The Machine Learning Canvas

Designed for: | Designed by: | Date: | Iteration: .

Decisions

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

historical data is used to build a mathematical model that captures important trends. That predictive model is then used on current data to predict what will happen next, or to suggest actions to take for optimal outcomes.

ML task

Input, output to predict, type of problem.

The ML task is to build mathematical model that will help predict ‘the future’ and therefore make more profit

Value Propositions

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

We are trying to facilitate the predictions to the end user(s) and help him(them) classify new data easily

Data Sources

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

all kinds of sources :

image

texte

sound

question answering data

biological data ...

Collecting Data

How do we get new data to learn from (inputs and outputs)?

we can collect data from different kinds of sources:

Social media

Vidéos

financial websites

Clients

Surveys

Making Predictions

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

Once the model(s) is(are) complete and already tested, then we can proceed to making predictions.

Offline Evaluation

Methods and metrics to evaluate the system before deployment.

In order to evaluate the system before deployment, we use the confusion matrix.

Features

Input representations extracted from raw data sources.

Many types of features can be done from the data :

Graphs(Line,

Plots

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?

We update the model when the scoring given by the initial one is far from expectations.

Live Evaluation and Monitoring

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

Crossvalidation, Root mean square error, Kmeans, Area under the curve...