

PythonCafe(8回目) 最終課題と最終レポート

2019.10.18 Tadashi Mori

1. PythonCafeに参加した目的やゴール

1. アルムナイ・リレーションシップという言葉が気に入って、MMCに貢献できそうなテーマを取り組んでみたかった。
2. インプットしたことをアプトプットすることで知識の定着と理解を深める。
3. 人に教えるスキルを磨く。

2. PythonCafeで学んだこと、理解できたこと

※間違っているかも知れない

- 次の言葉を理解した。（最低5つは答えてほしい）
 - AI/機械学習/ディープラーニングの違いは何か？
->
 - データサイエンティストには3つの能力が必要。それは何か？
->
 - 機械学習は3つに分けられる。それは何か？
->
 - 教師あり学習と教師なし学習の特徴は何か？それぞれどのようなビジネスシーンで活用できるか？（業務ニーズ・課題に対応できるか？）
->
 - データ分析プロセスを何と呼んでいるか？（IBM用語なので一般的ではないが、考え方は同じプロセス）
->
 - 教師あり学習の回帰と分類で使われるアルゴリズムをそれぞれ1つ以上答えよ
-> 回帰：
-> 分類：
 - 過学習とは何か？過学習を防ぐ方法は何か？
->
 - 回帰と分類で使われる評価指標をそれぞれ1つ以上答えよ
-> 回帰：
-> 分類：
- 次のモデル（プログラム）を説明する。（どれか一つ選択）
 - アンケート結果から統計学アプローチにて改善点を考察する
 - ボストン物件価格データから回帰モデルを構築したので説明する
 - アヤメデータから分類モデルを構築したので説明する
 - MNISTデータから分類モデルを構築したので説明する
 - オープンデータを使って回帰モデルを構築したので説明する

3. データサイエンティストスキルチェック結果

※見せなくても良い

-R ? A"

```
from [4] XLSR←HMTPE] 略import 略QEKI
-QEKI 略7GM I R略LSX←TRK_
```

3YX? A"



4. 感想（今後の展望など）

- 相手のレベルに沿ってや目的に応じて、カリキュラムを設定し、話をする事の難しさを知ることができた。
- 相手の質問の対応力の難しさや自分の知識の不十分さを思い知った。
- 相手にアウトプットすることで、言葉や理論、Pythonの使い方を説明できる能力が高まったと思える。上記スキル評価結果を見ても、サイエンス力やエンジニア力が上がっていることが確かめられた。
- 僕に機会を与えてくれた深谷さんやマネージャーの方、この活動に付いて来てくれた田中さん、楠元さん、影山さん、ITサポートしてくれた古河さんに感謝します。
- 次のステップとしては、故障診断データ（XML）からデータの可視化をしたり（ある値の傾向や特徴を確認する）、機械学習モデルを構築してエラーコードの異常検知やドメイン（地域や車種ごとなど）ごとでデータをクラスタリングして見るのにチャレンジして見るのも面白いかと思います。

END

-R ? A"

APPENDIXS

Markdown記法 サンプル集

<https://qiita.com/tbpg/items/989c6badeff69377da7> (<https://qiita.com/tbpg/items/989c6badeff69377da7>)

図の挿入方法

```
from IPython.display import Image
Image('ファイル名')
```

モデル（プログラム）の説明の仕方

- ①どんなタスク（分類or回帰）を行うモデルなのか？そのためにどんなデータを使ったのか？
- ②どんなライブラリが必要なのか？
- ③どうやってデータを作ったか？説明変数（特徴量・パラメータ）と目的変数（正解）はどれか？
- ④どんなアルゴリズムを使ったか？
- ⑤どうやって評価したか？
- ⑥結果（モデル）の考察

```

F"4(*スへほハくwは捕甫 JF" FV"
← *MPI_(S[ RPSEH W4(*THU FV"
  ♀ THJ糞鴨 s yは寒居醒拇 j 肥融えち b FV"
⇒ JMPI_TMRX

```

-R ? A"