THE MACHINE LEARNING CANVAS

Designed for: Designed by: Date: Iteration: .

PREDICTION TASK

Type of task? Entity on which predictions are made? Possible outcomes? Wait time before observation?

DECISIONS

How are predictions turned into proposed value for the end-user? Mention parameters of the process / application that does that.

VALUE PROPOSITION

Who is the end-user? What are their objectives? How will they benefit from the ML system? Mention workflow/interfaces.

DATA COLLECTION

Strategy for initial train set & continuous update. Mention collection rate, holdout on production entities, cost/constraints to observe outcomes.

DATA SOURCES

Where can we get (raw) information on entities and observed outcomes? Mention database tables, API methods, websites to scrape, etc.

IMPACT SIMULATION

Can models be deployed? Which test data to assess performance? Cost/gain values for (in)correct decisions? Fairness constraint?

MAKING PREDICTIONS

When do we make real-time / batch pred.? Time available for this + featurization + post-processing? Compute target?

BUILDING MODELS

How many prod models are needed? When would we update? Time available for this (including featurization and analysis)?

FEATURES

Input representations available at prediction time, extracted from raw data sources.

MONITORING

Metrics to quantify value creation and measure the ML system's impact in production (on end-users and business)?