

入賞候補者 モデリング詳細

モデルの再現性及び内容について確認するため以下にご回答下さい。ご協力のほどよろしくお願いいたします。

***必須**

1. メールアドレス *

2. SIGNATEでの登録ユーザー名 *

【OSのバージョン、使用ソフトウェア等】

3. OS *

例) CentOS Linux release 7.4.1708 (core)

4. 言語 *

例) Python 3.6.3

5. ライブラリ *

例) numpy 1.13.3

6. GPU

GPUを使った場合は、GPUについて説明ください。例) NVIDIA Tesla K80

【データ及びモデリング内容】

※ご提出物内容の理解の為に大変重要となりますので、詳細な記述にご協力下さい。

7. 今回のデータや問題の特徴について *

どんな特徴のあるデータだったか？ どういうところが難しい問題だったか？ 何を精度向上のポイントとなると仮定したか？ その上でどんなアプローチをとったのか？

8. 予測ロジック *

例) 1. トレンド考慮の為に、時系列モデルを利用し予測 2. 1で求めた値を特徴量に加え、Stacking (RandomForest、Xgboost) を用いて最終的な予測値を求める

9. 採用したモデリング手法 *

例) Stacking (RandomForest + Xgboost)

10. 特徴量の工夫点 *

モデルに入力する特徴量について、工夫した点等を記述して下さい。

11. モデリングの工夫点 *

ご提出いただいたモデルの特徴や工夫点、優れていると考えられる点も合わせて記述下さい。

12. 分析・モデリングから得られた示唆 *

なぜ高い予測精度が出せたかやどうすれば更に予測精度が上がると考えられるか等。もし他にも試されたが上手くいかなかった手法や苦戦したことなどありましたら合わせて記述下さい。

【提出ソースコード】

13. ソースコード概要 *

データ加工・前処理から予測出力までのソースコードの流れを示してください。関数定義等されていたら、名前も記述頂きながら、何を実行されているのかも合わせて記述下さい。

14. スコアを再現する為の手順を示してください *

リーダーボード上の最終評価のスコアを再現するための手順を示してください

15. 学習部分と予測部分とを独立して実行する手順 *

16. 学習部分のおおよその実行時間 *

17. 予測部分のおおよその実行時間 *

18. ハイパーパラメータについて *

ハイパーパラメータとしてチューニングした変数名や、チューニング方法等を記述下さい。例)
RandomForestのn_estimatorとmax_depthをチューニング変数とし、Gridsearchにてn_estimator:
[100,200]、max_depth:[5,10]でcv=3で探索し、最も精度の良かったn_estimator=100, max_depth=5を
採用した。他の変数についてはデフォルト値をそのまま採用した。

19. 乱数の取り扱い方法（シード値等）についてご説明ください *

シード値が必要な関数や設定したその具体的なシード値等を記述下さい。

20. ルールを遵守できていますか？ *

※コンペによって順守いただくルールは異なります

1つだけマークしてください。

☐ はい

☐ その他: _____

21. 再現性の検証にあたり、事務局での経済的負担や、権利侵害がないことを保証いただけますか？ *

1つだけマークしてください。

☐ はい

☐ その他: _____

【その他アンケート】

ご協力をよろしくお願いいたします。

22. 今回のコンペにおける実質の作業時間はおよそどれくらいでしたか？ *

計算時間を除いた実際の作業時間を数値のみで記載下さい 例) 5時間半⇒5.5

23. SIGNATEをどこでお知りになりましたか？ *

web検索などでお知りになった方はその他にて何を調べていた時かや覚えておりましたらその検索クエリ、記事等のリンクからの場合はその記事名を教えてください。

24. 今回のコンペへの参加理由を教えてください *

- 25。 今回のコンペの感想・ご意見、SIGNATEに期待することを教えてください *
些細なことでも構いません

このコンテンツは Google が作成または承認したものではありません。

Google フォーム