مشروع نهاية الفصل 4

مقرر Deep Learning — Regularization — أثر الـ weight decay

الهدف: تشوف الـ overfitting على داتا تدريب صغيرة، وتجرّب weight decay بقيم مختلفة.
الأدوات: Python مع numpy وscikit-learn، وتقدر تشغّل الكود في Google Colab المجاني.

الداتا: sklearn.datasets.load_digits (١٧٩٧ صورة ٨×٨، ١٠ أرقام). مصدرها UCI. مدمجة في scikit-learn.

الأسئلة العملية

1 جهّز (Prepare)

- قسّم ٨٠/٢٠، وخد أول ٢٠٠ مثال بس للتدريب عشان الـ overfitting يظهر.

2 شبكة كبيرة (Big network)

- درّب MLPClassifier بـ (128,) و alpha=0 واطبع دقة الـ train والـ test.

3 Weight decay (Weight decay)

- جرّب alpha = 0.001, 0.1, 1 واطبع الدقتين.

4 اختار (Choose)

- اختار alpha بالـ cross-validation على الـ training بس (cv=3).

بونس

- بونس: جرّب early_stopping=True.
- بونس: جرّب (16,) بدل (128,).

أسئلة تفكير (اكتب إجابة في سطرين)

1. إيه علامة الـ overfitting؟
2. ليه نختار alpha بالـ cross-validation مش بالـ test؟

ماذا تسلّم؟

- ملف الكود (أو notebook) بيشتغل من الأول للآخر.
- فقرة قصيرة (٥ أسطر) بتلخّص النتيجة وإيه اللي اتعلمته.

الحل النموذجي

جرّب بنفسك الأول، وبعدين قارن. الأرقام الحقيقية في النتيجة تحت.

الكود

```
# dataset ships inside scikit-learn (no download needed)
import numpy as np
from sklearn.datasets import load_digits
from sklearn.model_selection import train_test_split

X, y = load_digits(return_X_y=True)
X = X / 16.0
Xtr, Xte, ytr, yte = train_test_split(X, y, test_size=0.2, random_state=0)
Ytr = np.eye(10)[ytr]
from sklearn.neural_network import MLPClassifier
from sklearn.model_selection import cross_val_score

Xs, ys = Xtr[:200], ytr[:200]
for alpha in (0.0, 0.001, 0.1, 1.0):
    m = MLPClassifier((128,), alpha=alpha, max_iter=500, random_state=0).fit(Xs, ys)
    print('alpha', alpha, 'train', round(m.score(Xs, ys), 3), 'test', round(m.score(Xte,
yte), 3))
best = max((0.0, 0.001, 0.1, 1.0), key=lambda a: cross_val_score(MLPClassifier((128,),
alpha=a, max_iter=500, random_state=0), Xs, ys, cv=3).mean())
print('chosen alpha by CV:', best)
```

النتيجة

```
alpha 0.0 train 1.0 test 0.931
alpha 0.001 train 1.0 test 0.933
alpha 0.1 train 1.0 test 0.936
alpha 1.0 train 1.0 test 0.933
chosen alpha by CV: 1.0
```

ملاحظات على الإجابات

- الفجوة بين دقة الـ train والـ test هي العلامة. الفروق الصغيرة بين قيم alpha ممكن تكون ضوضاء، شوف الأرقام الحقيقية فوق.
- الـ test لازم يفضل للتقييم النهائي بس.

Credits: topic follows the matching chapter of Dive into Deep Learning (d2l.ai, CC BY-SA 4.0).
Questions, code and explanations are original.