一、資料集分析

1. 資料集：DNRTI (Dataset for Named Entity Recognition in Threat Intelligence) 命名實體辨識資料集
2. 每行為一個字與對應標籤，標註格式為 BIOES
   i. BIOES 標註方案 (Beginning, Inside, Outside):
      - B-TYPE: 一個命名實體的開始。TYPE 指的是實體的類型
      - I-TYPE: 一個命名實體的內部。表示這個詞是前面同類型命名實體的延續
      - O: 不屬於任何已定義的命名實體
      - E- (End): 標示一個命名實體的結束詞 (最後一個詞)
      - S- (Single): 標示由單個詞組成的命名實體

      // 註：有將原本的 BIO 標註方案，改成 BIOES 標註方案

   ii. 實體類型（Entity Types），共計 13 類
      - SamFile: 檔案，如:Vietnam.exe 、malicious file、decoy documents
      - Way: 攻擊方法，如:spear、emails、lure
      - SecTeam: 網路安全團隊，FireEye、Arbor、information security community
      - Time: 時間，如: August 2015
      - HackOrg: 駭客組織, 如: admin@338、cyber threat groups、threat actors
      - Purp: 目的, 如: steal information、intelligence-gathering
      - Tool: 工具, 如: CORESHELL、SPLM、JHUHUGIT
      - Features: 功能, send POST requests、contain information、connect to a command and control、persistence functionality
      - Org: 一般組織, 如: International Civil Aviation Organization、Bitcoin users
      - Exp: 漏洞利用, 如: CVE-2015-8651、CVE-2016-1019、CVE-2016-4117
      - Area: 地區, 如: Turkish、Turkey
      - OffAct: 攻擊行為, 如: malicious cyber activity、watering hole attacks、cyberespionage attacks

- Idus: Industrial , 如: government、enterprises、businesses、telecommunications、shipping、car、manufacturers

二、模型設計(使用 SecBERT + BiLSTM + CRF)

1. 架構：SecBERT + BiLSTM + CRF

```
[Input Text]
     ↓
[SecBERT Embedding]
     ↓
[BiLSTM]
     ↓
[CRF Layer]
     ↓
[Predicted Labels]
```

i. SecBERT（Security BERT）：以資安文本為訓練語料的預訓練語言模型，負責輸出 contextual embedding

- tokenizer = AutoTokenizer.from_pretrained("jackaduma/SecBERT")

ii. BiLSTM：雙向 LSTM 捕捉序列資訊，提升辨識前後文關係

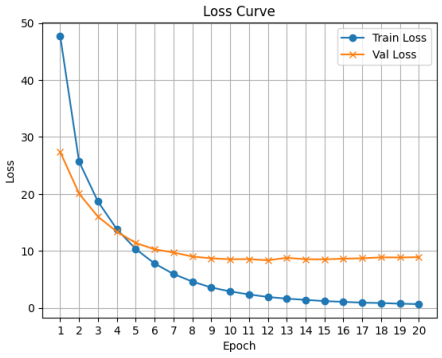
iii. CRF（條件隨機場）：用於標籤解碼，考慮標籤間的轉移關係以提升序列標註準確性

三、超參數設定

1. parameters['lower'] = True：將字改為小寫
2. parameters['zeros'] = True：將所有數字改為 0
3. Optimizer：Adam
4. Learning Rate：3e-5
5. Batch Size：32
6. Epochs：20
7. Padding Token：[PAD] 以 label = -3 表示，訓練與評估時忽略

四、評估指標與結果分析

1. Loss 圖：Train 和 Validation 都有穩定下降



2. Validation

Entity	Precision	Recall	F1	Support
B-Area	0.8113	0.9773	0.8866	44
B-Exp	0.9600	1.0000	0.9796	24
B-Features	0.8814	1.0000	0.9369	52
B-HackOrg	0.6897	0.7752	0.7299	129
B-Idus	0.8611	0.9688	0.9118	32
B-OffAct	0.6591	0.6591	0.6591	44
B-Org	0.7636	0.7000	0.7304	60
B-Purp	0.7586	0.9851	0.8571	67
B-SamFile	0.9714	0.9444	0.9577	72
B-SecTeam	0.8936	0.7925	0.8400	53
B-Time	0.8312	0.8312	0.8312	77
B-Tool	0.6179	0.6387	0.6281	119
B-Way	0.8293	0.8718	0.8500	78

Entity	Precision	Recall	F1	Support
I-Area	0.5000	0.7500	0.6000	4
I-Exp	1.0000	1.0000	1.0000	17
I-Features	0.8140	1.0000	0.8974	35
I-HackOrg	0.7778	0.5833	0.6667	12
I-Idus	1.0000	1.0000	1.0000	5
I-OffAct	0.6667	0.3810	0.4848	21
I-Org	0.5000	0.5000	0.5000	8
I-Purp	0.8182	1.0000	0.9000	63
I-SamFile	0.9444	1.0000	0.9714	17
I-SecTeam	0.7500	0.7500	0.7500	12
I-Time	0.3500	0.7778	0.4828	9
I-Tool	0.6774	0.5676	0.6176	37
I-Way	0.9643	0.9310	0.9474	29

Entity	Precision	Recall	F1	Support
E-Area	0.8600	0.9773	0.9149	44
E-Exp	0.9600	1.0000	0.9796	24
E-Features	0.9455	1.0000	0.9720	52
E-HackOrg	0.6643	0.7364	0.6985	129
E-Idus	0.8611	0.9688	0.9118	32
E-OffAct	0.6739	0.7045	0.6889	44
E-Org	0.7742	0.8000	0.7869	60
E-Purp	0.7500	0.9851	0.8516	67
E-SamFile	0.8933	0.9306	0.9116	72
E-SecTeam	0.8696	0.7547	0.8081	53
E-Time	0.7356	0.8312	0.7805	77
E-Tool	0.6638	0.6471	0.6553	119
E-Way	0.8250	0.8462	0.8354	78

Entity	Precision	Recall	F1	Support
S-Area	0.8901	0.8100	0.8482	100
S-Exp	0.9906	0.9813	0.9859	107
S-Features	0.9184	1.0000	0.9574	45
S-HackOrg	0.7654	0.8350	0.7987	297
S-Idus	0.9286	0.9155	0.9220	71
S-OffAct	0.6471	0.7857	0.7097	56
S-Org	0.6923	0.5192	0.5934	52
S-Purp	0.1899	1.0000	0.3191	15
S-SamFile	0.8829	0.8235	0.8522	119
S-SecTeam	0.8862	0.8720	0.8790	125
S-Time	0.9263	0.8302	0.8756	106
S-Tool	0.8207	0.6010	0.6939	198
S-Way	0.8621	0.8065	0.8333	31

Entity	Precision	Recall	F1
Macro avg	0.7917	0.8336	0.8015
Micro avg	0.7882	0.8202	0.8039

3. Test

Entity	Precision	Recall	F1	Support
B-Area	0.7818	0.8958	0.8350	48
B-Exp	1.0000	1.0000	1.0000	30
B-Features	0.9438	1.0000	0.9711	84
B-HackOrg	0.7833	0.7581	0.7705	124
B-Idus	0.7941	1.0000	0.8852	27
B-OffAct	0.9032	0.8000	0.8485	70
B-Org	0.7191	0.7191	0.7191	89
B-Purp	0.8515	1.0000	0.9198	86
B-SamFile	0.9589	0.9589	0.9589	73
B-SecTeam	0.8696	0.8511	0.8602	47
B-Time	0.9048	0.9620	0.9325	79
B-Tool	0.7561	0.7623	0.7592	122
B-Way	0.8974	1.0000	0.9459	70

Entity	Precision	Recall	F1	Support
I-Area	0.6667	0.5000	0.5714	4
I-Exp	1.0000	1.0000	1.0000	19
I-Features	0.9437	1.0000	0.9710	67
I-HackOrg	0.5000	0.2222	0.3077	9
I-Idus	1.0000	1.0000	1.0000	9
I-OffAct	0.8125	0.6190	0.7027	21
I-Org	0.2857	0.5000	0.3636	12
I-Purp	0.8429	1.0000	0.9147	59
I-SamFile	1.0000	1.0000	1.0000	22
I-SecTeam	0.5294	1.0000	0.6923	9
I-Time	0.7895	0.7500	0.7692	20
I-Tool	0.8056	0.7436	0.7733	39
I-Way	0.9500	1.0000	0.9744	19

Entity	Precision	Recall	F1	Support
E-Exp	0.9677	1.0000	0.9836	30
E-Features	0.9438	1.0000	0.9711	84
E-HackOrg	0.7917	0.7661	0.7787	124
E-Idus	0.7941	1.0000	0.8852	27
E-OffAct	0.9048	0.8143	0.8571	70
E-Org	0.7849	0.8202	0.8022	89
E-Purp	0.8431	1.0000	0.9149	86
E-SamFile	0.9452	0.9452	0.9452	73
E-SecTeam	0.9111	0.8723	0.8913	47
E-Time	0.8675	0.9114	0.8889	79
E-Tool	0.7876	0.7295	0.7574	122
E-Way	0.9091	1.0000	0.9524	70

Entity	Precision	Recall	F1	Support
S-Area	0.8968	0.8274	0.8607	168
S-Exp	0.9714	1.0000	0.9855	102
S-Features	0.9697	1.0000	0.9846	32
S-HackOrg	0.8256	0.8694	0.8469	245
S-Idus	0.9691	0.9216	0.9447	102
S-OffAct	0.8889	0.8000	0.8421	80
S-Org	0.6809	0.6667	0.6737	48
S-Purp	0.3671	1.0000	0.5370	29
S-SamFile	0.8902	0.8800	0.8851	175
S-SecTeam	0.9216	0.8952	0.9082	105
S-Time	0.9222	0.9222	0.9222	90
S-Tool	0.8625	0.7150	0.7819	193
S-Way	0.9667	0.9667	0.9667	30

Entity	Precision	Recall	F1
Macro avg	0.8413	0.8700	0.8479
Micro avg	0.8531	0.8713	0.8621

4. 結論

i. 可以看到大多 Entity 的表現都良好，Precision/Recall/F1 都很高，且 macro F1、micro F1 都在 0.8 以上， 代表模型對於整體實體辨識已不錯準確。

ii. 然而依舊有些 Entity 表現不佳，表現不佳的大多只有少量樣本，如 I-HackOrg, I-Org。但並非所有只有少量樣本的 Entity 都表現差，如 I-Idus。對於這類只有少量樣本且表現不佳的 Entity，可以使用資料增強手段來平衡資料分布。