國立聯合大學

112 學年度第 2 學期

深度學習

期末報告

授課教師：陳世杰 博士

學　　號：U0923009

姓　　名：卓緯晶

日　　期：112/6/12

目錄

一. Classification And Explanation of Different Internet of Things (IoT) Network Attacks Using Machine Learning, Deep Learning

i. 資料集

- Train_Test_Network.csv

ii. 修改

- 預處理方式

1.檢查每筆資料是否有重複或空缺值並刪除

2.文字轉數字

若使用 One hot encoding 則會使欄位過多，從而使模型太過於複雜。所以只針對 TYPE。

特徵值做 Label encoding	目標值做 one hot encoding

3.刪除高度相關性的特徵值欄位

Delete

dns_rejected

dns_RA

ssl_established

ssl_issuer

http_method

http_version

http_status_code

weird_notice

```
('dns_AA', 'dns_rejected')
('dns_RD', 'dns_RA')
('dns_RD', 'dns_rejected')
('dns_RA', 'dns_rejected')
('ssl_resumed', 'ssl_established')
('ssl_subject', 'ssl_issuer')
('http_trans_depth', 'http_method')
('http_trans_depth', 'http_version')
('http_trans_depth', 'http_status_code')
('http_method', 'http_version')
('http_method', 'http_status_code')
('http_version', 'http_status_code')
('weird_name', 'weird_notice')
```

4.正規化

將值轉換為 0-1 以降低離群值的產生

5.PCA 特徵值篩選

主成分 1 貢獻率: 0.1211
對應的欄位名稱: ts

主成分 2 貢獻率: 0.1005
對應的欄位名稱: src_port

主成分 3 貢獻率: 0.0726
對應的欄位名稱: dst_port

主成分 4 貢獻率: 0.0701
對應的欄位名稱: proto

主成分 5 貢獻率: 0.0572
對應的欄位名稱: service

主成分 6 貢獻率: 0.0554
對應的欄位名稱: duration

主成分 7 貢獻率: 0.0524
對應的欄位名稱: src_bytes

主成分 8 貢獻率: 0.0454
對應的欄位名稱: dst_bytes

主成分 9 貢獻率: 0.0345
對應的欄位名稱: conn_state

主成分 10 貢獻率: 0.0312
對應的欄位名稱: missed_bytes

主成分 11 貢獻率: 0.0301
對應的欄位名稱: src_pkts

主成分 12 貢獻率: 0.0300
對應的欄位名稱: src_ip_bytes

主成分 13 貢獻率: 0.0292
對應的欄位名稱: dst_pkts

主成分 14 貢獻率: 0.0277
對應的欄位名稱: dst_ip_bytes

主成分 15 貢獻率: 0.0273
對應的欄位名稱: dns_query

主成分 16 貢獻率: 0.0257
對應的欄位名稱: dns_qclass

主成分 17 貢獻率: 0.0242
對應的欄位名稱: dns_qtype

主成分 18 貢獻率: 0.0221
對應的欄位名稱: dns_rcode

主成分 19 貢獻率: 0.0207
對應的欄位名稱: dns_AA

主成分 20 貢獻率: 0.0159
對應的欄位名稱: dns_RD

主成分 21 貢獻率: 0.0157
對應的欄位名稱: ssl_version

主成分 22 貢獻率: 0.0155
對應的欄位名稱: ssl_cipher

主成分 23 貢獻率: 0.0148
對應的欄位名稱: ssl_resumed

主成分 24 貢獻率: 0.0109
對應的欄位名稱: ssl_subject

主成分 25 貢獻率: 0.0096
對應的欄位名稱: http_trans_depth

主成分 26 貢獻率: 0.0086
對應的欄位名稱: http_uri

主成分 27 貢獻率: 0.0073
對應的欄位名稱: http_request_body_len

主成分 28 貢獻率: 0.0060
對應的欄位名稱: http_response_body_len

主成分 29 貢獻率: 0.0047
對應的欄位名稱: http_user_agent

主成分 30 貢獻率: 0.0043
對應的欄位名稱: http_orig_mime_types

主成分 31 貢獻率: 0.0037
對應的欄位名稱: http_resp_mime_types

主成分 32 貢獻率: 0.0032
對應的欄位名稱: weird_name

主成分 33 貢獻率: 0.0025
對應的欄位名稱: weird_addl

取前 30 大約可高達:98.62%

主成分的累積貢獻率: [0.12113329 0.22167661 0.29426457 0.36433384 0.42155248 0.47693914
 0.52929599 0.57472819 0.60921296 0.64043465 0.6704919 0.70044191
 0.72963665 0.7573417 0.78460162 0.81031849 0.83449619 0.8565932
 0.87728339 0.89318648 0.90890268 0.92436207 0.93919323 0.95011942
 0.95975341 0.96838513 0.97565485 0.98161962 0.98627983 0.99054973
 0.99424435 0.99747368 1.   ]

- 成果

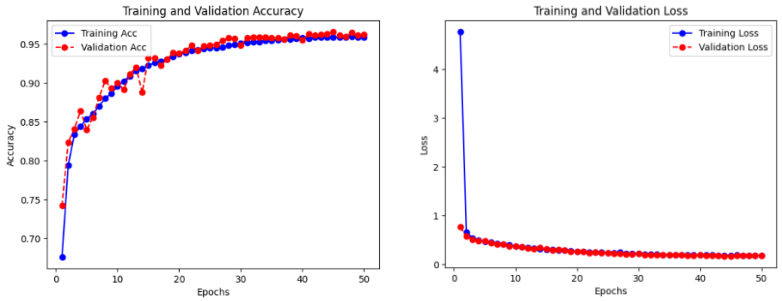
XGBOOST

```
XGBoost Accuracy: 0.9998288626809033
XGBoost Precision: 0.9998810187468155
XGBoost Recall: 0.9998511849399159
XGBoost F1 Score: 0.9998660546051443
```

MLP

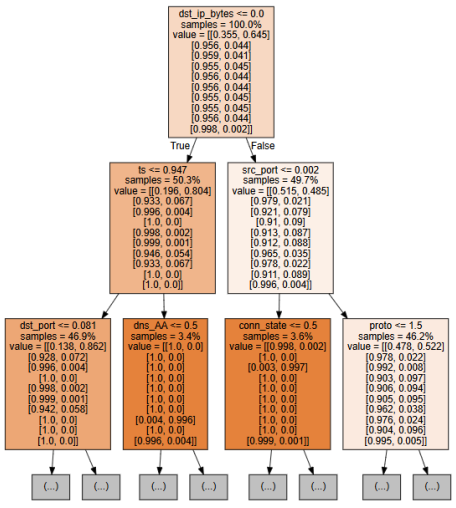
```
4200/4200 [==============================] - 6s 1ms/step
MLP Accuracy: 0.9478105584285129
MLP Precision: 0.9669408245458935
MLP Recall: 0.9478105584285129
MLP F1 Score: 0.9536314587682603
```

無 overfitting 的產生

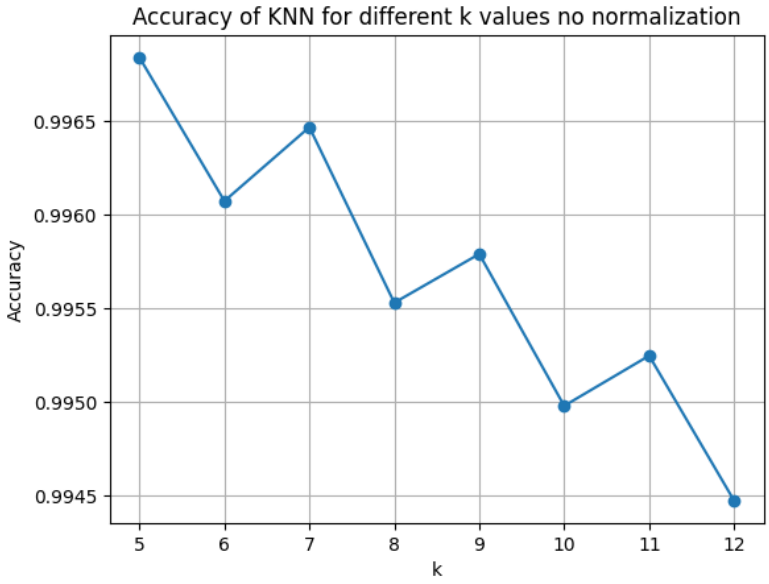


RandomForestClassifier：10 trees to decide

```
DF Accuracy: 0.9992708062055881
DF Precision: 0.9998362134795806
DF Recall: 0.9992708062055881
DF F1 Score: 0.9995530973363224
```



KNN



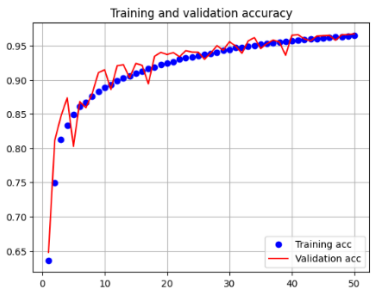
KNN Accuracy: 0.9966665426541166
KNN Precision: 0.9971269501358682
KNN Recall: 0.9966665426541166
KNN F1 Score: 0.9968843197954497

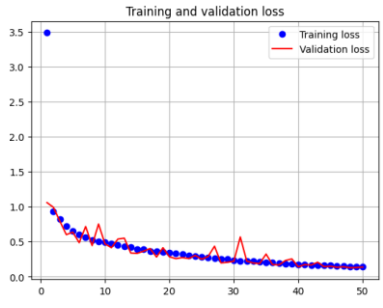
when k = 5

CNN

CNN Accuracy: 0.9581234420923398
CNN Precision: 0.9687171825958013
CNN Recall: 0.9581234420923398
CNN F1 Score: 0.9626252915712921

無 overfitting 的產生





二. Fine-Tuning LLaMA 2

i. 環境介紹

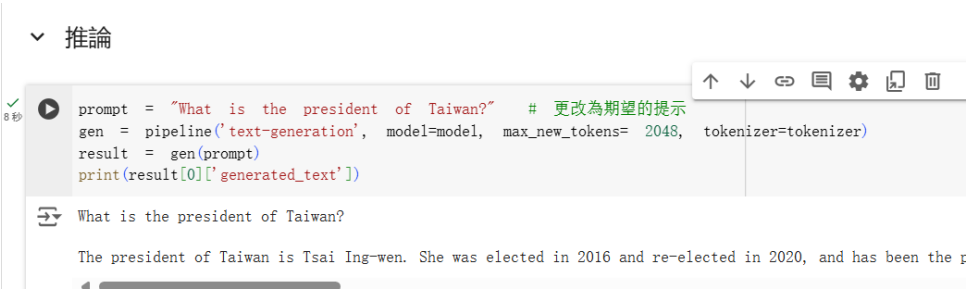
- **LLaMA 模型微調**：使用 Hugging Face 的 accelerate 以及 transformers 庫來在不同設備上運行原始 PyTorch 訓練腳本。
- **Q-LoRA**：QLoRA 引入了 4 位 NormalFloat 量化和 Double 量化，改進了傳統量化方法，有效壓縮模型資訊並節省記憶體，同時保持整體性能。這對於在有限硬體資源上進行深度學習模型的訓練和微調非常有用。此外，QLoRA 利用 NVIDIA Unified Memory 實現了 CPU 和 GPU 之間的無縫頁到頁傳輸，確保了 GPU 處理過程中無錯誤。
- **PEFT 庫**：PEFT（Parameter-Efficient Fine-Tuning）庫可有效地適應各種下游應用，而無需微調模型的所有參數。
- **bitsandbytes 庫**：展示了如何使用 bitsandbytes 庫實現 8 位優化器、矩陣乘法和量化功能。
- **TRL 庫**：說明了如何使用 TRL 庫通過監督微調（SFT）、獎勵建模（RM）和近似策略優化（PPO）等方法微調和對齊轉換器語言模型。

ii. 修改

- Change dataset to finetune：mlabonne/guanaco-llama2-1k => Fine_tuning_LLMs_Dataset
- 將訓練資料切分為 train_dataset and valid_dataset 分別查看其 loss

Step	Training Loss	Validation Loss
5	0.797600	0.663555
10	0.560000	0.516613
15	0.463100	0.459853
20	0.408000	0.449419

- 將微調完的模型儲存合併並載入進行推論



iii. 來源資料

[1] https://github.com/LlamaFamily/Llama-Chinese

[2] https://huggingface.co/meta-llama/Llama-2-7b-chat-hf

[3] Dataset 1 2

三. Alpaca + Llama-3 8b

i. 介紹

- Unsloth 預訓練模型：使用 Unsloth 預訓練模型，其中包括 Llama-3 8b、Mistral、Phi-3 等多種模型。

1 A100 40GB	Hugging Face	Flash Attention	Unsloth Open Source	Unsloth Pro
Alpaca	1x	1.04x	1.98x	**15.64x**
LAION Chip2	1x	0.92x	1.61x	**20.73x**
OASST	1x	1.19x	2.17x	**14.83x**
Slim Orca	1x	1.18x	2.22x	**14.82x**

Unsloth supports	Free Notebooks	Performance	Memory use
Llama 3 (8B)	Start for free	2x faster	60% less
Mistral v0.3 (7B)	Start for free	2.2x faster	73% less
Phi-3 (medium)	Start for free	2x faster	50% less
Phi-3 (mini)	Start for free	2x faster	50% less
Gemma (7B)	Start for free	2.4x faster	71% less
ORPO	Start for free	1.9x faster	43% less
DPO Zephyr	Start for free	1.9x faster	43% less
TinyLlama	Start for free	3.9x faster	74% less

Current memory	After training：final memory and time state
GPU = Tesla T4. Max memory = 14.748 GB. 5.594 GB of memory reserved.	471.5364 seconds used for training. 7.86 minutes used for training. Peak reserved memory = 7.535 GB. Peak reserved memory for training = 1.941 GB. Peak reserved memory % of max memory = 51.092 %. Peak reserved memory for training % of max memory = 13.161 %

- **LoRA 訓練參數設置**: 說明了如何添加 LoRA 適配器以更新模型的部分參數。
- **Alpaca 數據集和格式化提示**
- **模型訓練**: 使用 Huggingface TRL 的 SFTTrainer 進行模型訓練，並展示了如何設置訓練參數。
- **保存為不同格式**：例如 float16 或 int4，以及如何轉換為 GGUF / llama.cpp 格式。

- **模型測試與保存**: 演示了如何測試訓練效果，以及如何保存和加載微調後的模型。

ii. 來源資料

[1] Unsloth Finetune Llama 3, Mistral, Phi-3 & Gemma 2-5x faster with 80% less memory