

5.5

以下の命題の真偽を答えよ（真か偽で答える）。偽の場合、その根拠を簡潔に示せ。

1. 訓練誤差を評価する訓練サンプルとテスト誤差を評価するテストサンプルは同じ分布からサンプリングしなければならない

回答：真

2. k-fold 交差検証のための事例を k 分割した時に、一度モデルの訓練に用いた訓練事例を、別のモデルの訓練のためならば再び訓練事例として用いてはならない

回答：偽

理由：交差検証を使ったモデル選択では、一度モデルの訓練に用いた訓練事例を別のモデルの訓練のために訓練事例として用いるため。

3. サポートベクターマシンにおいて、精度と再現率のバランスは、モデルパラメータの符号をコントロールすることで調整可能である

回答：偽

理由：線形分類器の場合、正解率を最大にする分類器を表す超平面と同じ法線ベクトルを持つ平面について、「正解率を最大にする分類器からの距離」をコントロールすることで、精度と再現率のバランスをコントロールすることができるため。

4. LASSO による特徴選択は訓練誤差を小さくする効果を持つ

回答：真

5. 正則化パスによって最適な正則化パラメータを得ることはできない

回答：真

6. 一度モデルの訓練に用いた訓練事例は、テスト誤差の評価のためにテスト事例として用いてはならない

回答：真

7.2 ノルムの単位円は 1 ノルムの単位円を包含する

回答：真

8. 重回帰は常に解析解を持つ

回答：偽

理由：線形回帰モデルが凸関数で微分可能であれば解析的に解を求めることができるが、そうではない場合には厳密解は求められず、最急降下法などを使って近似解を求めることになるため。

9. 確率的勾配降下法では、目的関数が凸関数であるならば、目的関数値は更新毎に単調減少する

回答：偽

理由：目的関数が凸関数であっても、更新ごとに必ず勾配方向に向かうとは限らないため。

10. L1 正則化項+二乗損失からなる目的関数の劣勾配は、一意に定まらない場合がある

回答：真

11. リッジ回帰の正則化パラメータはマージンを最大化することによってチューニングする

回答：偽

理由：リッジ回帰の正則化パラメータは、交差検定を用いてチューニングをするため。

12. 線形分類器はソフトマージンが最大化されたからといって、訓練誤差がゼロになるとは限らない

回答：真

13. 目的関数に正則化項を持つ線形回帰では、L1 正則化を用いた回帰 (Lasso) であろうが L2 正則化を用いた回帰 (Ridge 回帰) であろうが、解を解析的に求めることはできない。

回答：偽

理由：Lasso では解析解を求めることができないが、Ridge 回帰では解を解析的に求めることができる場合があるため。

14. 多項式回帰におけるモデルの複雑さは、L1 正則化であっても L2 正則化であっても、コントロールすることは可能である

回答：偽

理由：L2 正則化ではモデルの複雑さ(パラメータの絶対値)を制御することができるが、L1 正則化は特徴選択をするが、モデルの複雑さをコントロールすることはできないため(特徴選択をすると、モデルの解釈性を向上させるという点では、複雑さをコントロールしていると言える)。

15. サポートベクターの数は常に少ないとは限らない

回答：真