

5.3 訓練データとテストデータ

- 訓練データはネットワークの学習に用いるが、テストデータは学習結果の検証に用いる。
- 訓練データとテストデータは、多数の入力と正解の組み合わせにより構成されている。
- 入力と正解のペアをサンプルと呼ぶことにする。
- 通常は、訓練データのサンプル数をテストデータのサンプル数より多くする。
- 訓練データで学習したネットワークがテストデータでも良い結果を出せれば、ネットワークが未知のデータにも対応できていることになる。
- テストデータで良い結果が出なければ、ネットワーク自体や学習方法に何らかの問題があることになる。

- 回帰問題において、正解は例えば以下のような数値が並んだベクトルで表すことができる：

[0.54 -0.34 1.05 0.21 -0.84]

- この場合、出力層には5個のニューロンがあり、それぞれの出力が上記の正解値に近づくようにネットワークが学習することになる。

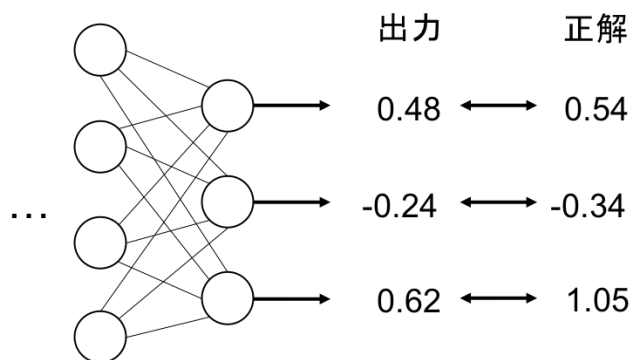
- 分類問題における正解値は、例えば以下のような1つだけ1で残りが0のベクトルで表される：

[0 1 0 0 0]

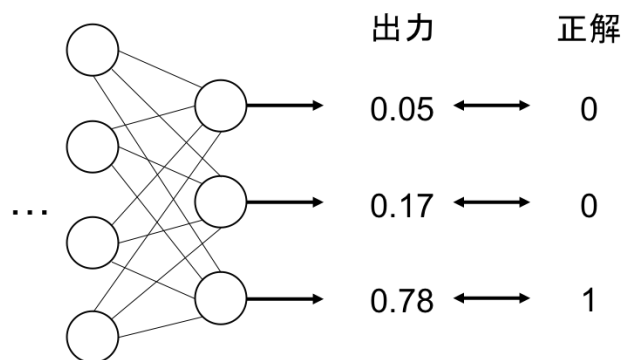
- この場合、出力層に5個のニューロンが並んでおり、それぞれの出力が0もしくは1に近づくようにネットワークが学習することになる。

- 次の図は、回帰と分類それぞれにおける出力と正解のイメージ：

回帰



分類



□出力と正解

- 回帰、分類いずれのケースでも、すべての訓練データで出力が正解に近づくように、ニューラルネットワークは学習していく。