Machine Learning Canvas

	PREDICTIONS	OBJECTIVES	DATA
IDEA	**Context** Who will use the predictive system / who will be affected by it? Provide some background.	**Value Proposition** What are we trying to do? E.g. spend less time on X, increase Y...	**Data Sources** Where do/can we get data from? (internal database, 3rd party API, etc.)
SPECS	**Problem** Question to predict answers to (in plain English) Input (i.e. question "parameter") Possible outputs (i.e. "answers") Type of problem (e.g. classification, regression, recommendation...) **Baseline** What is an alternative way of making predictions (e.g. manual rules based on feature values)?	**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)	**Dataset** How do we collect data (inputs and outputs)? How many data points? **Features** Used to represent inputs and extracted from data sources above. Group by types and mention key features if too many to list all.
DEPLOYMENT	**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 predictive models** When do we create/update models? With which 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

Machine Learning Canvas v0.1