S6-ESE-AI

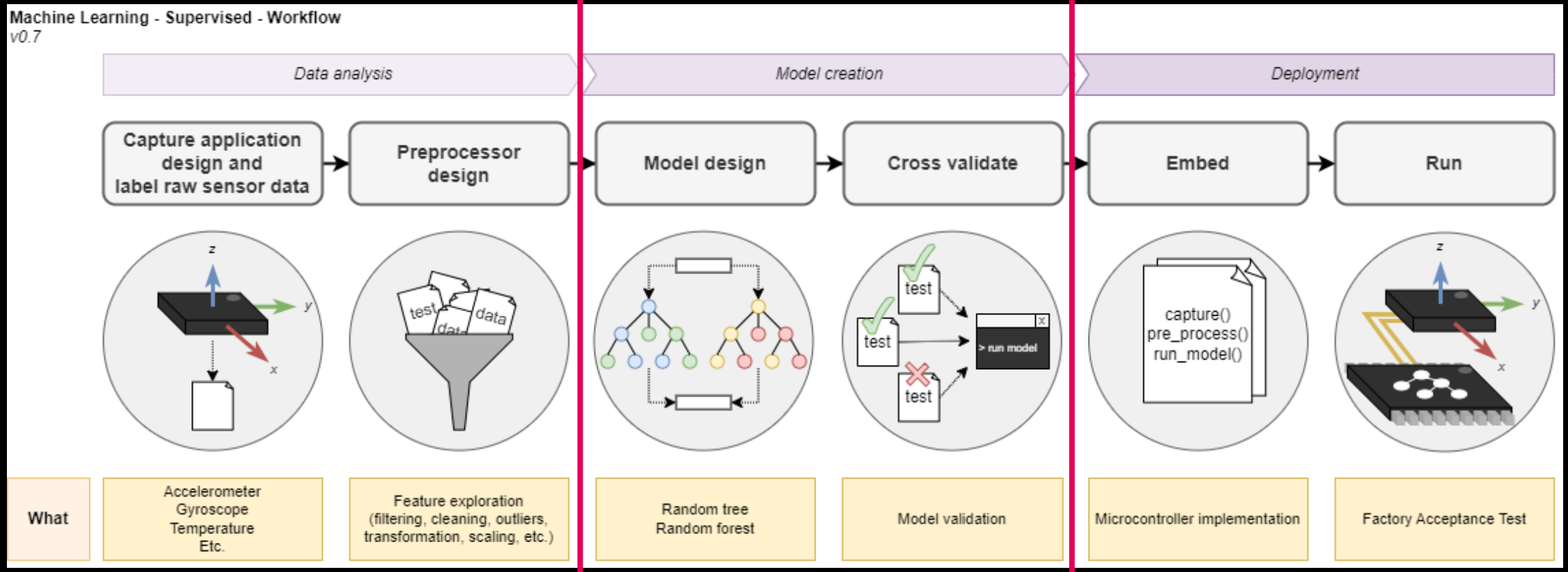
PERFORMANCE

JEROEN VEEN
HUGO ARENDS

HAN_UNIVERSITY
OF APPLIED SCIENCES

WORKFLOW

WORKFLOW

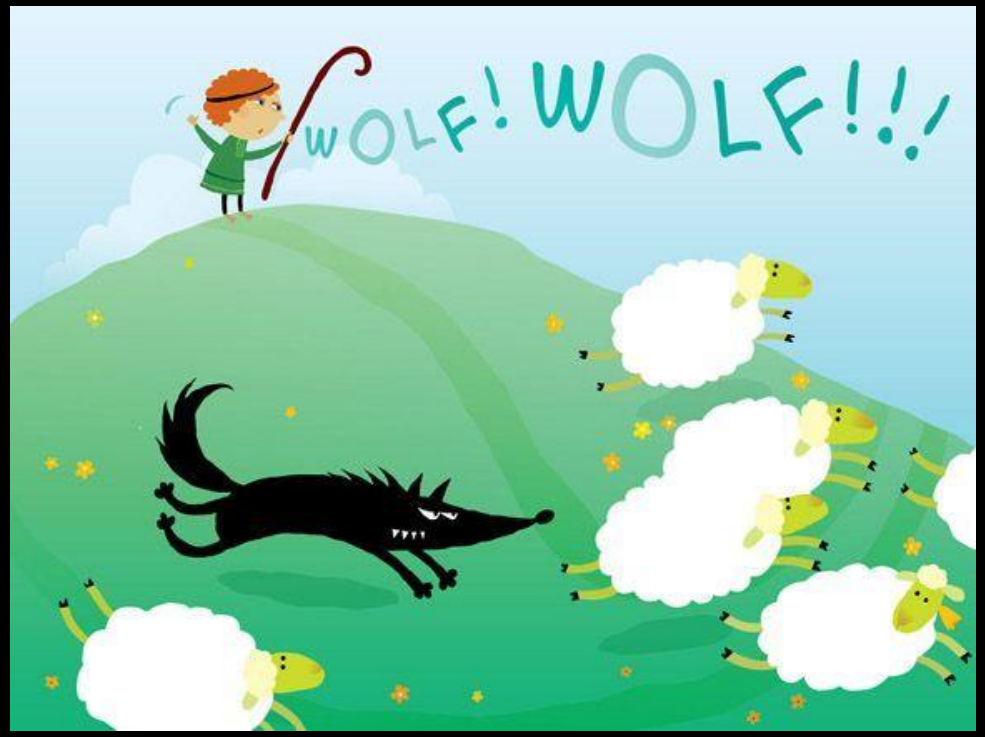

Machine Learning - Supervised - Workflow
v0.7
Data analysis
Model creation
Deployment
Capture application design and label raw sensor data
Preprocessor design
Model design
Cross validate
Embed
Run
test
data
> run model
capture()
pre_process()
run_model()
What
Accelerometer
Gyroscope
Temperature
Etc.
Feature exploration
(filtering, cleaning, outliers, transformation, scaling, etc.)
Random tree
Random forest
Model validation
Microcontroller implementation
Factory Acceptance Test

AGENDA

- Confusion matrix
- Accuracy
- Cross-validation
- Classification report
- Generalization

THE BOY WHO CRIED WOLF

- "Wolf" is a positive class.
- "No wolf" is a negative class
- An Aesop's Fable ~620 BCE



Source: Sam Taplin

CONFUSION MATRIX

ACTUAL

PREDICTED

True Positive (TP)
Reality: A wolf threatened.
Shepherd said: "Wolf."
Outcome: Shepherd is a hero.

(Type I error)

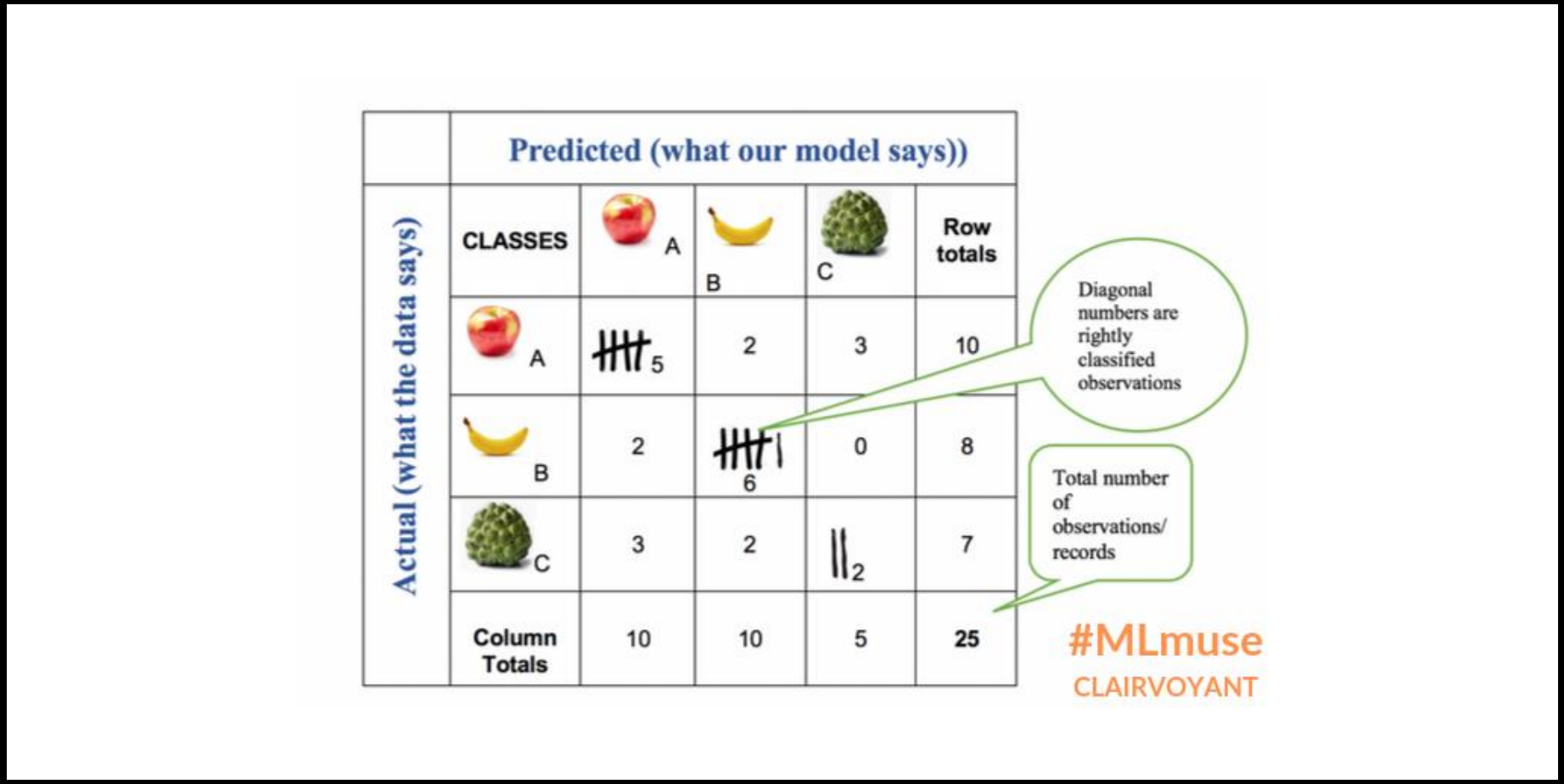
False Positive (FP)
Reality: No wolf threatened.
Shepherd said: "Wolf."
Outcome: Villagers are angry at shepherd for waking them up

False Negative (FN)
Reality: A wolf threatened.
Shepherd said: "No wolf."
me: The wolf ate all the sheep

(Type II error)

True Negative (TN)
Reality: No wolf threatened.
Shepherd said: "No wolf."
Outcome: Everyone is fine.

CONFUSION MATRIX

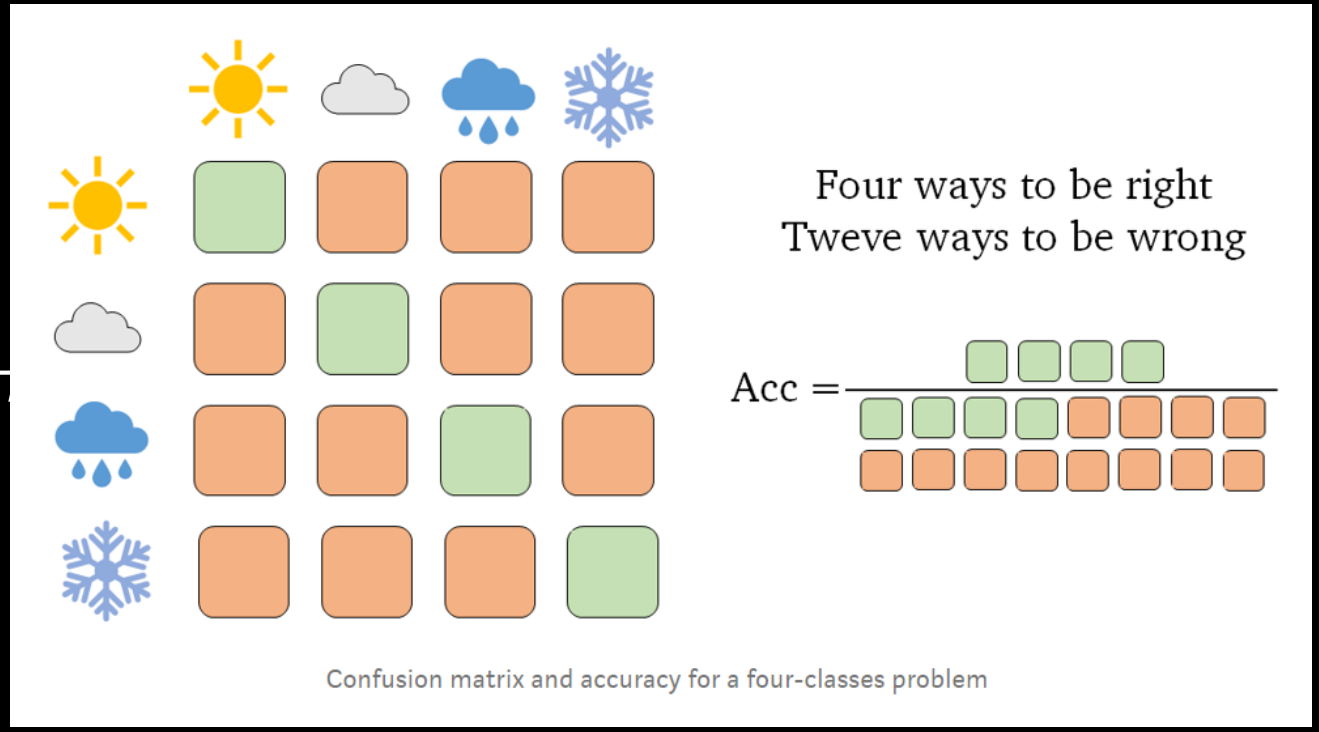


MEASURING PERFORMANCE

- Accuracy

$$= \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$

$$= \frac{}{TP +}$$



Confusion matrix and accuracy for a four-classes problem

Source: https://towardsdatascience.com/the-illustrated-guide-to-classification-metrics-the-basics-cf3c2e9b89b2

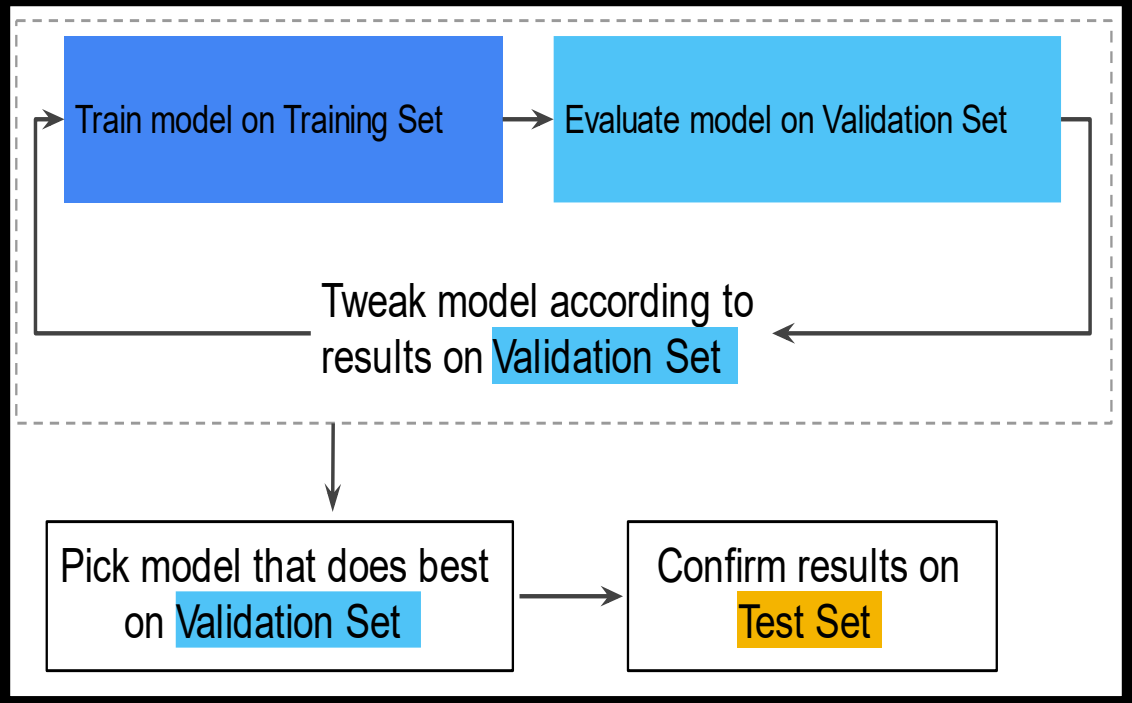
SPLITTING DATA

- Slice data into three subsets: Training, validation and test data

- Make sure that your subsets meet the following conditions:
  - Large enough to yield statistically meaningful results.
  - Representative of the data set as a whole.
    E.g. don't pick a test set with different characteristics than the training set.

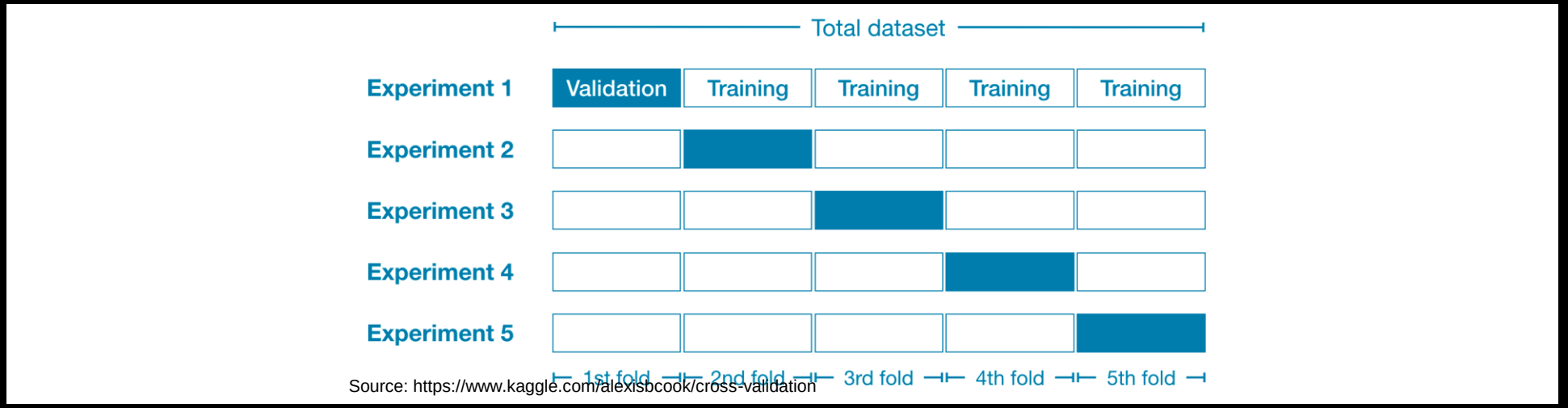
TRAINING, VALIDATION, TESTING

- Never train on test data!



CROSS-VALIDATION

- Estimate of a model's generalization performance
- Break the data into folds



Source: https://www.kaggle.com/alexisbcook/cross-validation

- For small datasets, where extra computational burden isn't a big deal, you should run cross-validation.

SKLEARN CLASSIFICATION REPORT

```
>>> from sklearn.metrics import classification_report
>>> y_true = [0, 1, 2, 2, 2]
>>> y_pred = [0, 0, 2, 2, 1]
>>> target_names = ['class 0', 'class 1', 'class 2']
>>> print(classification_report(y_true, y_pred, target_names=target_names))
              precision    recall  f1-score   support

     class 0       0.50      1.00      0.67         1
     class 1       0.00      0.00      0.00         1
     class 2       1.00      0.67      0.80         3

    accuracy                           0.60         5
   macro avg       0.50      0.56      0.49         5
weighted avg       0.70      0.60      0.61         5

>>> y_pred = [1, 1, 0]
>>> y_true = [1, 1, 1]
>>> print(classification_report(y_true, y_pred, labels=[1, 2, 3]))
              precision    recall  f1-score   support

           1       1.00      0.67      0.80         3
           2       0.00      0.00      0.00         0
           3       0.00      0.00      0.00         0

   micro avg       1.00      0.67      0.80         3
   macro avg       0.33      0.22      0.27         3
weighted avg       1.00      0.67      0.80         3
```

https://scikit-learn.org/stable/modules/generated/sklearn.metrics.classification_report.html

GENERALIZATION