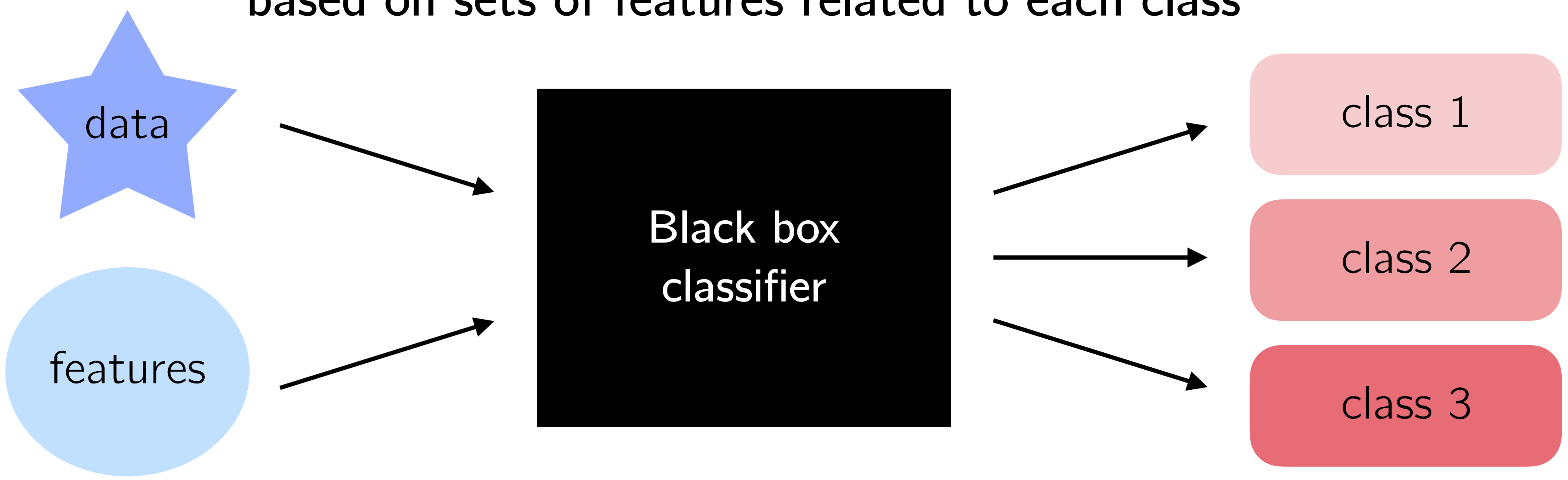
Generative Classification

Katie Chamberlain

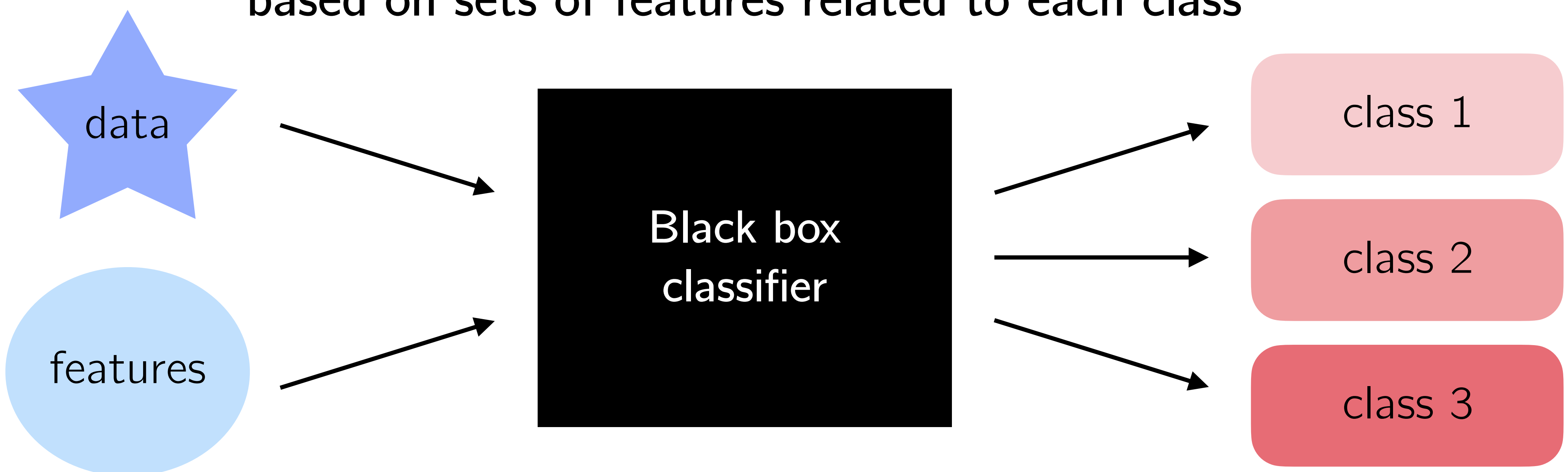
What is a classifier?

categorizes data into predefined sets of classes
based on sets of features related to each class

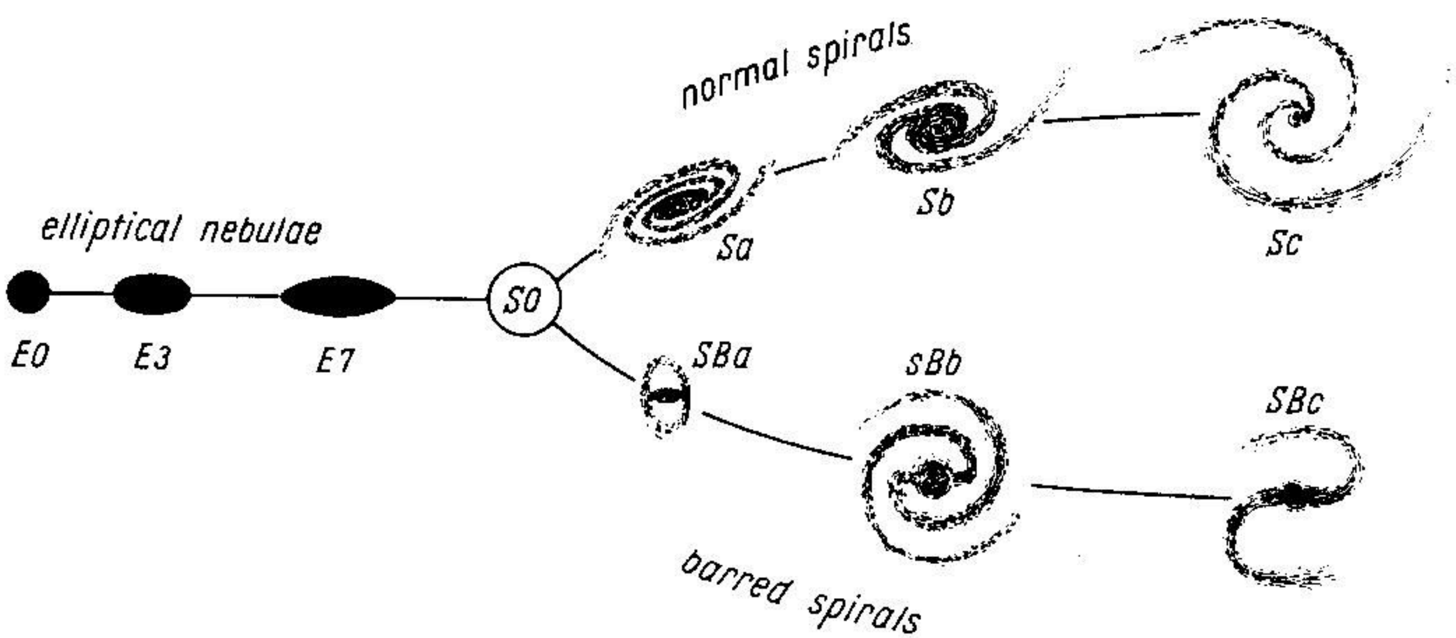


What is a classifier?

categorizes data into predefined sets of classes
based on sets of features related to each class



Edwin Hubble was a classifier!

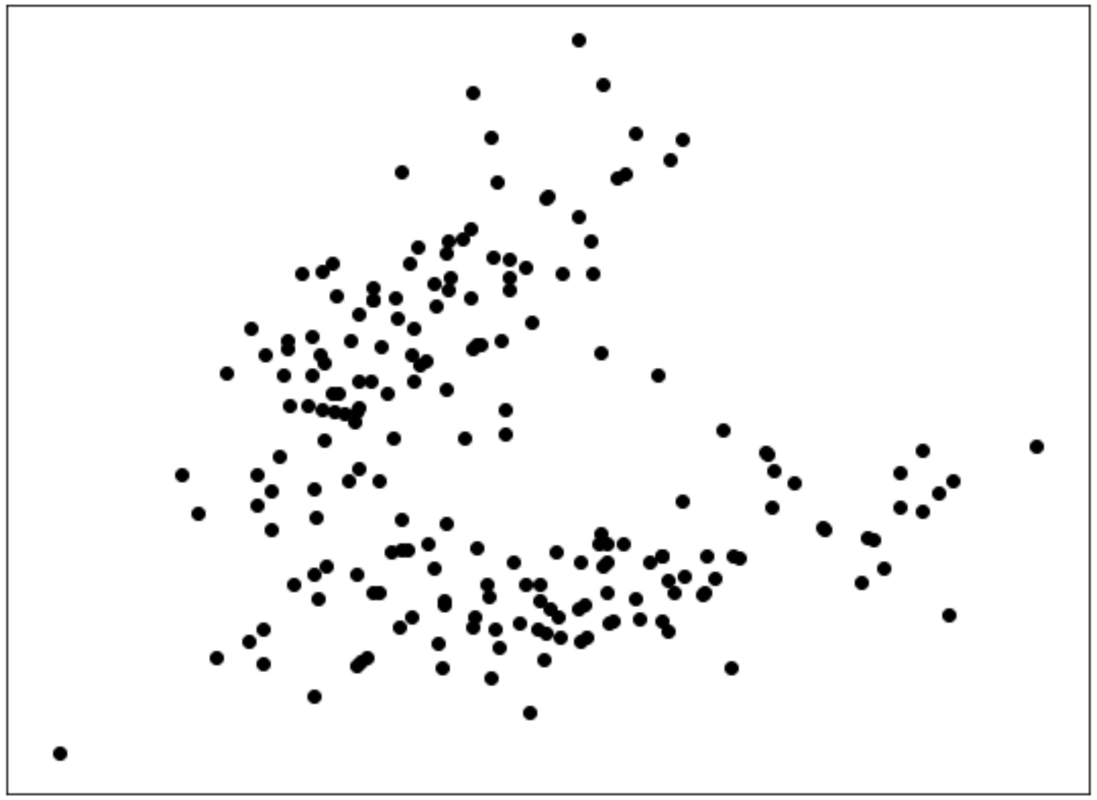


He separated observed galaxies into different classes

(spiral, elliptical, etc)

Unsupervised vs. supervised classification

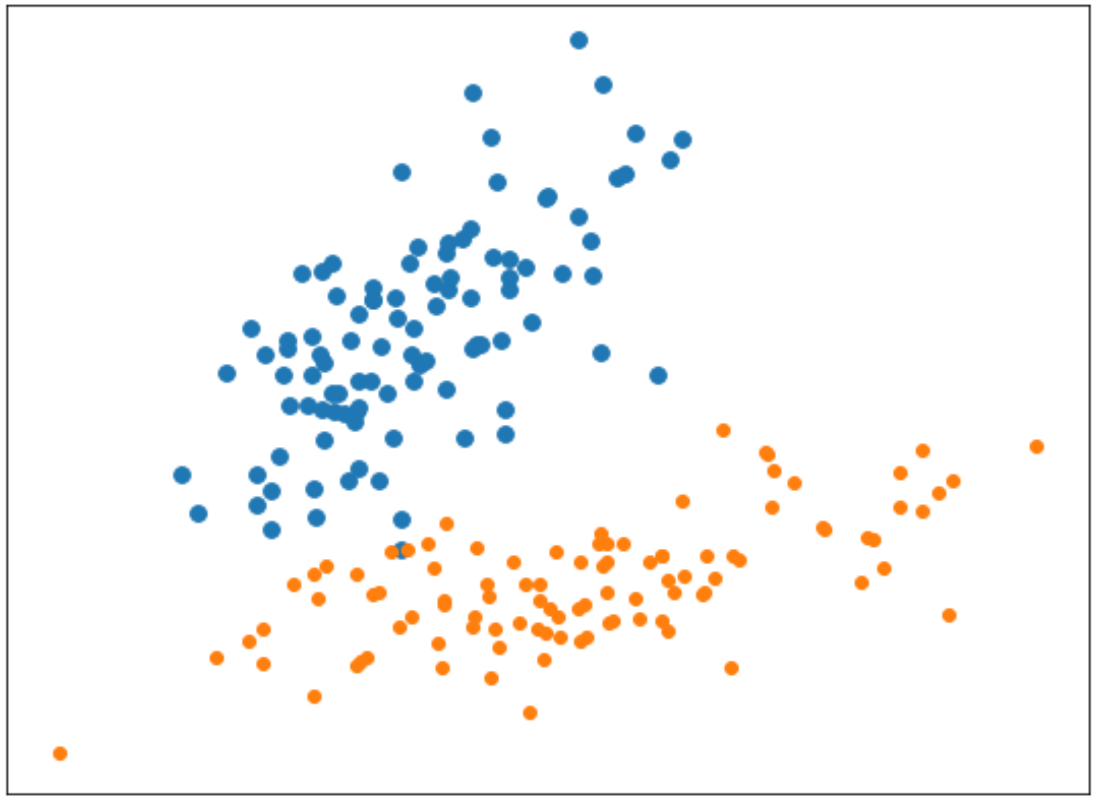
Unsupervised



Supervised

Unsupervised vs. supervised classification

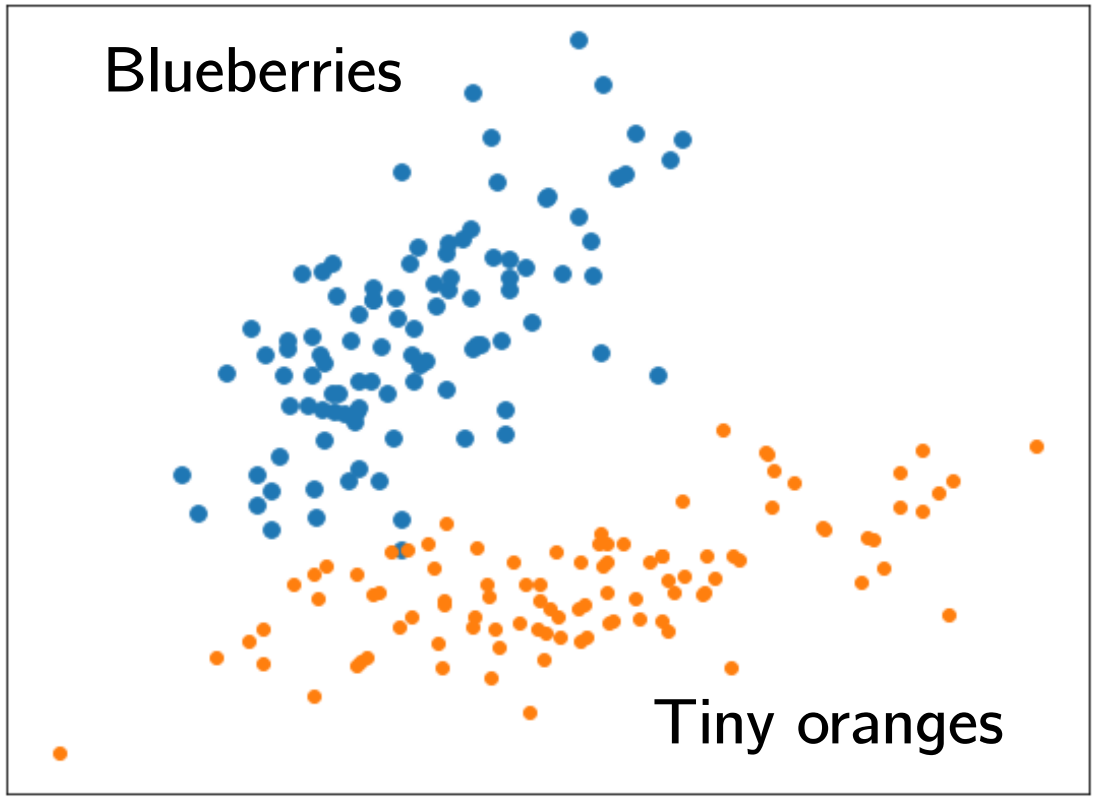
Unsupervised



Supervised

Unsupervised vs. supervised classification

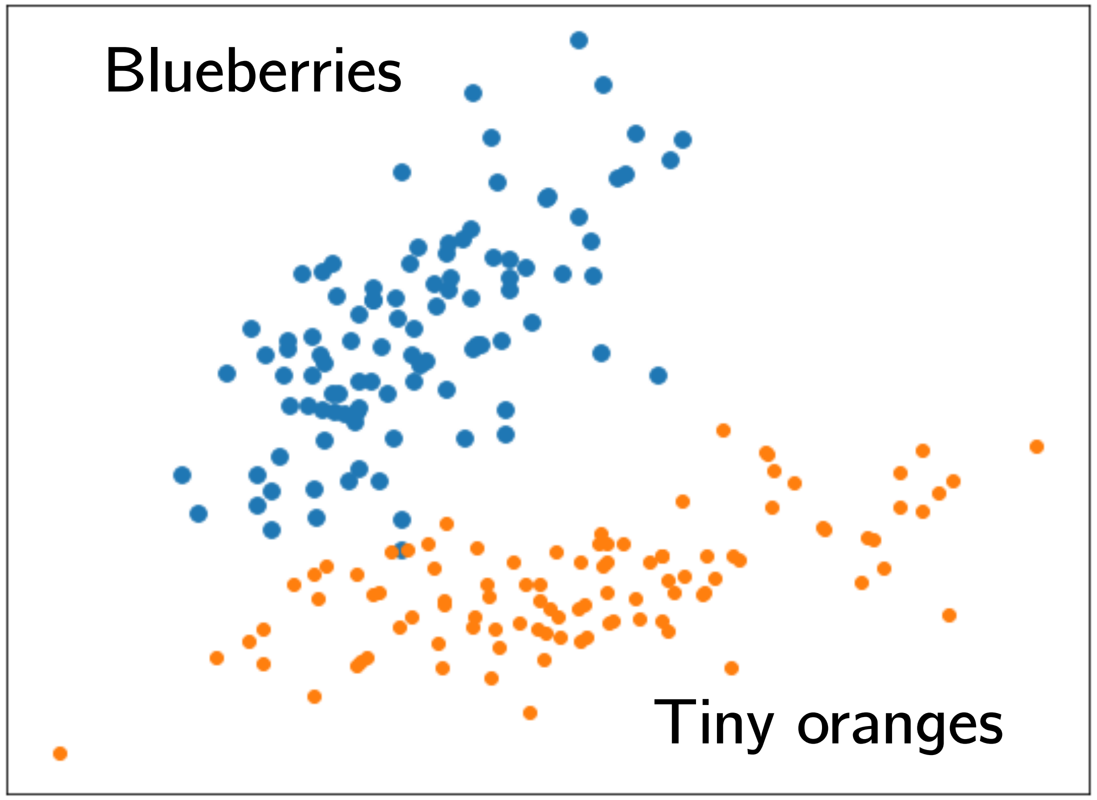
Unsupervised



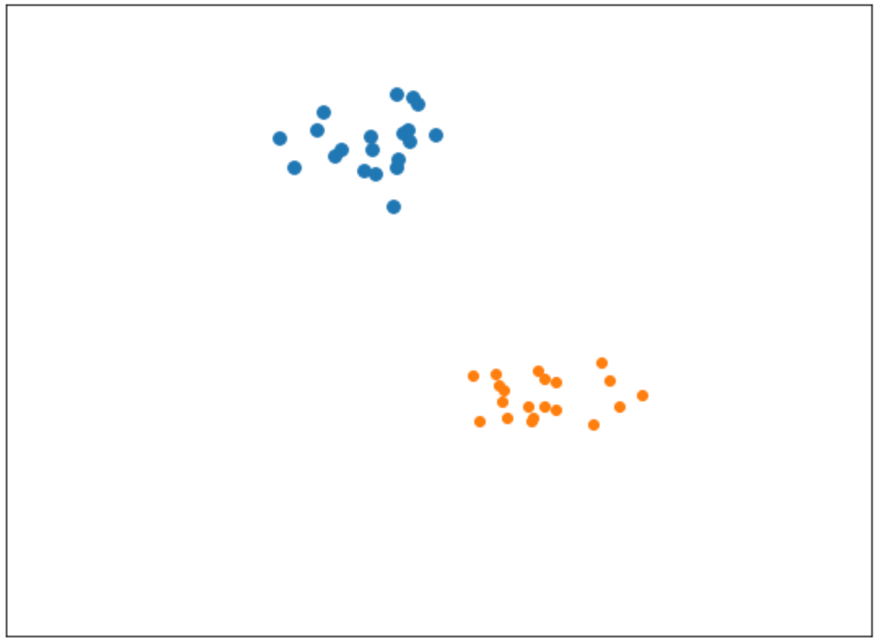
Supervised

Unsupervised vs. supervised classification

Unsupervised

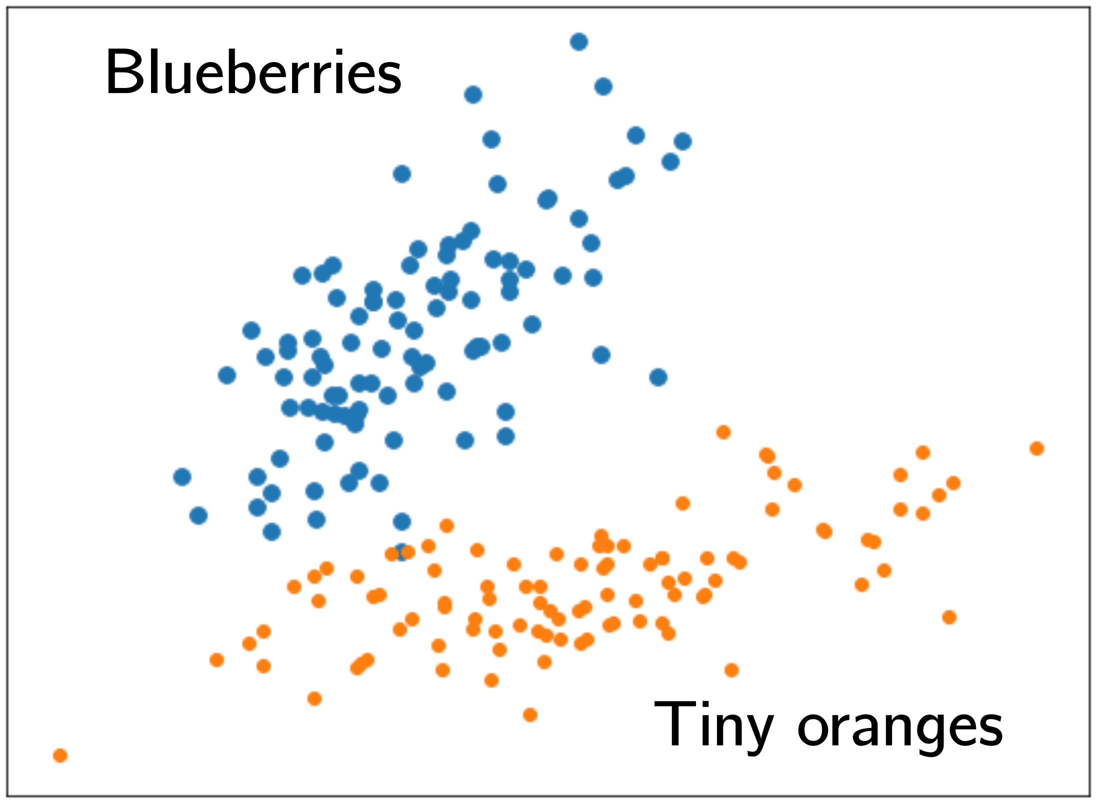


Supervised

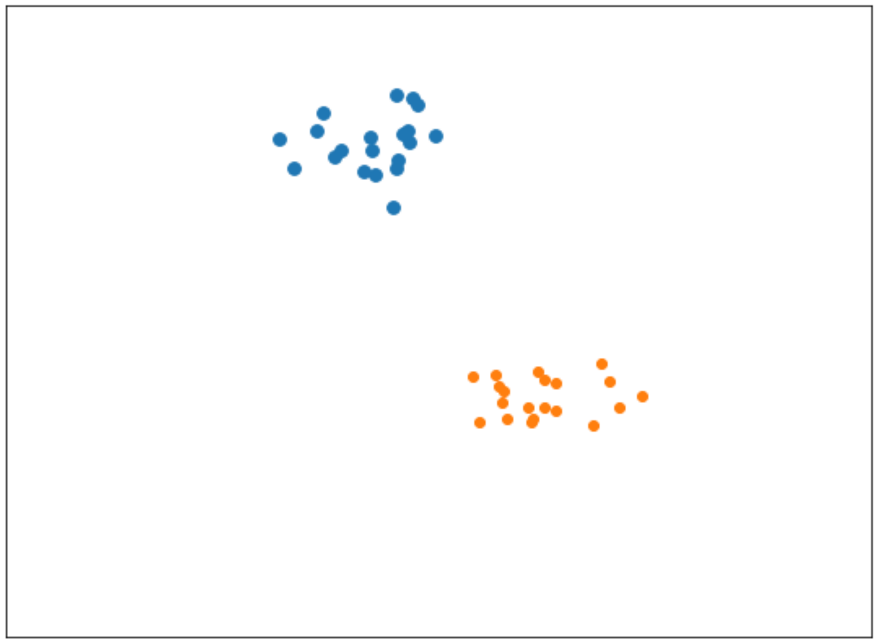


Unsupervised vs. supervised classification

Unsupervised

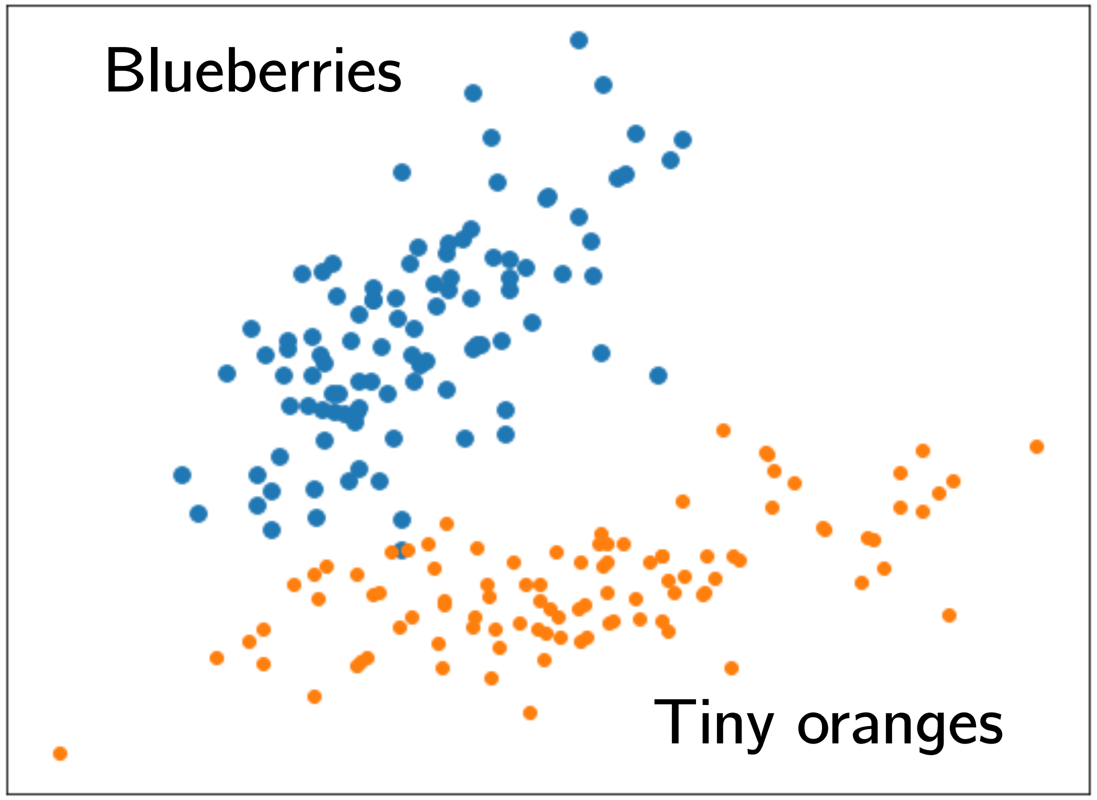


Supervised

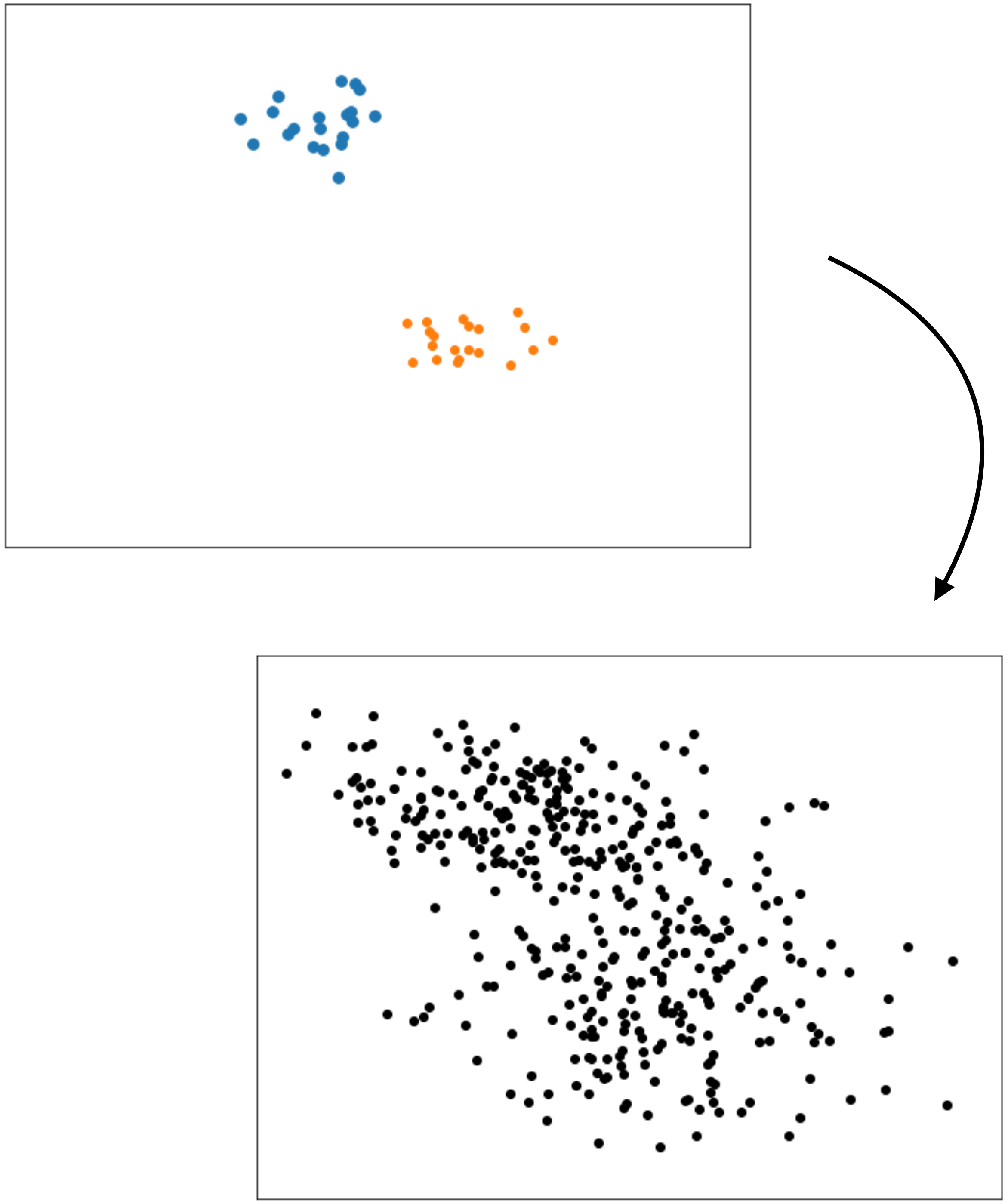


Unsupervised vs. supervised classification

Unsupervised

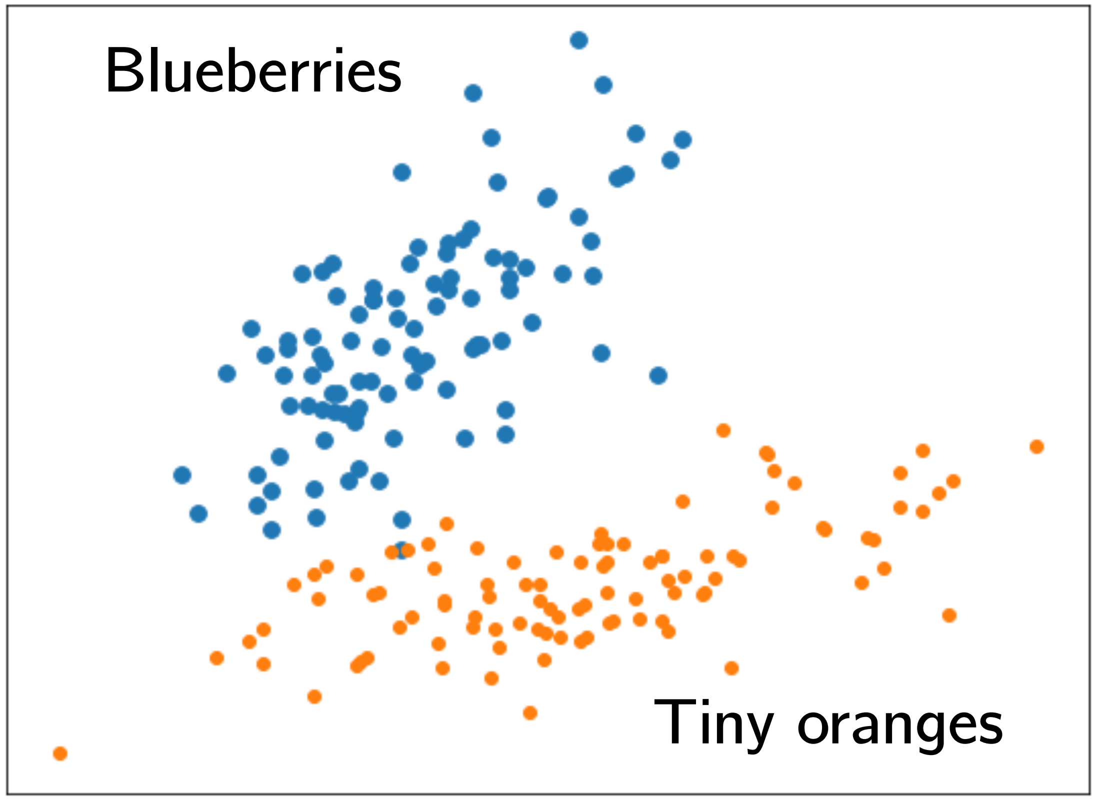


Supervised

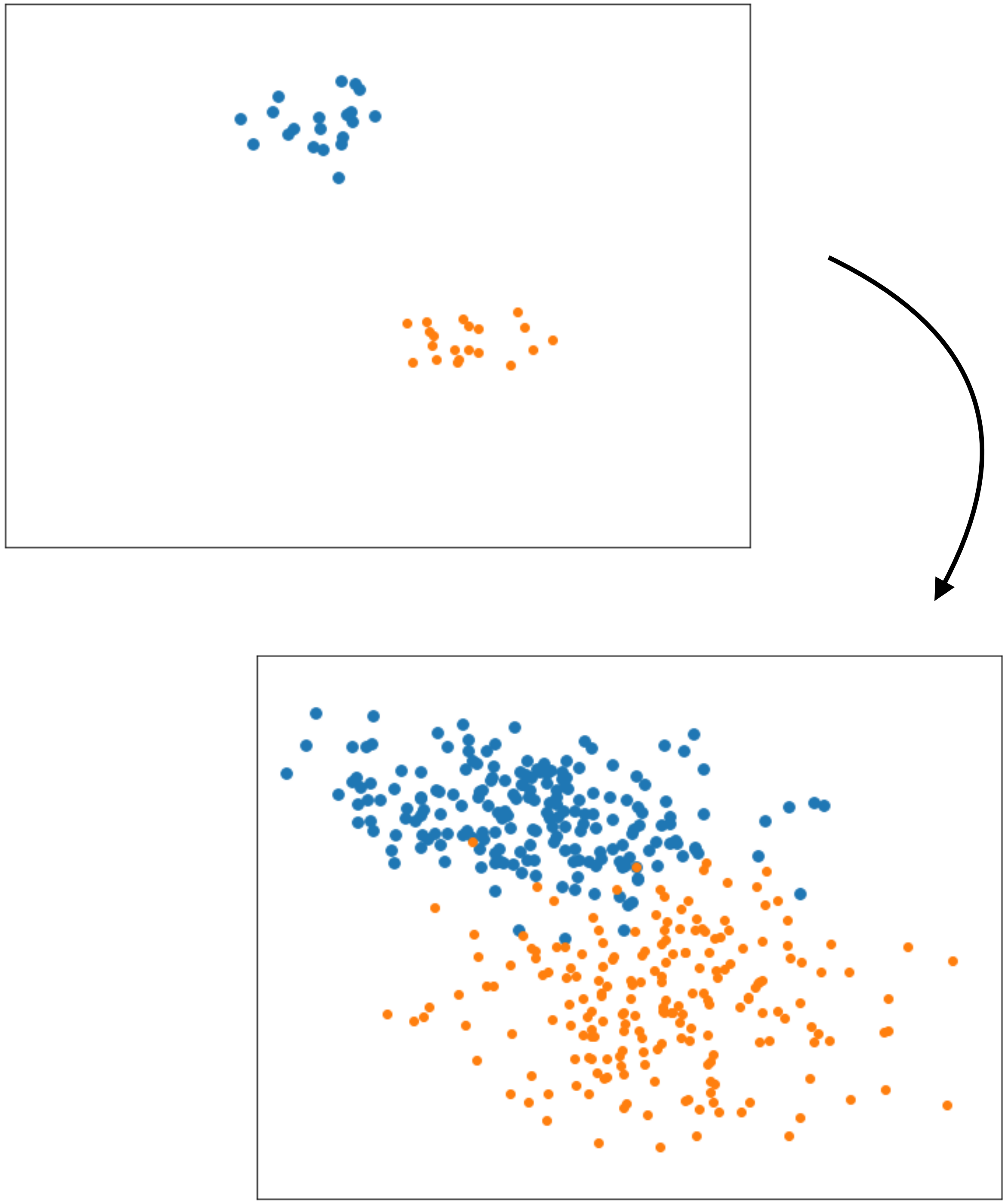


Unsupervised vs. supervised classification

Unsupervised



Supervised



Unsupervised vs. supervised classification

Unsupervised

Supervised

<u>Our goal</u>

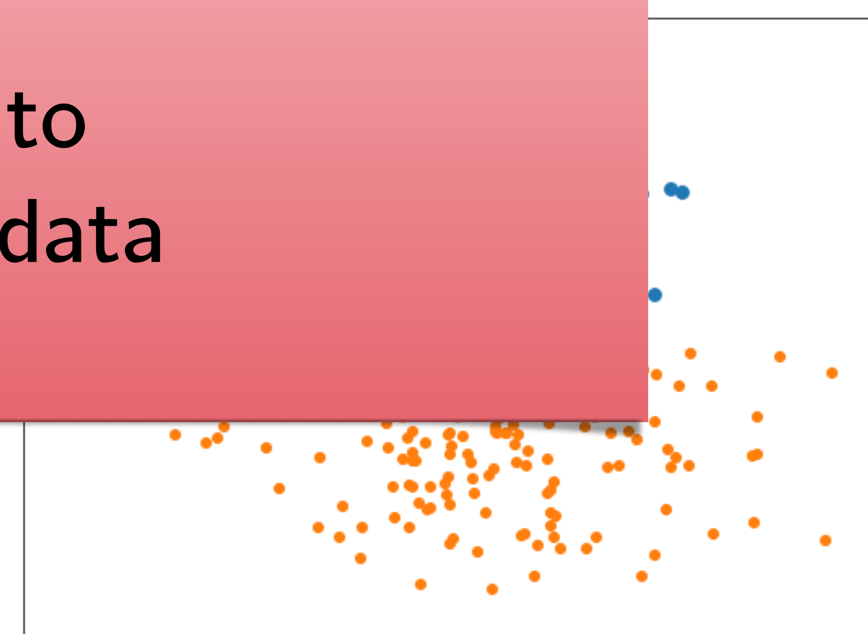
characterize the relation between:

• features in the data

• their classes

apply classifications to

large set of unlabeled data



What is a generative model?

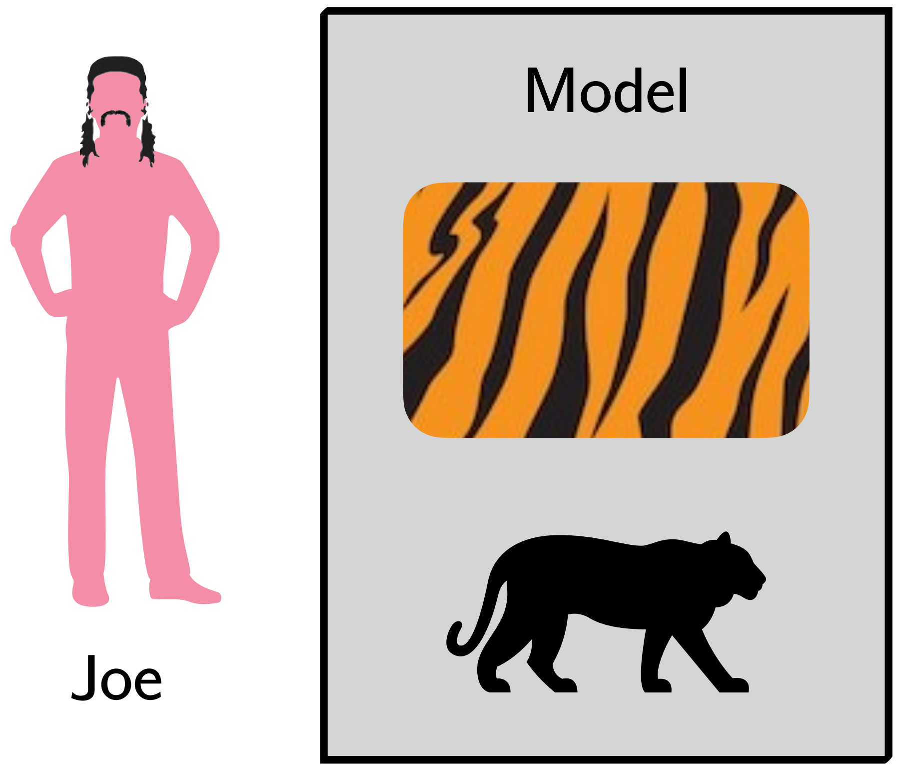
Generative classification is based on density estimation:

This just means that we have a model that
describes how data can be generated from each class

What is a generative model?

Generative classification is based on density estimation:

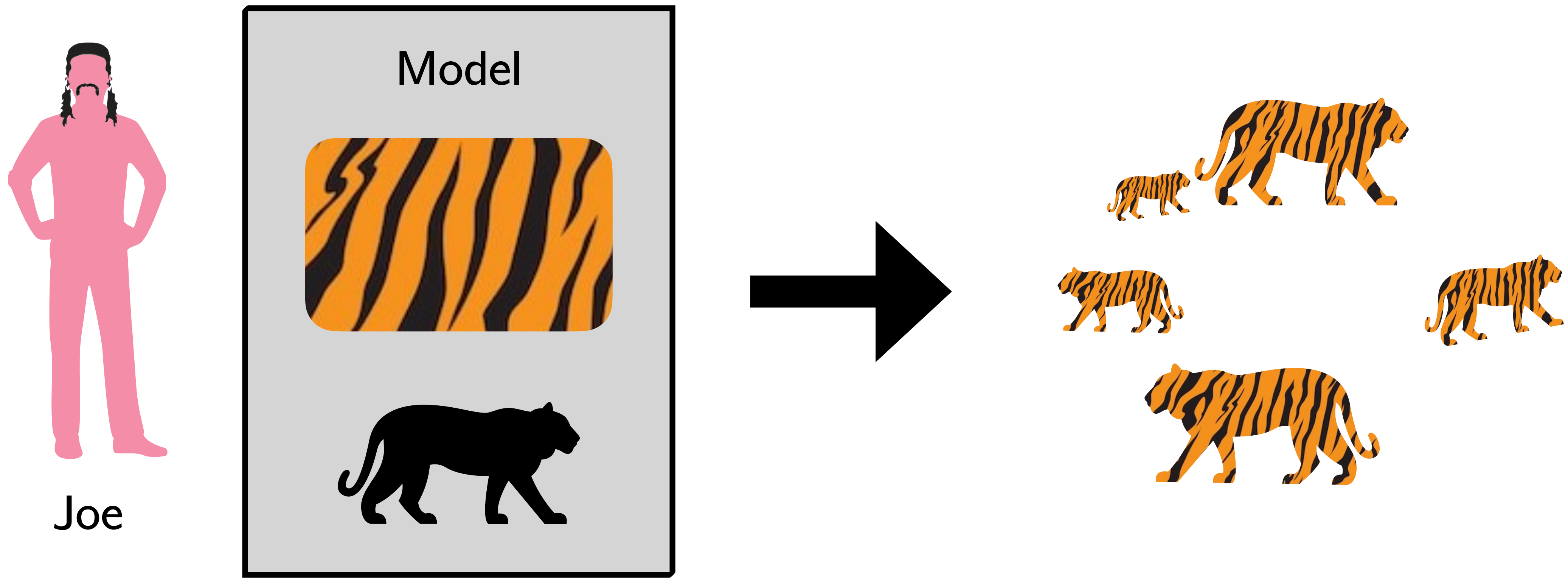
This just means that we have a model that
describes how data can be generated from each class



What is a generative model?

Generative classification is based on density estimation:

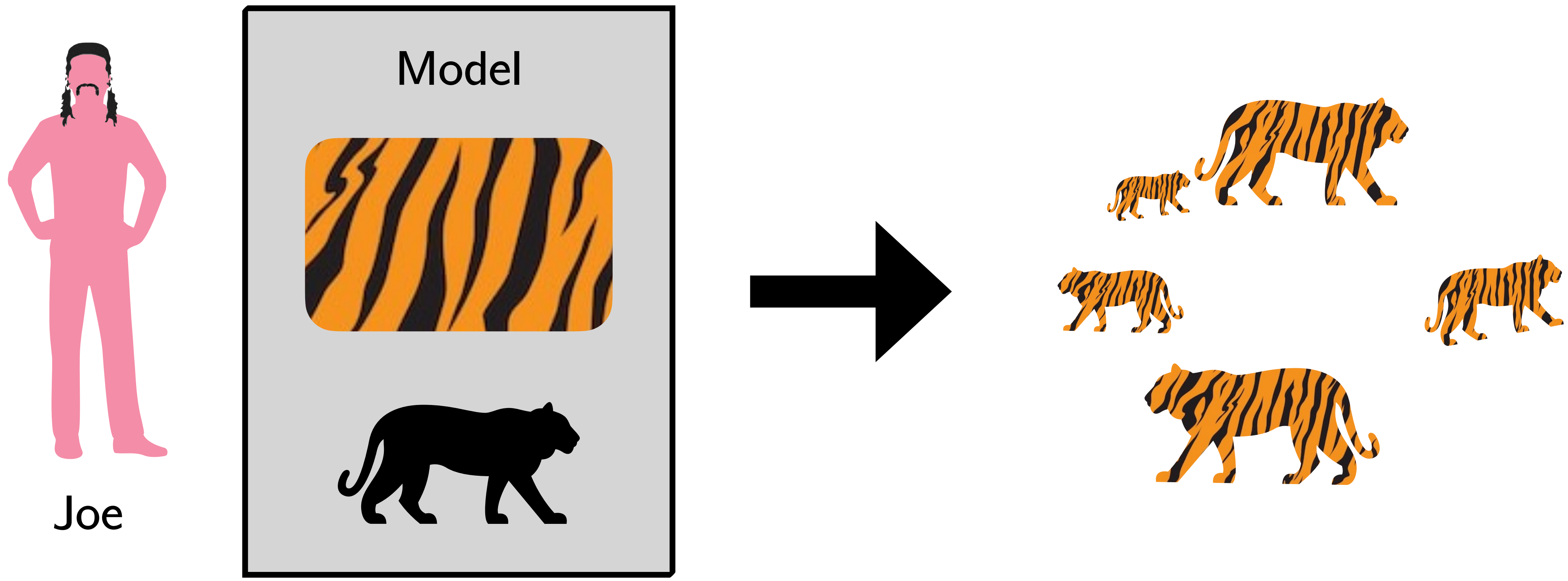
This just means that we have a model that
describes how data can be generated from each class



What is a generative model?

Generative classification is based on density estimation:

This just means that we have a model that
describes how data can be generated from each class



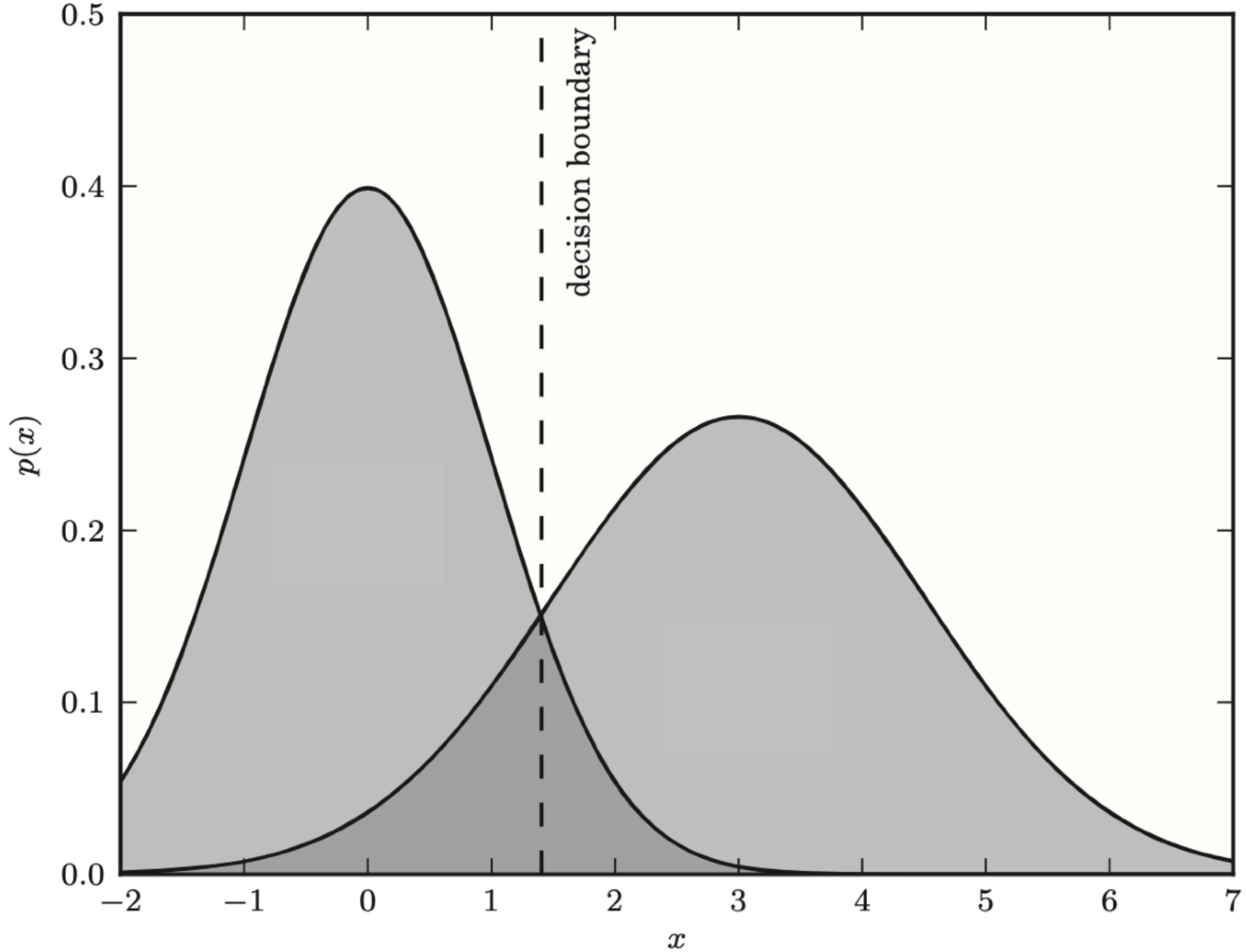
Generative classifiers are:

classifiers that utilize a generative model

Bayes Classifier and Models

<u>Bayes classifier:</u>

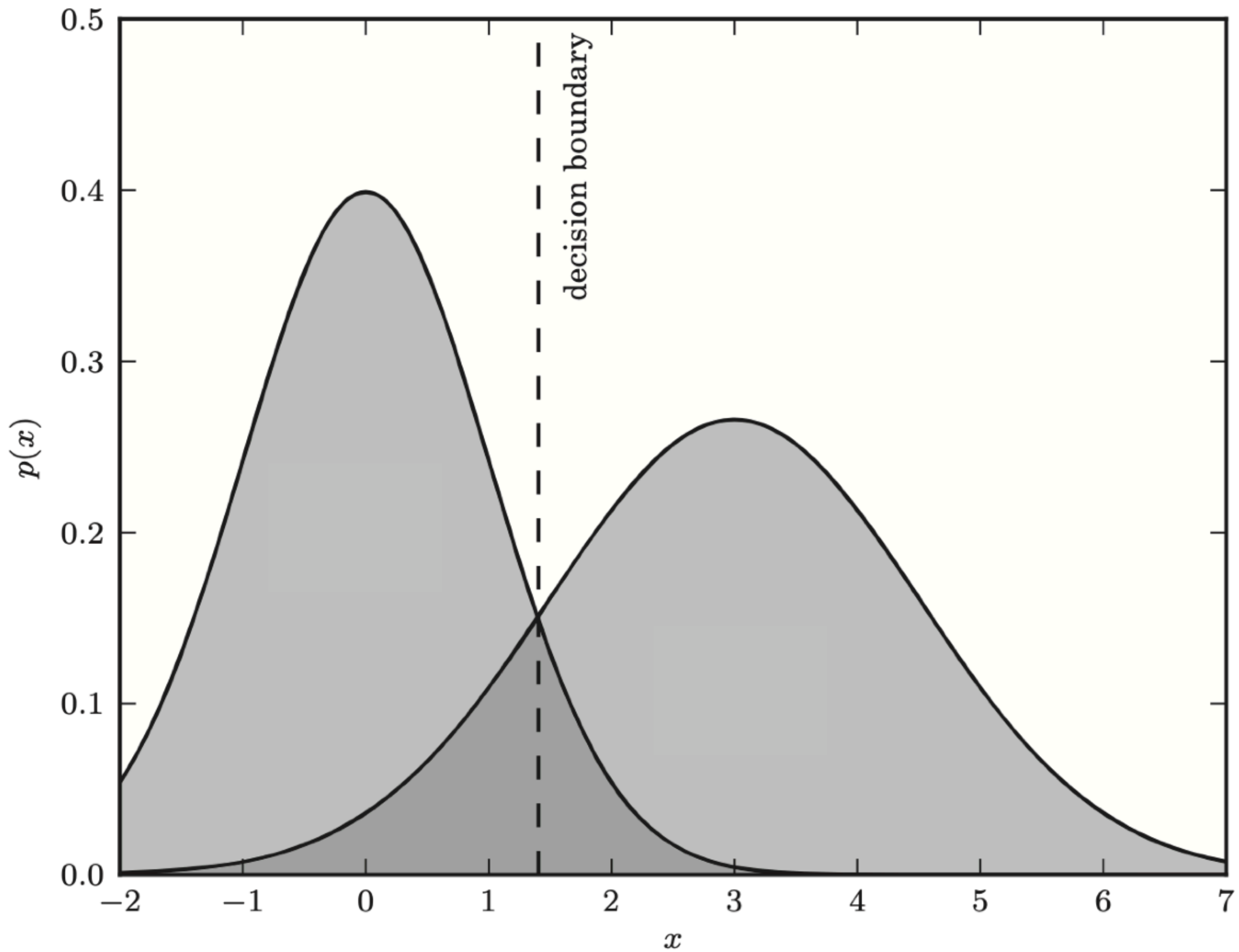
how we determine
which class an object belongs to



Bayes Classifier and Models

<u>Bayes classifier:</u>

how we determine
which class an object belongs to



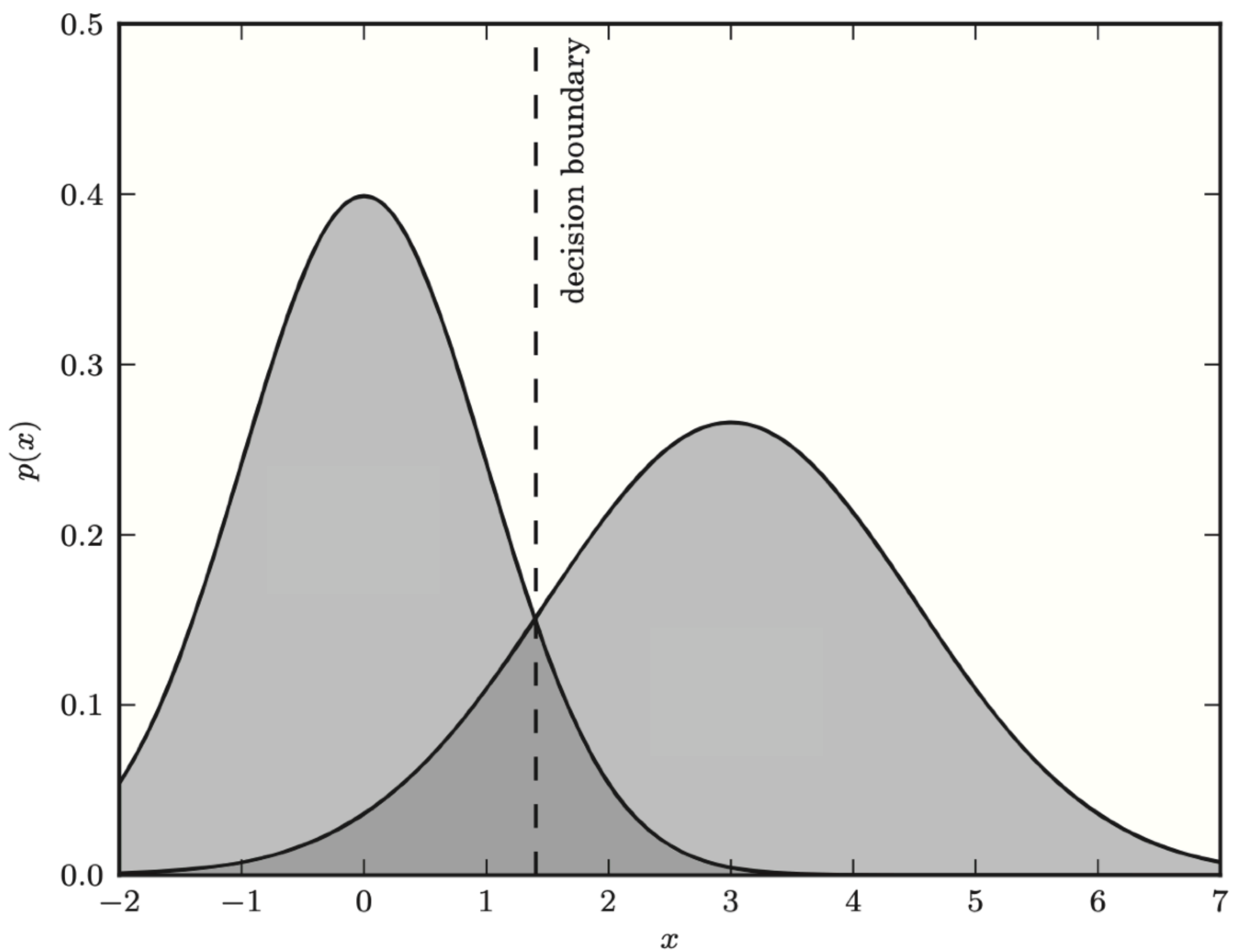
Example

Two classes:

Bayes Classifier and Models

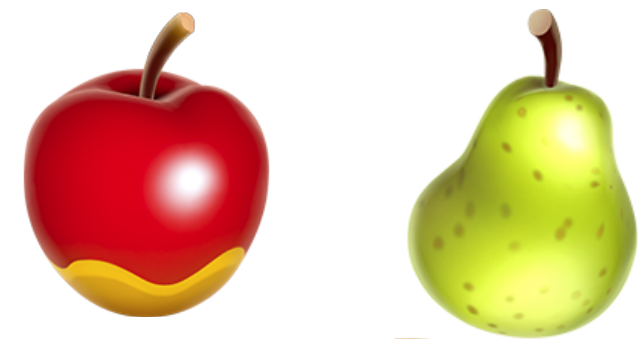
<u>Bayes classifier:</u>

how we determine
which class an object belongs to



Example

Two classes:



$$\widehat{y} = \begin{cases} 1 & \text{if } \pi_1 p_1(\mathbf{x}) > \pi_0 p_0(\mathbf{x}) \\ 0 & \text{otherwise} \end{cases}$$

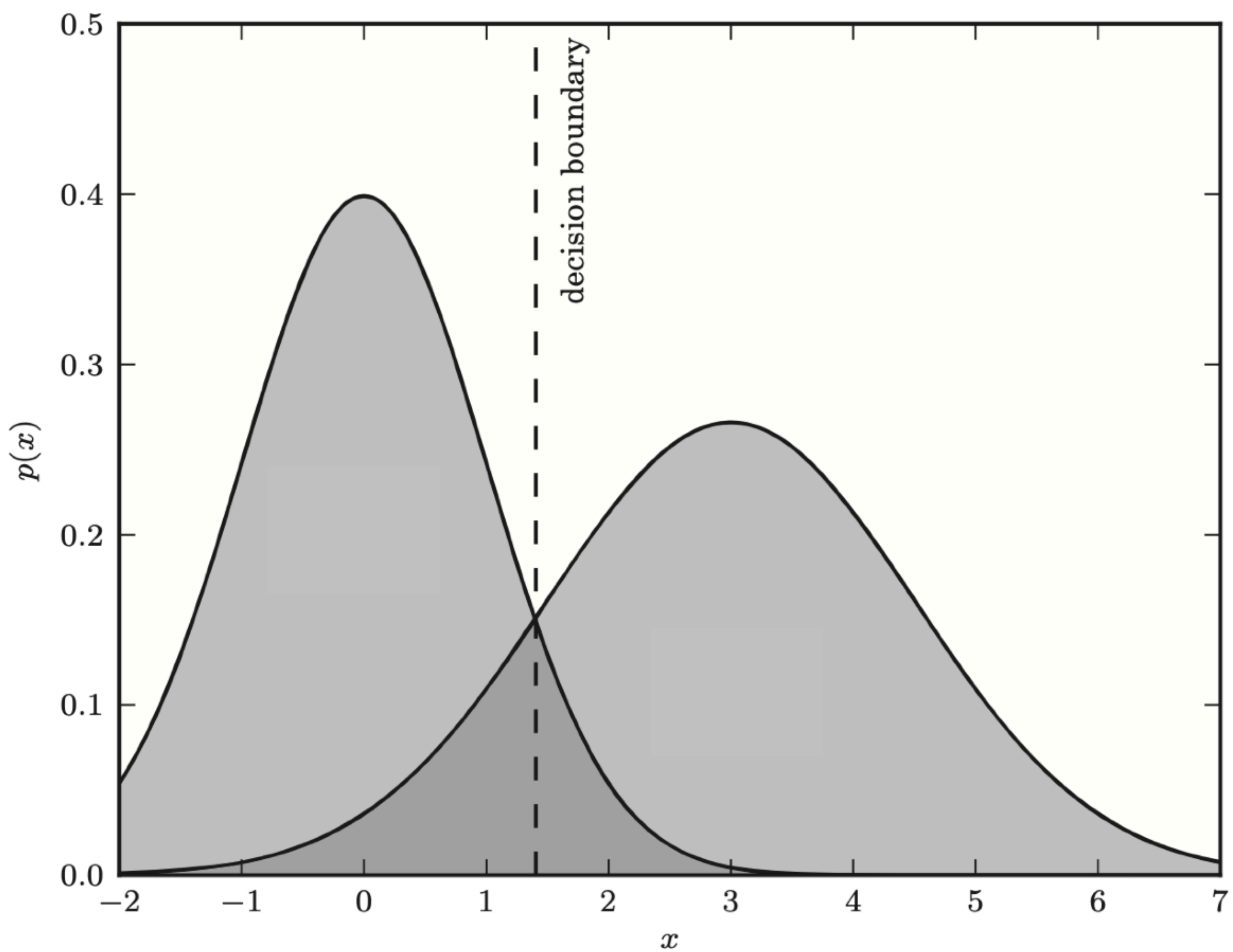
π_1 - probability of being apple

$p_1(x)$ - probability of finding x in your bushel of apples

Bayes Classifier and Models

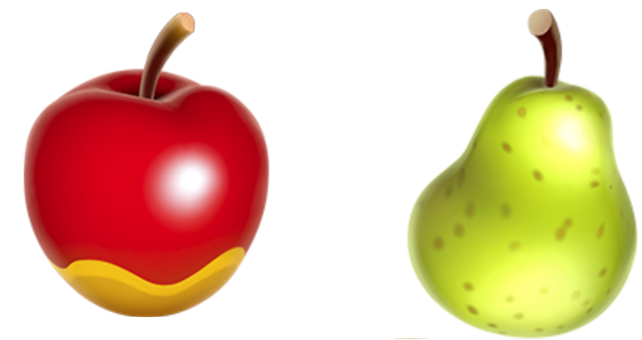
<u>Bayes classifier:</u>

how we determine
which class an object belongs to



Example

Two classes:



$$\hat{y} = \begin{cases} 1 & \text{if } \pi_1 p_1(\mathbf{x}) > \pi_0 p_0(\mathbf{x}) \\ 0 & \text{otherwise} \end{cases}$$

π_1 - probability of being apple

$p_1(x)$ - probability of finding x in your bushel of apples

$\hat{y} = 1$ - best fit is class 'apple'

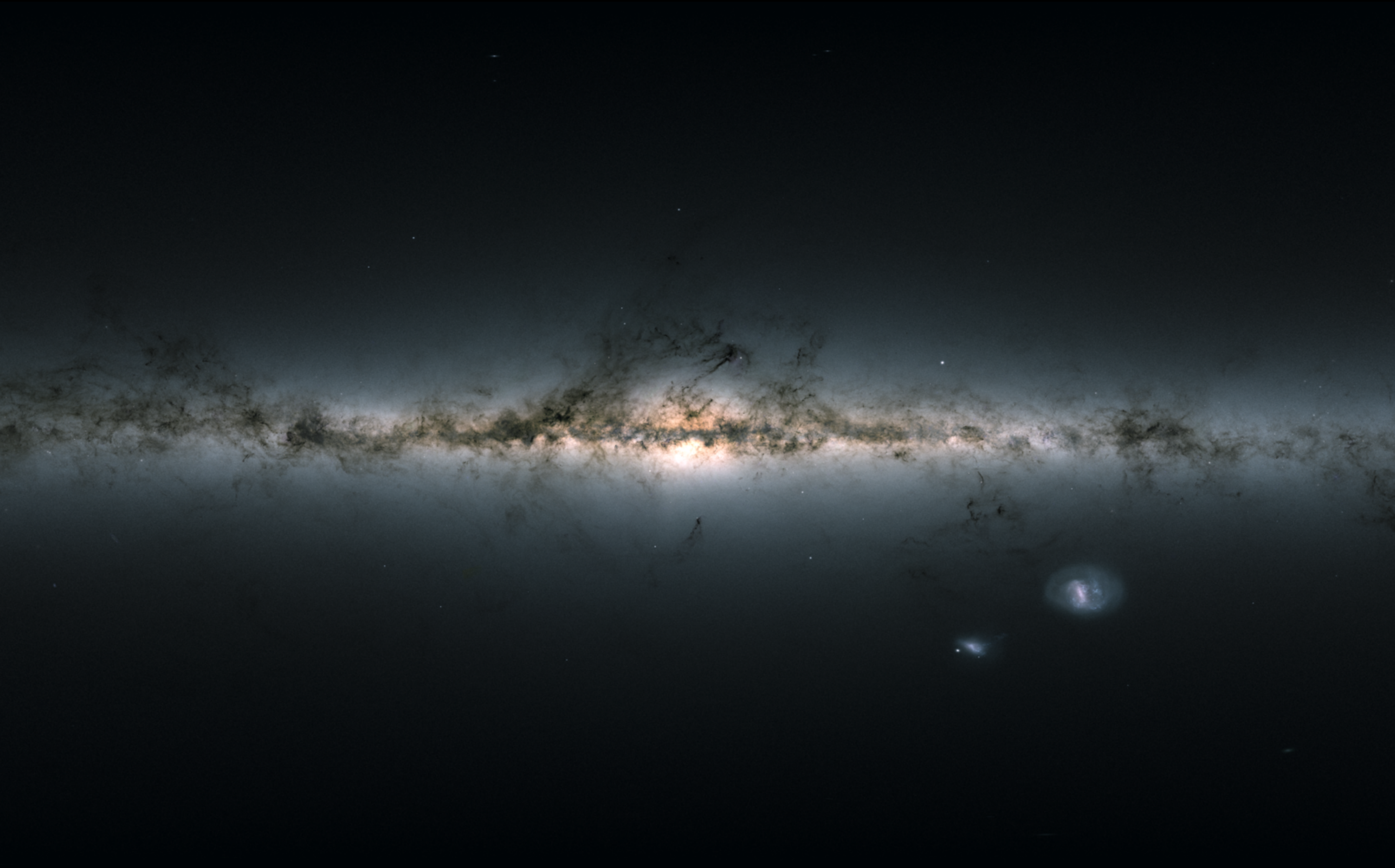
$\hat{y} = 0$ - best fit is class 'pear'

Brief intro to my research

Dwarf galaxies!

They are numerous!

High redshift dwarfs are the progenitors of z=0 galaxies like ours

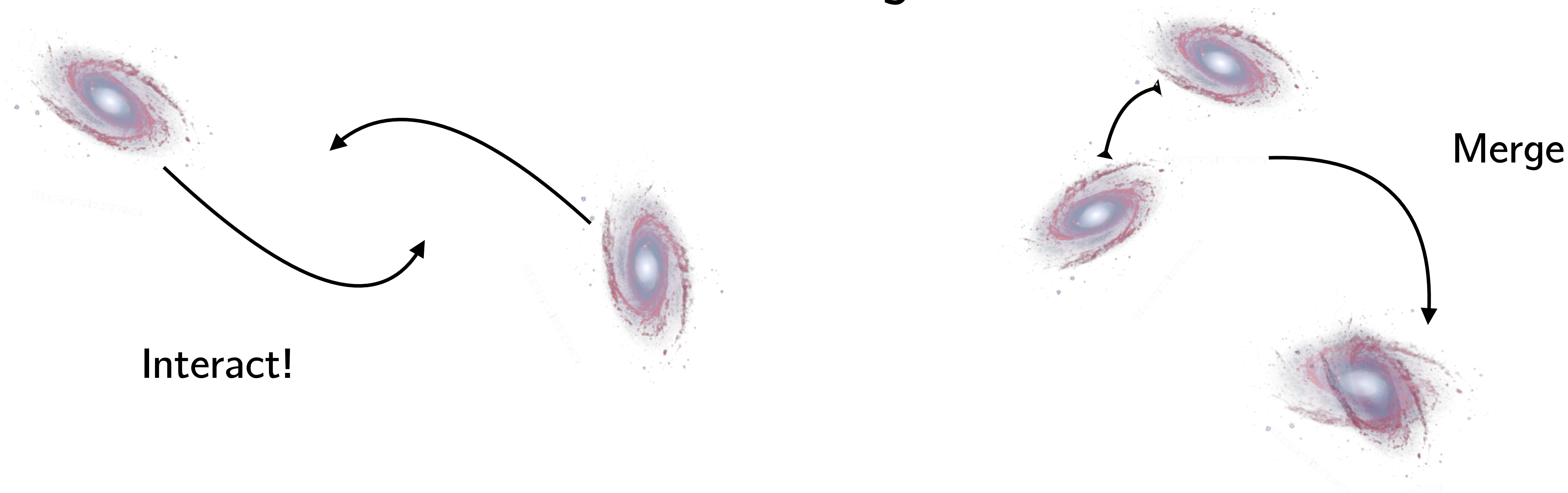


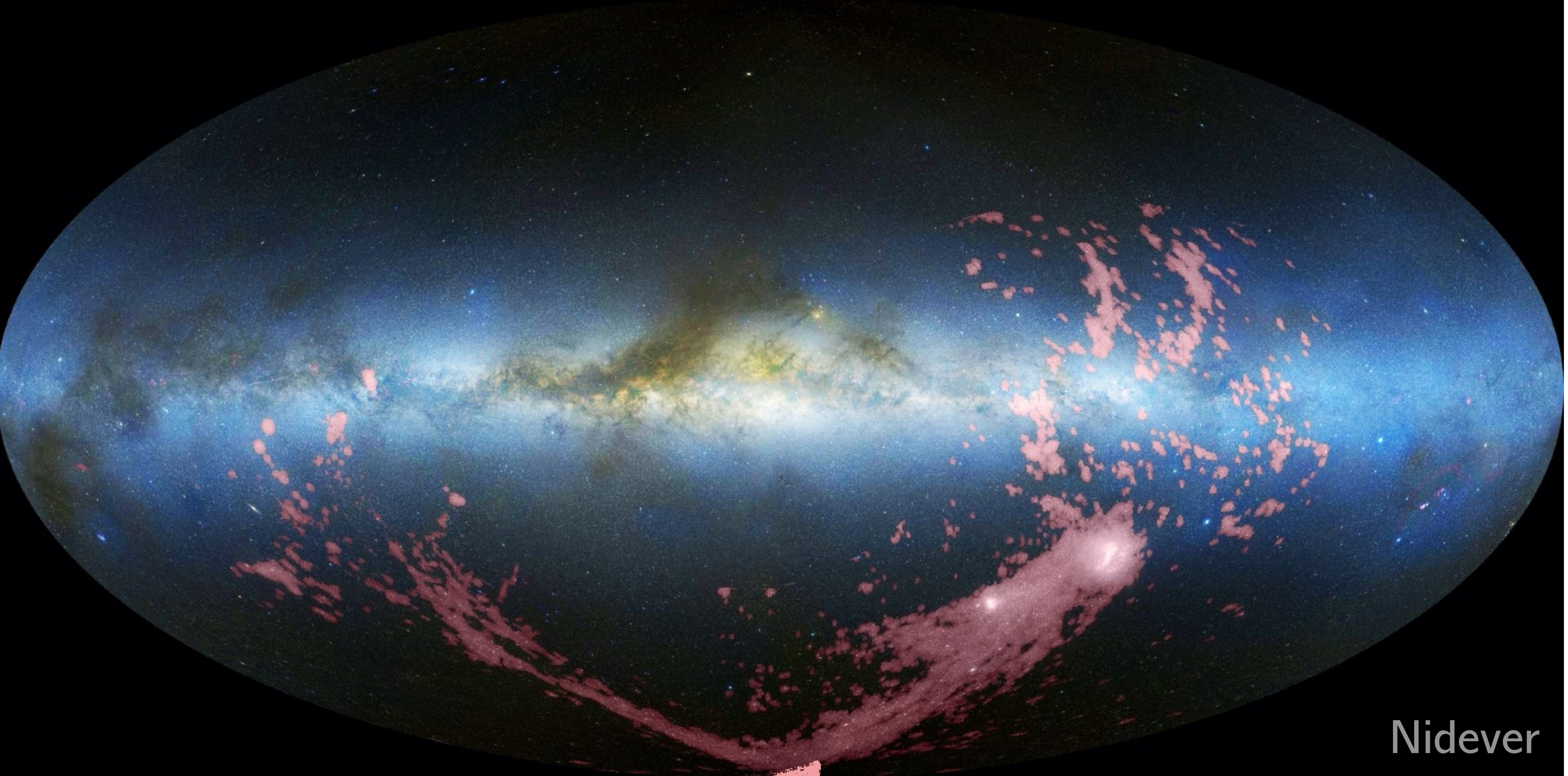
Brief intro to my research

Dwarf galaxies!

- They are numerous!
- High redshift dwarfs are the progenitors of z=0 galaxies like ours

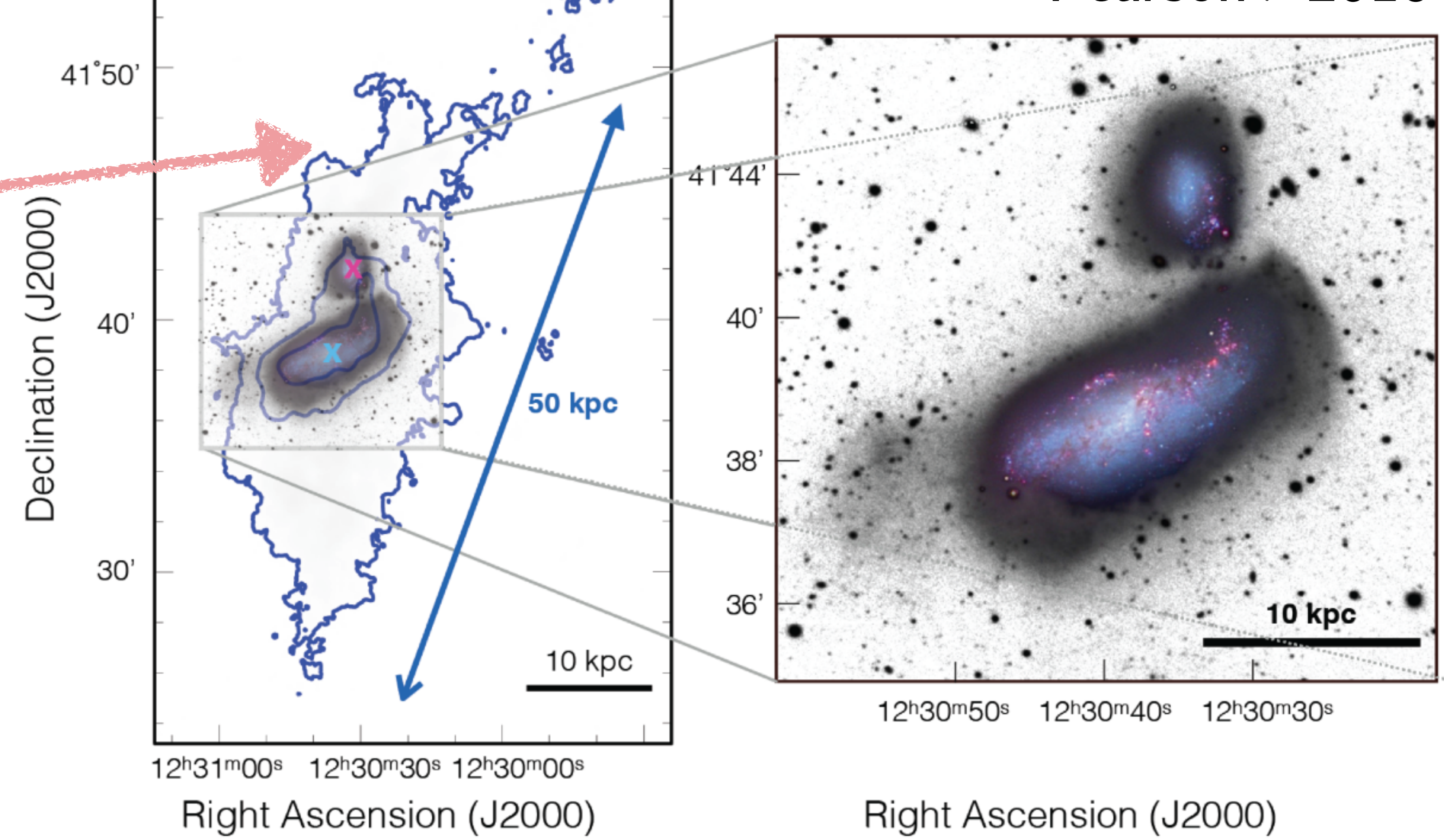
Pairs of dwarf galaxies





Extended
gas features in HI

Pearson+ 2018



Naïve Bayes Classifier + Example

Say we track 8 dwarf galaxies in a simulation, and find the following properties for each one!

Star Formation	Gas Envelope	Asymmetry	Recent Merger?
Extremely enhanced	No Envelope	Asymmetric	Yes
Extremely enhanced	Envelope	Asymmetric	No
Enhanced	Envelope	Asymmetric	Yes
Enhanced	Envelope	Symmetric	No
Enhanced	No envelope	Asymmetric	Yes
Normal	Envelope	Asymmetric	Yes
Normal	Envelope	Symmetric	Yes
Normal	No envelope	Asymmetric	No

Naïve Bayes Classifier + Example

Say we track 8 dwarf galaxies in a simulation, and find the following properties for each one!

Star Formation	Gas Envelope	Asymmetry	Recent Merger?
Extremely enhanced	No Envelope	Asymmetric	Yes
Extremely enhanced	Envelope	Asymmetric	No
Enhanced	Envelope	Asymmetric	Yes
Enhanced	Envelope	Symmetric	No
Enhanced	No envelope	Asymmetric	Yes
Normal	Envelope	Asymmetric	Yes
Normal	Envelope	Symmetric	Yes
Normal	No envelope	Asymmetric	No

If we find a galaxy that has:

- Extremely enhanced SFRs
- A large gas envelope
- Symmetry

What can we say about whether our galaxy has likely been through a merger or not?

Naïve Bayes Classifier + Example

Say we track 8 dwarf galaxies in a simulation, and find the following properties for each one!

Star Formation	Gas Envelope	Asymmetry	Recent Merger?
Extremely enhanced	No Envelope	Asymmetric	Yes
Extremely enhanced	Envelope	Asymmetric	No
Enhanced	Envelope	Asymmetric	Yes
Enhanced	Envelope	Symmetric	No
Enhanced	No envelope	Asymmetric	Yes
Normal	Envelope	Asymmetric	Yes
Normal	Envelope	Symmetric	Yes
Normal	No envelope	Asymmetric	No

- **Independent**
- **Equally important**

If we find a galaxy that has:

- Extremely enhanced SFRs
- A large gas envelope
- Symmetry

What can we say about whether our galaxy has likely been through a merger or not?

Naïve Bayes Classifier + Example

Probabilities by trait

Star Formation	Yes	No	P(Yes)	P(No)
Extremely Enhanced	1	1	1/5	1/3
Enhanced	2	1	2/5	1/3
Normal	2	1	2/5	1/3

Gas Envelope	Yes	No	P(Yes)	P(No)
Envelope	3	2	3/5	2/3
No envelope	2	1	2/5	1/3

Asymmetry	Yes	No	P(Yes)	P(No)
Asymmetric	4	2	4/5	2/3
Symmetric	1	1	1/5	1/3

$$\mathrm{P}(\mathrm{Yes}|[\mathrm{Ext.Enhanced}, \mathrm{Envelope}, \mathrm{Symmetric}]) = \frac{\mathrm{P}(\mathrm{Ext.Enhanced}|\mathrm{Yes}) * \mathrm{P}(\mathrm{Envelope}|\mathrm{Yes}) * \mathrm{P}(\mathrm{Symmetric}|\mathrm{Yes}) * \mathrm{P}(\mathrm{Yes})}{\mathrm{P}([\mathrm{Ext.Enhanced}, \mathrm{Envelope}, \mathrm{Symmetric}])}$$

Naïve Bayes Classifier + Example

Probabilities by trait

Star Formation	Yes	No	P(Yes)	P(No)
Extremely Enhanced	1	1	1/5	1/3
Enhanced	2	1	2/5	1/3
Normal	2	1	2/5	1/3

Gas Envelope	Yes	No	P(Yes)	P(No)
Envelope	3	2	3/5	2/3
No envelope	2	1	2/5	1/3

Asymmetry	Yes	No	P(Yes)	P(No)
Asymmetric	4	2	4/5	2/3
Symmetric	1	1	1/5	1/3

$$\mathrm{P}(\mathrm{Yes}|[\mathrm{Ext.Enhanced}, \mathrm{Envelope}, \mathrm{Symmetric}]) =$$

$$\frac{\mathrm{P}(\mathrm{Ext.Enhanced}|\mathrm{Yes}) * \mathrm{P}(\mathrm{Envelope}|\mathrm{Yes}) * \mathrm{P}(\mathrm{Symmetric}|\mathrm{Yes}) * \mathrm{P}(\mathrm{Yes})}{\mathrm{P}([\mathrm{Ext.Enhanced}, \mathrm{Envelope}, \mathrm{Symmetric}])}$$

Naïve Bayes Classifier + Example

$$\mathrm{P}(\mathrm{Yes}|[\mathrm{Ext.Enhanced}, \mathrm{Envelope}, \mathrm{Symmetric}]) = 0.35$$

$$\mathrm{P}(\mathrm{No}|[\mathrm{Ext.Enhanced}, \mathrm{Envelope}, \mathrm{Symmetric}]) = 0.65$$

According to this model, it's likely that this new galaxy has NOT recently undergone a merger!

Naïve Bayes Classifier + Example

$$\mathrm{P}(\mathrm{Yes}|[\mathrm{Ext.Enhanced}, \mathrm{Envelope}, \mathrm{Symmetric}]) = 0.35$$

$$\mathrm{P}(\mathrm{No}|[\mathrm{Ext.Enhanced}, \mathrm{Envelope}, \mathrm{Symmetric}]) = 0.65$$

According to this model, it's likely that this new galaxy has NOT recently undergone a merger!

Great, except a lot of variables in astronomy are *not* discrete!
Such as: colors, ages, masses, redshift, etc.

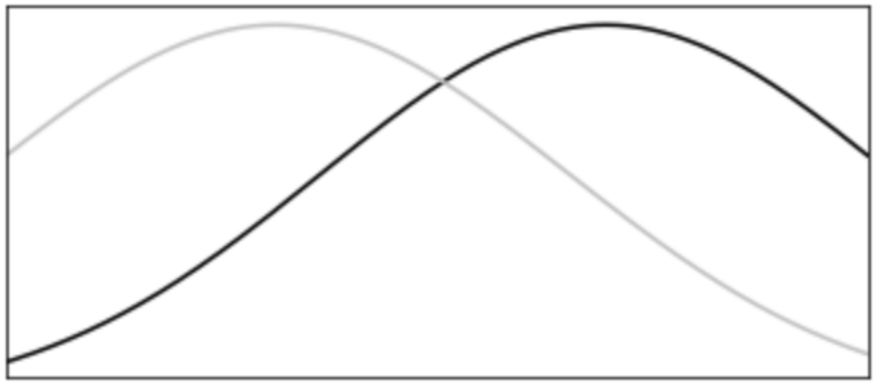
→ *What about continuous variables?* ←

Gaussian Naïve Bayes Classifier

This is a frequentist technique which:

Models the distribution with a Gaussian!

(instead of discrete probabilities
in the previous example)

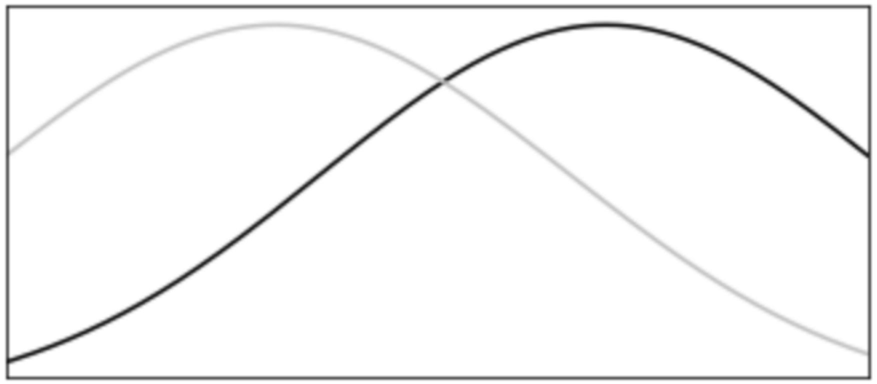


Gaussian Naïve Bayes Classifier

This is a frequentist technique which:

Models the distribution with a Gaussian!

(instead of discrete probabilities
in the previous example)



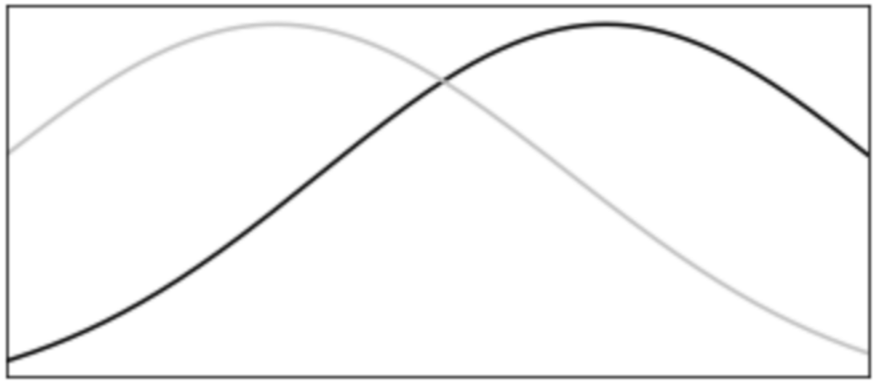
As in the discrete case, this again assumes perfect independence between variables

Gaussian Naïve Bayes Classifier

This is a frequentist technique which:

Models the distribution with a Gaussian!

(instead of discrete probabilities in the previous example)



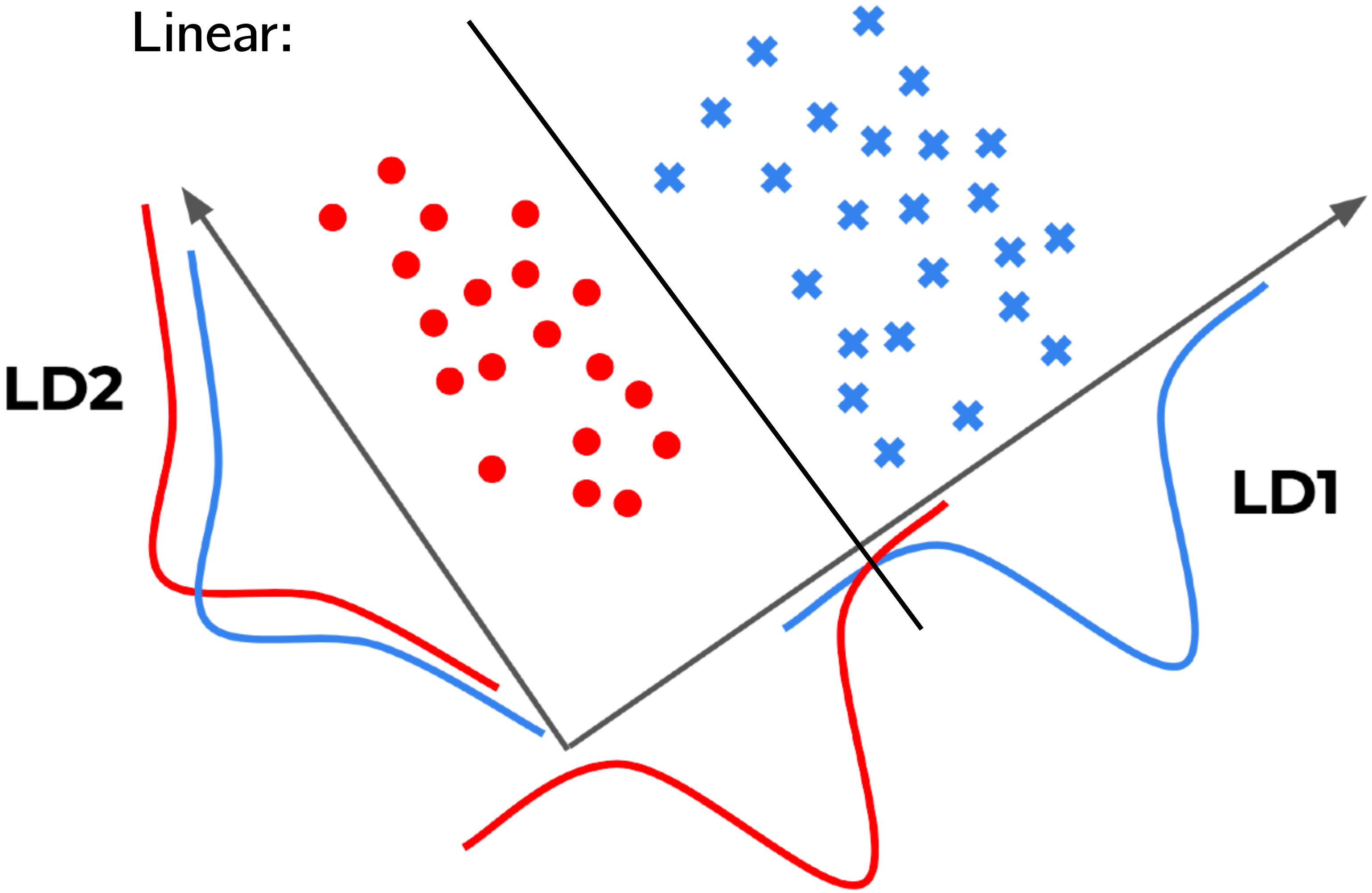
As in the discrete case, this again assumes perfect independence between variables

---

Note: if there are covariances between the variables in your distributions, then you need to use the *Gaussian Bayes Classifier* which accounts for those dependencies.

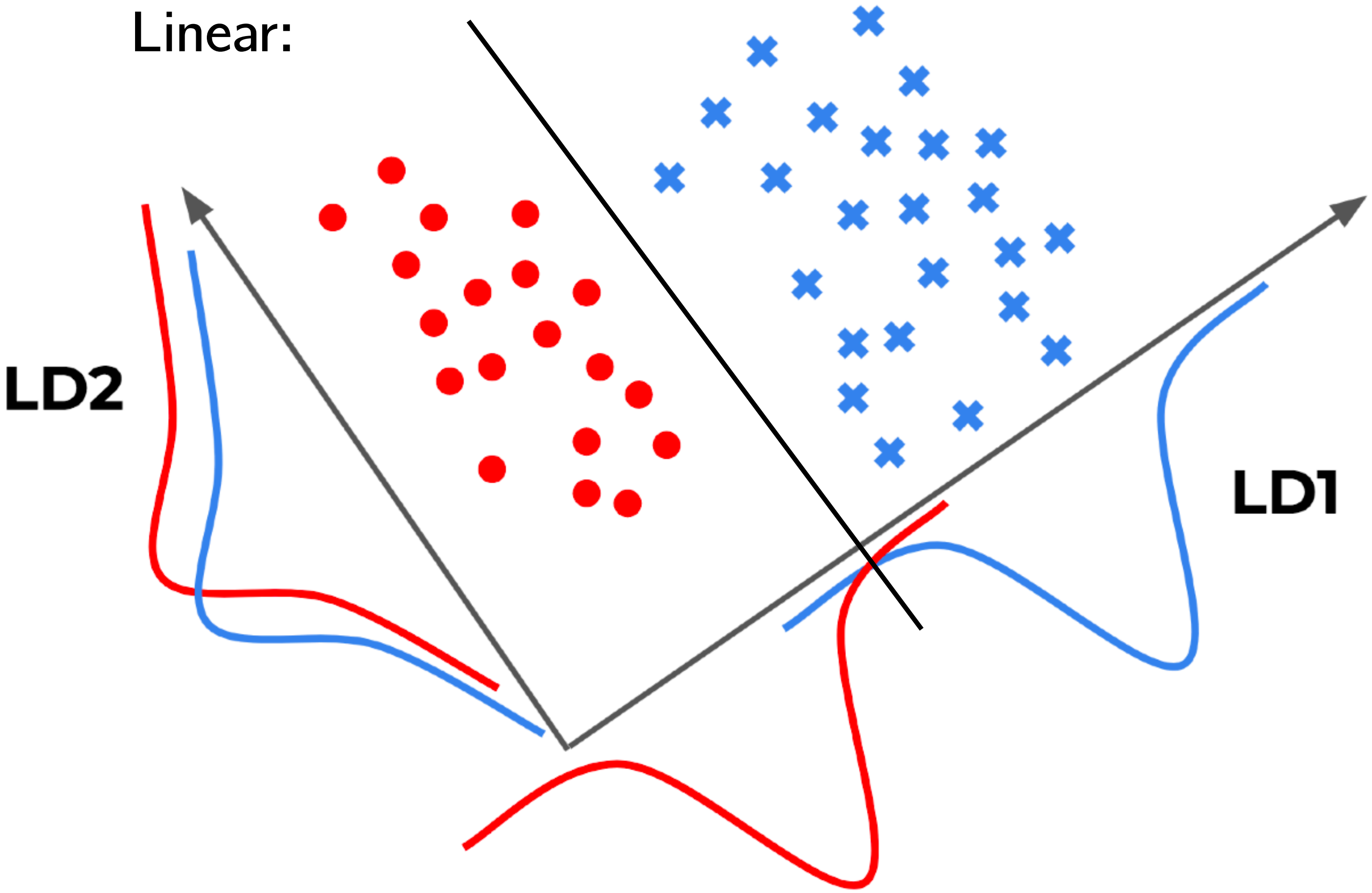
Linear/Quadratic Discriminant Analysis

Linear:



Still uses Gaussian models!

Linear/Quadratic Discriminant Analysis



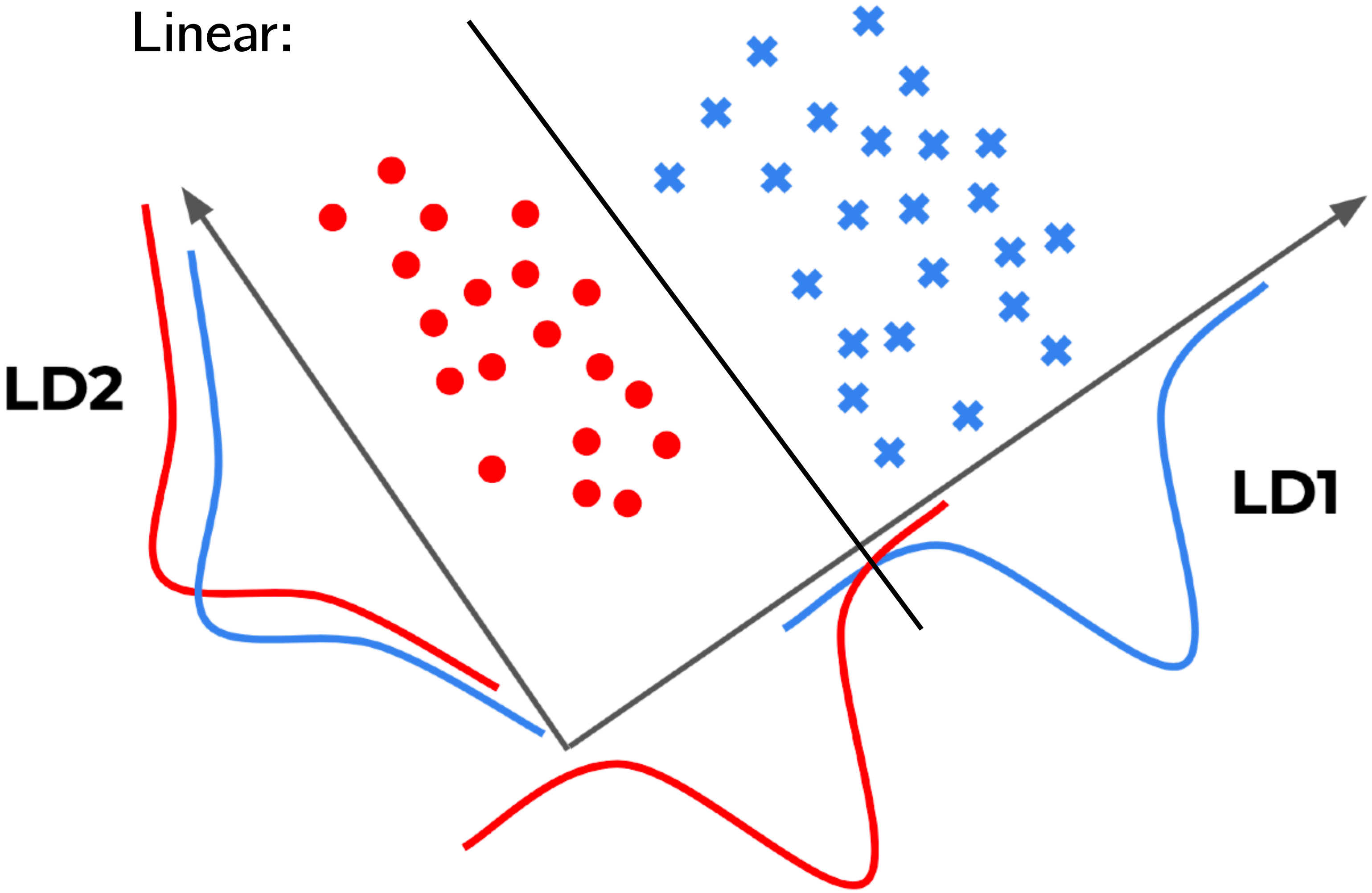
Still uses Gaussian models!

Decreases the dimensionality of the data by:

- Maximizing the separation between gaussians
- Minimizing the spread of each gaussian

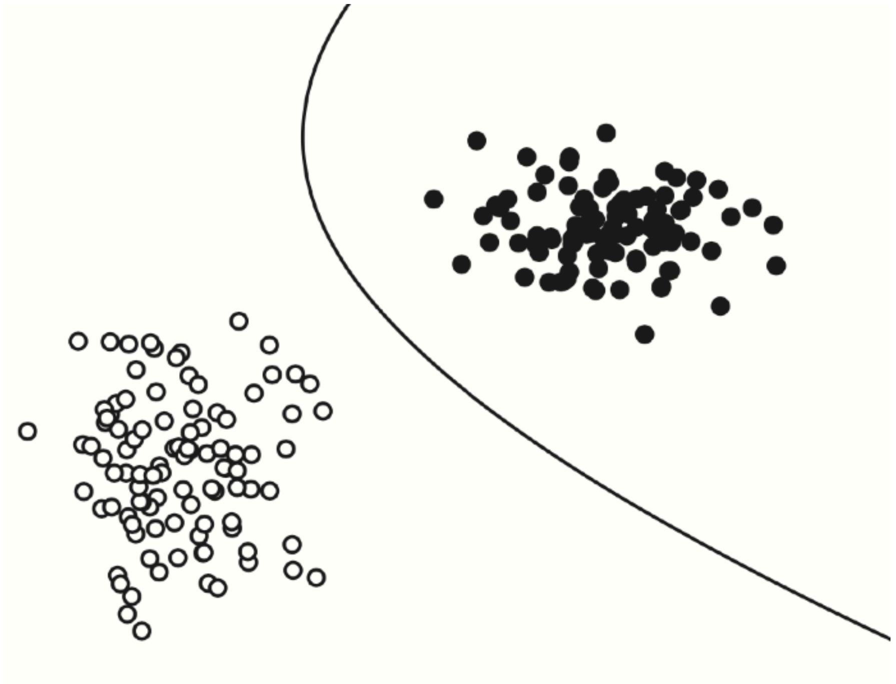
Linear/Quadratic Discriminant Analysis

Linear:



Still uses Gaussian models!

Quadratic:

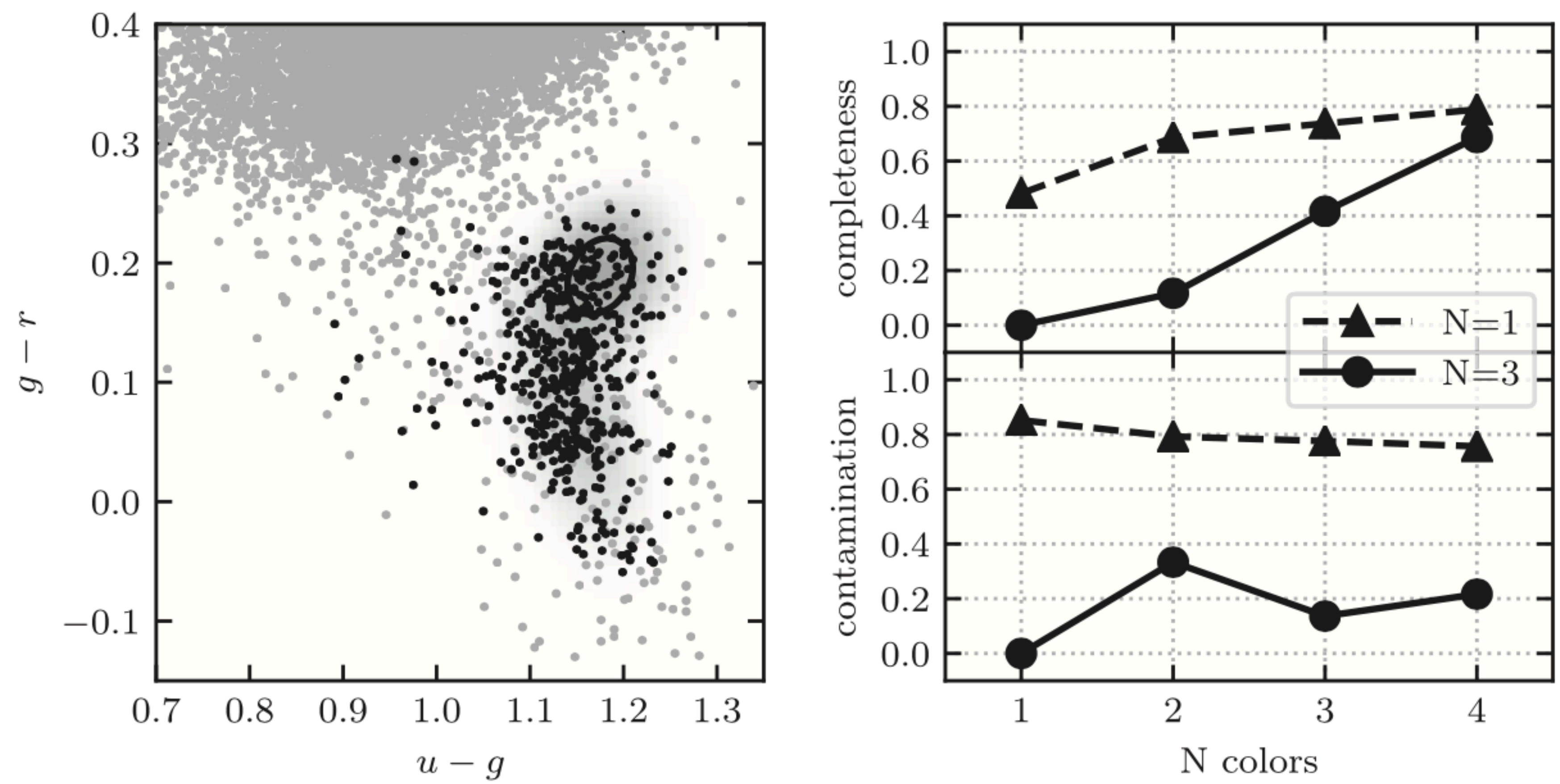


Decreases the dimensionality
of the data by:

- Maximizing the separation between gaussians
- Minimizing the spread of each gaussian

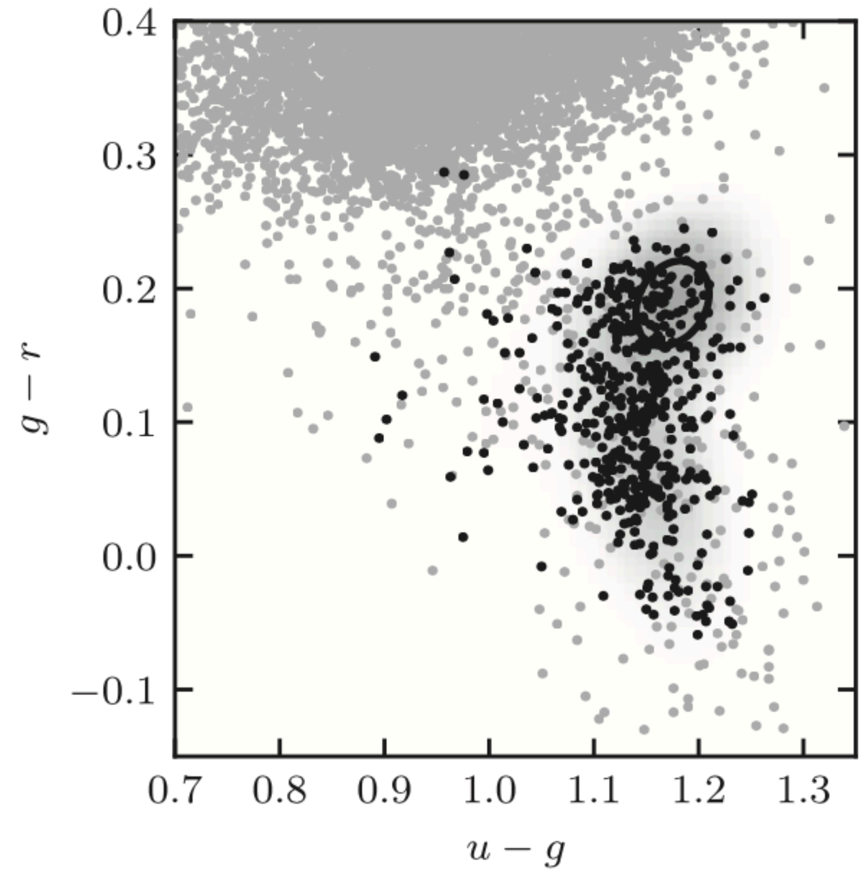
Gaussian Mixture Model Bayes Classifier

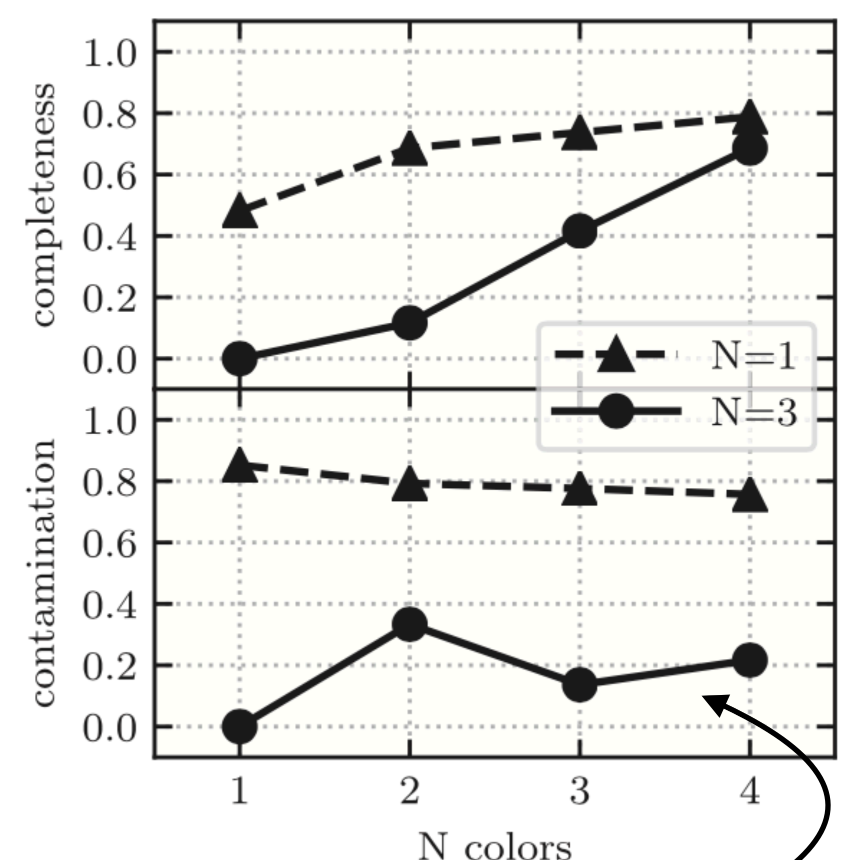
A step up from the
Gaussian Bayes classifier:
Use Gaussian mixture model



Gaussian Mixture Model Bayes Classifier

A step up from the
Gaussian Bayes classifier:
Use Gaussian mixture model

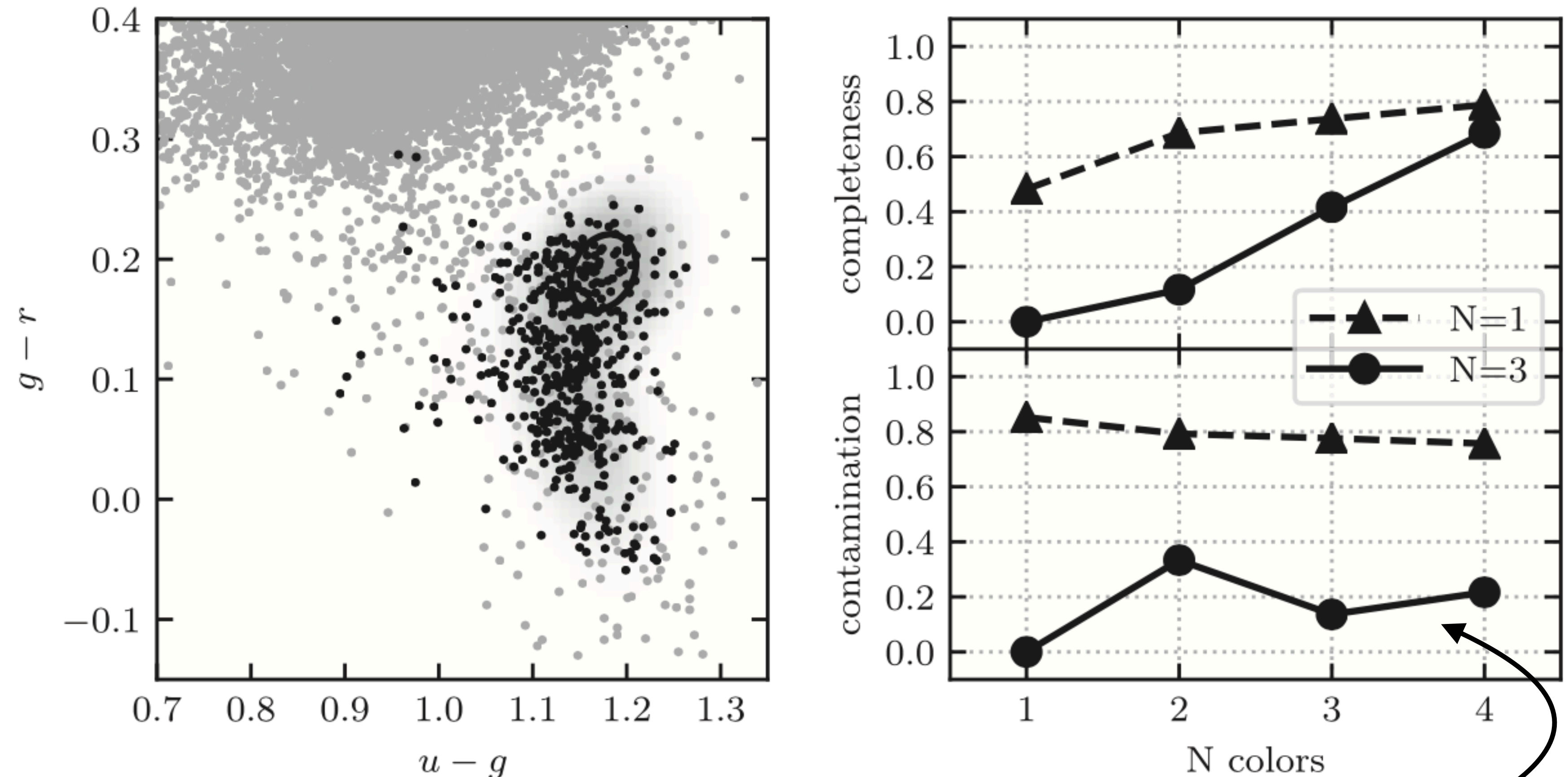




In examples from the book,
the GMM Bayes Classifier has significantly less contamination

Gaussian Mixture Model Bayes Classifier

A step up from the
Gaussian Bayes classifier:
Use Gaussian mixture model



In examples from the book,
the GMM Bayes Classifier has significantly less contamination

Pick your classifier based on what's most important to you:
Completeness or low contamination