1zSa46OzOCjo2YUNym6thlDD0F0zc)

Learning Journal (https://www.adasci.org/jou 35309407)

Continuous Learning (https://www.adasci.org/co learning)

Job Board (https://www.adasci.org/job

Membership (https://www.adasci.org/me

Discussion Board (https://www.adasci.org/forum)

Community (https://www.adasci.org/community)

(https://analyticsindiama s=top+10+data+scientists

Emerging Startups (https://analyticsindiama s=10+Emerging+Analytics

Trends (https://analyticsindiama s=top+trends&category=4

PeMa Quadrant (https://analyticsindiama

Analytics Service Providers (https://analyticsindiama

ABOUT US(HTTPS://ANALYTICSINDIAMAG.COM/ABOUT/)

ADVERTISE(HTTPS://ANALYTICSINDIAMAG.COM/ADVERTISE-WITH-US/)

WRITE FOR US(HTTPS://ANALYTICSINDIAMAG.COM/WRITE-FOR-US/)

COPYRIGHT(HTTPS://ANALYTICSINDIAMAG.COM/COPYRIGHT-TRADEMARKS/)

PRIVACY(HTTPS://ANALYTICSINDIAMAG.COM/PRIVACY-POLICY/)

TERMS OF USE(HTTPS://ANALYTICSINDIAMAG.COM/TERMS-USE/)

CONTACT US(HTTPS://ANALYTICSINDIAMAG.COM/CONTACT-US/)

(https://analyticsindiamag.com/)

(HTTPS://FACEBOOK.COM/ANALYTICSINDIAMAGAZINE)

(HTTPS://TWITTER.COM/ANALYTICSINDIAM)

(HTTPS://INSTAGRAM.COM/ANALYTICSINDIAMAGAZINE)

(HTTPS://PINTEREST.COM/ANALYTICSINDIAM)

(HTTPS://YOUTUBE.COM/CHANNEL/UCALWRSGEJAVG1VW9QSFOUMA)

(HTTPS://MEDIUM.COM/ANALYTICS-INDIA-MAGAZINE)

(HTTPS://WWW.LINKEDIN.COM/COMPANY/ANALYTICS-INDIA-MAGAZINE)

(HTTPS://T.ME/NJLXNHZB7GKX3CPVJS9QGQ)

Scroll To Top

Scroll To Top