

I. 英語の問題:英語には対応する日本語を、日本語には対応する英語を書け

- | | |
|-------------------------|---------------------|
| (1) 訓練データ | (2) 検証データ |
| (3) テストデータ | (4) サポートベクターマシン |
| (5) logistic regression | (6) overfitting |
| (6) estimator | (7) predictor |
| (8) transformer | (9) nominal feature |
| (10) ordinal feature | |

II. 3章で学んだことと4章の予習の成果の検証

以下のことを Jupyter notebook を用いて行い、コメントをつけて提出する

- (1) Python 機械学習プログラミングの p.116 に従い、wine データを取り込む
- (2) 特徴量として Proline と Flavanoids を選び、ラベルとして Wine のクラス(Class)を選ぶ。
訓練データとテストデータにわけよ。割合は 8:2 とする
- (3) 訓練データを標準化せよ。それを以降の訓練データとする。この標準化に従ってテストデータも標準化すること。
- (4) logistic regression モデルの訓練を行え(正則化なし)。p.70 にしたがって、決定領域、訓練データサンプル、テストサンプルを表示せよ。ただしパラメータは適切に定めよ。
- (5) L1 正則化を用いて、(4)と同じことを行え。
- (6) L2 正則化を用いて(4)と同じことを行え。
- (7) p.79 にしたがって、サポートベクターマシンで訓練を行え。その結果の決定領域、訓練データサンプル、テストサンプルを表示せよ。ただしパラメータは適切に定めよ。
- (8) Wine の標準化していない訓練データを用い、最大深さ=4、エントロピー条件を用いた決定木モデルを学習させよ。またこの結果の決定領域、訓練データサンプル、テストサンプルを表示せよ。