Bayesian Neural Networks

Diego Cesar Villa Almeyda

Master of Science

August, 2025

School of Mathematics
The University of Edinburgh

Dissertation Presented for the Degree of MSc in Statistics with Data Science

Table of contents

Preface

This is a Quarto book.

To learn more about Quarto books visit https://quarto.org/docs/books.

Executive summary

1 Introduction

This is a book created from markdown and executable code.

2 Methods

2.1 The UNSW-NB15 Network Dataset

The UNSW-NB15 dataset (1,2) was created to overcome the limitations of earlier benchmark datasets such as KDD99 and NSL-KDD, which have been criticised for outdated attack types, unrealistic normal traffic, and inconsistent distributions between training and testing sets. In contrast, UNSW-NB15 combines modern real-world network activity with synthetically generated attack behaviours, making it highly suitable for evaluating contemporary Network Intrusion Detection Systems (NIDSs). The dataset contains 49 features encompassing both flow-level host interactions and deep packet inspection metrics, enabling effective discrimination between normal and malicious traffic. It includes nine categories of contemporary cyberattacks alongside updated profiles of normal network behaviour. Statistically, UNSW-NB15 is more complex than its predecessors (2).

The full dataset comprises 2,540,044 records, of which 2,218,761 (approximately 87%) correspond to normal traffic, resulting in a highly imbalanced class distribution that reflects real-world network conditions (3). The training and testing subsets were obtained directly from the UNSW website (4), consisting of 175,341 and 82,332 records, respectively. Statistical analysis has demonstrated that the training and test sets share similar non-linear and non-normal feature distributions. Furthermore, high statistical correlation between the two sets supports their appropriateness as benchmark data for evaluating statistical and machine learning models tasked with distinguishing complex attack patterns from normal traffic (2). Non-informative features were excluded from the distributed datasets, yielding a total of 42 usable predictors and two target variables: `label` (binary attack indicator) and `attack_cat` (attack category), as described in Table S4.1.

An initial examination of the training dataset revealed that it contains a disproportionate number of attack records (68.06%) compared to normal traffic (31.93%), which does not reflect realistic conditions. To create a more representative imbalanced subset for our analysis, we retained only the normal traffic and denial-of-service (DoS) attack instances. This resulted in a subset with 82.03% normal and 17.97% DoS traffic, closely aligning with the class distribution in the full dataset.

2.2 Feature Engineering

2.3 Feature Selection

2.4 Bayesian Neural Networks (BNN)

2.4.1 Priors

2.4.2 Inference Methods

2.4.2.1 Markov Chain Monte Carlo (MCMC)

2.4.2.2 Variational Inference (VI)

2.4.3 Convergence

2.5 Model Benchmarking

2.5.1 Prediction Accuracy

2.5.2 Calibration

2.5.3 Running Time

2.6 Interpretability Analysis

3 Results

4 Conclusions

References

1. Moustafa N, Slay J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In: 2015 military communications and information systems conference (MilCIS). IEEE; 2015. p. 1–6.

2. Moustafa N, Slay J. The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. Information Security Journal: A Global Perspective. 2016;25(1-3):18–31.

3. Zoghi Z, Serpen G. UNSW-NB15 computer security dataset: Analysis through visualization. arXiv preprint arXiv:210105067. 2021;

4. University of New South Wales (UNSW). UNSW-NB15 dataset [Internet]. 2015. Available from: https://research.unsw.edu.au/projects/unsw-nb15-dataset

Appendix

Supplementary Tables

Name	Type	Description
id	Integer	Record ID.
dur	Float	Record total duration.
proto	Nominal	Transaction protocol.
service	Nominal	Such as http, ftp, smtp, ssh, dns and ftp-data.
state	Nominal	Indicates to the state and its dependent protocol (such as ACC, CLO and CON).
spkts	Integer	Source to destination packet count .
dpkts	Integer	Destination to source packet count.
sbytes	Integer	Source to destination transaction bytes.
dbytes	Integer	Destination to source transaction bytes.
rate	Float	Ethernet data rates transmitted and received.
sttl	Integer	Source to destination time to live value .
dttl	Integer	Destination to source time to live value.
sload	Float	Source bits per second.
dload	Float	Destination bits per second.
sloss	Integer	Source packets retransmitted or dropped .
dloss	Integer	Destination packets retransmitted or dropped.
sinpkt	Float	Source interpacket arrival time (mSec).
dinpkt	Float	Destination interpacket arrival time (mSec).
sjit	Float	Source jitter (mSec).
djit	Float	Destination jitter (mSec).
swin	Integer	Source TCP window advertisement value.
stcpb	Integer	Source TCP base sequence number.
dtcpb	Integer	Destination TCP base sequence number.
dwin	Integer	Destination TCP window advertisement value.
tcprtt	Float	TCP connection setup round-trip time, the sum of ’synack’ and ’ackdat’.
synack	Float	TCP connection setup time, the time between the SYN and the SYN_ACK packets.
ackdat	Float	TCP connection setup time, the time between the SYN_ACK and the ACK packets.
smean	Integer	Mean of the flow packet size transmitted by the source.
dmean	Integer	Mean of the flow packet size transmitted by the destination.
trans_depth	Integer	Represents the pipelined depth into the connection of http request/response transaction.
response_body_len	Integer	Actual uncompressed content size of the data transferred from the server’s http service.
ct_srv_src	Integer	No. of connections that contain the same service and source address in 100 connections according to the last time.
ct_state_ttl	Integer	No. for each state according to specific range of values for source/destination time to live.
ct_dst_ltm	Integer	No. of connections of the same destination address in 100 connections according to the last time.
ct_src_dport_ltm	Integer	No of connections of the same source address and the destination port in 100 connections according to the last time.
ct_dst_sport_ltm	Integer	No of connections of the same destination address and the source port in 100 connections according to the last time.
ct_dst_src_ltm	Integer	No of connections of the same source and the destination address in in 100 connections according to the last time.
is_ftp_login	Binary	If the ftp session is accessed by user and password then 1 else 0.
ct_ftp_cmd	Integer	No of flows that has a command in ftp session.
ct_flw_http_mthd	Integer	No. of flows that has methods such as Get and Post in http service.
ct_src_ltm	Integer	No. of records of the srcip in 100 records according to the ltime.
ct_srv_dst	Integer	No. of connections that contain the same service and destination address in 100 connections according to the last time.
is_sm_ips_ports	Binary	If source and destination IP addresses equal and port numbers equal then, this variable takes value 1 else 0
attack_cat	Nominal	The name of each attack category.
label	Binary	0 for normal and 1 for attack records

Table S4.1