The Machine Learning Canvas (v0.4)

Designed for: **Regression Task** Designed by: Jia-Ying Lin Date: 2017/09/05 Iteration: 1 .

ML task

Input, output to predict, type of problem.

Type of problem: Regression

Input為房產相關資料

Output為房價

Decisions

How are predictions used to make decisions that provide the proposed value to the end-user?

房地產公司藉由歷史房市資訊所建立的模型,預測下一季的房產價格

Value Propositions

What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?

藉由波士頓房產資料建立房價回歸模型，以預測新的房產價格。

Data Sources

Which raw data sources can we use (internal and external)?

CMU Data Set

Collecting Data

How do we get new data to learn from ?

資料取出80%作為訓練檔

剩下的20%作為測試檔

Making Predictions

When do we make predictions on new inputs? How long do we have to featurize a new input and make a prediction?

每一季訂定銷售策略的時候做出預測

目標：預測房產價格,作為下一季銷售測略的參考

Offline Evaluation

Methods and metrics to evaluate the system before deployment.

將歷史房市資訊做預測, 用當時真實房價作為驗證, 誤差率必須低於1%

Features

Input representations extracted from raw data sources.

1 - CRIM
2 - ZN
3 - INDUS
4 - CHAS
5 - NOX
6 - RM
7 - AGE
8 - DIS
9 - RAD
10 - TAX
11 - PTRATIO
12 - B
13 - LSTAT
14 - MEDV

Building Models

When do we create/update models with new training data? How long do we have to featurize training inputs and create a model?

用Linear Regression演算法建立預測模型

Live Evaluation and Monitoring

Methods and metrics to evaluate the system after deployment, and to quantify value creation.

將每一季結束後, 用真實房價作為驗證, 誤差率必須低於1%