# Linear regression

**Univariate and multivariate linear regression formula**

**MAE**

**MSE**

**R²**

Not the formula but know the interpretation.

**Scaling**

Not the formulas but know the different options and know when you better use a specific scaler.

**Regularization**

Know the difference between L1 and L2 (absolute and squared)

# Logistic regression

**Logistic function**

Not the function itself, but know that values are fixed between 0 and 1 and classes are decides based on a certain threshold

**Be able to construct a confusion matrix**

**Accuracy**

**Recall** (also for negatives)

**Precision** (also for negatives)

**F1**

Not the formula, but know how to interpret a specific F1 score (is this a good or a bad score)

**ROC**

TPR (= recall) vs FPR

FPR =

What does it represent + AUC

# Naive bayes

**Posterior**

+ What does each part represent (marginal, prior, likelihood)

* If given a text, be able to calculate the posterior (aka slide 11 in 03 – Naive bayes ppt)

# Decision trees

**Entropy**

Not the formula, but know this is a measurement for the disorder of the dataset

**Gini impurity**

Not the formula, but know this is a measurement for the likelihood to incorrectly classify a new random sample

# Clustering & dimensionality reduction

No formulas to know here.

# Neural networks

**Transfer function**

**Activation functions**

Only know formulas for step, Adaline and Relu

Step: if x<0 : result=0 else result=1

Adaline: result=x

Relu: if x<0: result 0 else result = x

Still know the disadvantages and advantages of all activation functions discussed! When do we use which layer