Machine Learning Kickstarter Canvas

nexocode

EXERCISE 1

Formulate your problem, in plain English

Please explain in simple words what would you like a machine learning algorithm to do. At this point, it's important to ensure that the statement is of a qualitative nature. However, it's crucial to focus on articulating your primary objective clearly. This means avoiding the description of peripheral or secondary aims, and instead, directly addressing the main goal you intend to achieve. This approach will help in ensuring that the statement is not only qualitative but also accurately reflects your central purpose or intention.

EXAMPLE

We want the machine learned model to predict how popular a video just uploaded now will become in the future.

EXERCISE 2

The ideal outcome

Model (or algorithm) is intended to produce some desirable outcome. What is the result of this process, separate from the model's own characteristics? It's important to recognize that this outcome might vary significantly from the way you evaluate the model and its performance.

EXAMPLE

Our ideal outcome is to only transcode popular videos to minimize serving resource utilization.

Our ideal outcome is to suggest videos that people find useful, entertaining, and worth their time.

EXERCISE 3

Measure performance

Think twice and write down how we can measure the success (or failure) of the ML algorithm. How can you tell when the ML system has failed or succeeded?
You can phrase them independently of evaluation metrics for the model (e.g., don't talk about statistics and math like precision, recall, R2, RMSE, or AUC; talk about the anticipated outcomes, instead). You need to focus:

- **Are the metrics measurable? Are they stable?**
- **How will you measure them?**
  - It's okay if this is via a live experiment. Many success metrics can't be captured offline. When deciding on your metrics, think about the ideal outcome that you specified in the previous step.
- **When are you able to measure them?**
  - How long will it take for you to know whether your new ML system is a success or failure?
- **Consider engineering and maintenance costs over the long-term gain.**
- **Failure may not only be caused simply by non achievement of success metric.**
  - For example; a model may be able to predict whether they click on recommended videos very well, but it may always be recommending "click baity" videos

EXAMPLE

Our success metrics are CPU resource utilization.
Our KR for the success metrics are to achieve 35% less CPU cost for transcoding.
Our ML model is deemed unsuccessful if the CPU resource cost reduction is less than the CPU costs for training and serving the model.

Our success metrics are the number of popular videos properly predicted.
Our KR for the success metrics are to properly predict the top 95% 28 days after being uploaded.
Our ML model is deemed unsuccessful if the number of popular videos properly predicted is no better than current heuristics.

Reviewing the ML design

EXERCISE 4

Defining the output

Define the output that you want your ML model to produce. The output should be measurable in a way that a machine can understand. Consider the difference between "the user playing the game" and "the user will recommend the game to a friend" Can you acquire sample outputs for training data? What method and source will you use for this? You might need to design your output examples, similar to how the example above converts watch time into a percentile. If acquiring training examples is challenging, you may have to reevaluate your previous responses to reframe your problem and objectives. This reframing will enable you to effectively train a model with your data.

EXAMPLE

The output from our ML model will be one of the 3 classes of videos (very popular, somewhat popular, not popular). It is defined as the top {3, 7, 90}-percentile of watch time 28 days after being uploaded.

EXERCISE 5

Using the output

Write when your output must be obtained from the ML model, and how it is used in your product.

EXAMPLE

The prediction of a video's popularity will be made as soon as a new video is uploaded. The outcome will be used for determining the transcoding algorithm for the video.

EXERCISE 6

Heuristics

Write all the tips and tricks you are using now, and what would have done to replace ML if there was no other way. Think about anything interesting and non-trivial about your data, that can help us in this task.

EXAMPLE

If we didn't use ML, we would assume new videos uploaded by creators who had uploaded popular videos in the past will become popular again.

EXERCISE 7

Cast your problem as a simpler one

When first starting out, simpler problem formulations are easier to reason about and implement. Take your given problem and state it as a binary classification or a unidimensional regression problem (or both).

EXAMPLE

We will predict whether uploaded videos are most likely to become very popular (binary classification).

We will predict how popular an uploaded video will be in terms of the number of views it will receive within a 28 day window (regression).

EXERCISE 8

Explainability of the model

Think whether it is important for you to understand WHY the model produces such a result (not technical details about how it is working). Can you agree and trust its outcome, even if it seems wrong or risky at first glance? If not, think about possible reasons that could convince you that the result is correct.

EXAMPLE

Our ML model predicts that the leg of a patient who is sick should be amputated. A decision of this caliber should be well justified to the patient...

Our ML model predicts that videos from relatively new and unknown creator will be popular. When asked why, it points that this creator has partnered with a very popular creator and is receiving heavy advertising from him.

EXERCISE 9

Dataset

Write about the dataset that can be utilized by the ML model. Do not limit yourself to the data used by existing solutions, as it is possible that the new model will be able to use seemingly unrelated features. Consider the types of data, its volume, recency, and any legal issues that may require anonymization. If some data are artificially produced, emphasize this fact and separate it from the rest. At this point, the numbers do not have to be exact.

EXAMPLE

The dataset consists of 5-10k videos in the span of ~12 months. Each video has its own statistics in the form of a short description and a time series of daily views. Additionally, the videos are linked to their creators and you can get statistics about them as well.