Waseda University Senior High

Microsoft Azure

高校生のための Azure Machine Learning

By M. Takezawa

機械学習 (Machine Learning)とは

機械学習とは、機械にデータを学習させ、データに潜むパターンや特性を発見し予測させることです。

Microsoft Azure Machine Learningとは

Microsoft社が提供するAzureの機能の一つであり、機械学習を用いたデータ分析をプログラミング手法 を意識せず、視覚的に構築できるクラウドサービスです。

ただし、RやPythonのスクリプトを記述するためのモジュールも用意してあります。

ここでは、回帰分析を例に、Rでのデータ分析と機械学習を用いたデータ分析を比較します。

Microsoft Azure Machine Learningとは



XYZW

:Module

AzureでのMachine Learningの概要(英語版)



AzureでのMachine Learningの概要

XYZW :モジュール名

学習実験:機械に学習させる 予測実験:計算を実行する 学習用データ 組込アルゴリズム 評価用データ モデルの学習 モデルの予測 モデルの評価

学習実験(Training Experiment)





よく使う操作 : 実行:RUNボタン 実行結果の可視化:モジュールを右クリックでVisualize



AzureでのMachine Learningのログイン方法

Azure portalへのログイン方法

http://portal.azure.com/

- もしくは検索ワード入力:MLstudio



Azure Machine Learningのログイン方法

https://studio.azureml.net/

Login ID : *****@manavi-i.com

Microsoft Azure Machine Learningの操作



新規のExperiment作成

Blank Experiment クリック



Microsoft Azure Machine Learningの操作

Waseda University Senior High

空のExperimentの出現

空の Experimentには

・メニュー ・キャンバス

の2画面構成

この空のキャンバスにいろい ろなモジュールを配置する。

次のスライドを参考のこと。



Waseda University Senior High



[DATASETS] →[New] →アップロード











回帰分析 2 (学習用・評価用データ: Automobile price data(RAW))





精度の評価精度の評価のために「ホールドアウト法」を用いる。



```
XYZW :モジュール名
[ xyzw ]:メニュー名
```

[Data Transformation]→[Sample and Split]

学習用データと評価用データを分離する Split Dataを選択 →Fraction of Row in the First output Dataset →分離率の%を代入(0.7など) →RUN





回帰分析 2 (学習用・評価用データ:Automobile price data(RAW))			Waseda University Senior High	
学習実験 Automobile price data(RAW):自動車のデータ分析 Evaluate Modelを右クリー 予測精度の確認				
Automobile price1 > Evaluate Model > Evaluation results Metrics		ults 平均絶対誤差:	平均絶対誤差:0に近いほど分析精度は高い	
Mean Absolute Error Root Mean Squared Error	1955.418471 3068.602079	二乗平均平方根	限誤差:0に近いほど分析精度は高い 0に近いほど分析精度は高い(0∼1)	
Relative Absolute Error Relative Squared Error	0.272215 0.111516	相対二乗誤差:	相対二乗誤差:0に近いほど分析精度は高い(0~1)	
Coefficient of Determination	0.888484		近いほど分析精度は高い(0~1)	
Error Histogram				

回帰分析・デシジョンフォレスト回帰分析との比較

Automobile price data(RAW):自動車のデータを例に



決定木 (Decision Tree)とは

説明変数の値をある基準にもとづいて分岐(木構造)させ、判別や予測 のモデルを構築する。

If-Thenのルールで表すことが出来る。精度はやや落ちる。

◎デシジョンフォレスト回帰分析とは:

通常の多項式に基づいた回帰分析とは異なり、決定木(デシジョンツリー)を基礎概念として回帰分析するアルゴリズムである。ビッグデータ時代に相応しい機械学習の方法として最近注目されている。

参考: ランダムフォレスト (Random Forest)



Rデータセットのcarsより

回帰分析3 (異なるアルゴリズムで比較する)

Waseda University Senior High

決定木 (Decision Tree)とは



回帰分析3 (異なるアルゴリズムで比較する)

