

深度学习是机器学习领域的一个子集。

深度神经网络 (DNN)

Keras 是一个用 Python 编写的高级神经网络 API，能够以 TensorFlow 作为后端运行。

Keras 的核心数据结构是 model，一种组织网络层的方式。最简单的模型是 Sequential 顺序模型，它由多个网络层线性堆叠。

```
model = tf.keras.Sequential()

model.add(tf.keras.layers.Dense(128, activation='relu', input_dim=10))
model.add(tf.keras.layers.Dense(128, activation='relu'))
model.add(tf.keras.layers.Dense(128, activation='relu'))
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))

model.compile(optimizer="adam",
              loss="binary_crossentropy",
              metrics=["acc"])
```

卷积神经网络 (CNN)

```
model = tf.keras.Sequential()

model.add(tf.keras.layers.Conv1D(100, 2, padding='valid', activation='relu', input_shape=(10,1)))
model.add(tf.keras.layers.MaxPool1D(pool_size=2))
model.add(tf.keras.layers.Conv1D(100, 2, padding='valid', activation='relu'))
model.add(tf.keras.layers.MaxPool1D(pool_size=2))
model.add(tf.keras.layers.Dropout(0.25))
model.add(tf.keras.layers.Flatten())
model.add(tf.keras.layers.Dropout(0.5))
model.add(tf.keras.layers.Dense(100, activation='relu'))
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))

model.compile(optimizer="adam",
              loss="binary_crossentropy",
              metrics=["acc"])
```

循环神经网络 (RNN)

Simple RNN:

```
model = tf.keras.Sequential()

model.add(tf.keras.layers.SimpleRNN(128, activation='tanh', return_sequences=True, input_shape=(10,1)))
model.add(tf.keras.layers.SimpleRNN(128, activation='tanh', return_sequences=True))
model.add(tf.keras.layers.SimpleRNN(128, activation='tanh'))
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))

model.compile(optimizer="rmsprop",
              loss="binary_crossentropy",
              metrics=["acc"])
```

LSTM:

```
model = tf.keras.Sequential()

model.add(tf.keras.layers.LSTM(128, activation='tanh', return_sequences=True, input_shape=(10,1)))
model.add(tf.keras.layers.LSTM(128, activation='tanh', return_sequences=True))
model.add(tf.keras.layers.LSTM(128, activation='tanh'))
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))

model.compile(optimizer="rmsprop",
              loss="binary_crossentropy",
              metrics=["acc"])
```

GRU:

```
model = tf.keras.Sequential()

model.add(tf.keras.layers.GRU(128, activation='tanh', return_sequences=True, input_shape=(10,1)))
model.add(tf.keras.layers.GRU(128, activation='tanh', return_sequences=True))
model.add(tf.keras.layers.GRU(128, activation='tanh'))
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))

model.compile(optimizer="rmsprop",
              loss="binary_crossentropy",
              metrics=["acc"])
```

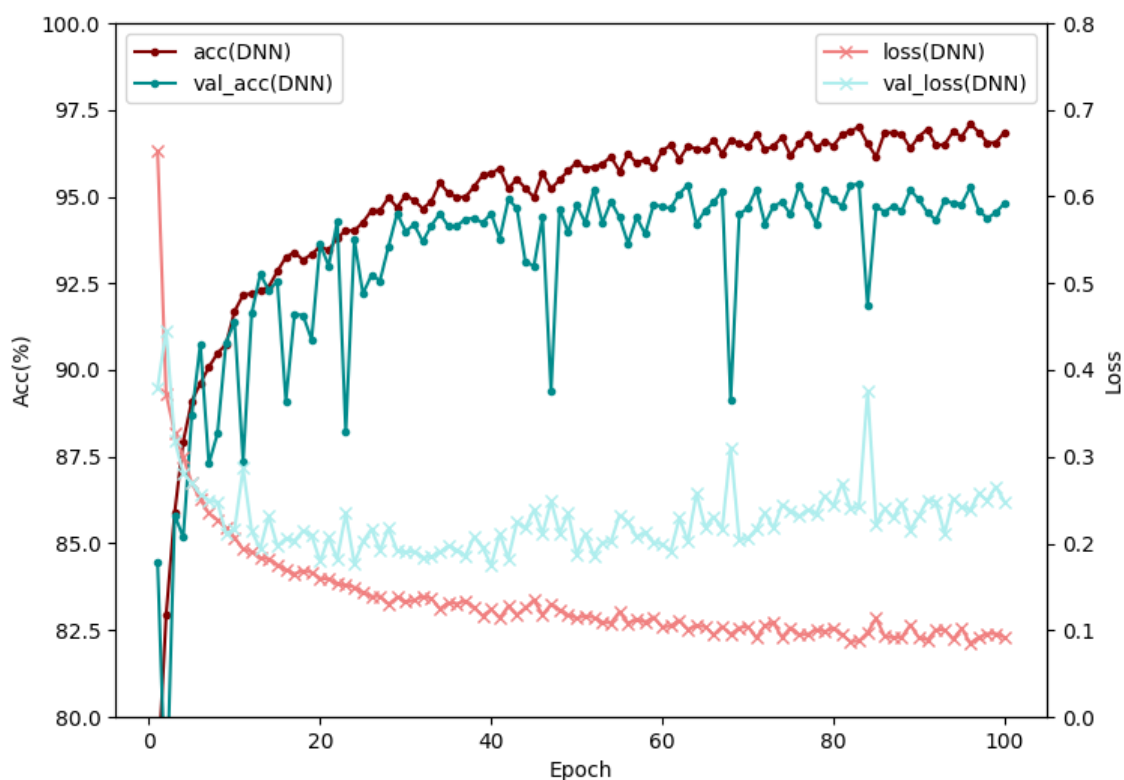
实验一：使用“长度”相关特征对普通流量和 Tor 流量进行二分类

原始数据集中，正样本数量为 13006，负样本数量为 7644。

```
rows_1 = [2000, 1225, 967, 2000, 2814, 2000, 2000]
rows_2 = [1797, 208, 273, 1725, 598, 1559, 1484]
```

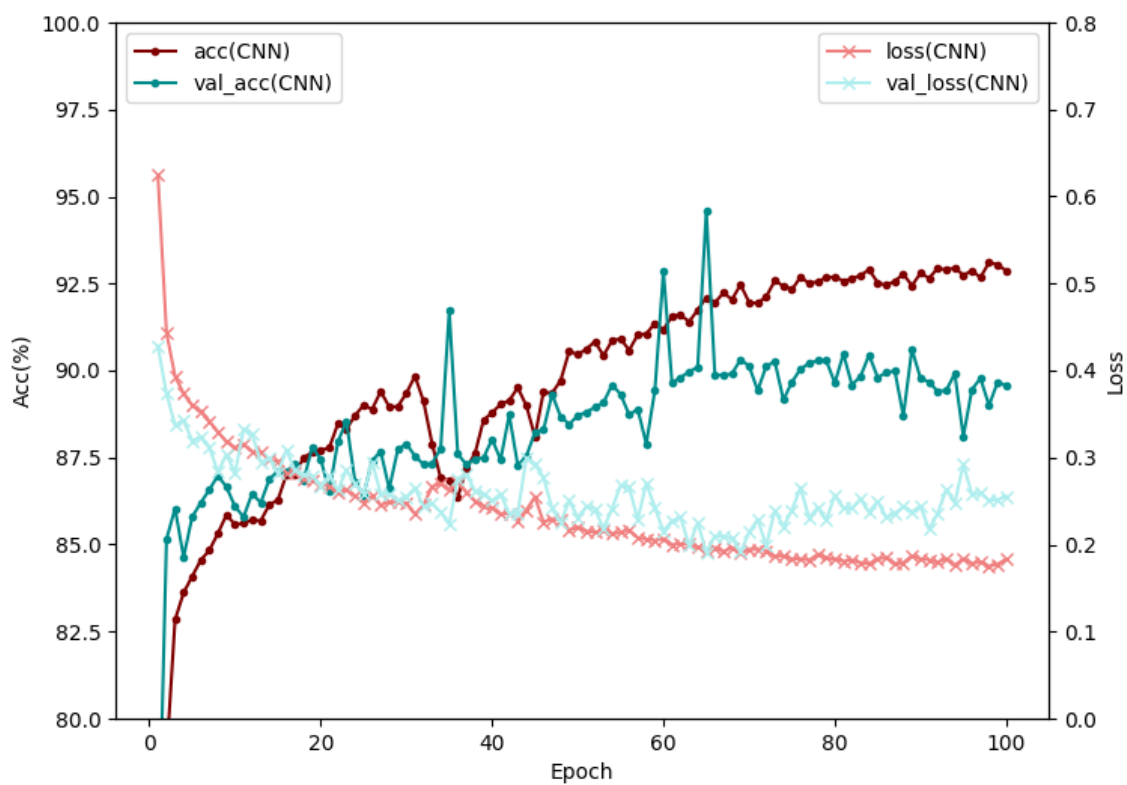
epoch = 100 / test_size = 0.1 / validation_split = 0.2

DNN 结果:



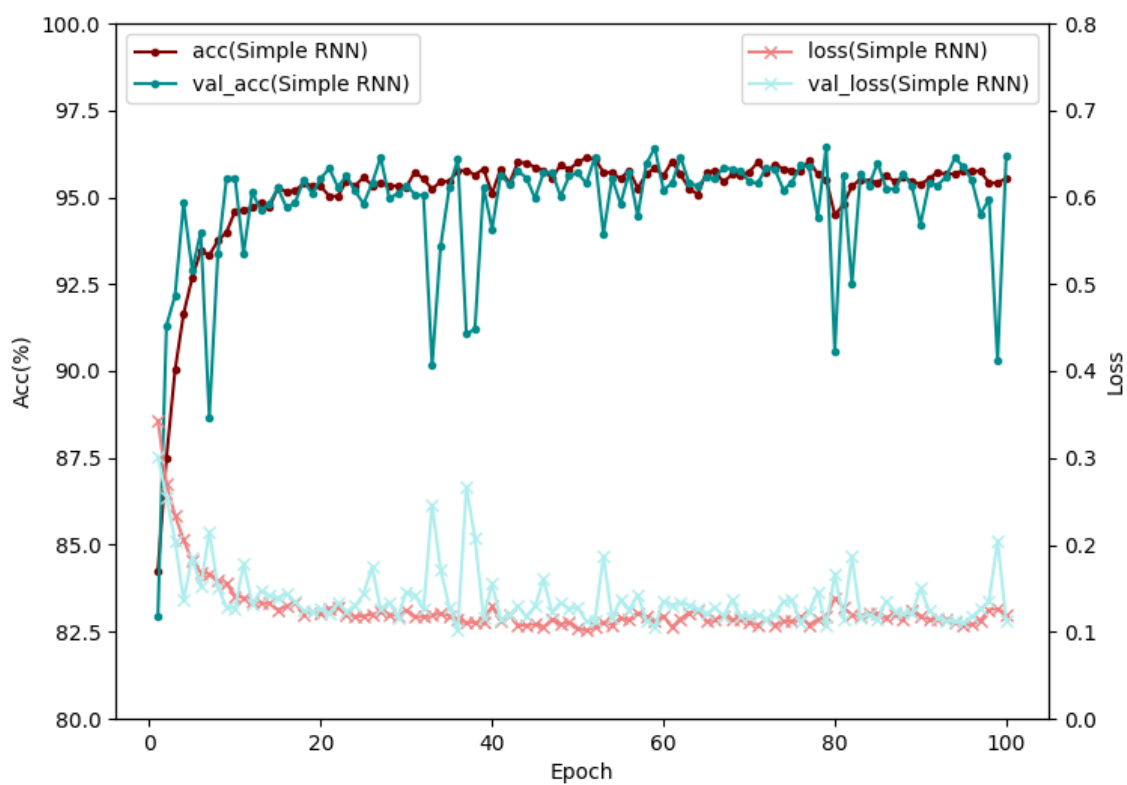
测试集准确率为 94.77%。

CNN 结果:



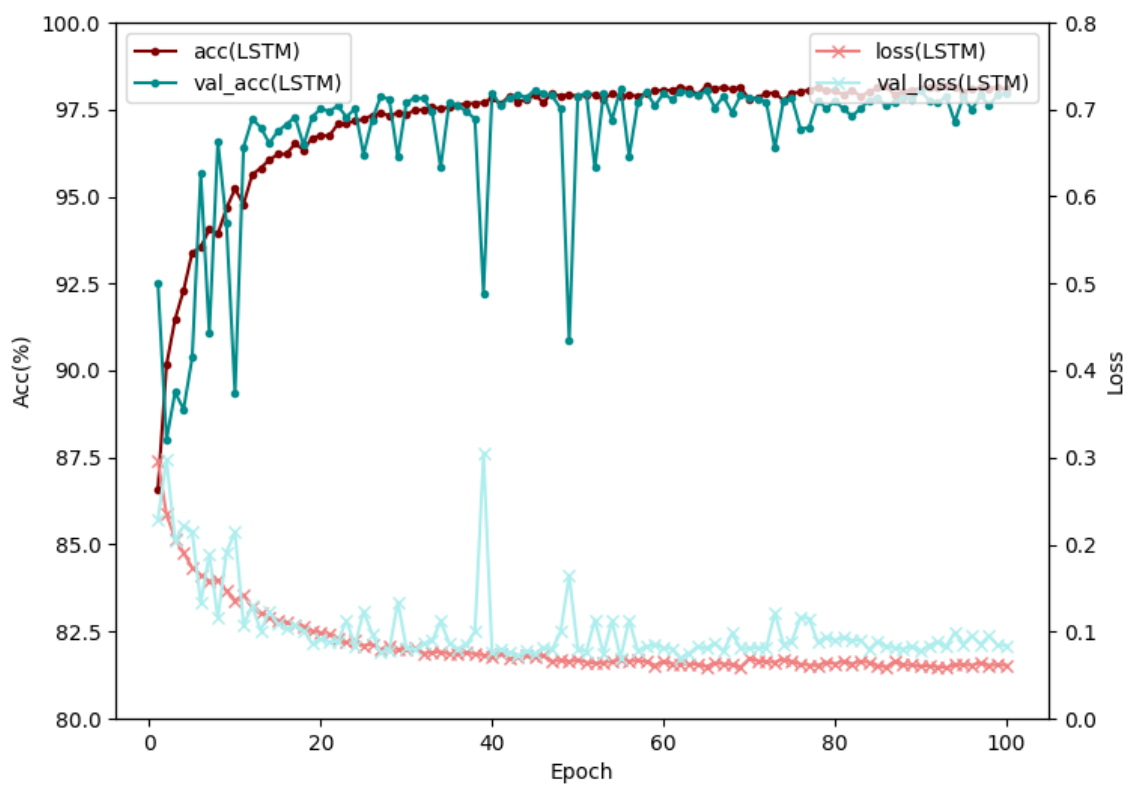
测试集准确率为 90.02%。

Simple RNN 结果：



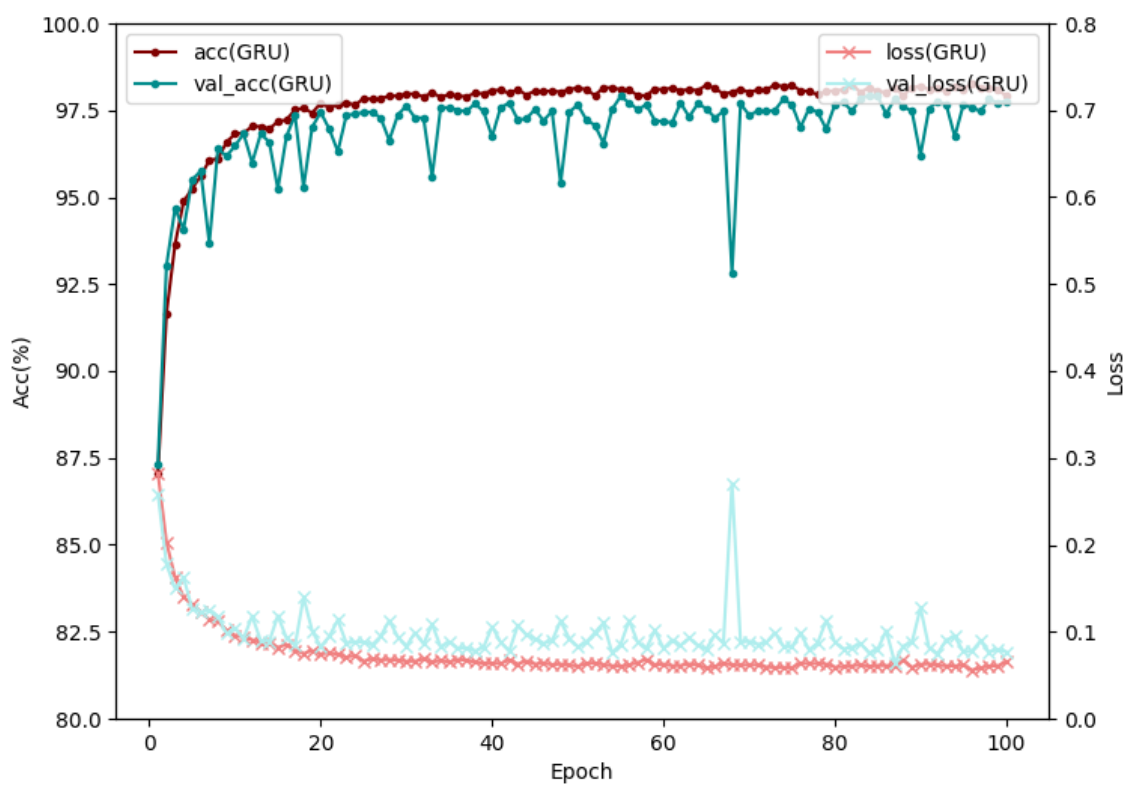
测试集准确率为 96.27%。

LSTM 结果：



测试集准确率为 97.68%。

GRU 结果:



测试集准确率为 97.38%。