

# Lesson 2

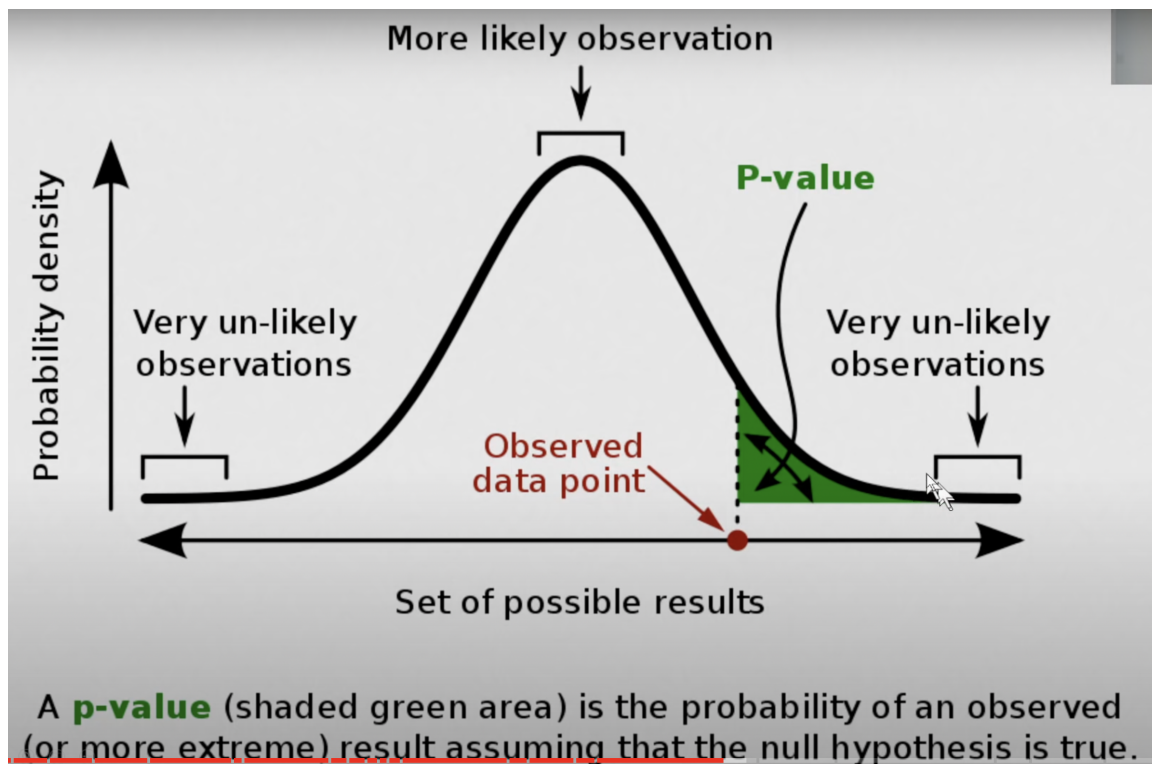
- Regression → Numeric quantities  
Classification → Discrete quantities
  - Overfitting
    - Train for too long
    - Not enough data
    - Too many parameters
  - loss v metric v error
    - metric → quality of the model's predictions using validation function, printed for each epoch
    - loss → measure of the performance of the model on the basis of which the model weights are updated
    - error → a kind of metric
    - metric might not be affected by change in function as the prediction category is not changing, but loss will still change
  - overfitting is the single most important and challenging issue with training
  - In overfitting, the model learns specific points of data instead of generalising
  - Ideally, to avoid overfitting to training AND validation, you should have a third dataset called a test set
  - For time series data, the validation should not be random, but should be chronologically end from the end
  - As you train,
    - train loss goes down, and the validation loss goes down
    - but as you over fit, train loss will still go down, and the validation loss will start going up
    - but focus on metric instead of loss
-

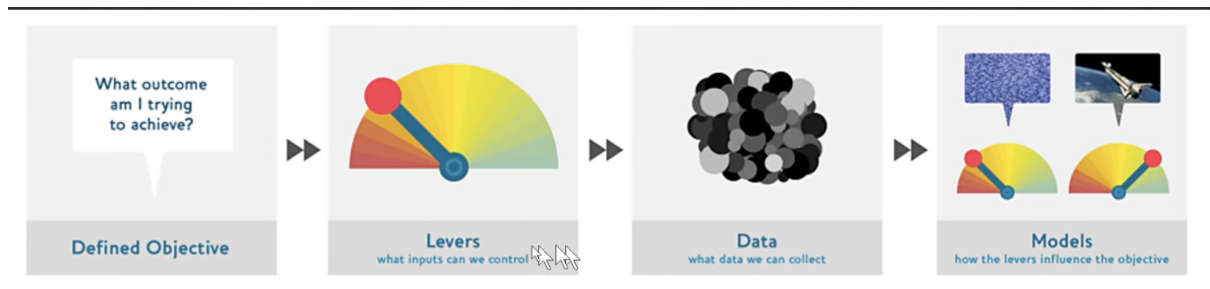
- Transfer Learning
  - Using a pre-trained model for a task different from what it was originally trained for
- Fine-tuning
  - A transfer learning technique where the weights of a pre-trained model are updated by training for additional epochs using a different task to that used for pre-training
- By default, the way fast.ai does it →
  - Use one epoch to get the *head* working
  - Use as many epochs as required to fit the model

Term	Meaning
label	The data that we're trying to predict, such as "dog" or "cat"
architecture	The <i>template</i> of the model that we're trying to fit; the actual mathematical function that we're passing the input data and parameters to
model	the combination of the architecture with a particular set of parameters
parameters	the values in the model that change what task it can do, and are updated through model training
fit	Update the parameters of the model such that the predictions of the model using the input data match the target labels
train	A synonym for <i>fit</i>
pretrained model	A model that has already been trained, generally using a large dataset, and will be fine-tuned
fine tune	Update a pretrained model for a different task
epoch	One complete pass through the input data
loss	A measure of how good the model is, chosen to drive training via SGD
metric	A measurement of how good the model is, using the validation set, chosen for human consumption
validation set	A set of data held out from training, used only for measuring how good the model is
training set	The data used for fitting the model; does not include any data from the validation set
overfitting	Training a model in such a way that it <i>remembers</i> specific features of the input data, rather than generalizing well to data not seen during training
CNN	Convolutional neural network; a type of neural network that works particularly well for computer vision tasks

- model zoo → pre-trained models

- null hypothesis → starting assumption
- gather data of independent and dependent variable
- What % of time would we see this relationship by chance





- What are the levers you can tweak in the predictive model you are building?

- ImageBlock → Independent Variable: the images of the bears  
CategoryBlock → Dependent Variable: the category of the bears