

確認問題 (1)

1. 数値データをカテゴリデータに変換する必要があるのはどのような場合か考察せよ。また、このような変換の具体例を示せ。
2. カテゴリデータをニューラルネットワークの入力とする際に必要な処理について考えよ。
(ヒント：都道府県の情報はどう扱えばよいか)

確認問題（１） 解答例

1. 別のデータと突き合わせることによって、個人が特定できてしまう場合がある。そのような場合は、数値をぼやかしてカテゴリで表現することによって、そのデータが表す個人を特定できないようにする。具体的には、数値データの値を特定の範囲で離散化する。たとえば年齢を表す数値データを 20 代、30 代 ... のようにカテゴリデータに変換する。

確認問題（１） 解答例

2. ニューラルネットワークは数値データを入力とすることが前提の学習アルゴリズムである。このような場合、0 または 1 の二値をとるダミー変数をカテゴリーの数だけ用意する。たとえば都道府県名を表すカテゴリデータは、47次元の one-hot ベクトル (1 つの次元の値だけが 1 で、残りは 0) となる。

確認問題 (2)

1. 決定木の学習に関する以下の記述の空欄 (A) ～ (J) を埋めよ。

事例番号	規模	収益	成長性	株価
1	小	普通	高	上昇
2	大	少ない	低	下降
3	大	少ない	高	下降
4	大	多い	高	上昇
5	小	少ない	高	下降
6	大	多い	普通	上昇
7	小	普通	普通	下降
8	大	普通	普通	下降

確認問題 (2) 解答例

A 0.95

B 0.92

C 0.97

D 0.95

E 0.34

F 収益

G 普通

H 成長性

I 低

J 下降

