

# 第一回レポート課題（回帰）

- Traffic volume prediction（交通量予測）
  - 学習データ数：37696（2013~2017年のデータ）
  - テストデータ数：7949（2018年のデータ）
  - 詳細は課題終了後解説（元データがばれるため…）
- 8種類の説明変数から、ある1時間における交通量を予測
- 説明変数
  - holiday: 祝日の種類（categorical）0時の行のみアノテーションされているので注意
  - temperature: 気温（real）
  - rain\_in\_hour: その1時間における降水量（real）
  - snow\_in\_hour: その1時間における降雪量（real）
  - clouds\_cover: 雲の割合(%)（real）
  - weather: 天気の種類（categorical）
  - weather\_detail: より詳細な天気の種類（categorical）
  - timestamp: 日付と時間
- 目的変数（9番目のカラム）
  - traffic\_volume: その1時間における交通量（real）

# コンペティション形式

- 評価指標

- Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} (y_i - \hat{y}_i)^2}$$

- ランキング

- スコアリングサーバ: <http://www.nlab.ci.i.u-tokyo.ac.jp/~nakayama/ds19/report1/index.php>

- ID: dslecture, PASS: tr@ffic

- 提出方法

- テストサンプルの推定結果を一行ずつ記載したテキストファイルを提出
- 名前に、ITC-LMSのログインIDを含めること（レポートとの紐づけを簡単にするため。複数のバージョンをつくるのはOK）
- 同じ名前で提出した結果は上書きされる。最後のものだけ保存されるので注意。
- 講義で扱った技術のみを使ってよいランキングと、何でもありのランキングに分けているので間違えないように
- 現在は、提出されたテストデータの推定結果のうち、所定の2000サンプルでスコアリングしている
- 最終的なスコアは、締め切り後に残りのサンプルも合わせて算出

# 課題詳細

- レポート内容
    - Traffic volume予測の問題に取り組み、以下の点を中心にA4用紙2~3枚程度にまとめよ。  
(※講義で扱っていない技術を用いても構わないが、その場合free-styleのランキングシステムへ提出すること。)
1. 予測性能を向上させるための自分なりの工夫点と結果、考察（必須）
    - 課題概要や、書いたコードそのものを記載する必要はない
  2. その他、自由にデータを分析した結果（+α）
    - 学習データ数と性能、正則化パラメータの関係
    - 各特徴の寄与の分析、など
  3. ここまでの講義の感想、要望など（任意）

# 課題詳細

- 評価の方針
  - あくまでレポートとしての出来を評価
  - アイデアや試行錯誤の過程を重視
  - スコアが悪くても、しっかり考察してくれればOK  
(もちろん良くなればプラスに評価しますが)
  - 面白い分析を期待します
- 提出先
  - ITC-LMS (レポート課題1)
- 締め切り
  - 1月8日の授業開始まで (10:25)
  - スコアリングサーバは1月5日24:00に締め切り

# 工夫できそうなところ（例）

- 回帰のオフセット（バイアス項）
- メタデータの扱い
- 手法・パラメータチューニング
  - 正則化項の入れ方（リッジ回帰）
  - クロスバリデーション
- 目的変数の性質をよく考える
  - どんな分布をしていそうか？
- 特徴の前処理
  - スケール変換など

# How to mount Google Drive in Colab



sample.ipynb ☆

ファイル 編集 表示 挿入 ランタイム ツール ヘルプ

+ コード + テキスト

Mount your Google Drive. In this notebook, we assume that 'report1' folder is placed directly under 'My Drive'.

```
from google.colab import drive
drive.mount('/content/drive')
```

① Run this code

... Go to this URL in a browser: [https://accounts.google.com/o/oauth2/auth?client\\_id=](https://accounts.google.com/o/oauth2/auth?client_id=)

Enter your authorization code:

② Open this URL in your browser

```
!ls /content/drive/MyDrive You should be able to see 'report1' folder by running this.
```

Google Drive File Stream が  
Google アカウントへのアクセス  
をリクエストしています

Google Drive File Stream に以下を許可します:

Google ドライブのすべてのファイルの表  
示、編集、作成、削除 ①

Google フォトの写真、動画、アルバムの表  
示 ①

プロフィールや連絡先などの Google ユーザ  
ー情報の表示 ①

Google ドライブのドキュメントの表示、編  
集、作成、削除 ①

Google Drive File Stream を信頼できることを確認

機密情報をこのサイトやアプリと共有する場合があります。Google Drive File Stream の利用規約とプライバシーポリシーで、ユーザーのデータがどのように取り扱われるかをご確認ください。アクセス権の確認、削除は、Google アカウントでいつでも行えます。

[リスクの詳細](#)

キャンセル

許可

③ Accept the access

Google

ログイン

このコードをコピーし、アプリケーションに切り替えて貼り付けてください。

④ Enter the code shown  
here in the notebook