

Progress

Mizuno Yasuaki

October 16, 2022

- アミノ酸ベクタの割り当て①
- 学習結果①
- アミノ酸ベクタの割り当て②
- 学習結果
- まとめ

アミノ酸ベクタの割り当て①

x 成分をアミノ酸のサイズ、y 成分を疎水性尺度とした。

Table 1: アミノ酸ベクタ

| アミノ酸 | (x, y) | アミノ酸 | (x, y) |
|------|--------------|------|--------------|
| A | (2.5, 0.02) | M | (6.0, 1.00) |
| C | (3.0, 0.77) | N | (5.0, -0.77) |
| E | (5.0, -1.14) | P | (5.5, -0.09) |
| F | (2.5, 1.35) | Q | (6.0, -1.10) |
| G | (0.5, -0.80) | S | (3.0, -0.97) |
| H | (6.0, 0.26) | T | (5.0, 1.13) |
| I | (5.5, 1.81) | V | (5.0, 1.13) |
| K | (7.0, -0.41) | W | (7.0, 1.71) |
| L | (5.5, 1.14) | Y | (7.0, 1.11) |

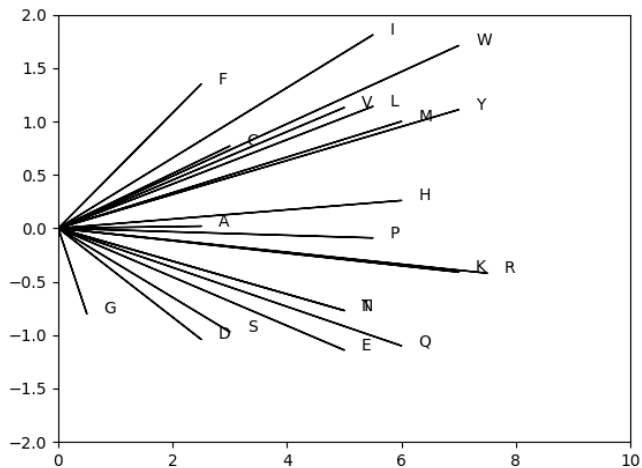


Figure 1: アミノ酸ベクタ①

モデル①

- 全結合層
 - 活性化関数 Relu
 - He の初期値
 - l2 正則化
- adam(learning_rate=0.001)
- batch_size = 128
- epochs = 20

Test accuracy: 0.9567307829856873

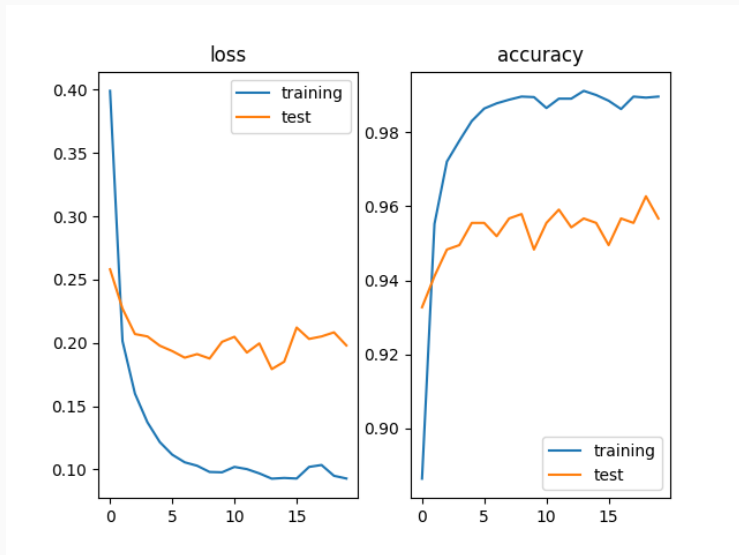


Figure 2: モデル①の学習結果①

Test accuracy: 0.9543269276618958

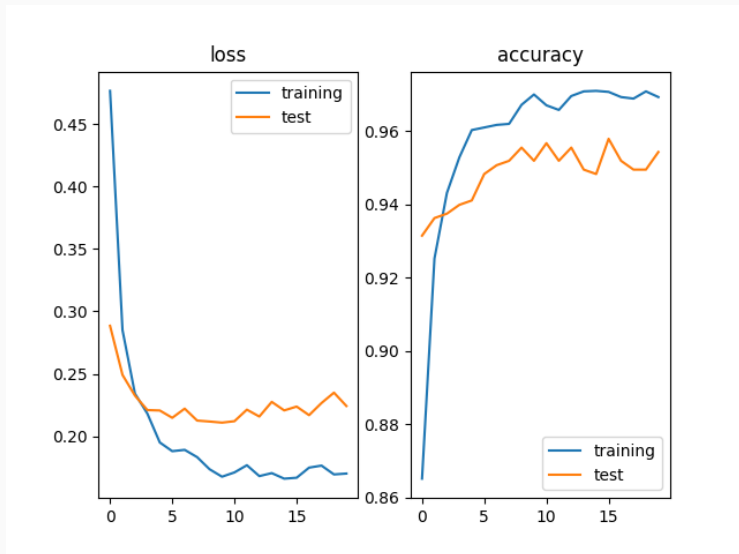


Figure 3: モデル①の学習結果

Table 2: 各クラスにおける正解率

| class | accuracy |
|-------|--------------------|
| A | 0.9718309640884399 |
| B | 0.7884615659713745 |
| C | 0.9567307829856873 |
| D | 0.0 |
| E | 1.0 |

アミノ酸ベクタの割り当て②

Table 3: アミノ酸の疎水性度

| アミノ酸 | 疎水性度 | アミノ酸 | 疎水性度 |
|-------|-------|-------|-------|
| Arg/R | -5.10 | Gly/G | -0.64 |
| Lys/K | -4.11 | Ser/S | -0.50 |
| Gln/Q | -3.68 | Trp/W | -0.46 |
| Asp/D | -3.60 | Ala/A | 1.10 |
| Asn/N | -3.50 | Met/M | 1.90 |
| His/H | -3.20 | Cys/C | 2.50 |
| Glu/E | -3.20 | Phe/F | 2.80 |
| Pro/P | -1.90 | Leu/L | 3.80 |
| Tyr/Y | -1.30 | Val/V | 4.20 |
| Thr/T | 0.70 | Ile/I | 4.50 |

アミノ酸ベクタの割り当て②

疎水精度を h とする。任意のアミノ酸 c におけるベクトルを以下に示す。

$$\begin{cases} arg = 180 - 170 \times \frac{h_c + |h_{min}|}{|h_{max}| + |h_{min}|} \\ x = \cos(arg) \\ y = \sin(arg) \end{cases}$$

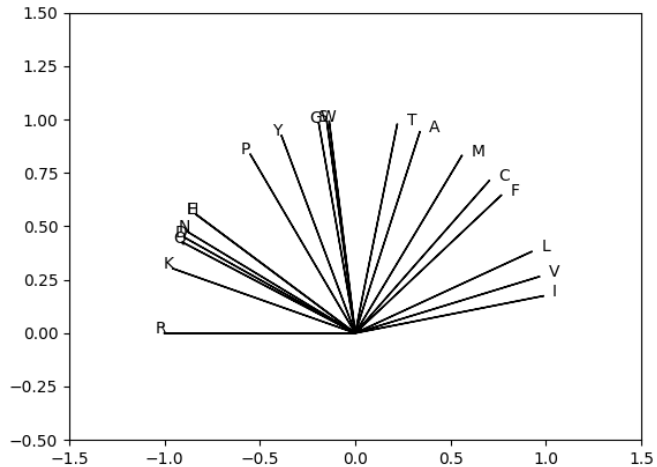


Figure 4: アミノ酸ベクタ②

Test accuracy: 0.9435096383094788

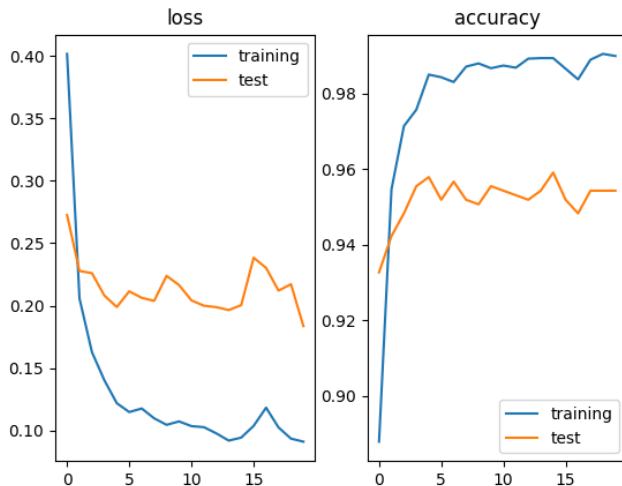


Figure 5: モデル①の学習結果②

Test accuracy: 0.9543269276618958

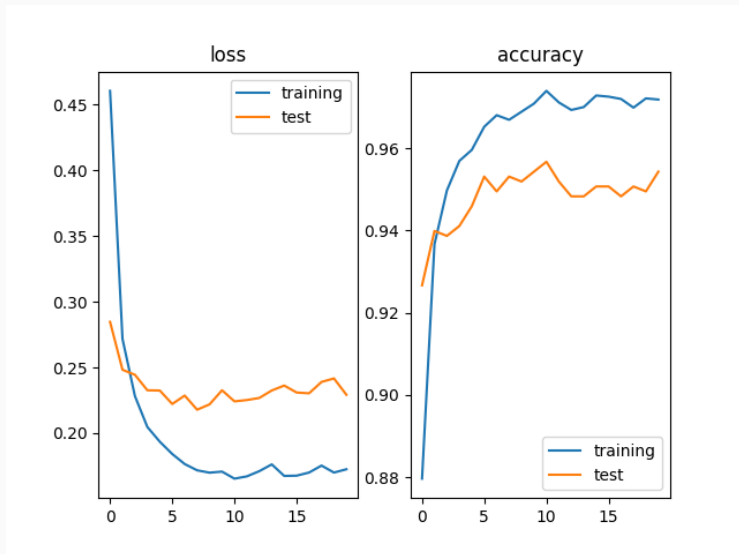


Figure 6: モデル①の学習結果②

Table 4: 各クラスにおける正解率

| class | accuracy |
|-------|-------------------|
| A | 0.966549277305603 |
| B | 0.807692289352417 |
| C | 0.961538434028625 |
| D | 0.0 |
| E | 1.0 |

- クラス D とクラス E の正答率が小さかった
- 単純なニューラルネットワークを試したので、より複雑なニューラルネットワークを使う
- COG のデータベースで試す