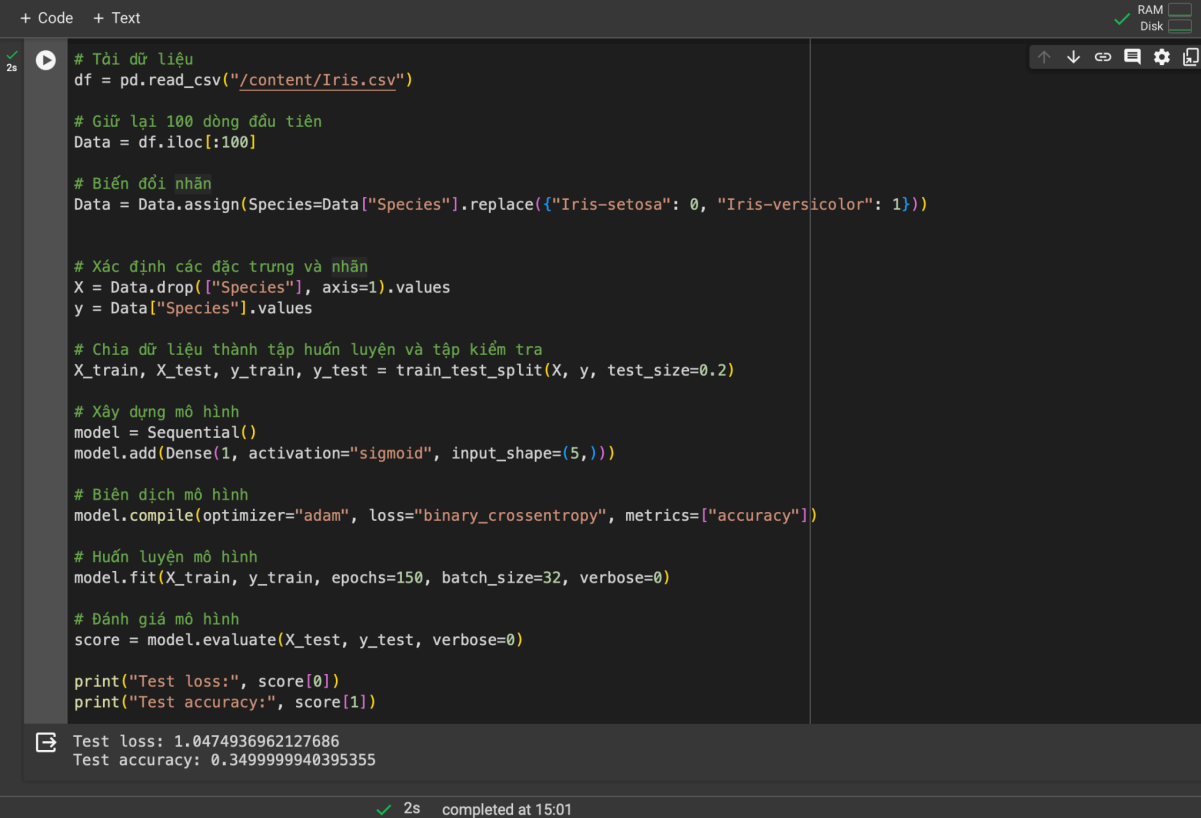


Họ và Tên: Nguyễn Gia Thịnh
MSSV: B2017081

Bài 1: Phân loại hoa iris (2 lớp)



```
+ Code + Text
# Tải dữ liệu
df = pd.read_csv("/content/Iris.csv")

# Giữ lại 100 dòng đầu tiên
Data = df.iloc[:100]

# Biến đổi nhãn
Data = Data.assign(Species=Data["Species"].replace({"Iris-setosa": 0, "Iris-versicolor": 1}))

# Xác định các đặc trưng và nhãn
X = Data.drop(["Species"], axis=1).values
y = Data["Species"].values

# Chia dữ liệu thành tập huấn luyện và tập kiểm tra
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

# Xây dựng mô hình
model = Sequential()
model.add(Dense(1, activation="sigmoid", input_shape=(5,)))

# Biên dịch mô hình
model.compile(optimizer="adam", loss="binary_crossentropy", metrics=["accuracy"])

# Huấn luyện mô hình
model.fit(X_train, y_train, epochs=150, batch_size=32, verbose=0)

# Đánh giá mô hình
score = model.evaluate(X_test, y_test, verbose=0)

print("Test loss:", score[0])
print("Test accuracy:", score[1])

Test loss: 1.0474936962127686
Test accuracy: 0.3499999940395355

2s completed at 15:01
```

Các bước thực hiện

- Tải dữ liệu iris
- Giữ lại 100 dòng đầu tiên
- Biến đổi nhãn
- Xáo trộn dữ liệu
- Chia dữ liệu thành tập huấn luyện (80%) và tập kiểm tra (20%)
- Xây dựng mô hình
 - 1 tầng (Dense) gồm 1 nơon duy nhất (2 lớp)
 - Input_shape: (4,)
 - Use bias: true (mặc định)
 - Activation: sigmoid
- Biên dịch mô hình
 - Optimizer: 'adam'
 - Loss: 'binary_crossentropy'
 - Metrics: ['accuracy']
- Huấn luyện mô hình
 - Epochs: 150
 - Batch size: 32
- Đánh giá mô hình
 - Test loss
 - Test accuracy

Kết quả

- Test loss: 1.047
- Test accuracy: 0.349

Giải thích

- Test loss là 1.047, nghĩa là mô hình ước lượng sai nhãn hoa iris trong tập kiểm tra trung bình là 1.047.
- Test accuracy là 0.349, nghĩa là mô hình phân loại đúng nhãn hoa iris trong tập kiểm tra trung bình là 34.9%.

Nhận xét

- Mô hình đạt được độ chính xác thấp, chỉ 34.9%.

Bài 2: Phân loại hoa iris (3 lớp)

```
+ Code + Text
2s
# Tải dữ liệu
df = pd.read_csv("/content/Iris.csv")

# Xử lý toàn bộ tập dữ liệu
# Biến đổi nhãn
df["Species"].replace({"Iris-setosa": 0, "Iris-versicolor": 1, "Iris-virginica": 2}, inplace=True)

# Xác định các đặc trưng và nhãn
X = df.drop(["Species"], axis=1).values
y = df["Species"].values

# Chia dữ liệu thành tập huấn luyện và tập kiểm tra
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25)

# Xây dựng mô hình
model = Sequential()
model.add(Dense(3, activation="softmax", input_shape=(5,)))

# Biên dịch mô hình
model.compile(optimizer="adam", loss="sparse_categorical_crossentropy", metrics=["accuracy"])

# Huấn luyện mô hình
model.fit(X_train, y_train, epochs=150, batch_size=32, verbose=0)

# Đánh giá mô hình
score = model.evaluate(X_test, y_test, verbose=0)

print("Test loss:", score[0])
print("Test accuracy:", score[1])

WARNING:tensorflow:5 out of the last 5 calls to <function Model.make_test_function.<locals>.test_function at 0x7ff617483f40>
Test loss: 1.8467010259628296
Test accuracy: 0.5526315569877625

New section
1s completed at 16:04
```

Các bước thực hiện

- Tải dữ liệu iris
- Biến đổi nhãn
- Xáo trộn dữ liệu
- Chia dữ liệu thành tập huấn luyện (80%) và tập kiểm tra (20%)
- Xây dựng mô hình
 - 1 tầng (Dense) gồm 3 nơon (3 lớp)
 - Input_shape: (4,)
 - Use bias: true (mặc định)
 - Activation: softmax
- Biên dịch mô hình
 - Optimizer: 'adam'
 - Loss: 'sparse_categorical_crossentropy'
 - Metrics: ['accuracy']

- Huấn luyện mô hình
 - Epochs: 150
 - Batch size: 32
- Đánh giá mô hình
 - Test loss
 - Test accuracy

Kết quả

- Test loss: 1.846
- Test accuracy: 0.55

Giải thích

- Test loss là 1.846, nghĩa là mô hình ước lượng sai nhãn hoa iris trong tập kiểm tra trung bình là 1.846.
- Test accuracy là 0.552, nghĩa là mô hình phân loại đúng nhãn hoa iris trong tập kiểm tra trung bình là 55.2%.

Nhận xét

- Mô hình đạt được độ chính xác trung bình, chỉ 55.2%.