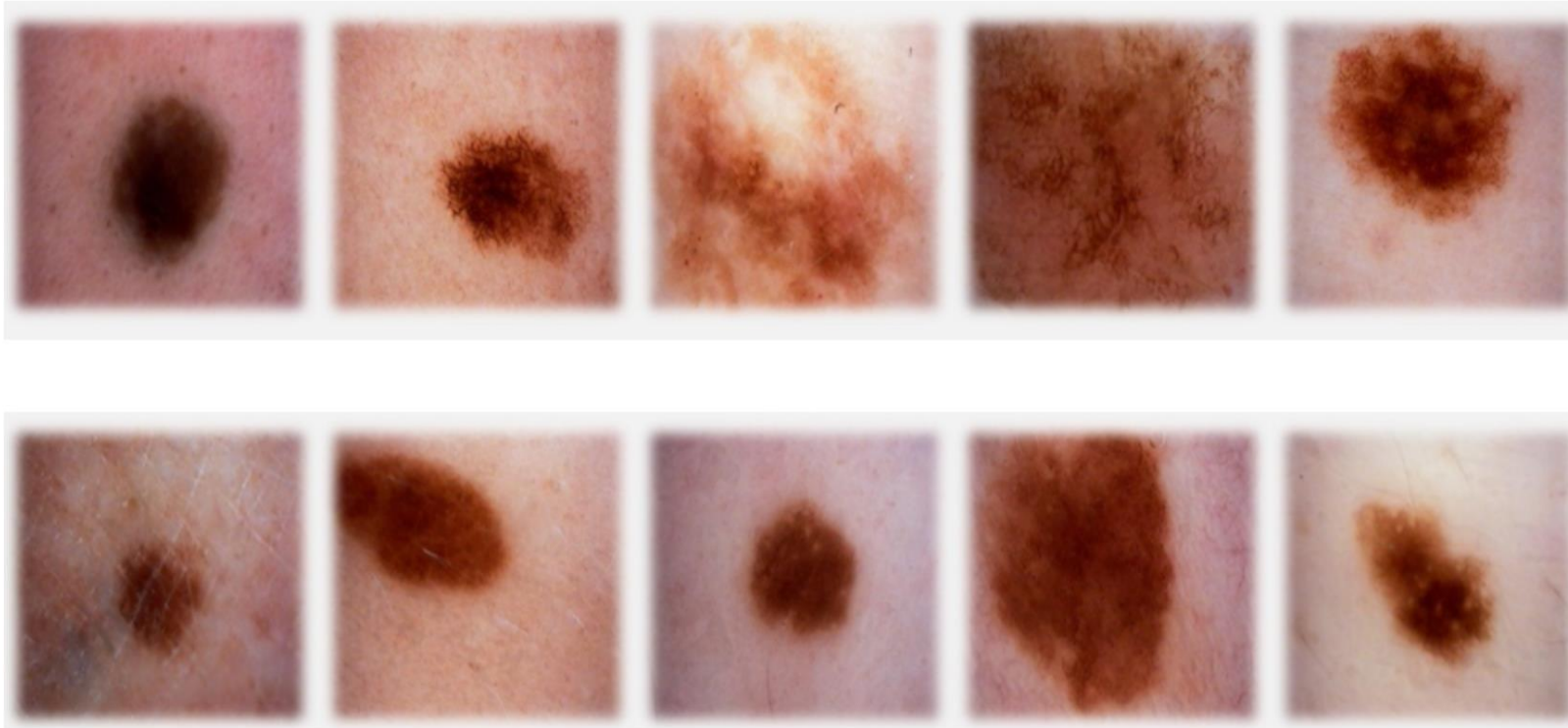


classification



sign of cancer

top row malignant











bottom row benign

classification

skinScan™

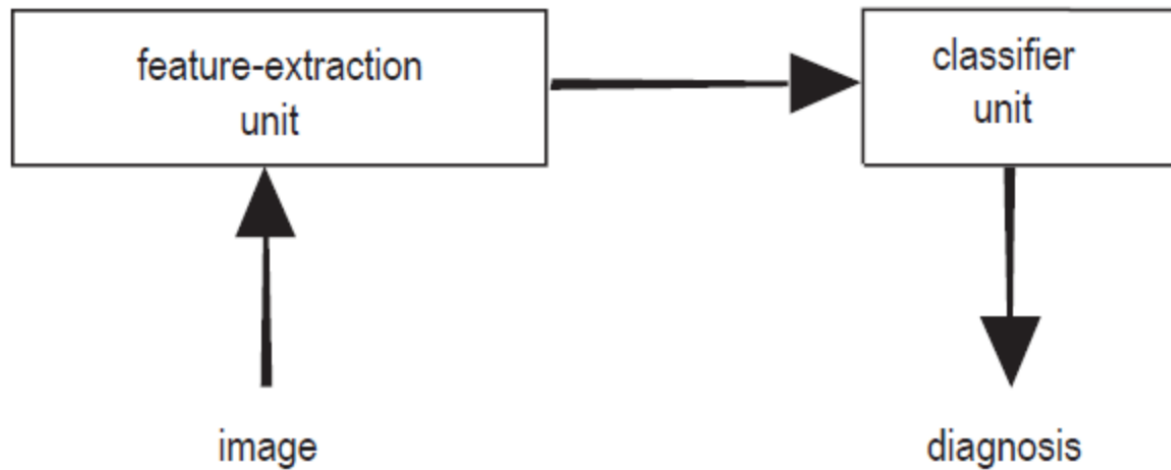
THE ABCDE SYSTEM OF MELANOMA DETECTION

The ABCDE criteria represent a commonly used clinical guide for early diagnosis of melanoma. The following features are considered suspicious:

A	Asymmetry: Moles that have asymmetrical appearance		
		Symmetrical	Asymmetrical
B	Border: A mole that has blurry and/or jagged edges		
		Smooth borders	Irregular borders
C	Color: A mole that has more than one colour		
		Single color	Multicolor
D	Diameter: Moles with a diameter larger than a pencil eraser (6 mm or 1/4 inch)		
		Smaller than 6mm/0.2in	Bigger than 6mm/0.2in
E	Evolution: A mole that has gone through sudden changes in size, shape or colour		
		No changes	Some changes

TeleSkin © 2013

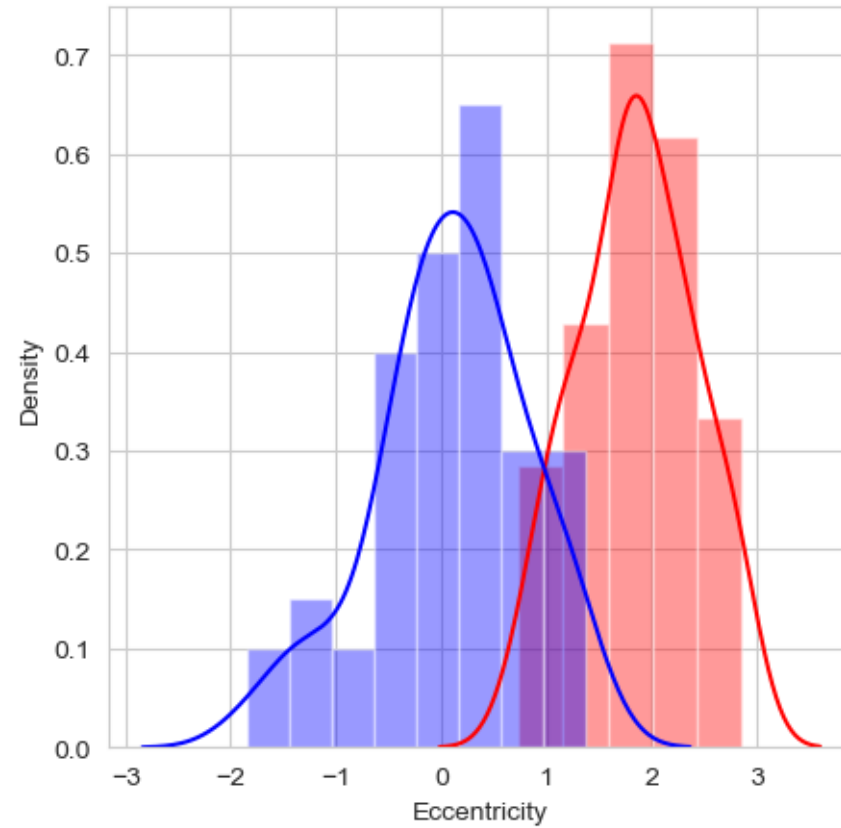
classification: terminology



feature extraction: features (a.k.a. properties or attributes)

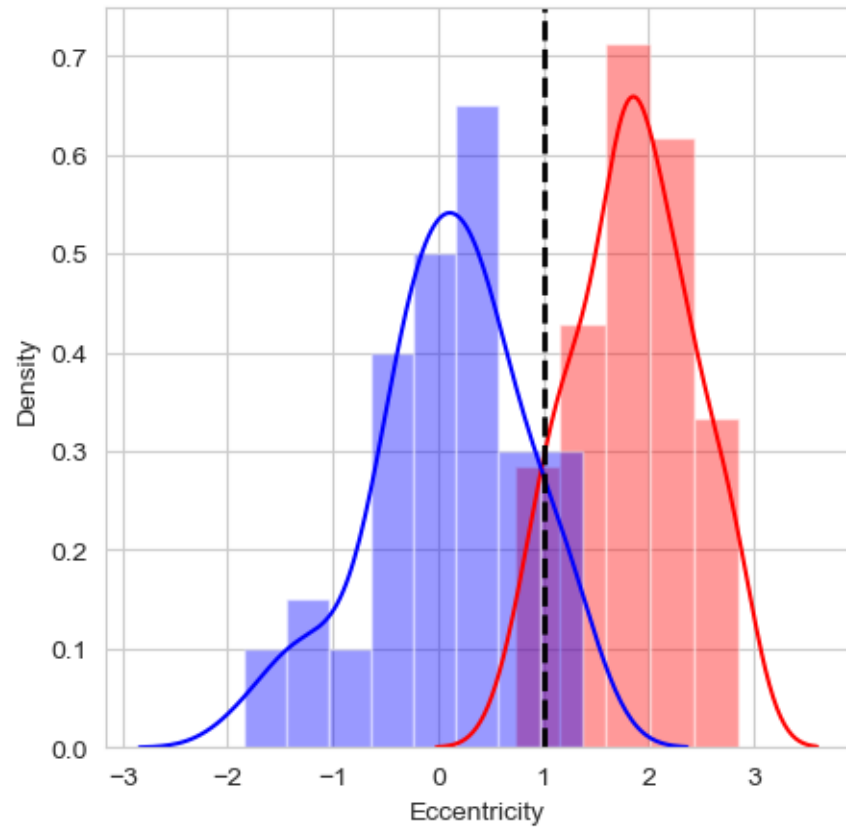
data set, sample (a.k.a. example, instance or data point), label (a.k.a. target)

classification: a feature



feature: eccentricity of lesion (how nearly circular the lesion is)

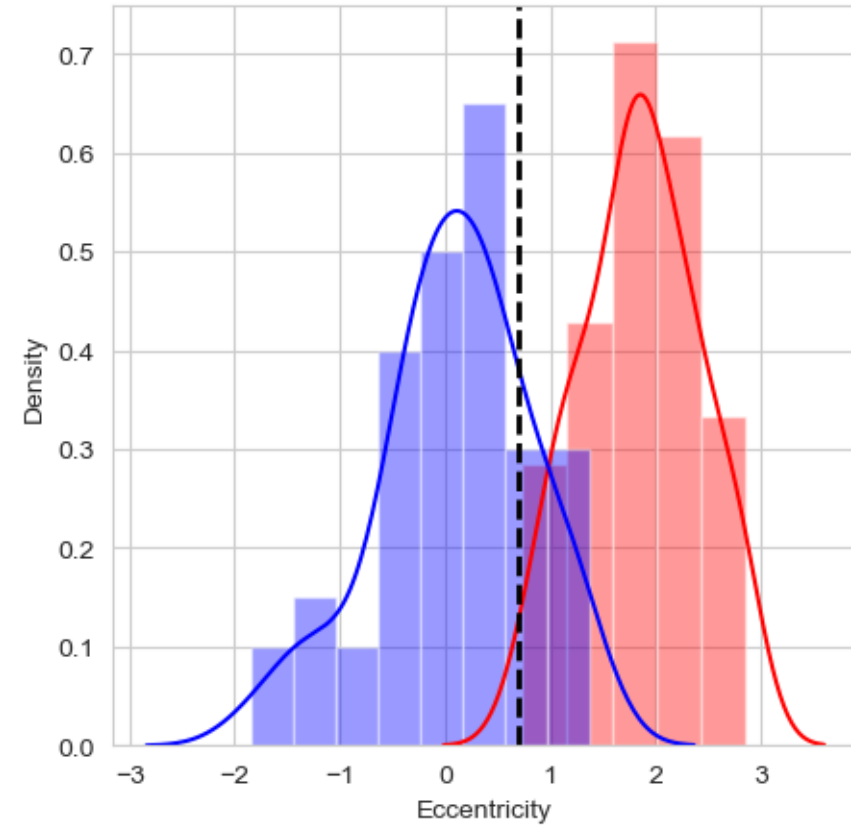
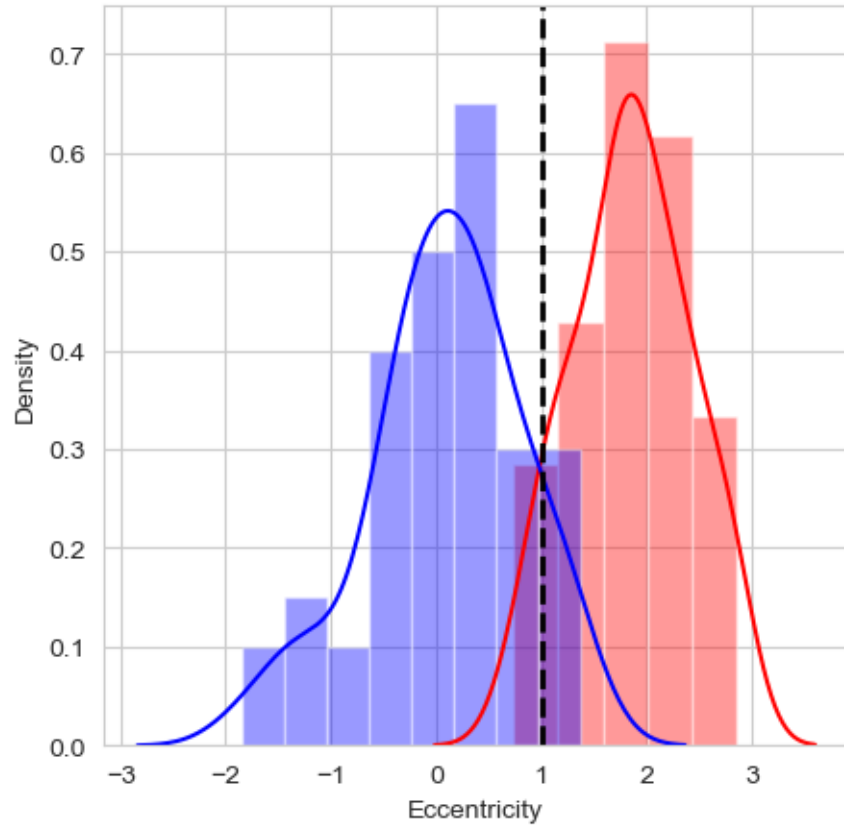
classification: the model



feature: eccentricity of lesion (how nearly circular the lesion is)

model: threshold t

classification: the model



feature: eccentricity of lesion (how nearly circular the lesion is)

model: threshold t : consequence of the predictions

classification: prediction errors

malignant: **positive** class

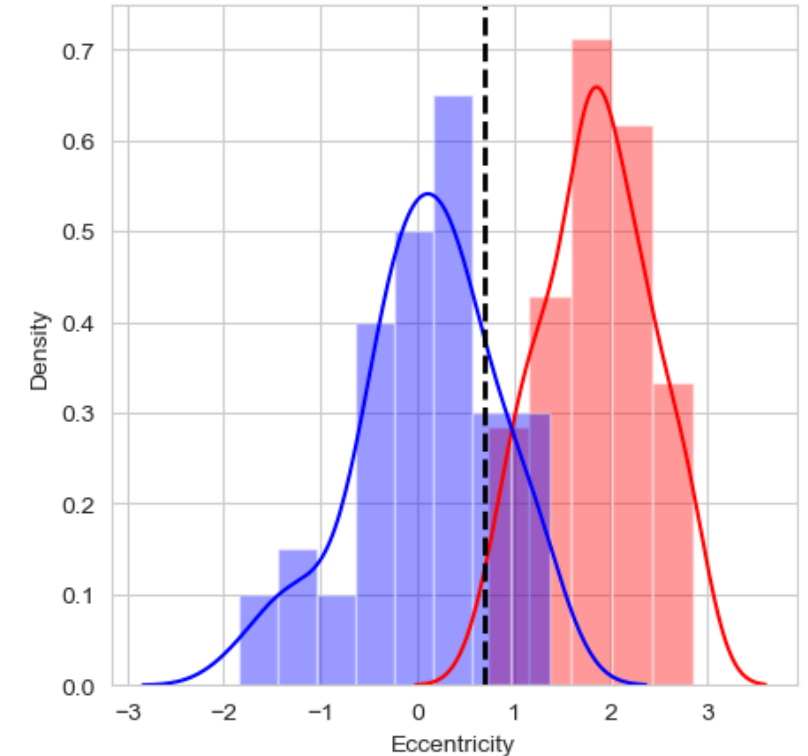
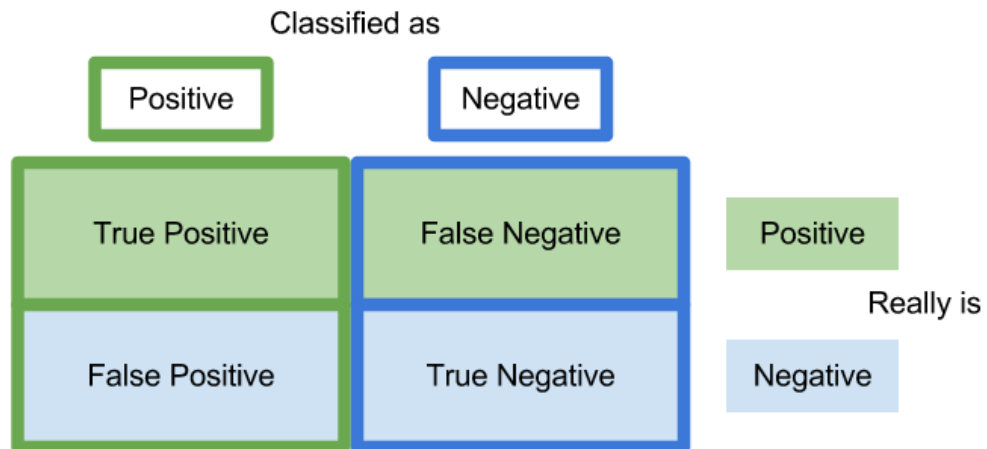
benign: **negative** class

count the number of malignant images with eccentricity value $\geq t$: **true positive** predictions (TP)

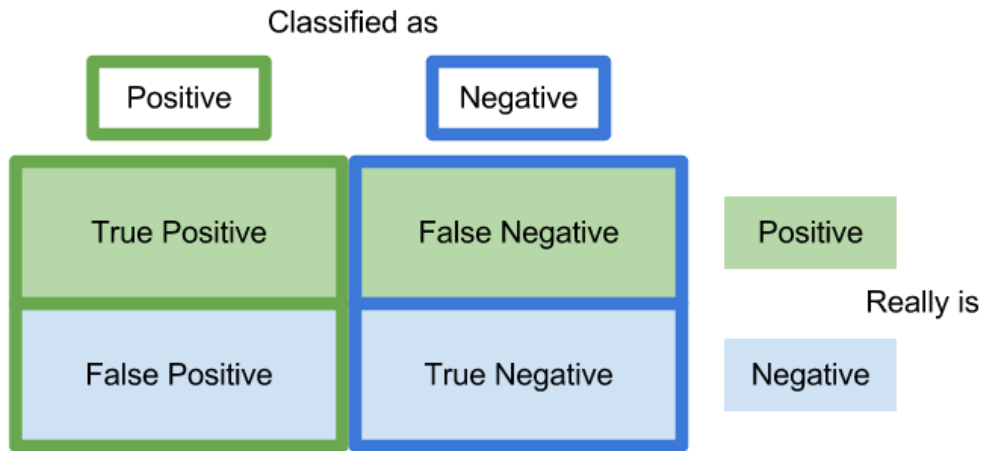
count the number of malignant images with eccentricity value $< t$: **false negative** predictions (FN)

count the number of benign images with eccentricity value $\geq t$: **false positive** predictions (FP)

count the number of benign images with eccentricity value $< t$: **true negative** predictions (TN)



classification: prediction errors

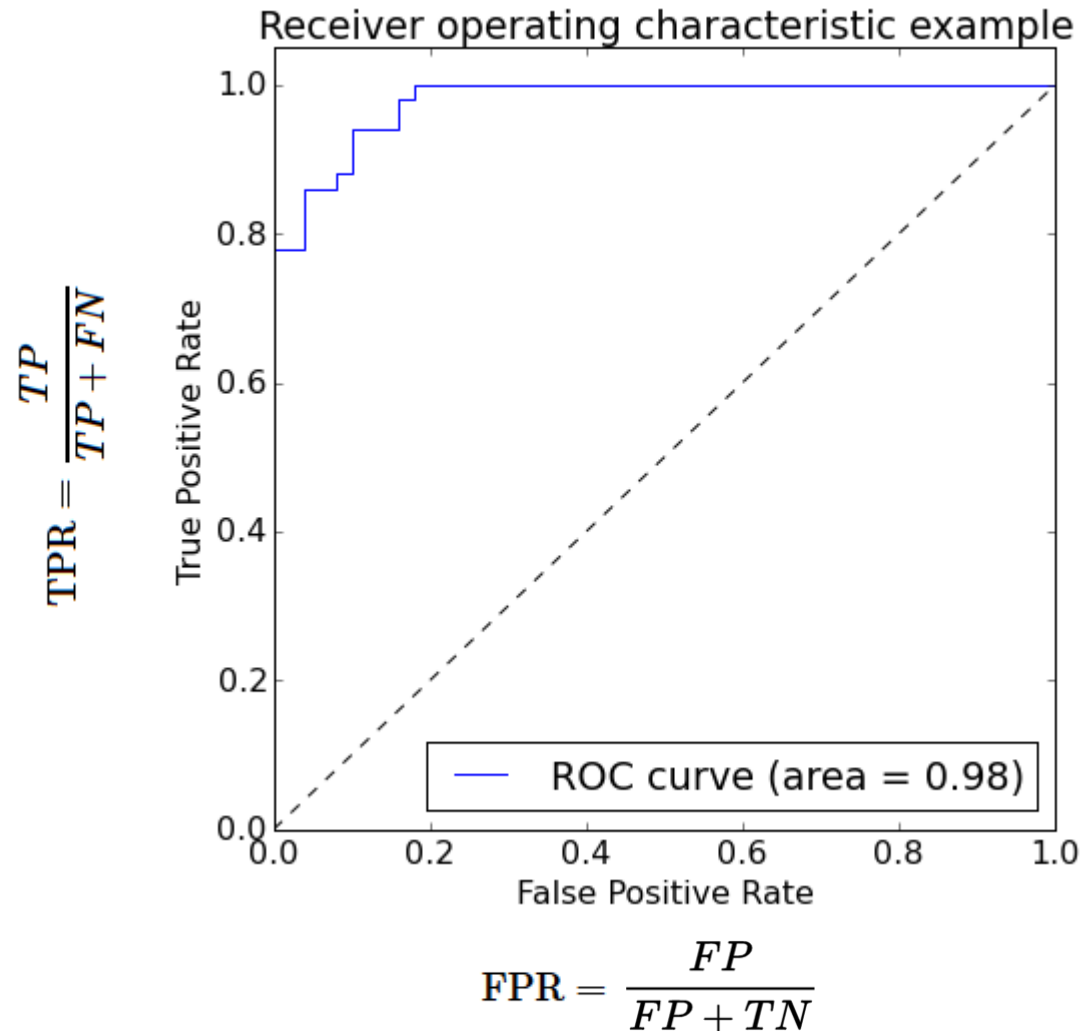


$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{TPR} = \frac{TP}{TP + FN}$$

$$\text{FPR} = \frac{FP}{FP + TN}$$

classification: prediction errors



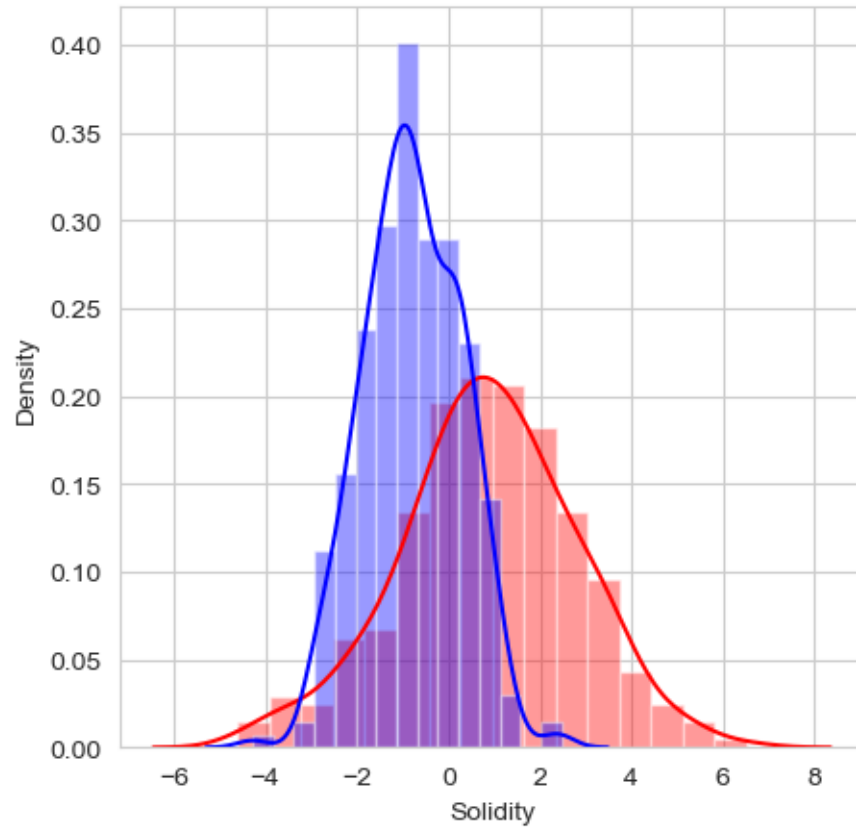
model that classifies all images as malignant:
TPR=1 and FPR=1

model that classifies all images a benign:
TPR=0 and FPR=0

vary threshold t

Area Under the Curve (AUC)

classification: multi-dimensional

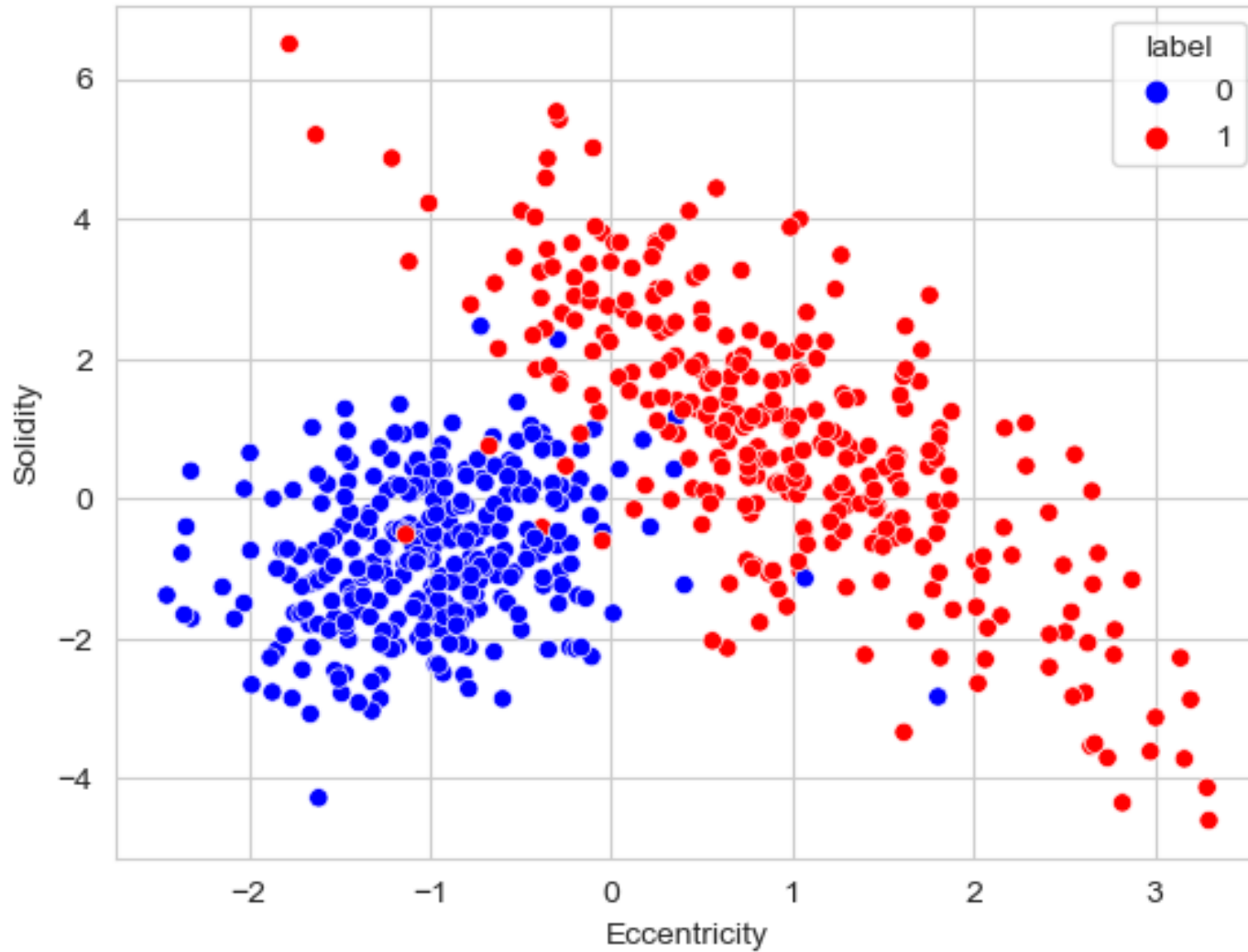


add another feature?

feature vector X

Euclidean vector space

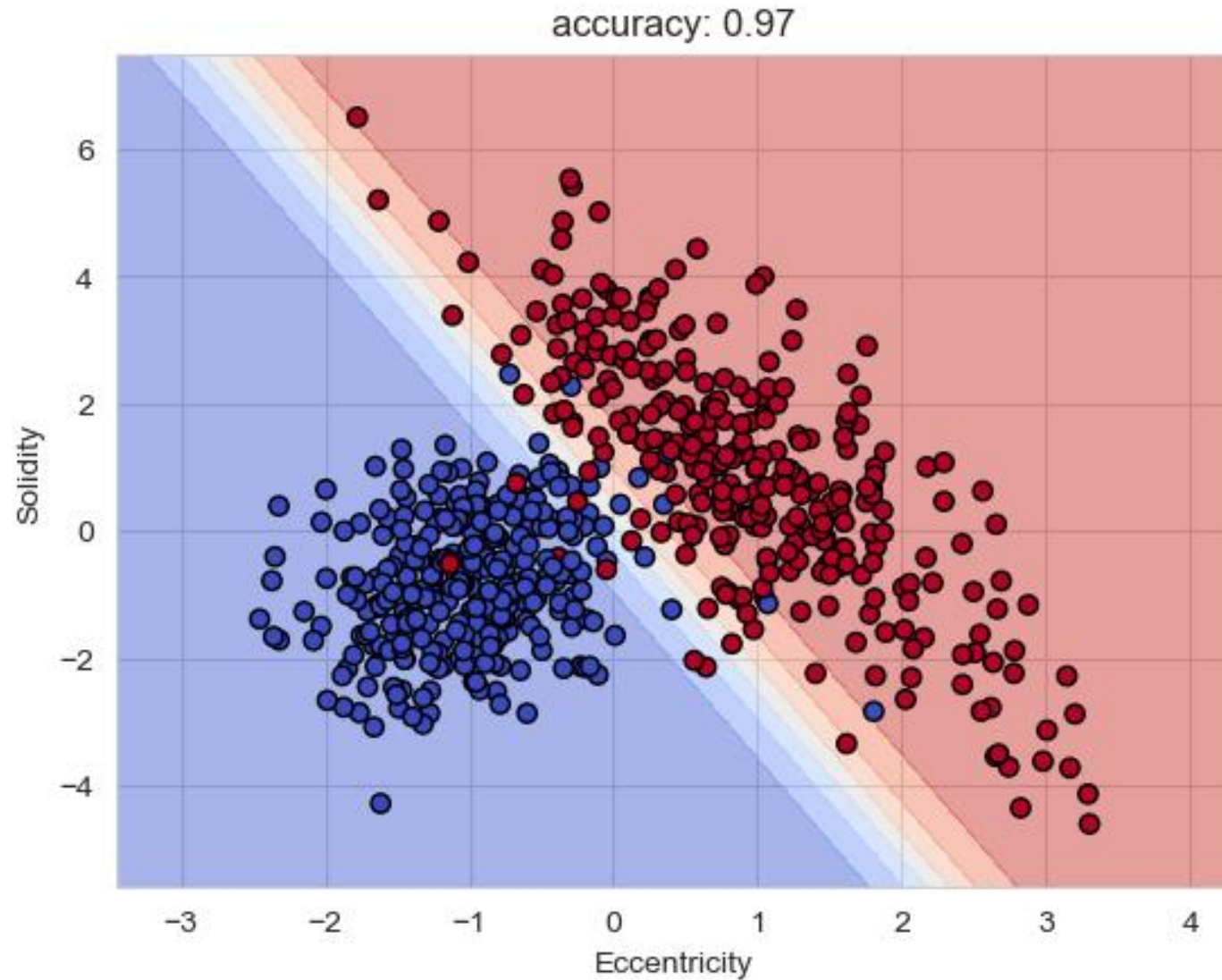
classification: multi-dimensional



feature vector X

Euclidean vector space

classification: multi-dimensional



linear decision boundary

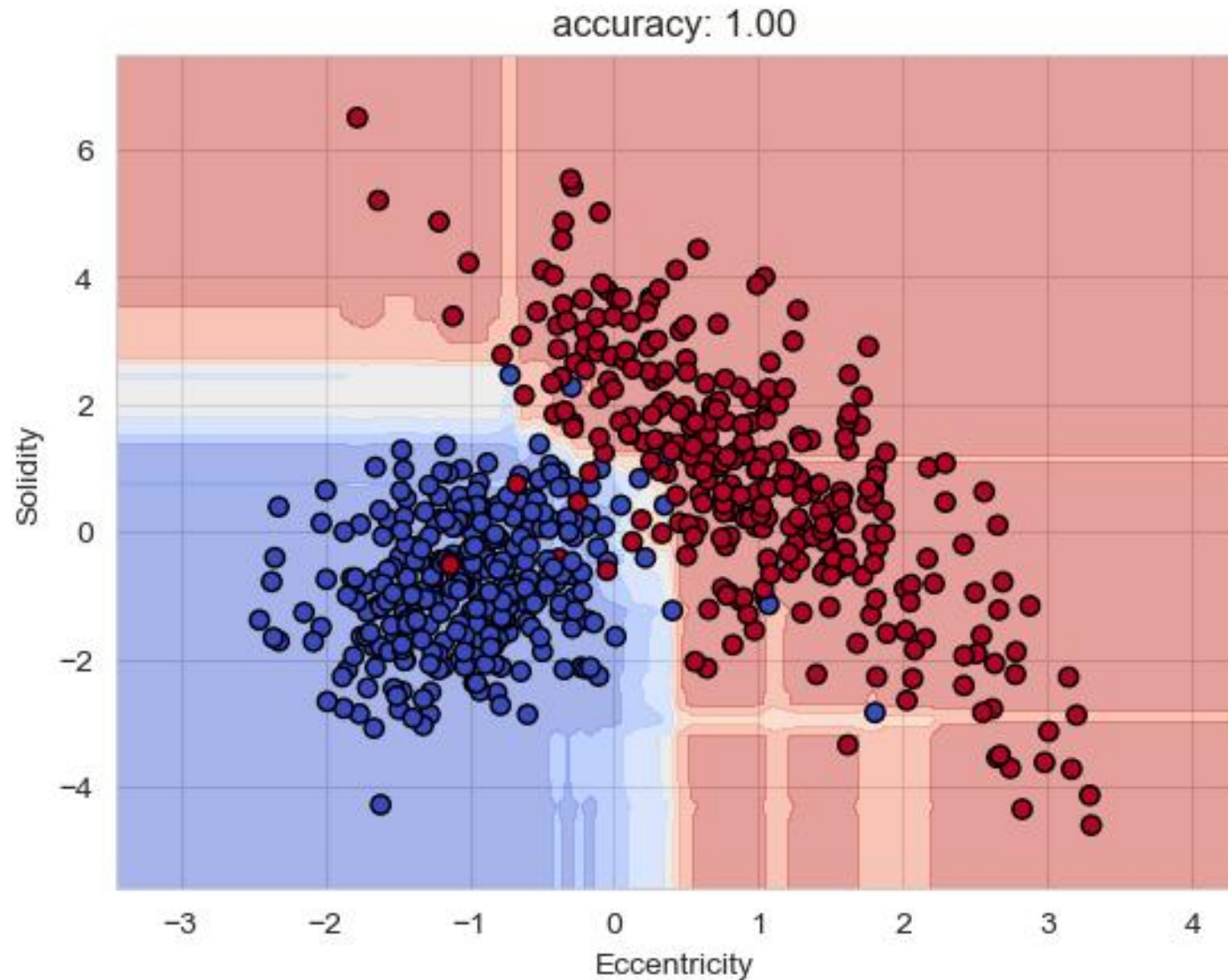
blue region malignant, red region benign

yet more features

can't look at the decision boundary

more complex

classification: model complexity

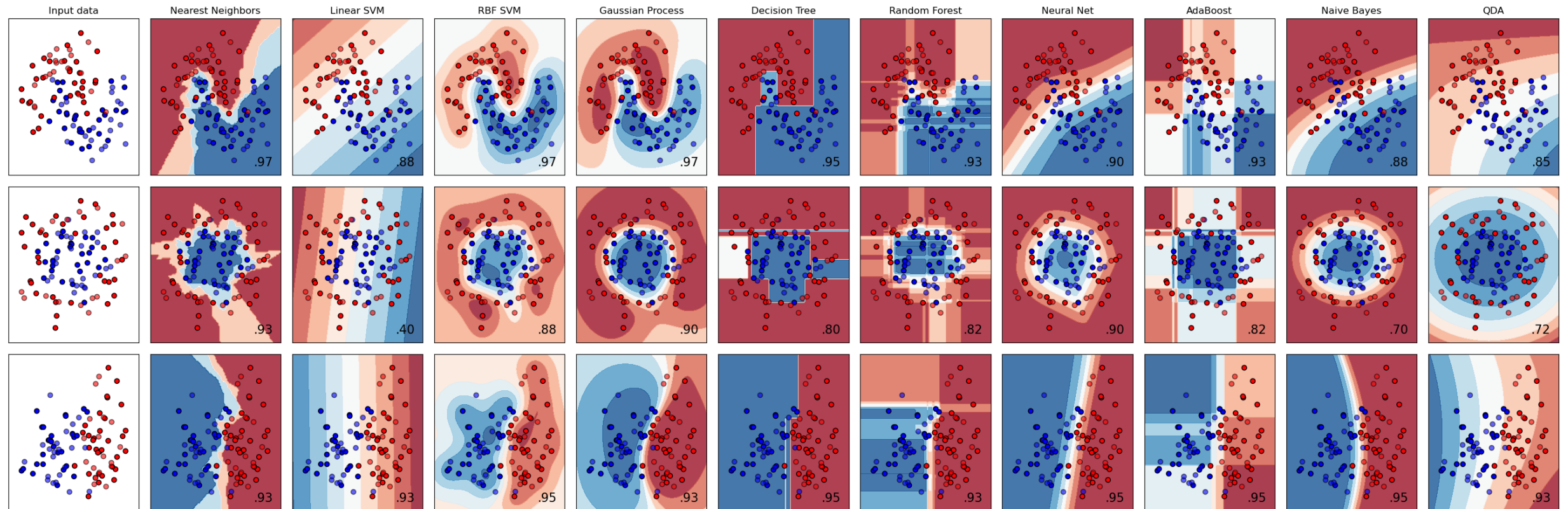


unseen external images

generalization

overfitting

scikit-learn



data normalization

make all features same scale

Eccentricity [0,100], Solidity [-5,7]

weights all features equally in their representation

standardization

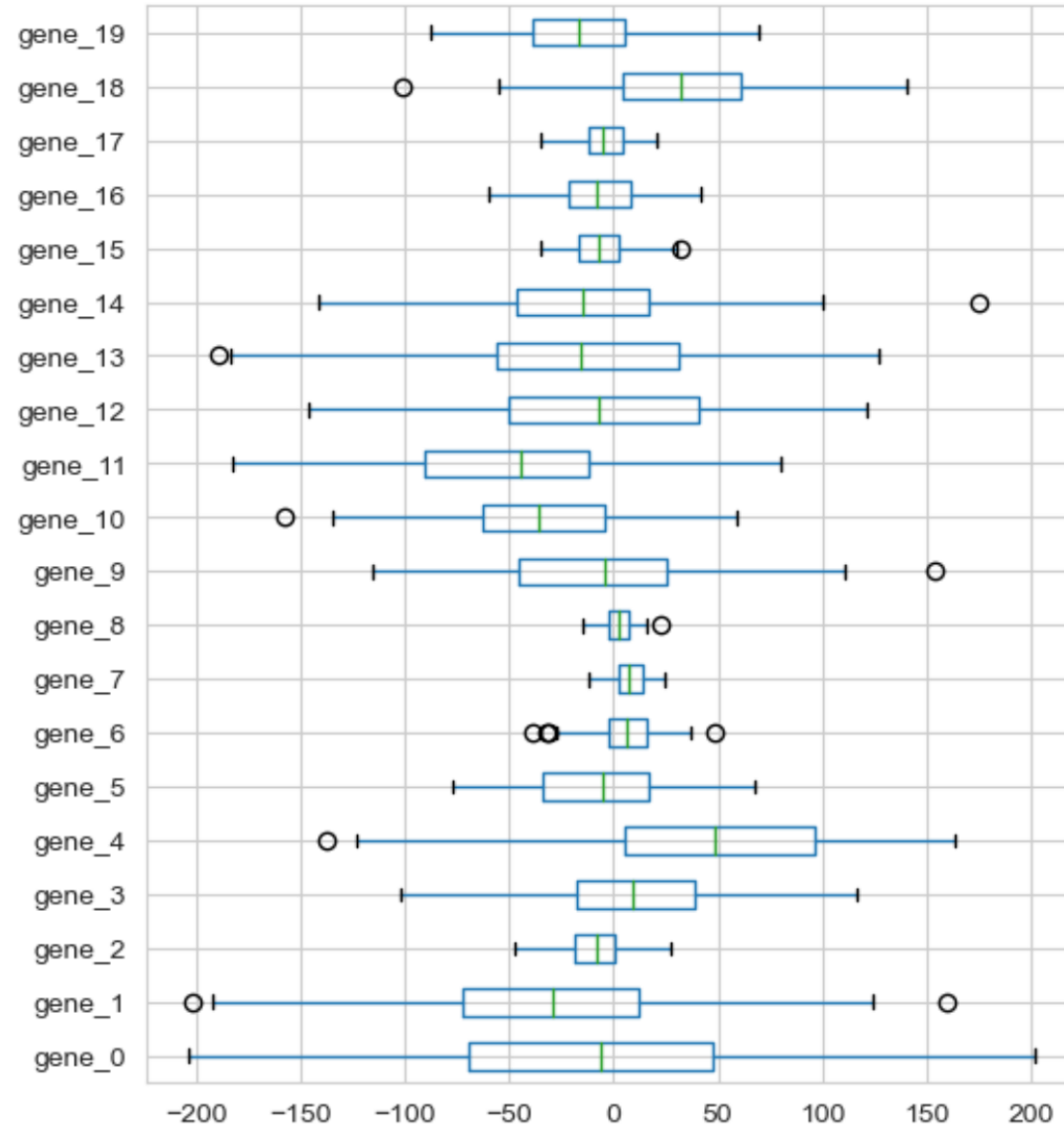
$$\mu = 0 \quad \sigma = 1$$

$$x_{norm} = \frac{x - \mu}{\sigma}$$

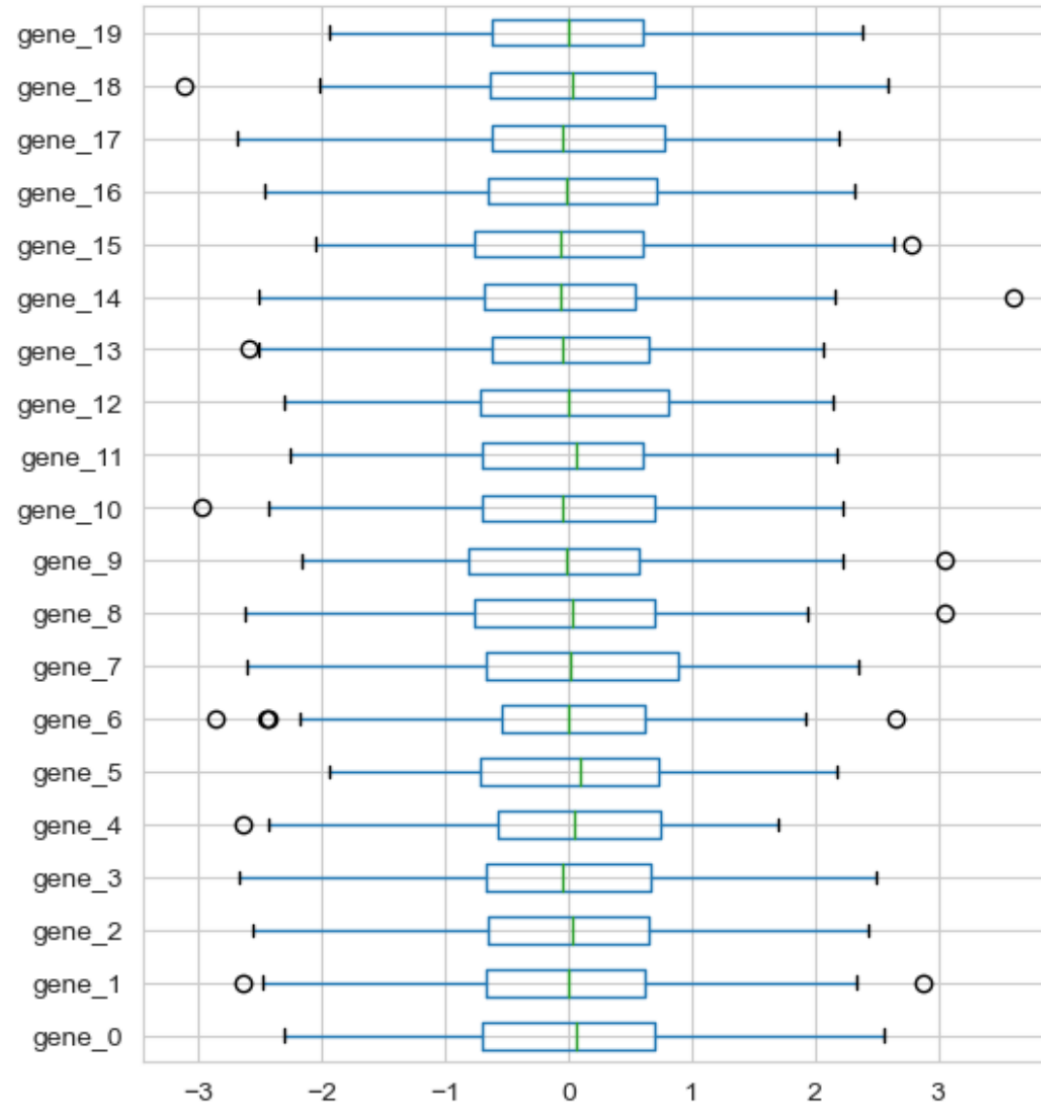
$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

min-max scaling: scale the features to a fixed range

data normalization

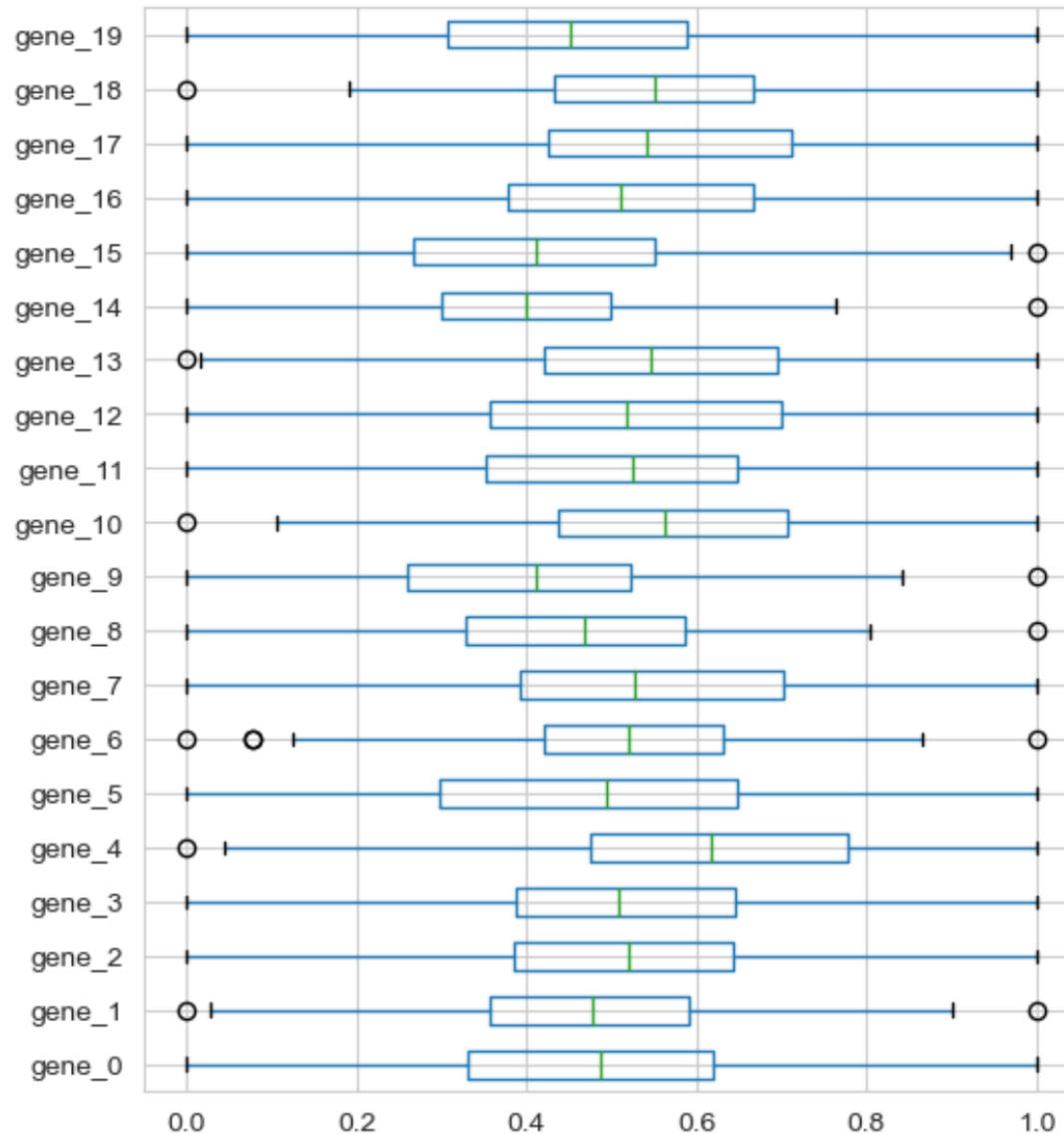


data normalization: standardization



$$x_{norm} = \frac{x - \mu}{\sigma}$$

data normalization: min-max scaling



$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$$