

実践！機械学習のプロセス  
徹底解説

ステップ6:  
機械学習モデルの交差検証

# 決定木とは？

## <特徴>

- ・ シンプルな構造で、推測結果を説明しやすい
- ・ 機械学習でもよく使われる手法

# 決定木とは？

表1：旅客等級順で並べたデータ

| 乗客乗員ID | 性別 | 年齢 | 旅客等級 | 生存 |
|--------|----|----|------|----|
| 2      | 女性 | 55 | 1    | 1  |
| 4      | 男性 | 20 | 1    | 1  |
| 8      | 女性 | 41 | 1    | 1  |
| 10     | 男性 | 32 | 1    | 1  |
| 1      | 女性 | 29 | 2    | 1  |
| 3      | 男性 | 46 | 2    | 0  |
| 7      | 女性 | 61 | 2    | 1  |
| 5      | 男性 | 15 | 3    | 1  |
| 6      | 女性 | 18 | 3    | 0  |
| 9      | 男性 | 22 | 3    | 0  |

旅客等級で生存率が分類できる

# 決定木とは？

表2：旅客等級[2]を抽出したテーブル

| 乗客乗員ID | 性別 | 年齢 | 旅客等級 | 生存 |
|--------|----|----|------|----|
| 1      | 女性 | 29 | 2    | 1  |
| 7      | 女性 | 61 | 2    | 1  |
| 3      | 男性 | 46 | 2    | 0  |

性別で生存率が分類  
できる

# 決定木とは？

表3：旅客等級[3]を抽出したテーブル

| 乗客乗員ID | 性別 | 年齢 | 旅客等級 | 生存 |
|--------|----|----|------|----|
| 5      | 男性 | 15 | 3    | 1  |
| 6      | 女性 | 18 | 3    | 0  |
| 9      | 男性 | 22 | 3    | 0  |

年齢で生存率が分類  
できる

# 決定木とは？

- データの特徴量を解析し、ターゲットに対する結論を導きだす

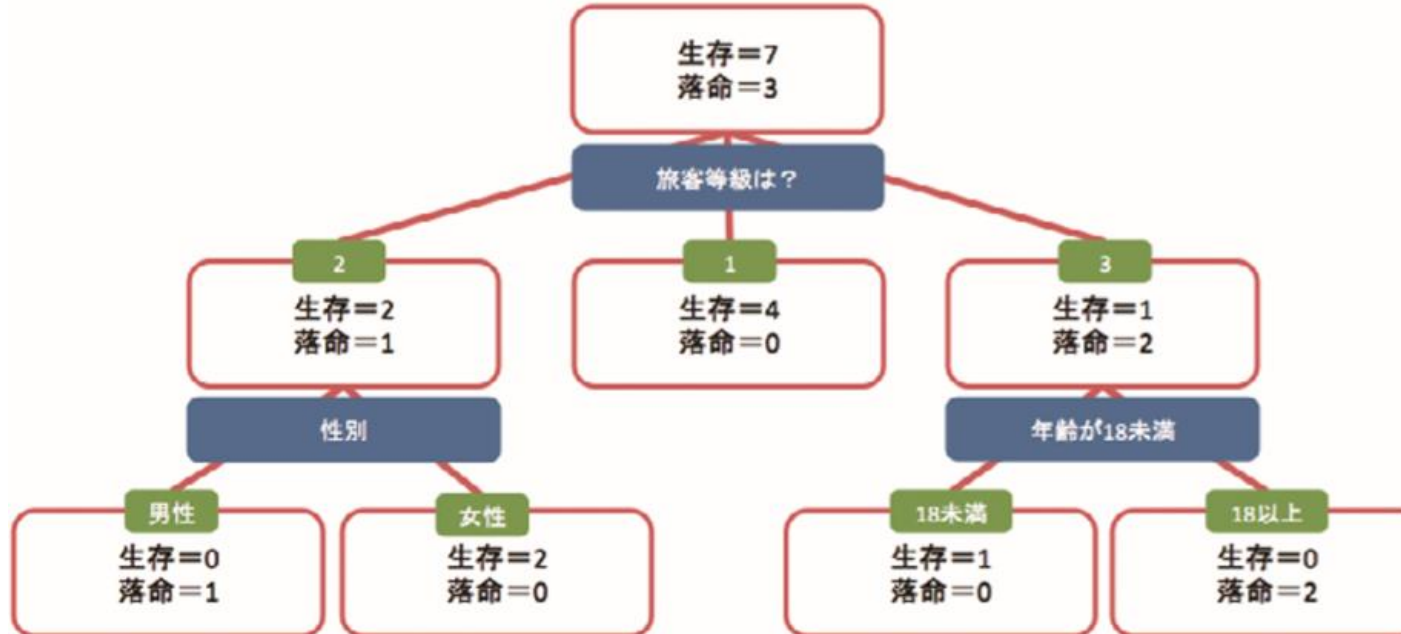


図1. 簡易データを決定木として表現