Machine Learning Canvas

PREDICTIONS	OBJECTIVES	DATA

BACKGROUND

End-user

Who will use the predictive application/who will be affected by it?

Value proposition

What are we trying to do for the system's users? (e.g. spend less time on X, increase Y...)

Hahhaha

Data sources

Where do/can we get data from? (internal database, 3rd party API, etc.)

ENGINE SPECS

Problem

Decision to be made (on behalf of user)

Input (i.e. question "parameter")

Possible outputs (i.e. "answers")

Applicable trade-offs (i.e. "time vs. quality")

Type of problem (e.g. classification, regression, recommendation...)

Baseline: simple, alternative way of making predictions (e.g. manual rules)

Performance evaluation

Domain-specific/bottom-line metrics for monitoring performance in production

Prediction accuracy metrics (e.g. MSE if regression; % accuracy, #FP for classification)

Offline performance evaluation method (e.g. cross-validation or simple training/test split)

Data preparation

How do we get training data (inputs, and outputs if supervised learning)? How many data points?

Input features (extracted from data sources). If too many, list types of features and mention key ones.

INTEGRATION

Using predictions

When do we make predictions and how many?

What is the time constraint for making those predictions?

How do we use predictions and confidence values?

Learning models

When do we create/update models? With what data/how much?

What is the time constraint for creating a model?

Criteria for deploying model (e.g. minimum performance value — absolute, relative to baseline or to previous model)

Reset Form