

J-Quants ファンダメンタルズ分析チャレンジ！ 目標15分で提出できる サブミッションチュートリアル

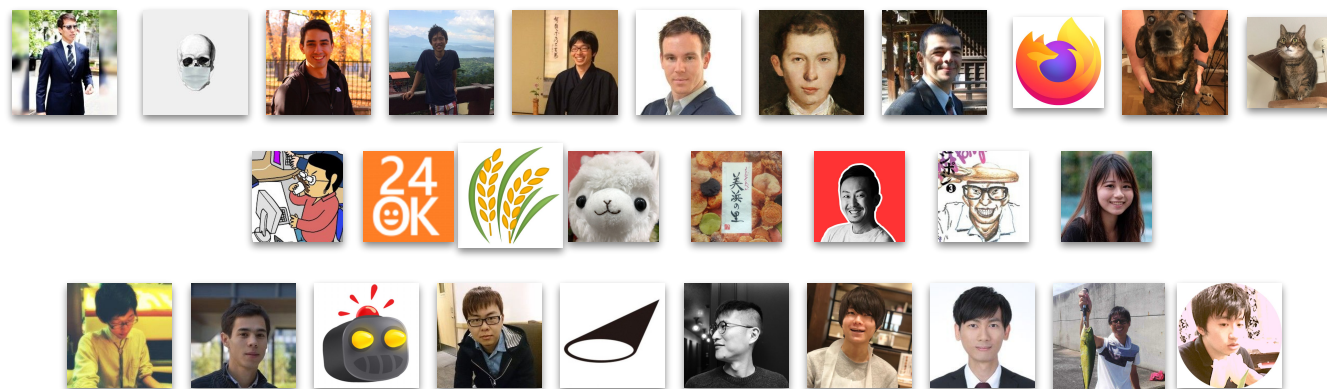
2021.03.10 // AlpacaJapan



Alpacaはデータサイエンス・機械学習を金融マーケットの様々なリアルな問題に適用することに強みを持つ会社です。

5人 → 30人

サイエンティスト/エンジニアの3年の増加数



Deutsche Bank



TWO SIGMA

Client's Testimonial from Quants

優れたテクノロジーは疑う余地が無い。
ただ、それだけではパートナーとして不十分。
成功までの道のりはトライアンドエラーの連続。
市場の謎解きへの純粋なモチベーション、あきらめない姿勢、そして結果への強いこだわり。
Alpacaを語る時にそうした要素は欠かせない。
彼らは真に信頼できる集団だ。

Client's Testimonial from Traders

“**金融市場の未来はAIと人とのコラボレーションにこそある**”というのはまさに私も強く共感するコンセプトです。世の先をいく貴社の皆様にはどんどん未来へ向かって進んでいって頂きたい。

コロナ禍の中で金融業界全体で今まで出来ないと思っていたことが、**実際にやってみると出来てしまうことが沢山出てきている**と感じています。その中で AIと人間のコラボレーションも益々増えていくでしょうし、そういう部分が **各企業の競争源泉**になっていくのだろーうと思います。

チュートリアルを紹介

J-QuantsではAlpacaはチュートリアル作成支援と問題作成支援を担当させていただきました。
本チュートリアルはAlpacaの機械学習をマーケットに適用するための様々な基礎知識をもりこみしましたので、お楽しみいただけたかと思います。

チュートリアルの2章の構成と本日説明したいこと

問題説明 & データセットの説明

データセットの可視化

特徴量の設計

モデルの構築・評価

モデルの提出

この部分を充実させたため、結果的に提出するモデルの複雑度が上昇。

提出するモデルの難易度が上がり、シンプルに投稿方法を学ぶことが難しくなっていました。今回はその点を重点的にケアしたセッションになります。

 本日は**15分**で新たに作成したサブミッションサンプルのライブコードを行います。

ライブコードには Google Colaboratory(<https://colab.research.google.com/>) を利用します。難しそうな箇所、
嵌りそうな箇所には適宜説明をいれていきます。

前提条件

以下の事項を前提条件として記載されています。

- ・ 「ファンダメンタルズ分析チャレンジ」にご参加いただいていること
- ・ Googleアカウントを保有されていること
- ・ Google Colaboratory を使用できること
- ・ インターネットに接続されていること
- ・ Pythonについて基礎的な知識をお持ちであること
- ・ Pandasについて基礎的な知識をお持ちであること

データの準備

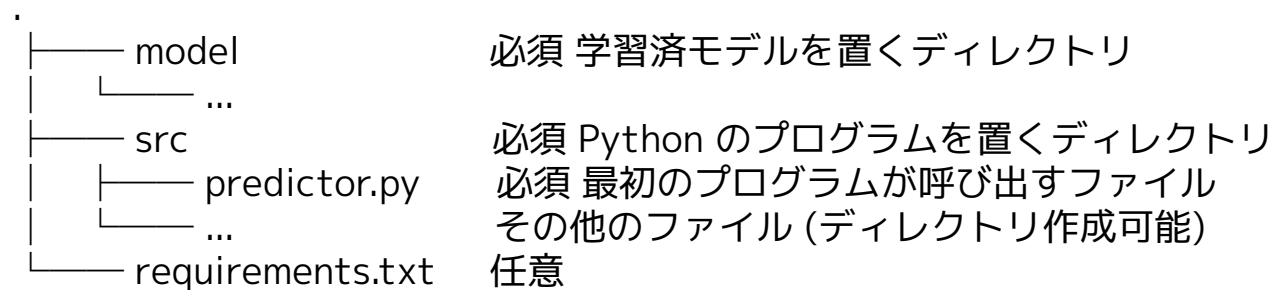
1. コンペティションページのデータタブより以下のデータをダウンロード
データタブ: <https://signate.jp/competitions/423/data>
stock_price (stock_price.csv.gz) をダウンロード
2. Google Drive にデータを配置します: Google Drive: <https://drive.google.com/>
My Drive 配下に JPX_competition フォルダを作成します。
JPX_competition フォルダにダウンロードした stock_price.csv.gz ファイルを配置します。

上まで準備できたら Google Colaboratory(<https://colab.research.google.com/>) を起動しましょう！

ここまででサブミッションのコードの準備が整いました。それでは、最後に投稿手前までパッケージをしましょう。

パッケージの構造

<https://signate.jp/features/runtime/detail> に記載がある通り、以下のディレクトリ構造である必要があります



作業手順

1. 先程のScoringServiceクラスをエディタなどでpredictor.pyファイルとして保存します。
2. modelディレクトリをzipファイルに含めるためにダミーファイルを作成します。
`$ touch model/dummy.txt`

コマンドラインに不慣れな方はエディタで空ファイルをmodel内に作成してください。

3. zipで圧縮します。
`$ ls`
`model src`
`$ zip -v submit.zip src/predictor.py model/dummy.txt`

コマンドラインに不慣れな方はOSのzip圧縮機能/zip圧縮ツールなどを利用し、srcとmodelディレクトリを選択した状態でzip圧縮機能を起動し、zipファイルをご作成ください。




Alpaca

STRICTLY CONFIDENTIAL