VALIDATION METHODS & PERFORMANCE METRICS

Lecture 6

MAL1, 2024

THE BIG PICTURE

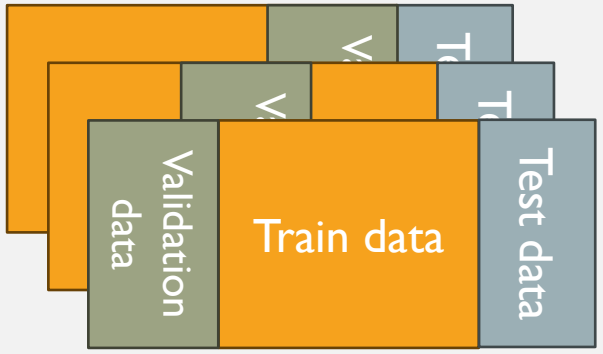
THE BIG PICTURE

VALIDATION METHODS

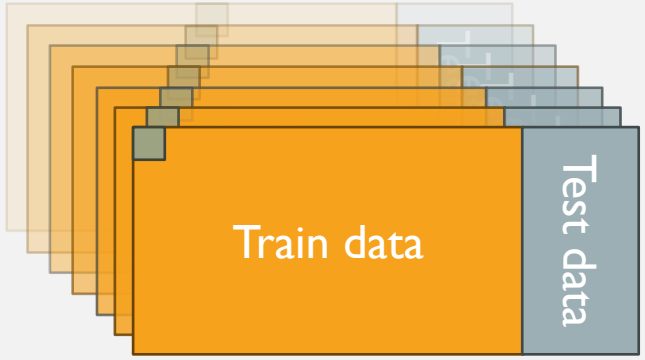
VALIDATION METHODS

Train-test methodology

Train-val-test methodology

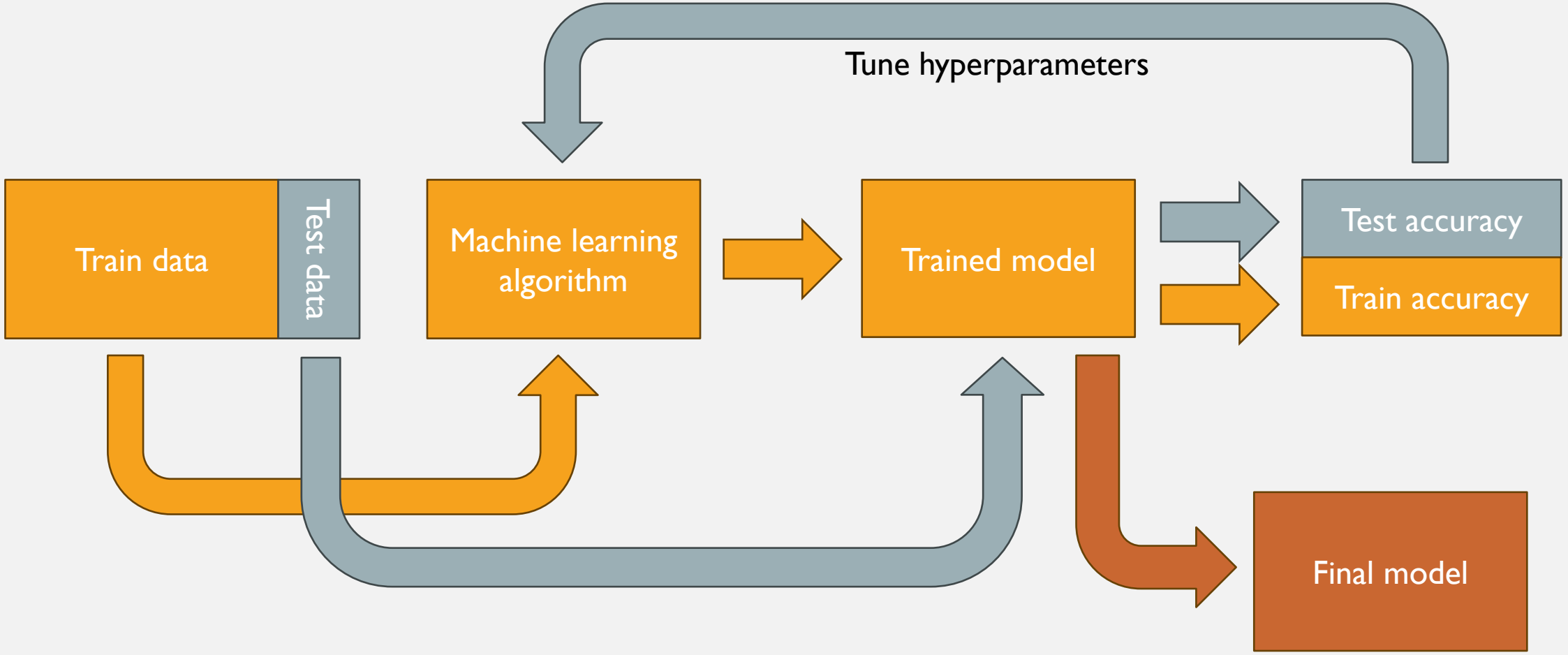


Cross-validation methodology

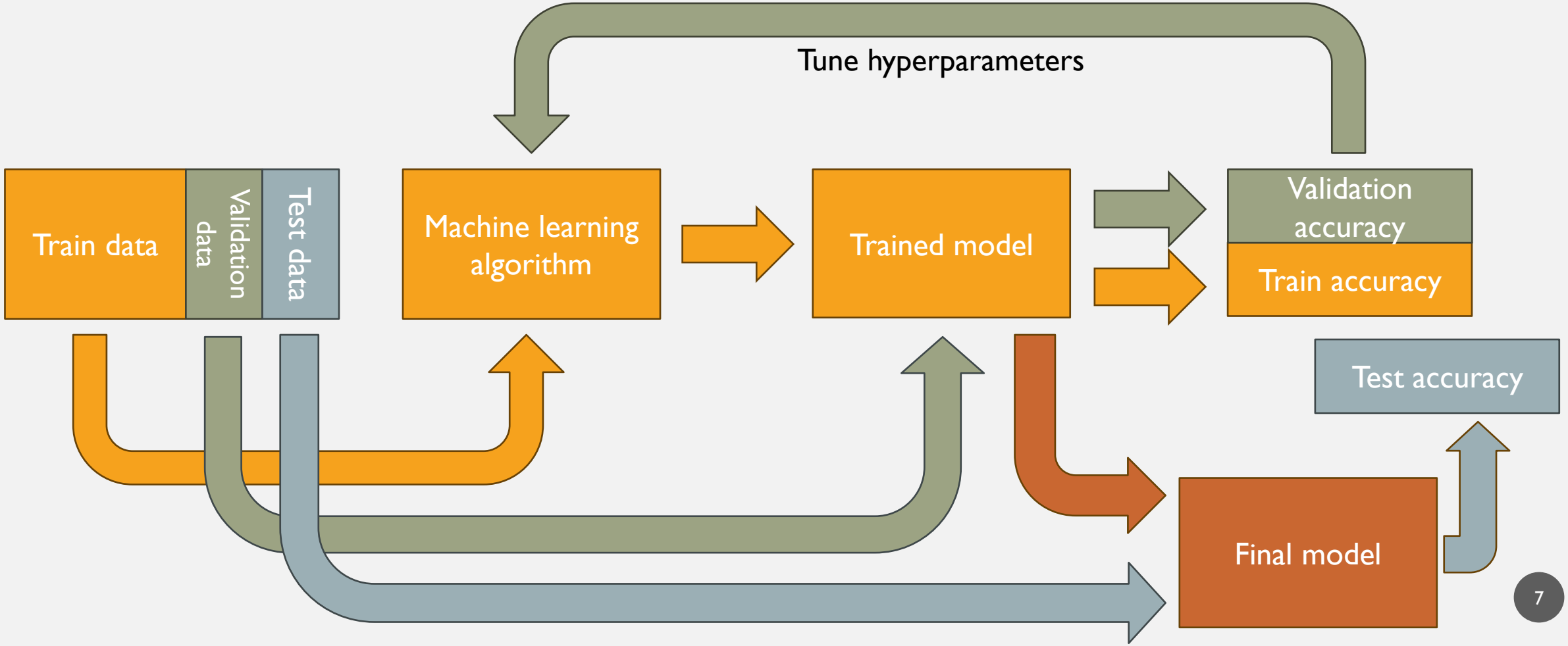


Leave-1-out cross validation methodology

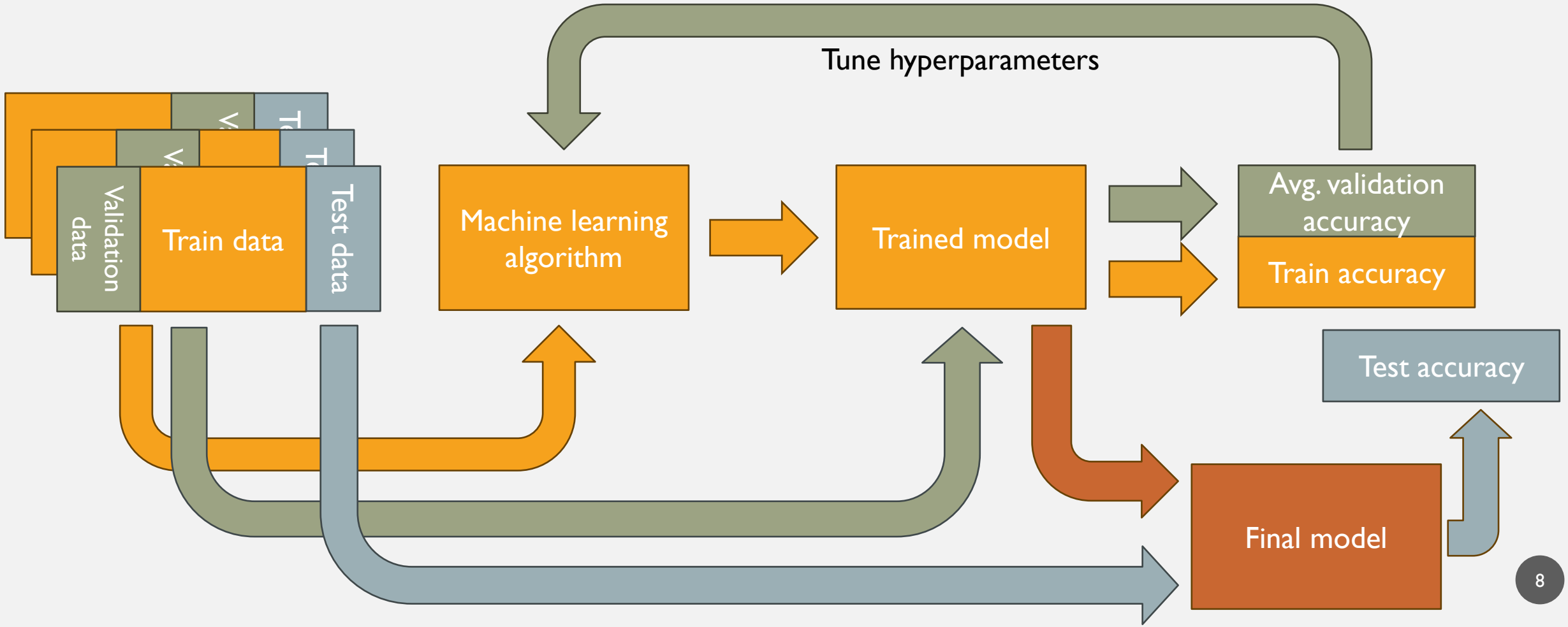
TRAIN-TEST METHODOLOGY



TRAIN-VAL-TEST METHODOLOGY

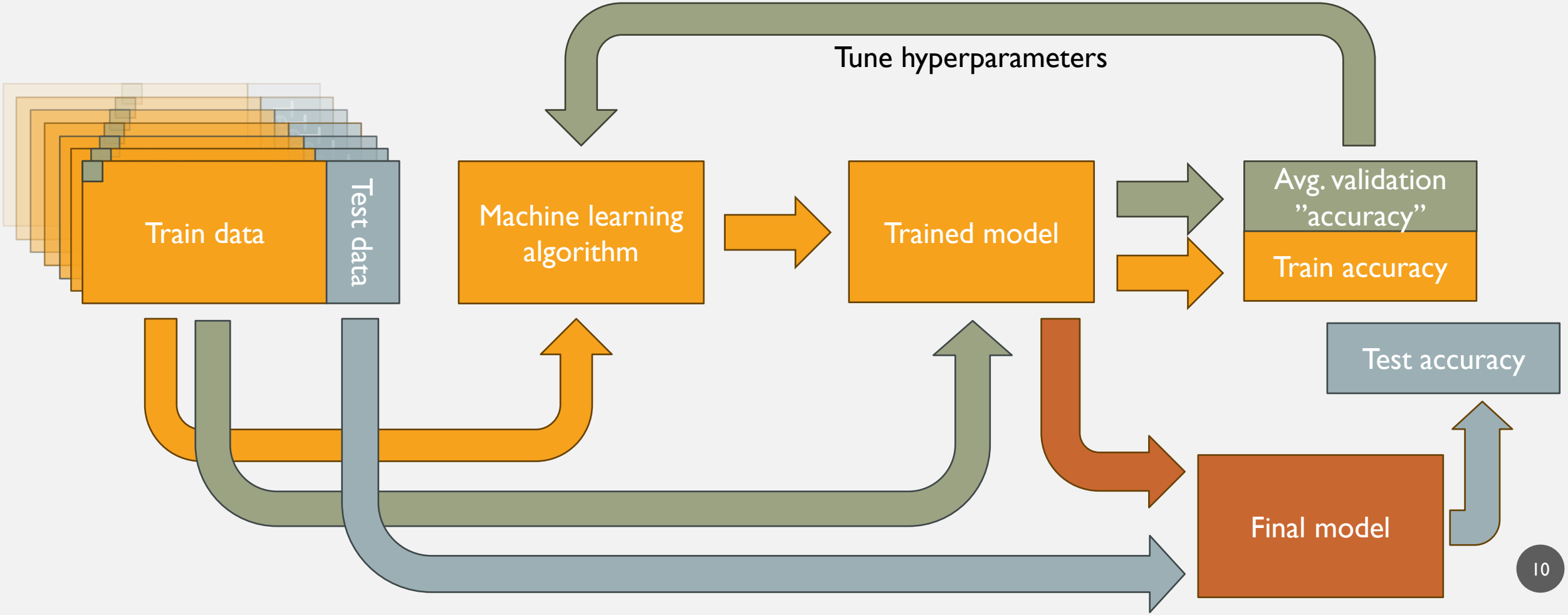


CROSS-VALIDATION METHODOLOGY



CROSS-VALIDATION METHODOLOGY

LEAVE-1-OUT CROSS-VALIDATION METHODOLOGY



CODE EXAMPLE

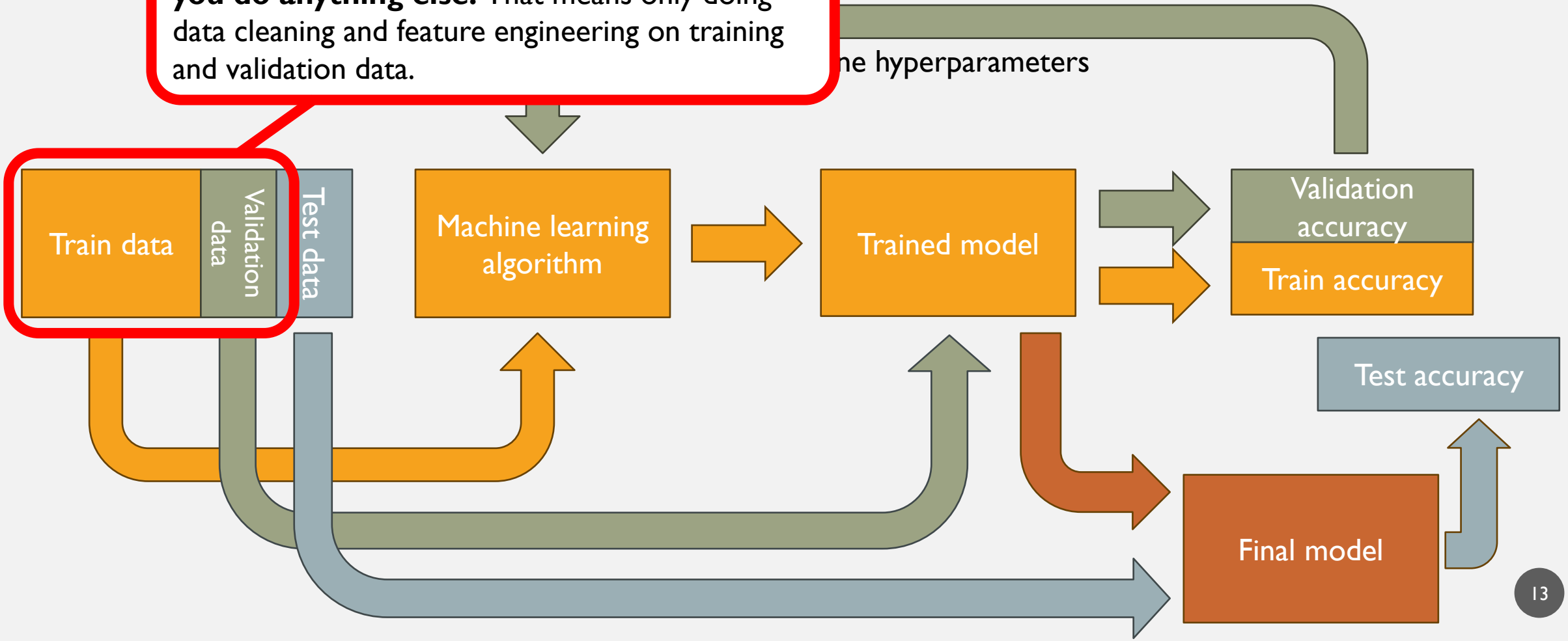
Jupyter Notebook **Validation methods**

A NOTE ON PREPROCESSING

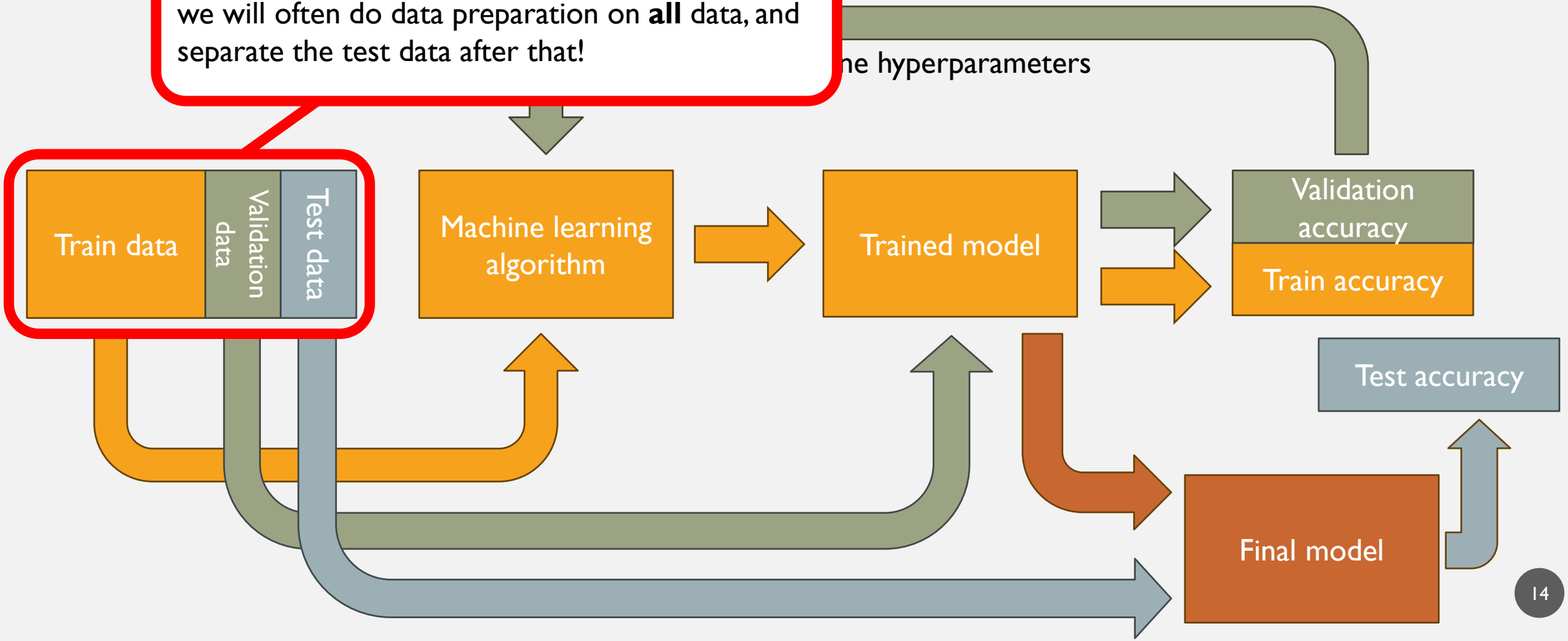
- Deal with outliers
- Deal with missing values
- Normalize/scale data
- One-hot encoding
- Representing text data
- …

A NOTE ON PREPROCESSING

Ideally you should isolate you testing data **before you do anything else!** That means only doing data cleaning and feature engineering on training and validation data.



A NOTE ON PREPROCESSING



PERFORMANCE METRICS

THE BIG PICTURE

TYPES OF ERRORS

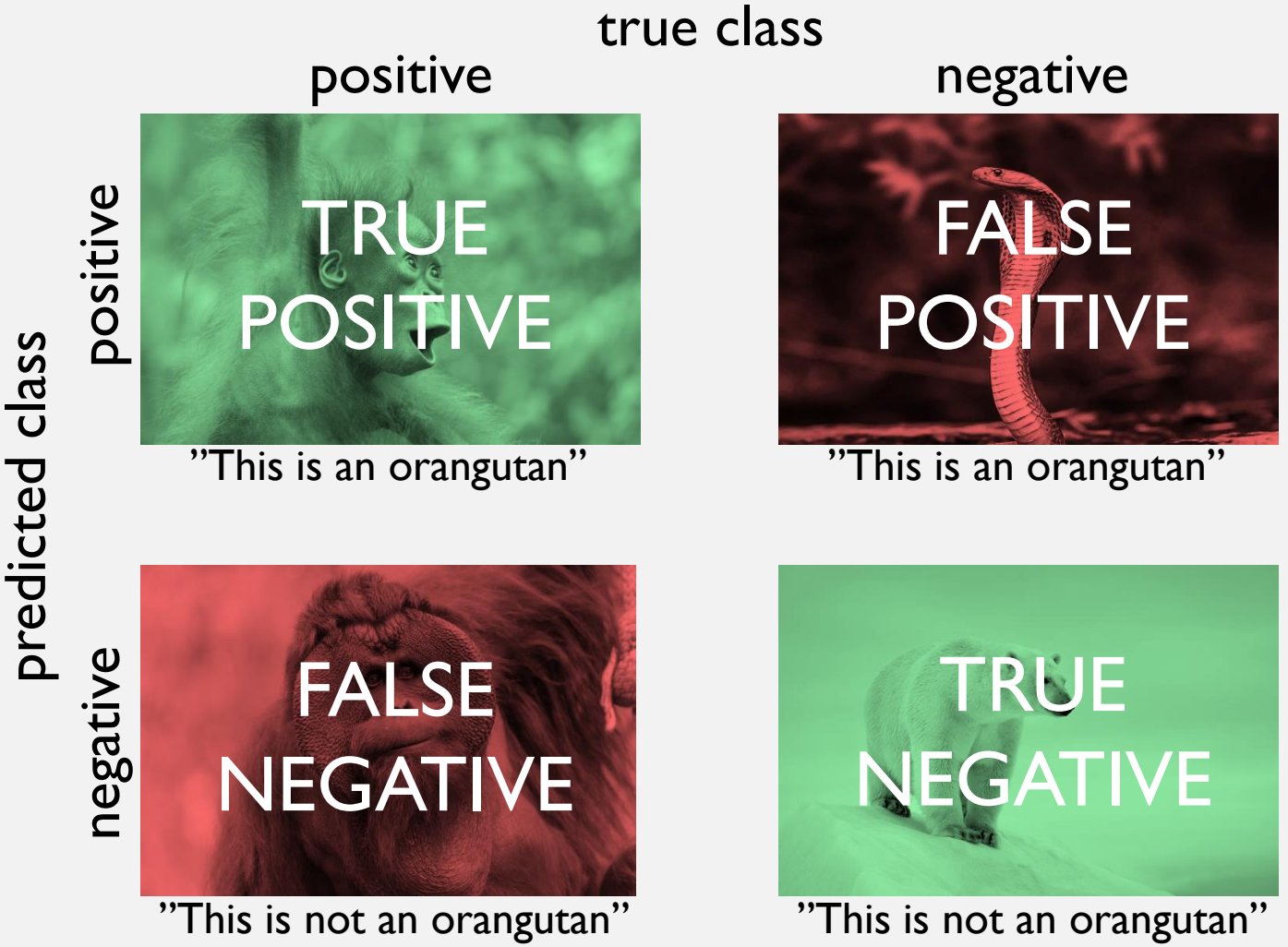
”This is an orangutan”

”This is an orangutan”

”This is not an orangutan”

”This is not an orangutan”

TYPES OF ERRORS



TYPES OF ERRORS

”orangutan”

”not orangutan”

”orangutan”

”not orangutan”


”orangutan”


”not orangutan”


”orangutan”


”orangutan”


”orangutan”


”not orangutan”


”not orangutan”


”orangutan”

predicted class	true class: positive	true class: negative
positive	TRUE POSITIVE	FALSE POSITIVE
negative	FALSE NEGATIVE	TRUE NEGATIVE

TYPES OF ERRORS

”orangutan”

”not orangutan”

”orangutan”

”not orangutan”

”orangutan”

”not orangutan”

”orangutan”

”orangutan”

”orangutan”

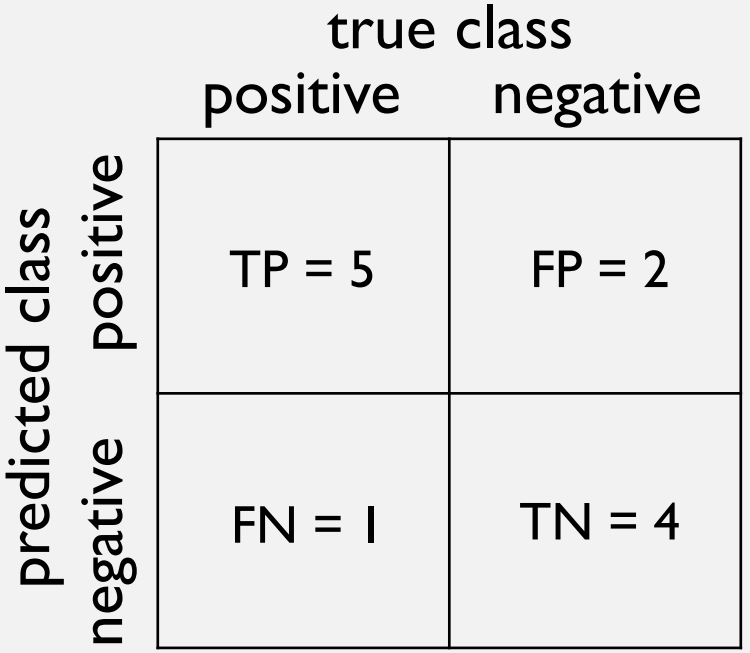
”not orangutan”

”not orangutan”

”orangutan”

predicted class	true class: positive	true class: negative
positive	TP = 5	FP = 2
negative	FN = 1	TN = 4

ACCURACY


true class
positive
negative
predicted class
positive
negative
TP = 5
FP = 2
FN = 1
TN = 4

WHY ACCURACY IS NOT GOOD ENOUGH

A model to predict whether or not someone is a terrorist:

*Everyone is **not** a terrorist.*

		true class: positive	true class: negative
predicted class	positive	TP = 0	FP = 0
	negative	FN = 1	TN = 9999

PERFORMANCE METRICS

- $\text{accuracy} = \dfrac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} = \dfrac{\text{correct predictions}}{\text{all predictions}}$

- $\text{precision} = \dfrac{\text{TP}}{\text{TP}+\text{FP}} = \dfrac{\text{correct positive predictions}}{\text{all positive predictions}}$

- $\text{recall} = \dfrac{\text{TP}}{\text{TP}+\text{FN}} = \dfrac{\text{correct positive predictions}}{\text{all positive instances}}$

USING RECALL INSTEAD

A model to predict whether or not someone is a terrorist:

*Everyone is **not** a terrorist.*

predicted class \ true class	positive	negative
positive	TP = 0	FP = 0
negative	FN = 1	TN = 9999

SOME EXAMPLES

- Determine whether someone is a terrorist
  - avoid false negatives – use recall!
- Determine whether you have COVID-19 during the pandemic
  -
- Determine whether a video is suitable for children to watch
  -
- Determine whether someone should be sentenced to life in prison
  -

To optimize hyperparameters on a particular metric:

```
GridSearchCV(clf, parameters, scoring="recall")
GridSearchCV(clf, parameters, scoring="precision")
```

THE PRECISION/RECALL TRADE-OFF

High recall usually means low precision – and vice versa

propability of orangutan

$$\text{precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad \text{recall} = \frac{\text{TP}}{\text{TP+FN}}$$

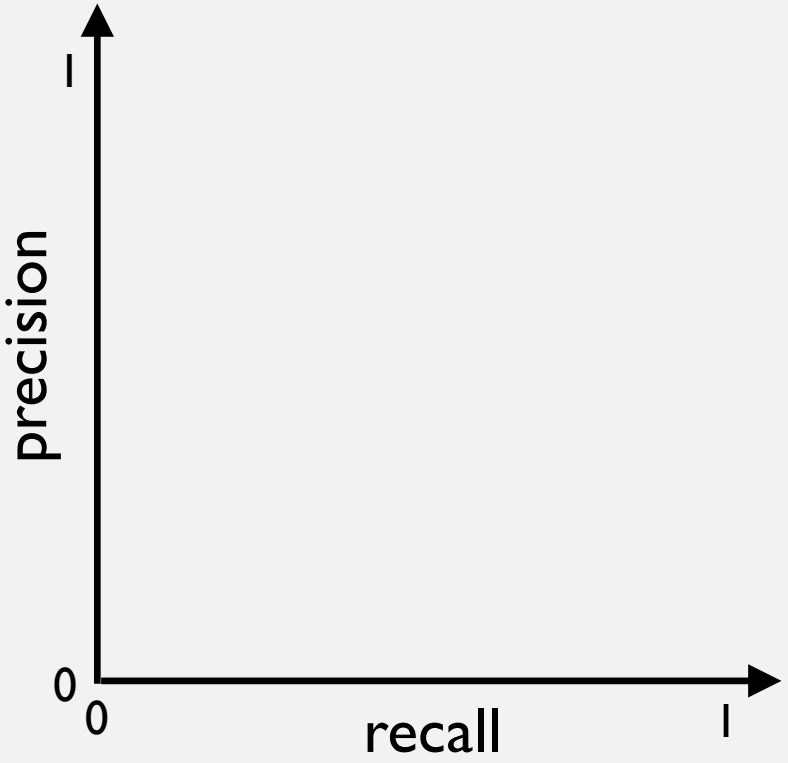
THE PRECISION/RECALL TRADE-OFF

High recall usually means low precision – and vice versa

$$\text{precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \qquad \text{recall} = \frac{\text{TP}}{\text{TP}+\text{FN}}$$

THE PRECISION-RECALL CURVE

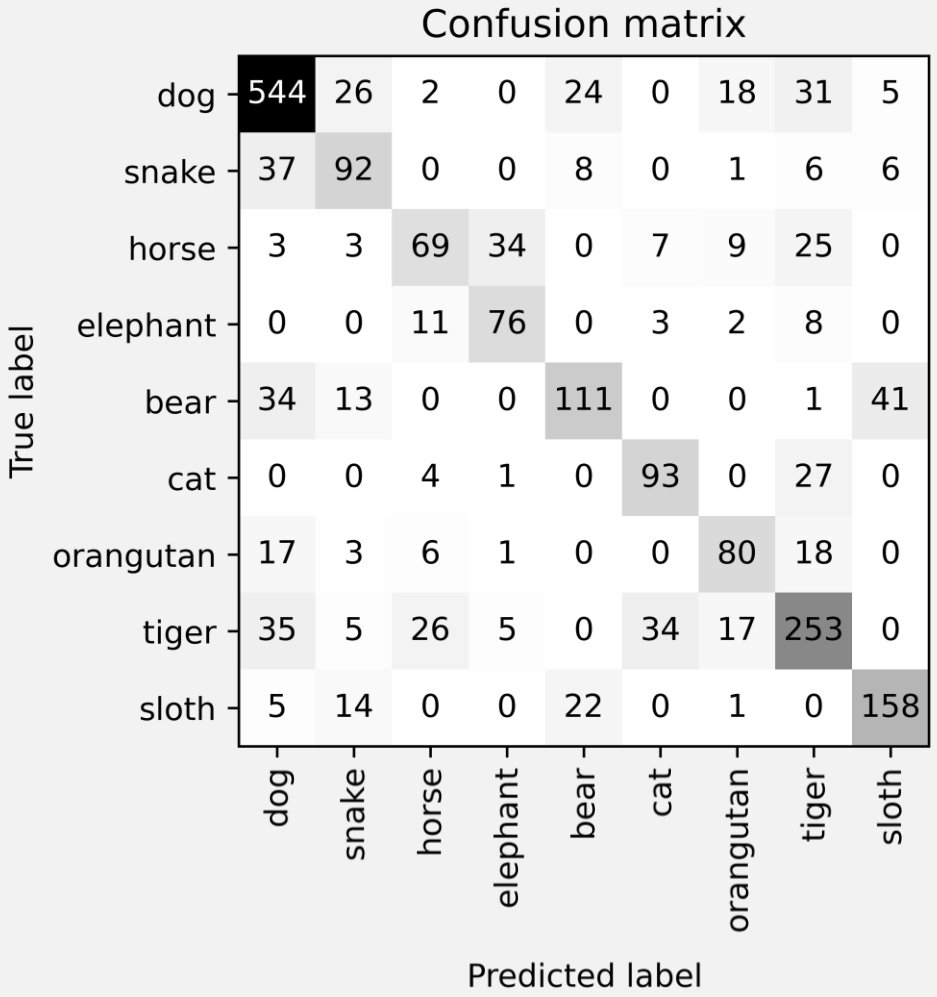
$$\text{precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad \text{recall} = \frac{\text{TP}}{\text{TP+FN}}$$



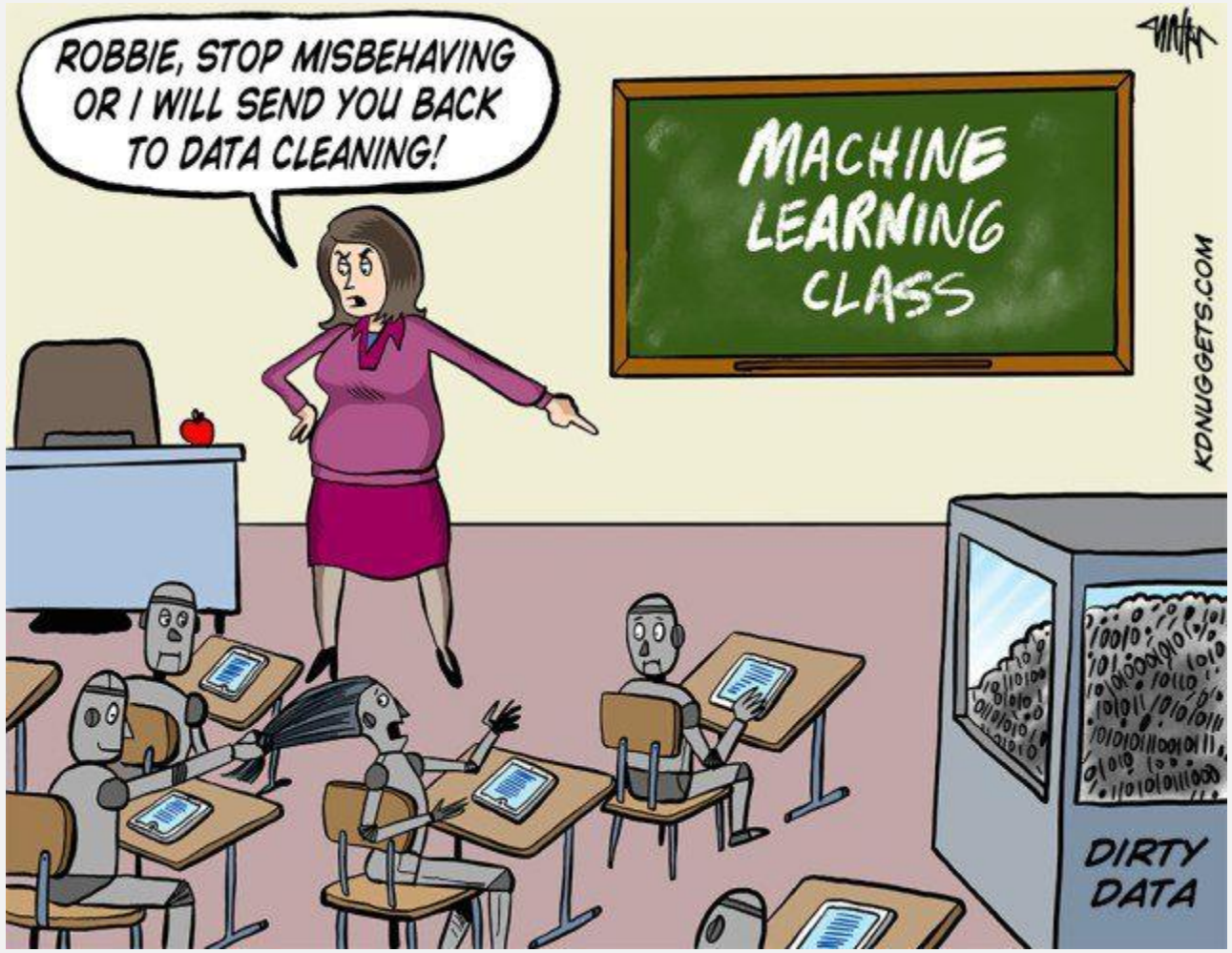
CODE EXAMPLE

Jupyter Notebook **Performance metrics**

METRICS IN MULTICLASS PROBLEMS



You can calculate all metrics for all classes, but


ROBBIE, STOP MISBEHAVING OR I WILL SEND YOU BACK TO DATA CLEANING!
MACHINE LEARNING CLASS
KDNUGGETS.COM
DIRTY DATA