

Kaggle の使い方

1 概要

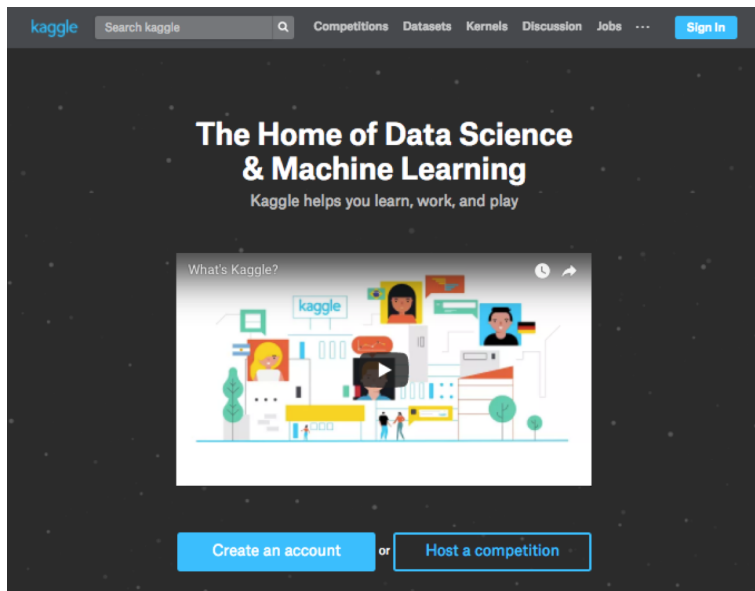
演習の課題では、自分の興味のある内容について、畳み込みネットワークあるいは LSTM を使って分析します。そのためのデータセットや分析手法を学習する際の参考として、今回は Kaggle を紹介します。

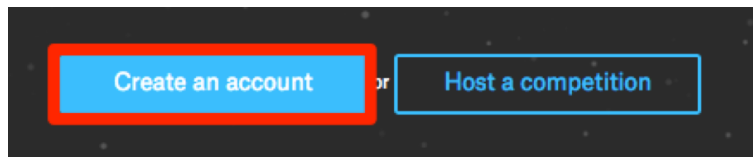
Kaggle は、世界最大のデータサイエンティストのコミュニティです。機械学習等のための多くのデータセットが公開され、また、それを分析する環境が無料で提供されています。データセットの分析に関するコンペティションも開催されています。

2 登録

Kaggle サイトにアクセスします。

<https://www.kaggle.com/>





Create an account をクリック

facebook Google Yahoo のアカウントを持っている場合にはそれぞれの画像をクリック
それ以外の場合には、「manually create an account」をクリック

「manually create an account」を押した場合

Create an Account With Your Email Address

ユーザ名 Username Used for login and your profile URL
Your profile URL will be kaggle.com/

表示名 Display Name John Doe
Shown on your public profile, leaderboards, etc.
Full name recommended.

メールアドレス Email Address Email Address

メールアドレス (確認) Confirm Email Re-enter Email Address

パスワード Password Minimum 7 Chars

パスワード (確認) Confirm Password Re-enter Password

☐ Email me news and updates

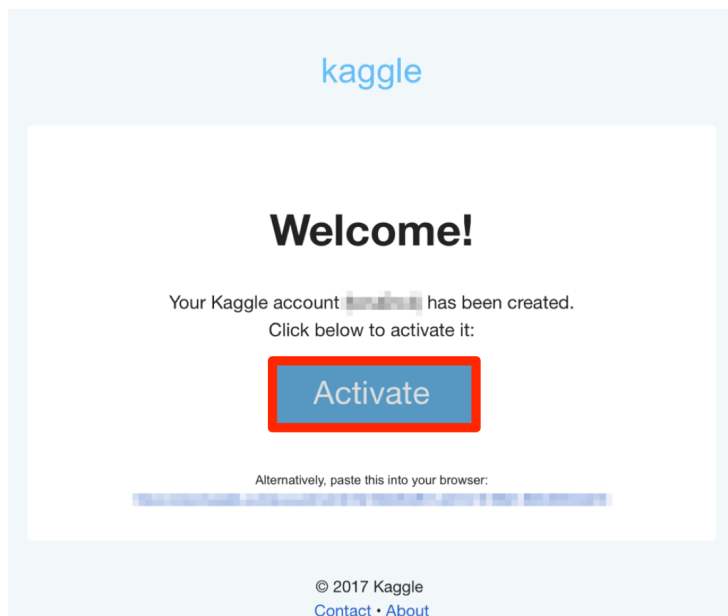
☐ By clicking this checkbox, you confirm that you accept our [Terms of Service](#) and [Privacy Policy](#).

利用規約とプライバシーポリシーに同意する場合にチェック

ニュースレターやアップデートお知らせを受信する場合にチェック

Get Started

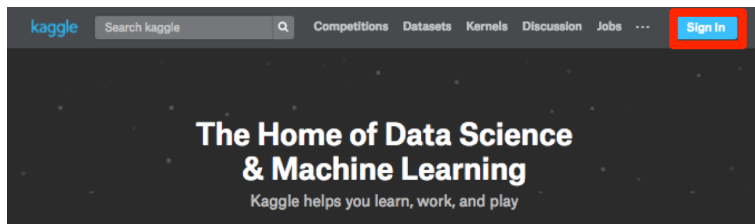
ユーザ名・表示名・パスワード等を入力
利用規約とプライバシーポリシーに同意する場合にチェック
「Get Started」をクリック



登録したメールアドレスに確認のメールが届くので、「activate」をクリックすると登録は完了です。

3 ログイン

Kaggle サイトにアクセスします。
<https://www.kaggle.com/>



「sign in」を押します。

A screenshot of the Kaggle sign-in/sign-up form. The form is titled 'Sign in or sign up with one click:' and includes a privacy notice: 'We won't share anything without your permission.' Below this are three social media login buttons: Facebook, Google, and Yahoo. An 'or' separator follows. The next section is titled 'Use your Kaggle username or email:' and includes a link to 'Manually create a new account'. There are two input fields: 'Username or Email' and 'Password'. A blue 'Sign in' button is below the fields. At the bottom, there is a checkbox for 'Remember me' and a link for 'Forgot Username / Password'. A footer section contains two paragraphs of text: 'One account per individual. e.g., If you're joining as a company, please create one account for each participant.' and 'Have an account but need a university address for InClass? Change your email account on your profile.'

facebook Google Yahoo のアカウントを持っている場合にはそれぞれの画像をクリック

それ以外の場合には、登録したメールアドレスとパスワードを入力し「Sign in」を押します。

4 カーネルを使う

サンプルのカーネルにアクセスします。

<https://www.kaggle.com/kenichinakatani/lstm-stock-prediction/>



`https://www.kaggle.com/kenichinakatani/lstm-stock-prediction/`

4.1 フォークする



自分で編集するためにフォークします。

4.2 プログラムの実行



- ①本文をクリックし
- ②実行ボタンを押して実行します。
- ③実行結果が下に表示されます。

このカーネルは LSTM を使って株価を予測するプログラムの解説です。カーネルの内容を順番に実行して、内容を確認しましょう。

4.3 ノートの追加

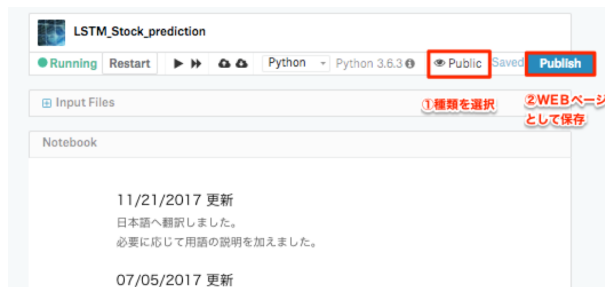


+↓ボタンで現在のノートの下に追加します。(+↑で上に追加します。)



文章を書くのか、プログラムを書くのかを選択します。

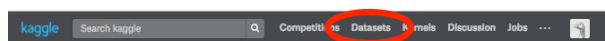
4.4 Publish



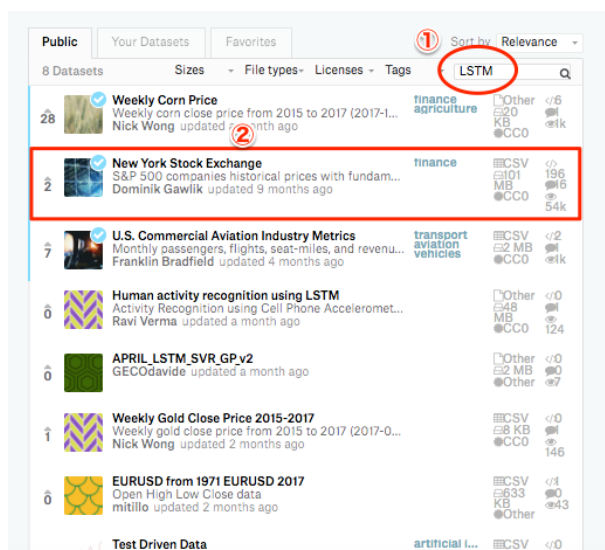
① Public(公開)/Private (非公開)を選択します。

② 「Publish」を押すと、WEBページとして出力されます。

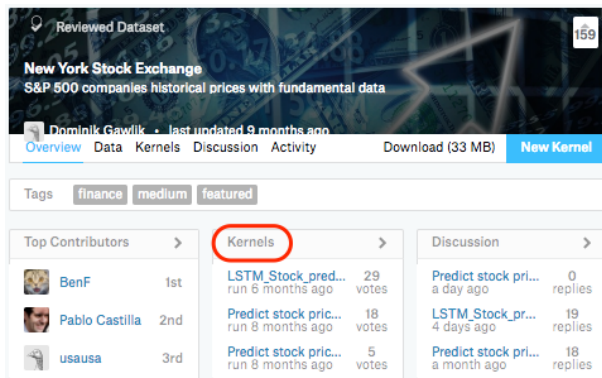
5 データセットを探す



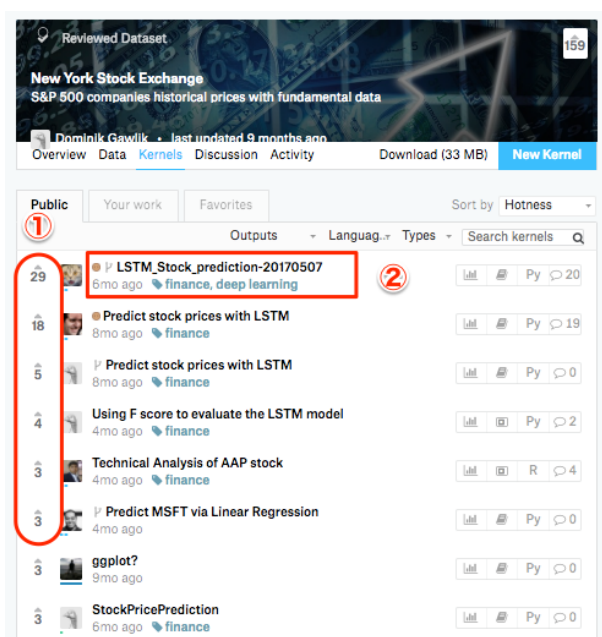
Datasets を押すとコンペティションのページに移動します。



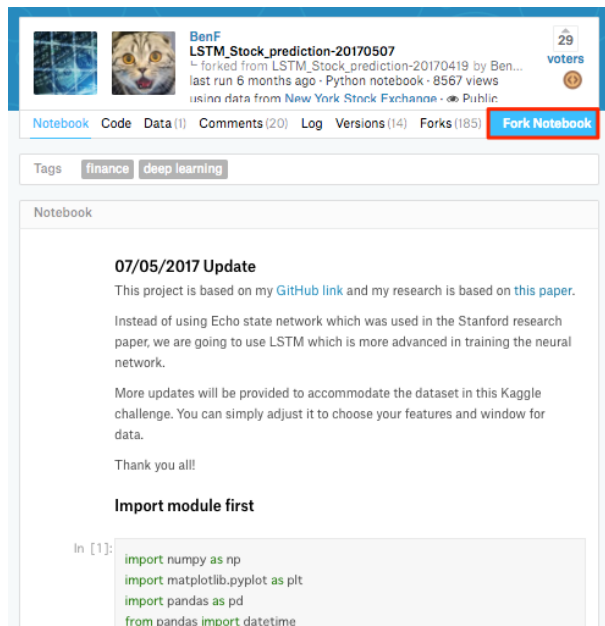
①検索ボックスにキーワードを入れると、関連するデータセットが表示されます。②興味のあるものをクリックしてみましょう。



「Kernels」をクリックすると、関連するカーネルが表示されます。



①興味のあるものや、評価の値などを参考に
②クリックして内容を確認してみましょう。



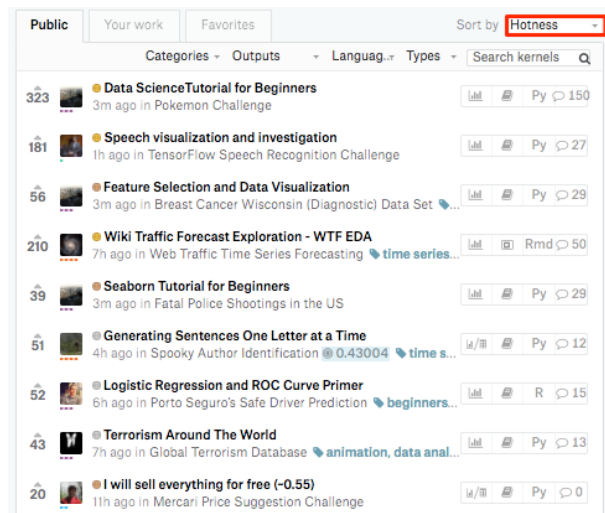
また、必要に応じてカーネルを
フォークして、実行してみまし
ょう。

6 カーネルを探す

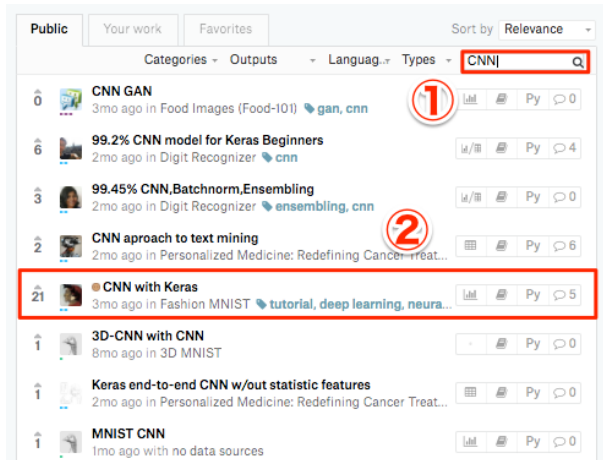
カーネルを直接検索することもできます。



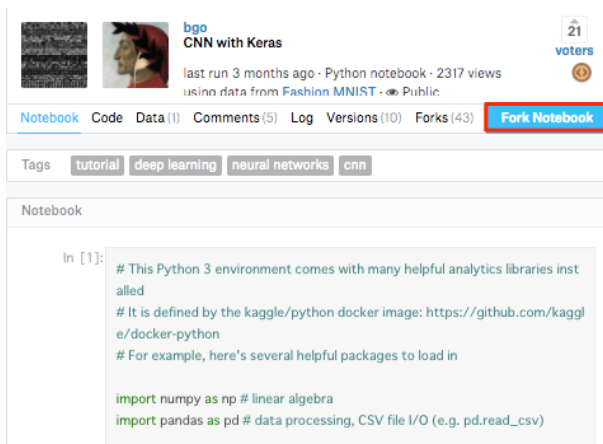
Kernels を押すとカーネルのページに移動します。



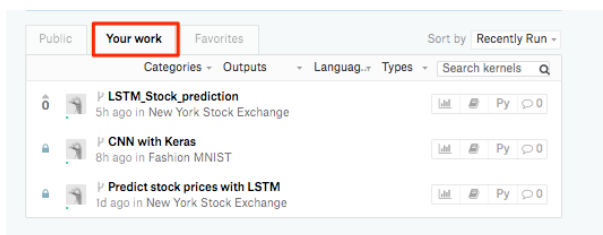
表示の際にソートする方法を設
定できます。



①キーワードを入力すると、関連するカーネルが表示されます。
②興味のあるものや、評価の値などを参考にクリックして内容を確認してみましょう



また、必要に応じてカーネルをフォークして、実行してみましょう。



自分がフォークしたカーネルの一覧は「Your work」タブをクリックすると確認・編集できます。

7 コンペティション（コンペティションへの参加は任意）



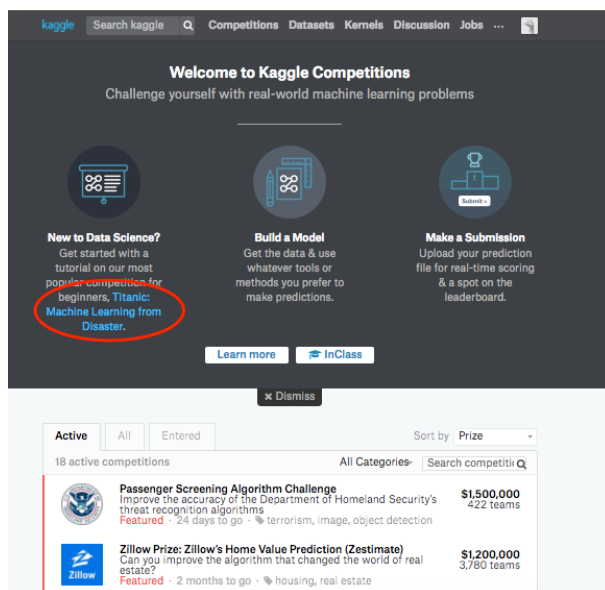
Competition を押すとコンペティションのページに移動します。開催されているコンペティションの一覧が表示されます。興味のあるコンペティションの内容を確認してみましょう。

7.1 機械学習初心者向けのコンペティション

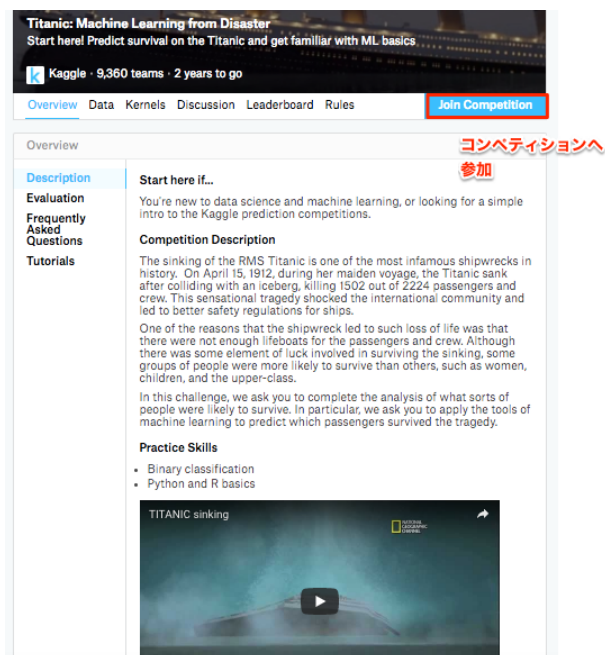
7.1.1 タイタニック

<https://www.kaggle.com/c/titanic>

コンペティションに興味のある方は、機械学習の初心者向けコンペティション「Titanic: Machine Learning from Disaster.」があります。どのような特性を持った人がタイタニック号の事故で生き残ったかを推定するモデルを作成するコンペティションです。



クリックして内容を確認



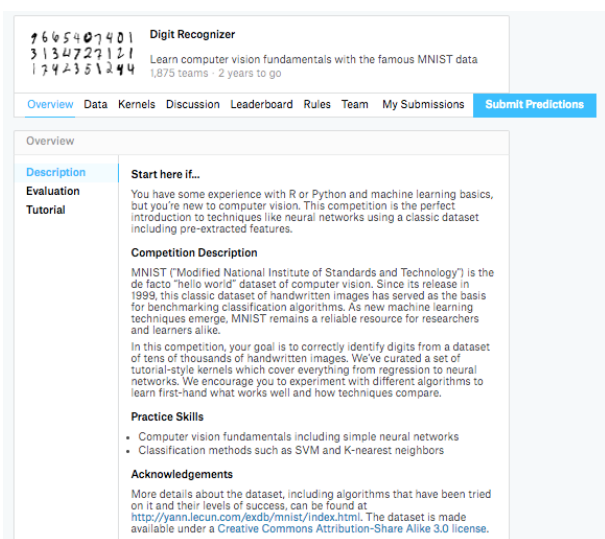
興味のある方は参加してみてください（任意です）。

このコンペティションで LSTM や畳み込みネットワークを利用する可能性は低いですが、コンペティションへの参加方法を確認するのによいと思います。また、他の LSTM や畳み込みネットワークを利用する可能性のあるコンペティションに参加することは、課題の参考になると思います。

7.1.2 Digit Recognizer

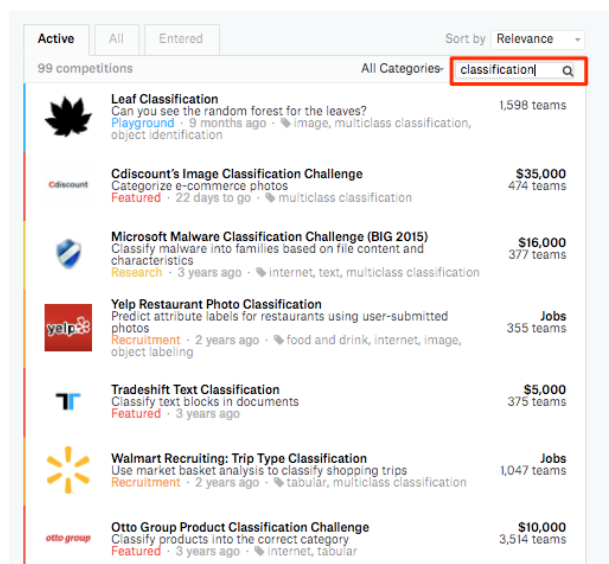
<https://www.kaggle.com/c/digit-recognizer>

手書きの 0 ～ 9 までの文字を認識のコンペティションです。機械学習の基本を学ぶことができます。

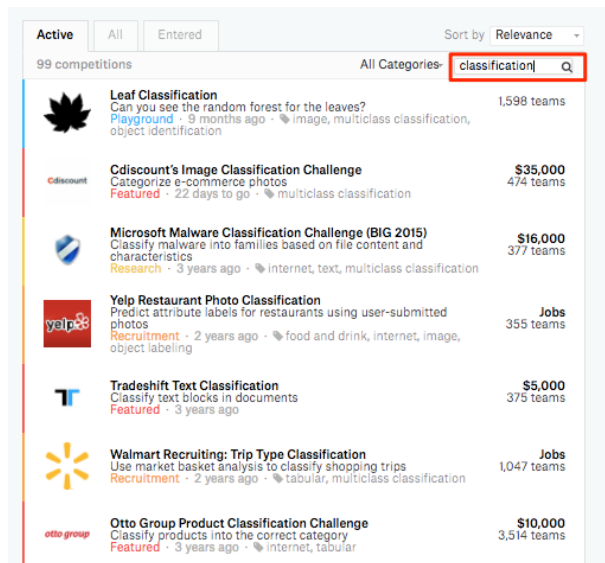


興味のある方は参加してみてください（任意です）。

7.2 コンペティションの検索方法



①キーワードを入力すると、関連するコンペティションが表示されます。（この例は「classification」）
②興味のあるものや、評価の値などを参考にクリックして内容を確認してみましょう



キーワード「time」で検索した例

Kaggle のコンペティションリスト (随時更新) <https://qiita.com/TomHortons/items/54b2121a4c5eb0908920> なども参考にしてみてください。

コンペティションの種類には、

Featured 賞金が高額で、スポンサーが提出した商業的な問題を解くもの

Masters 参加者が上位者や招待者など一定の者に限られているもの。賞金が高額であったり、取扱に注意を要するデータを扱う場合であったりする。

Recruiting スポンサー企業が採用活動の一環として行うもの

Kaggle Prospect 勝者をきめるのに、スコアボードをつかわず、コンペティションのゴールが予測モデルでなく、データの探索やビジュアル化等であるもの。ユーザの投票やコメント等を参考に審判する。

Research 目的が研究や社会に役立つものであるもの。モデルの解法をオープンソースとしてリリースすることが義務づけられている。

Playground 研究目的やビジネス目的でなく、楽しみ、独創性、アイデアに基づくもの。楽しみのために行うもので、賞金が目的でないもの。

Getting Started 学習をはじめたばかりの人が機械学習の感触をつかむために取り組むもの。賞金や期限がない。

があります。実際のソースコードを確認したい場合には、「Getting Started」や「Research」をキーワードにして検索してみてもよいでしょう。