



Checkpoint 2

Power and problems

As we already mentioned, unsupervised learning can work in a myriad of contexts some of which operate on their own and others which are great for working in concert with other approaches (like how PCA is a great first step to building a supervised model).

Outline

Evaluation

This is not without challenges, however. Probably the biggest difficulty in unsupervised learning is evaluation. Since there is no outcome variable, defining the accuracy of the model isn't really possible, since you have nothing to evaluate against. There are some approaches for dealing with this.

Firstly, you can look at the volatility of your model. If it stays reasonably stable to slight perturbations or the exclusion or addition of some data this is a relatively good thing.

Unsupervised learning also often has to simply be tested. If you're trying to create clusters, test what you intended to do with those clusters and see if it works as you'd hoped. Instead of working on this specific step

in the data pipeline, you evaluate how later steps perform and tweak the unsupervised step to see how it affects performance.

The truth remains, however, that many of these models have no absolute form of evaluation. It can be extremely hard to say whether an unsupervised model is a 'good' or 'bad' model.

Complexity

With the challenge of evaluation comes another: complexity. Since these models are hard to evaluate it is also often difficult to decide on a stopping condition. This leads to models that can get increasingly more complex and therefore more resource consuming to run. This is particularly true of neural networks, something we'll talk about in greater detail later.

Outline

This can be addressed through imposing some of your own limitations or keeping some of the structures simple, but it is definitely still a risk in this class of models.

There is also significant research, particularly in neural networks, that shows the advantages of some of the more complex models. Their performance can be quite strong.

Mark as read

[Report a typo or other issue](#)

[Next checkpoint](#)

[Go to Overview](#)

Outline