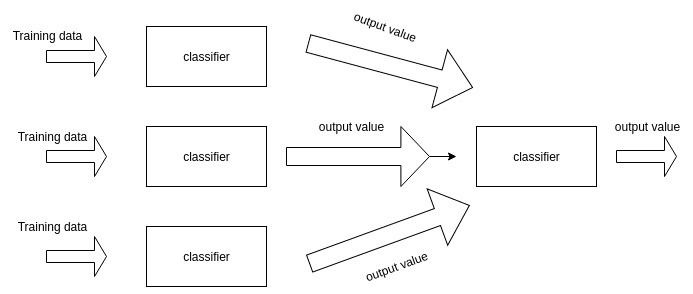
**Stacking**

<http://blog.kaggle.com/2016/12/27/a-kagglers-guide-to-model-stacking-in-practice/>

Stacking (also called meta ensembling) is a model ensembling technique used to combine information from multiple predictive models to generate a new model. Often times the stacked model will outperform each of the individual models due its smoothing nature and ability to highlight each base model where it performs best and discredit each base model where it performs poorly. For this reason, stacking is most effective when the base models are significantly different.

you can build multiple different learners and use them to build an intermediate prediction, one prediction for each learned model. Then you add a new model which learns from the intermediate predictions. This final model is said to be stacked on the top of the others, hence the name. Often you end up with a model which is better than any individual intermediate model.

The main point is that we’re using the **predictions** of the base models as **features** (i.e. meta features) for the stacked model.



In general, stacking produces small gains with a lot of added complexity – not worth it for most businesses. But Stacking is almost always fruitful so it’s almost always used in top Kaggle solutions.

averaging???