

Machine Learning Algorithms

Common Machine Learning Algorithms

Linear Regression

Logistic Regression

Support Vector Machine

Decision Tree

K-Nearest Neighbor

K-Means

$$\hat{f}(X)$$

The Prediction

$$\hat{y} = f(\hat{x})$$

output input

Regression
continuous

$$\hat{y} = 66.5$$

Classification probability

$\hat{y} = .85$

Classification prediction

$$\hat{y} = 1$$

Classification

prediction

$$\hat{y} = 0$$

Linear Regression

equation of a line

$$y = mx + b$$

equation of a line

$$y = mx + b$$

linear regression

$$y = \beta_0 + \beta_1 x$$

Simple Linear Regression

Input

\mathbf{x}_1

learned coefficients
(weights)

β_0, β_1

output

\mathbf{y}

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x}$$

16

12

8

4

0

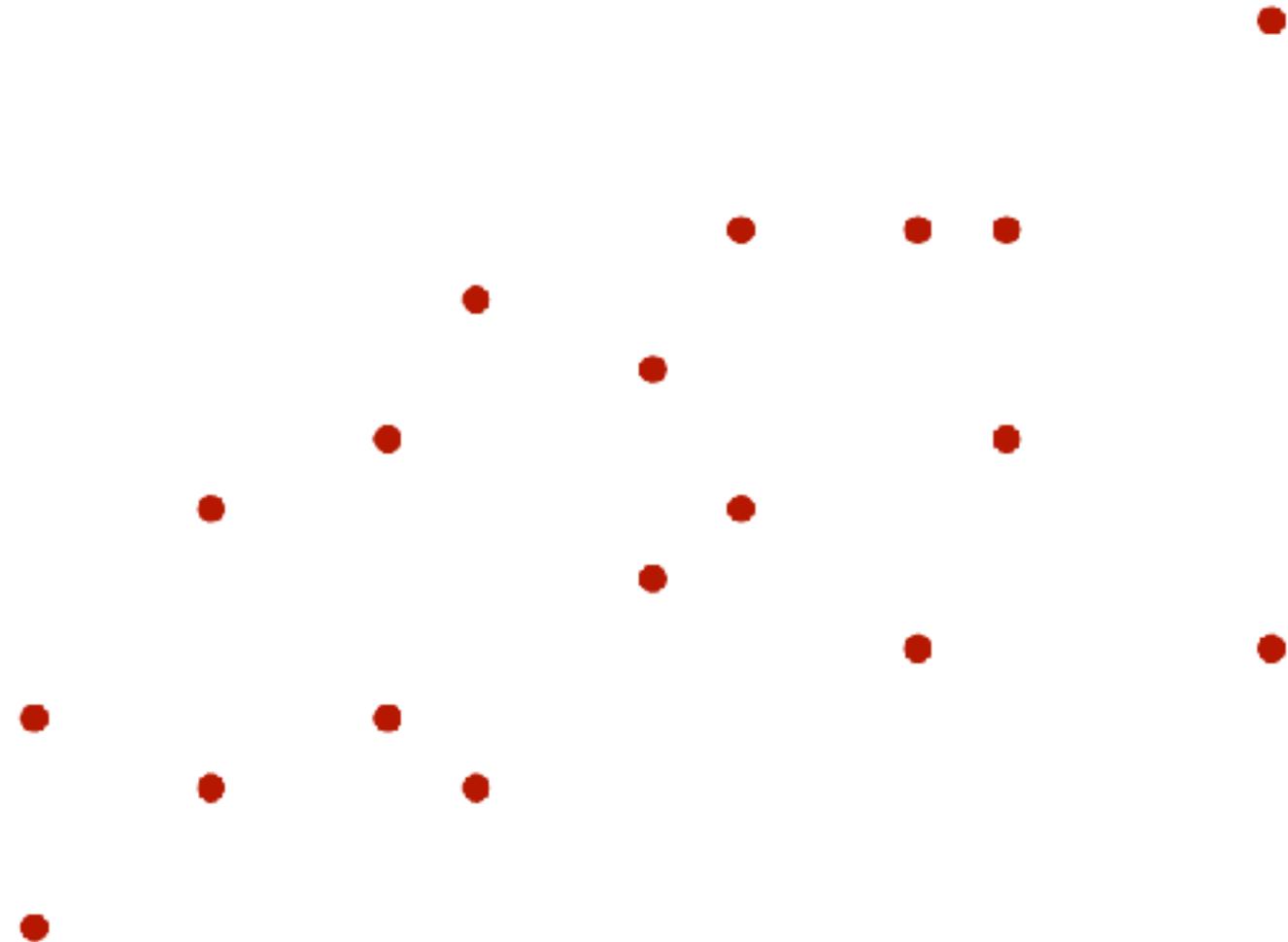
0

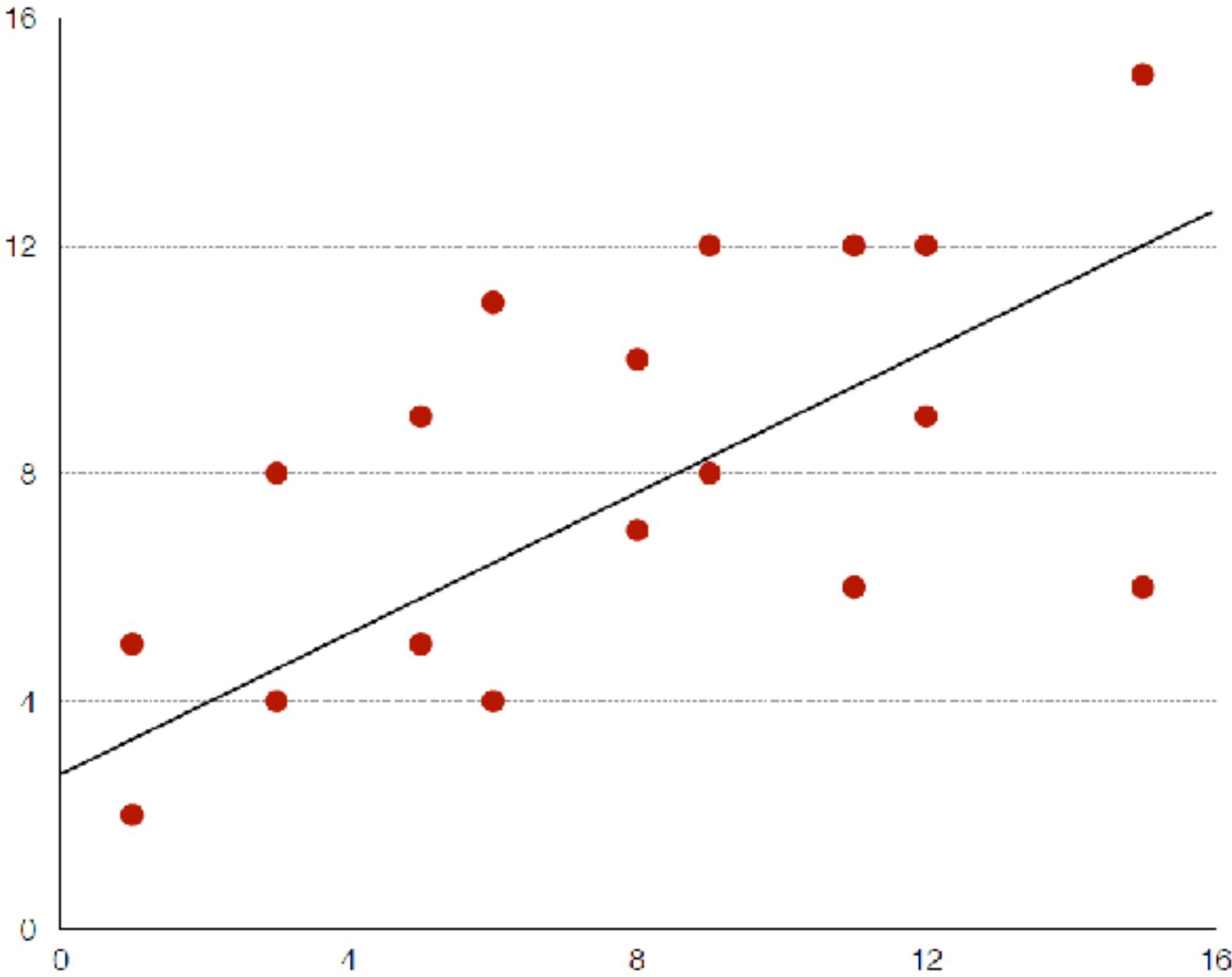
4

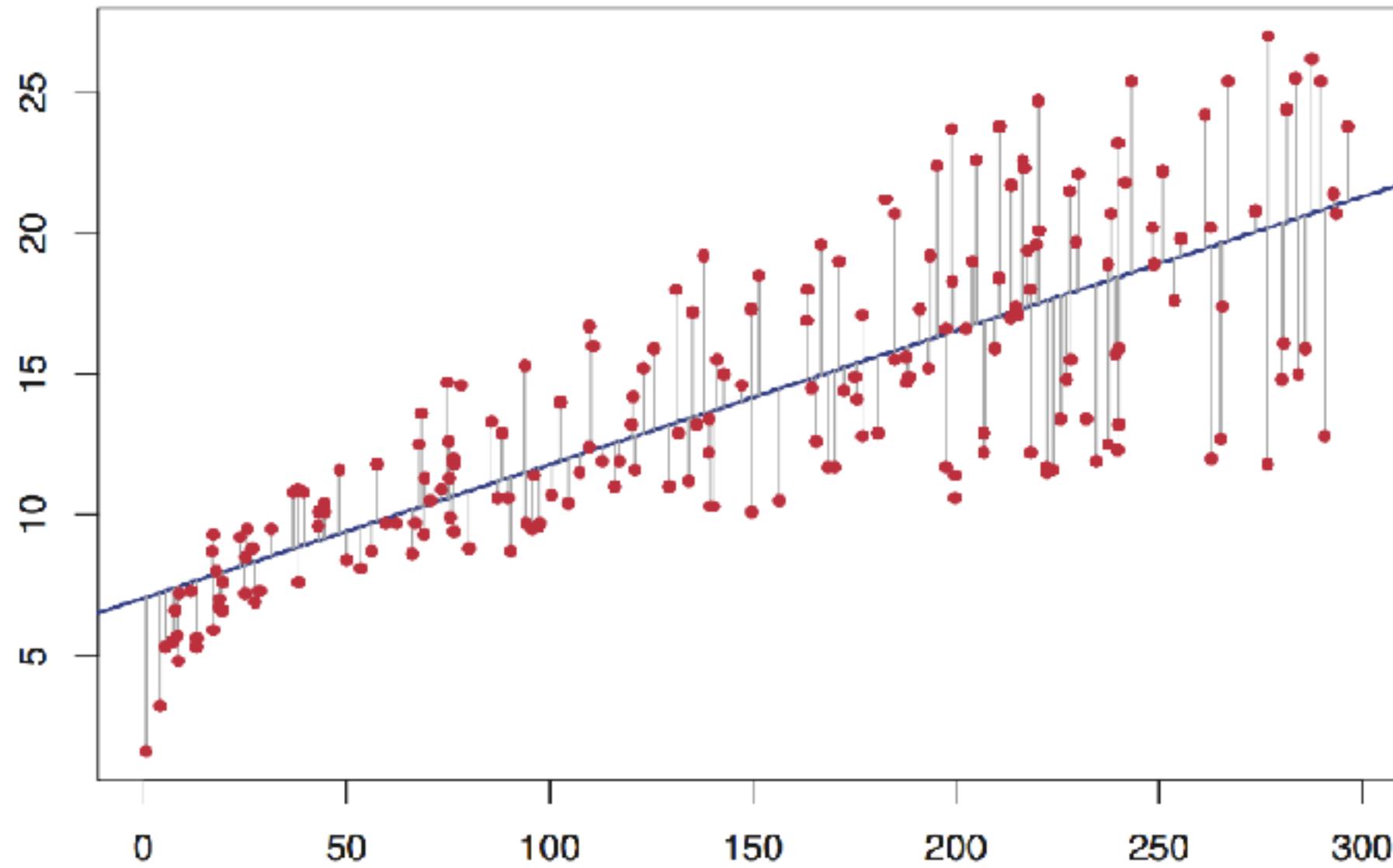
8

12

16







Weekly Hours Spent Studying

Gender	Number of Classes	Social Accounts	Hours Studying
Male	2	1	7.5
Male	4	3	12.25
Female	4	3	12.75
Female	3	4	7.75
Female	4	2	14
Male	2	3	5.75
Female	5	1	18.25

Weekly Hours Spent Studying

Features →

Gender	Number of Classes	Social Accounts	Hours Studying
Male	2	1	7.5
Male	4	3	12.25
Female	4	3	12.75
Female	3	4	7.75
Female	4	2	14
Male	2	3	5.75
Female	5	1	18.25

Weekly Hours Spent Studying

Features →

Gender	Number of Classes	Social Accounts	Hours Studying
0	2	1	7.5
0	4	3	12.25
1	4	3	12.75
1	3	4	7.75
1	4	2	14
0	2	3	5.75
0	5	1	18.25

Weekly Hours Spent Studying

Gender	Number of Classes	Social Accounts	Hours Studying	Target (y)
0	2	1	7.5	
0	4	3	12.25	
1	4	3	12.75	
1	3	4	7.75	
1	4	2	14	
0	2	3	5.75	
0	5	1	18.25	

Multiple Linear Regression

$$\text{Hrs. Studying} = 1.63 + 3.51\mathbf{x}_1 + .25\mathbf{x}_2 - 1.08\mathbf{x}_3$$

\mathbf{x}_1 = number of classes

\mathbf{x}_2 = gender (m=0, female=1)

\mathbf{x}_3 = number of social accounts





HEIGHT

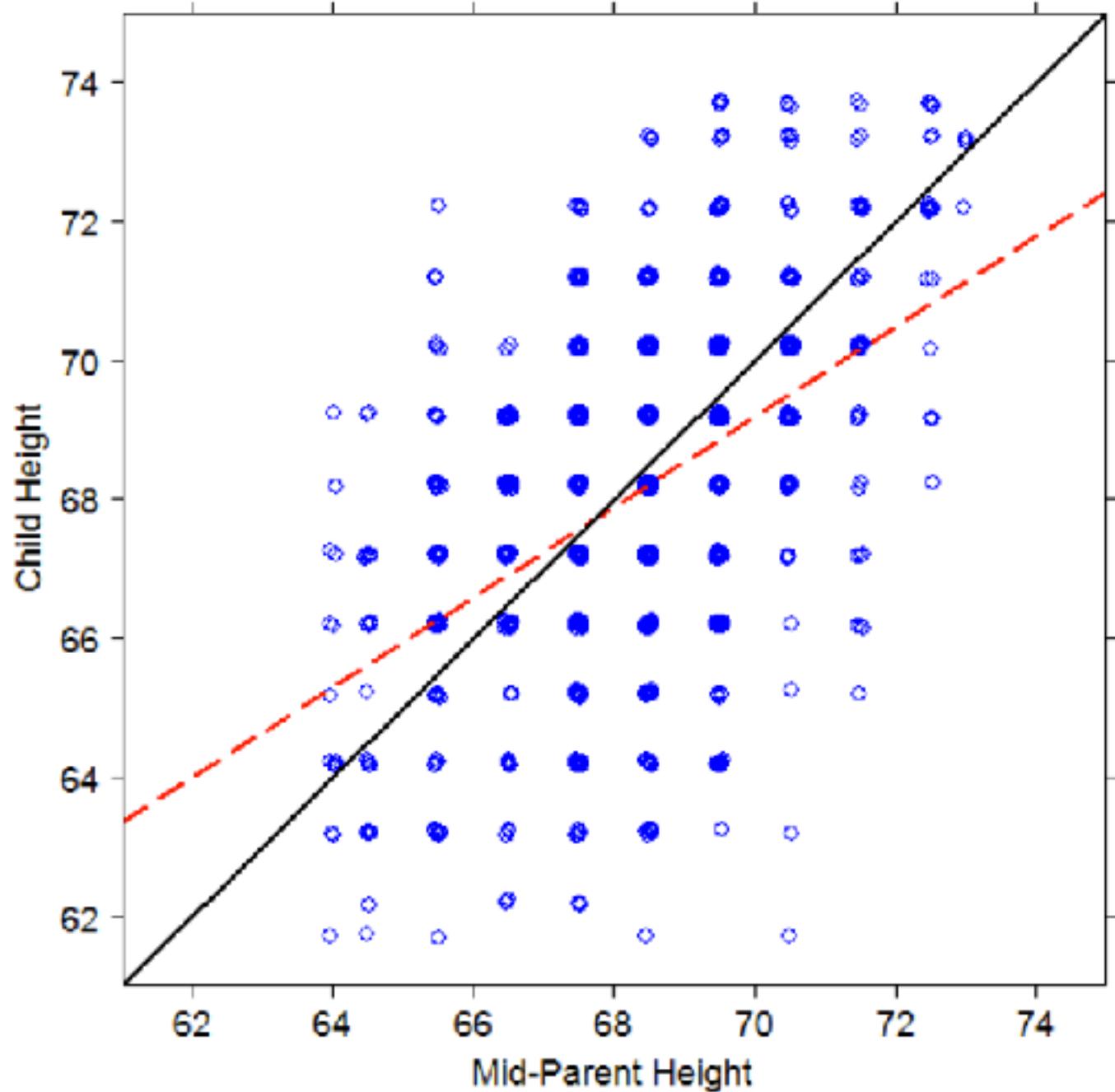


A MALE
WILL GROW TO
BE ABOUT THE
HEIGHT OF

THE MOTHER (INCHES)
+
FATHER (INCHES)
**+ 5 INCHES,
DIVIDED BY TWO.**

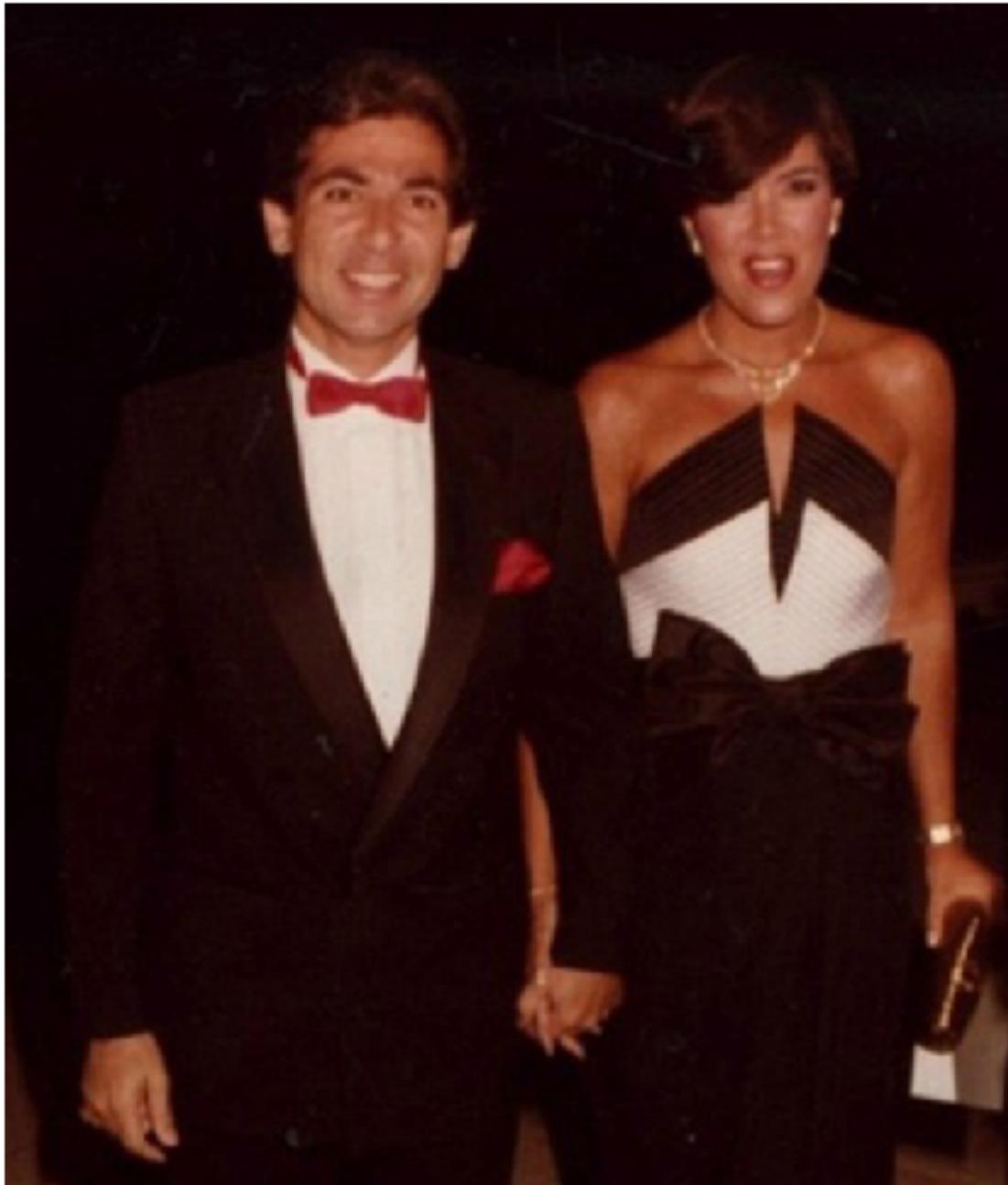
A FEMALE
WILL GROW TO
BE ABOUT THE
HEIGHT OF

THE MOTHER (INCHES)
+
FATHER (INCHES)
**- 5 INCHES,
DIVIDED BY TWO.**





5'7 —



5'7"

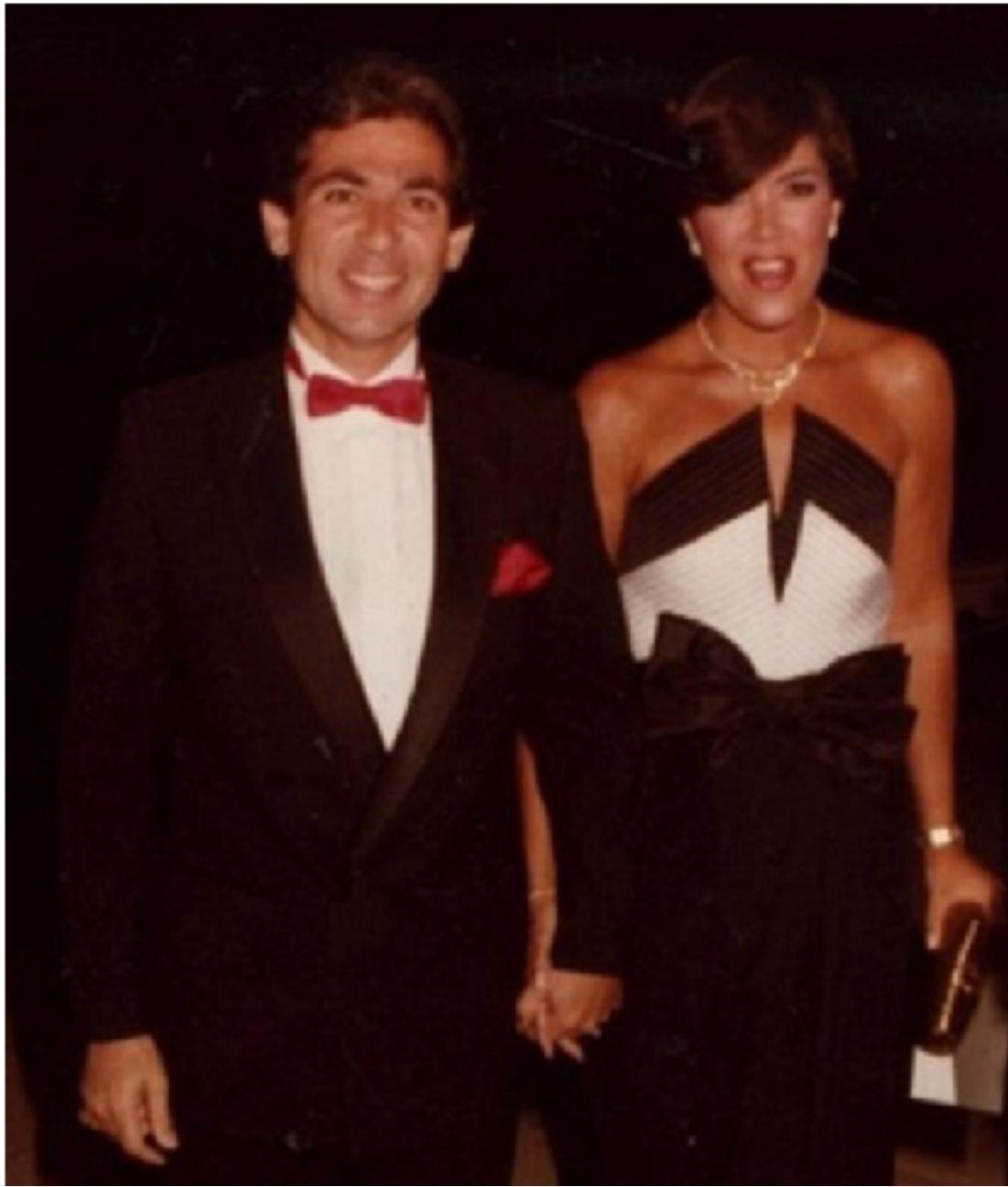


"5'6"

Mid-parent Height
5'6.5

5'7

5'6



5'0



5'3 —



5'10



5'10

5'3

5'0



Average
5'4¹/₃



6'1

5'3

5'0

5'10

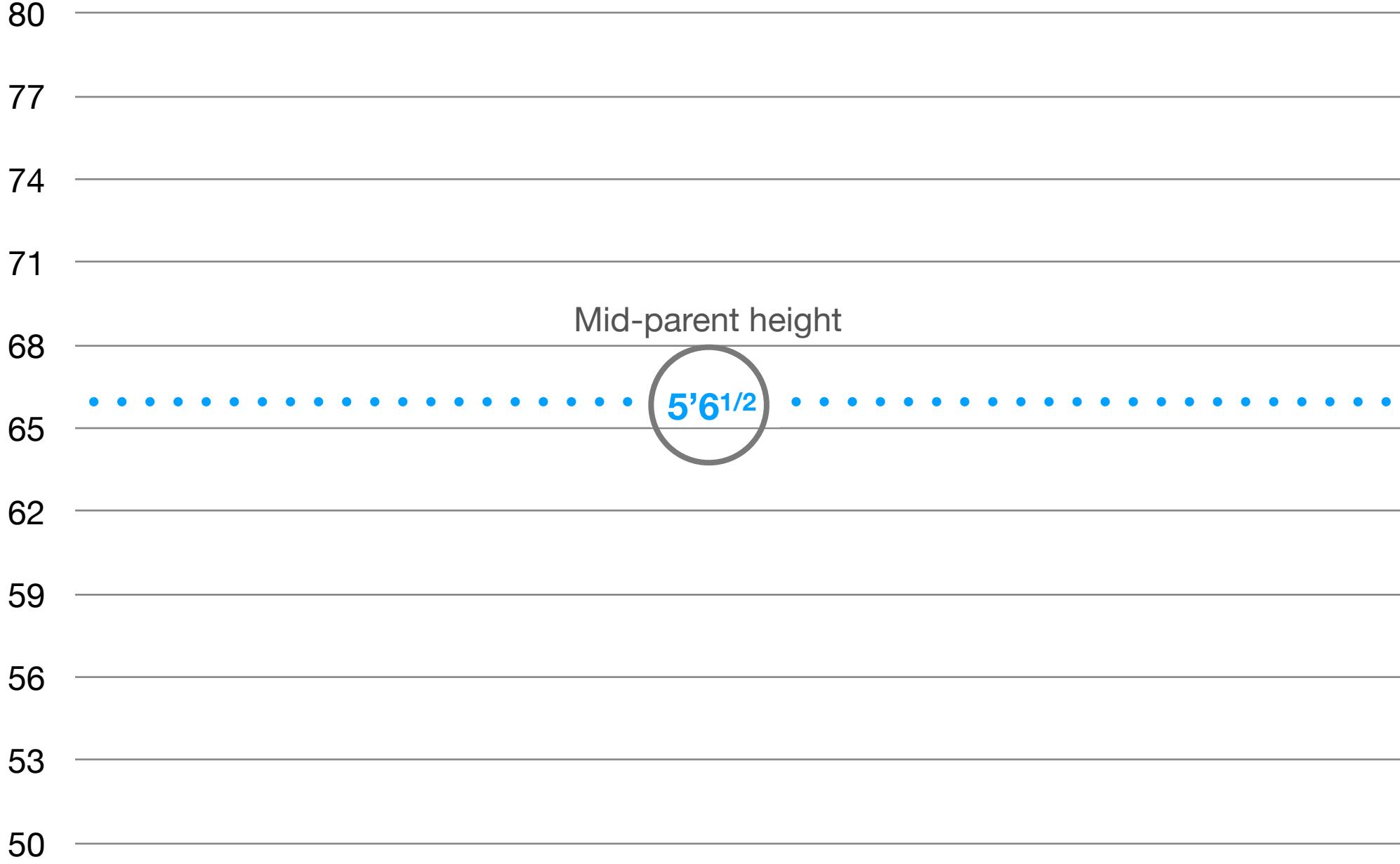


Mid-parent Height

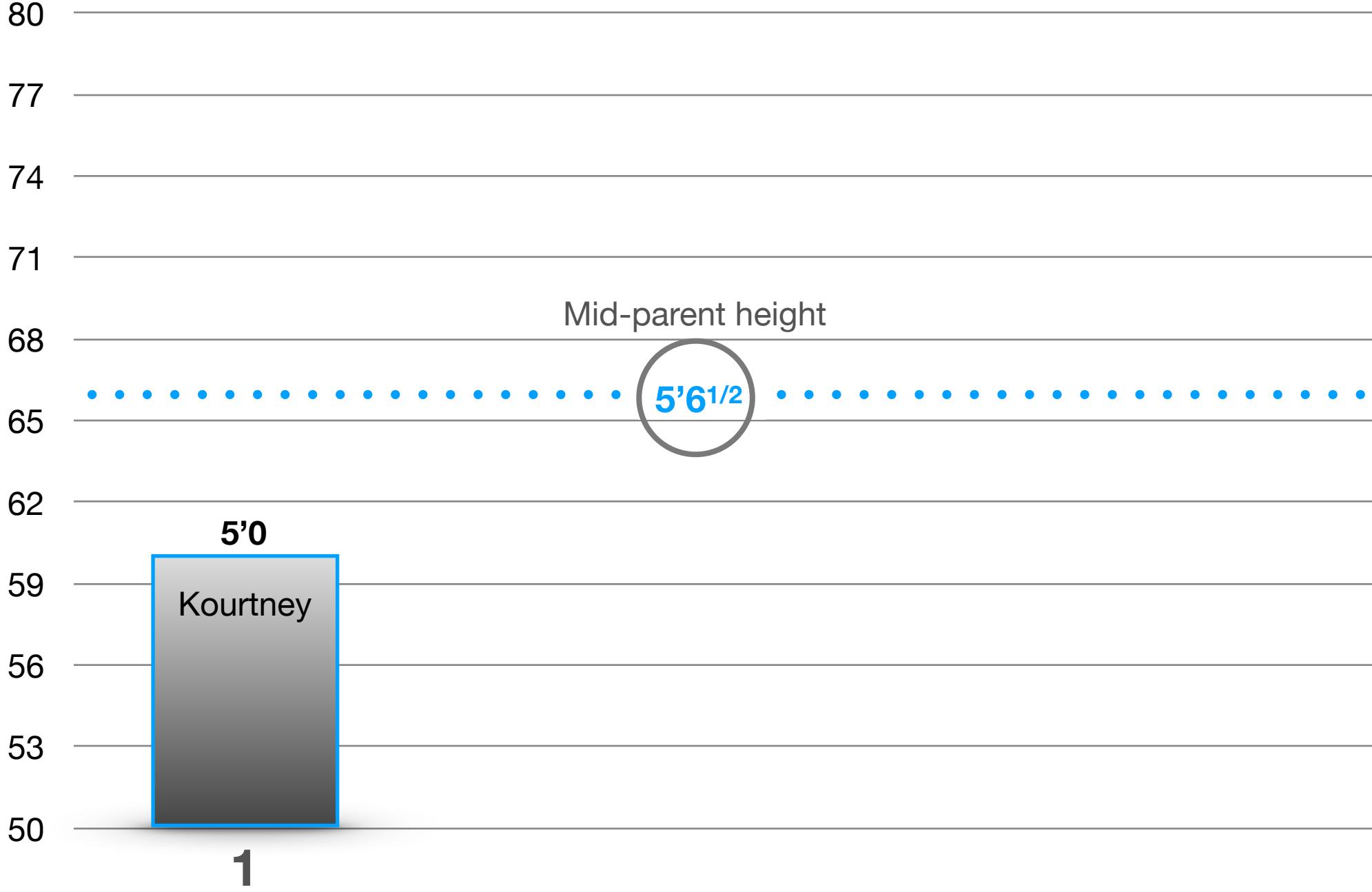
5'6.5



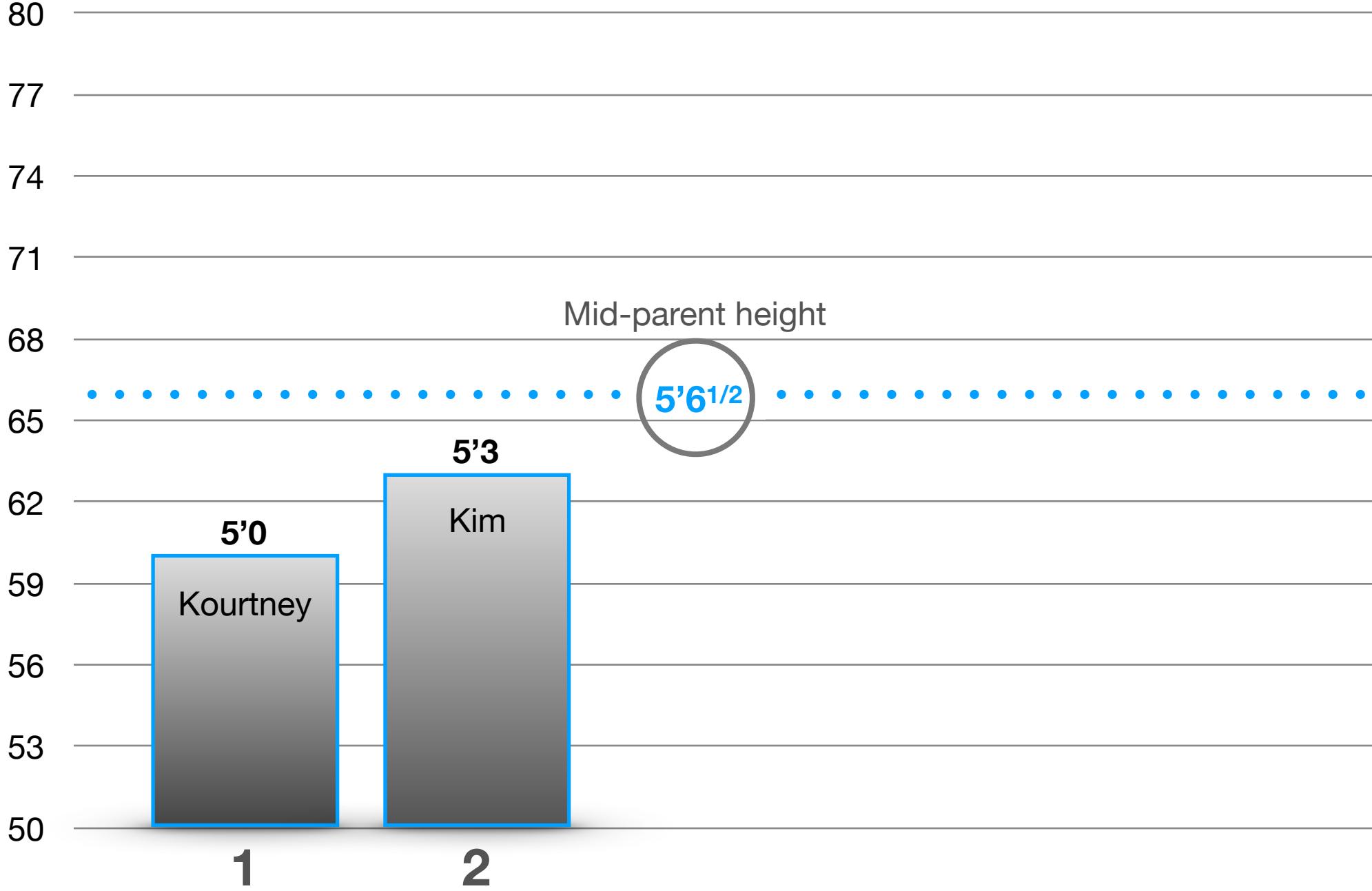
Child's Adult Height



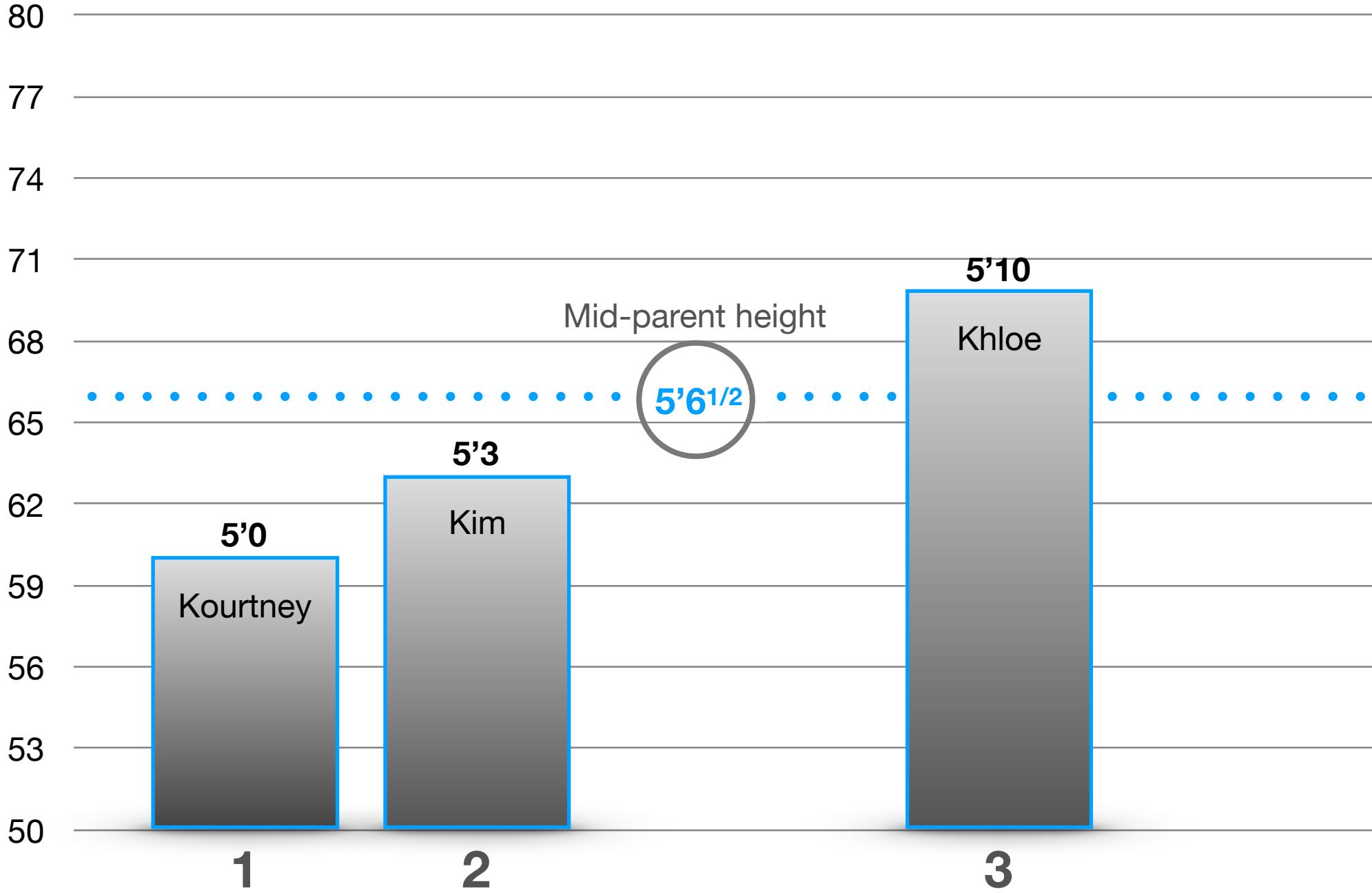
Child's Adult Height

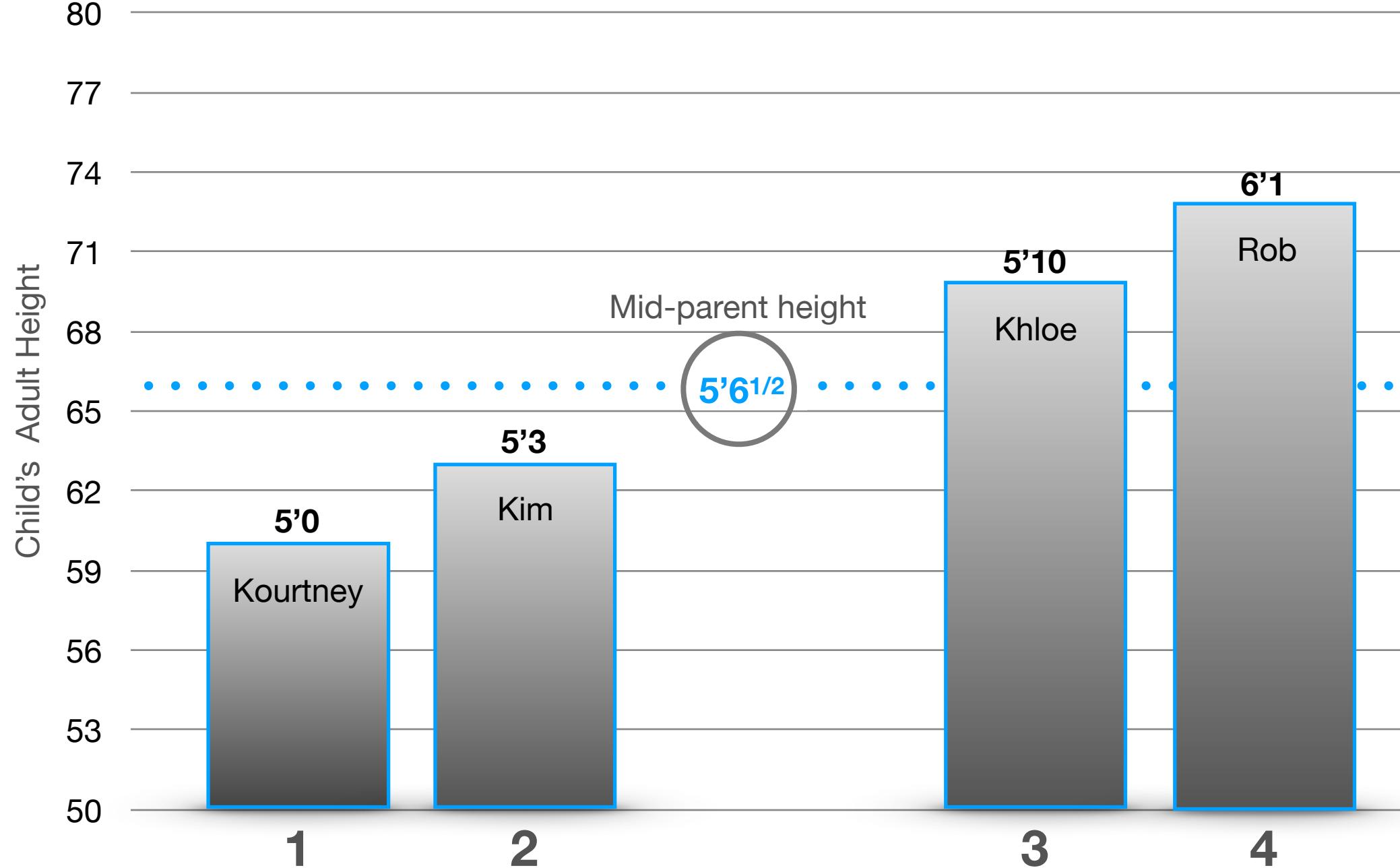


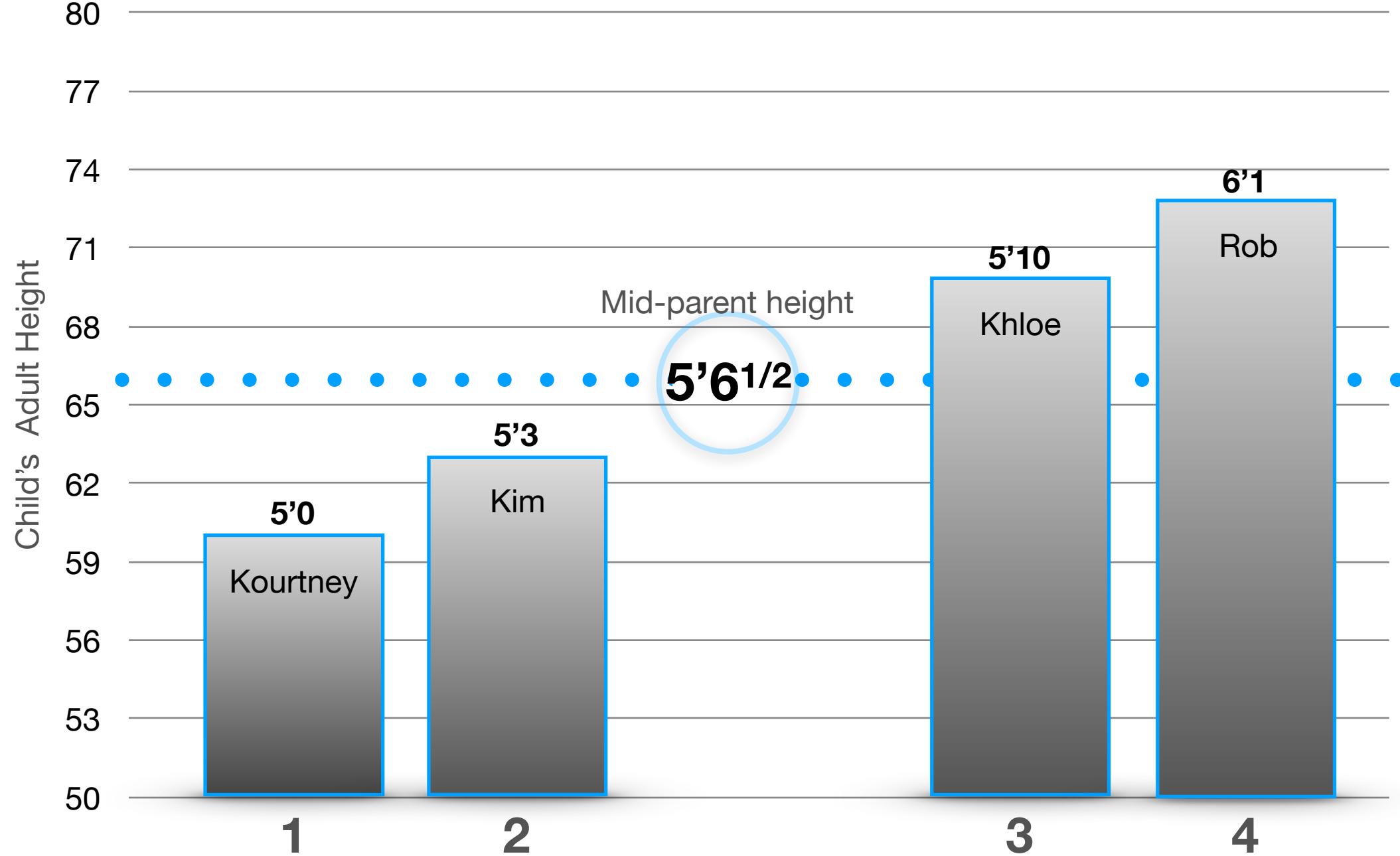
Child's Adult Height



Child's Adult Height

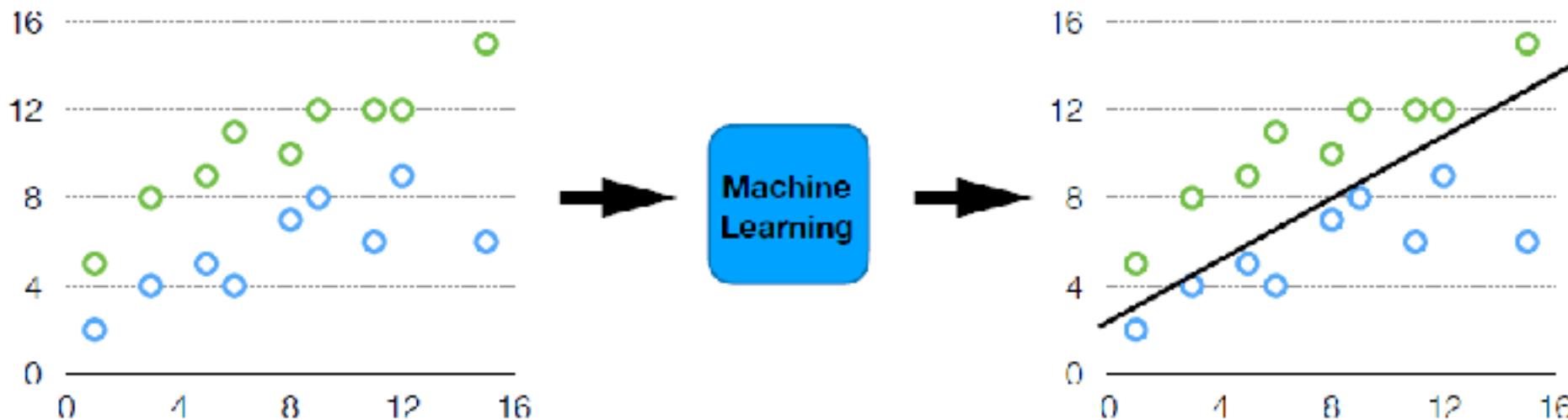


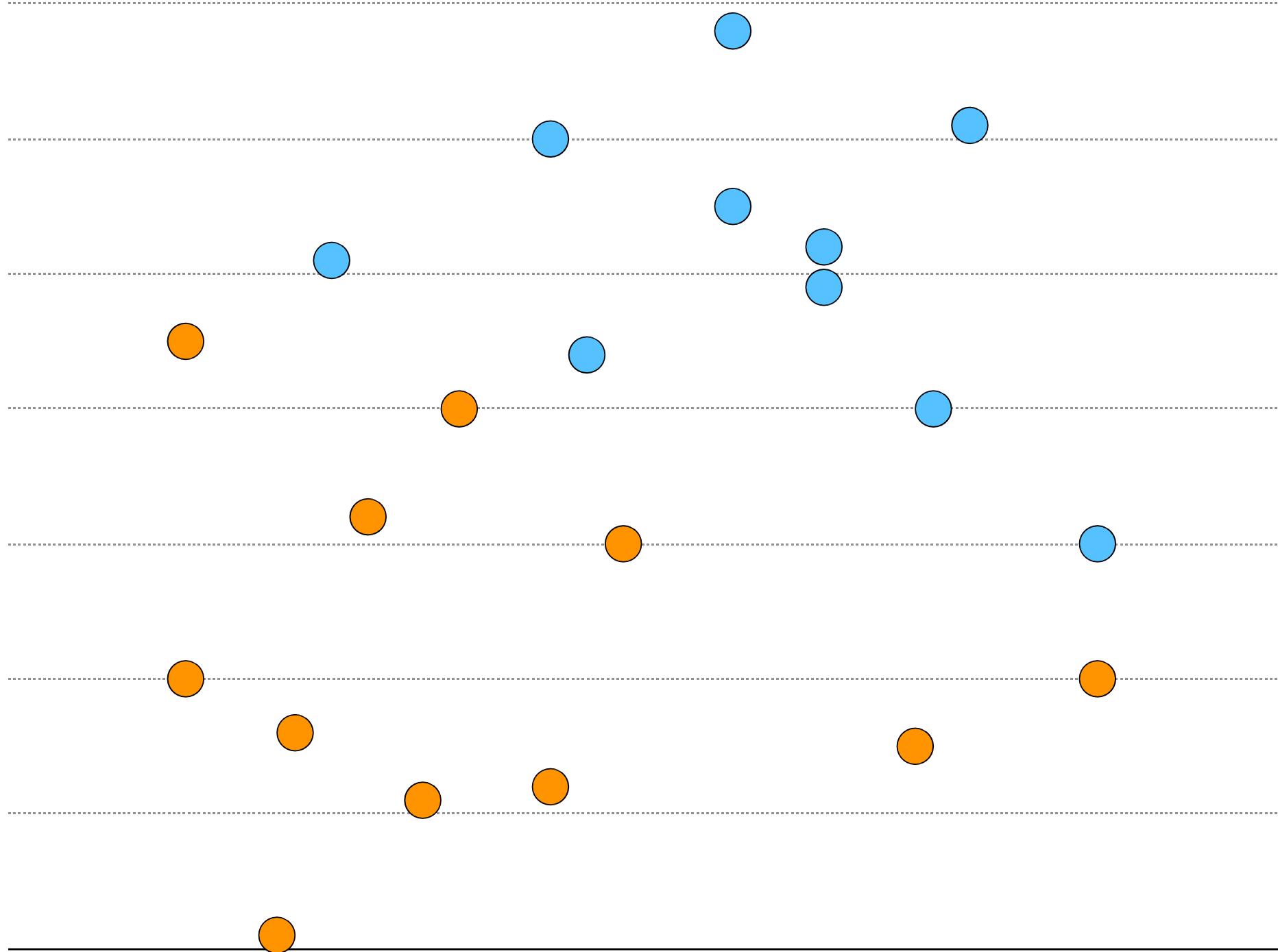


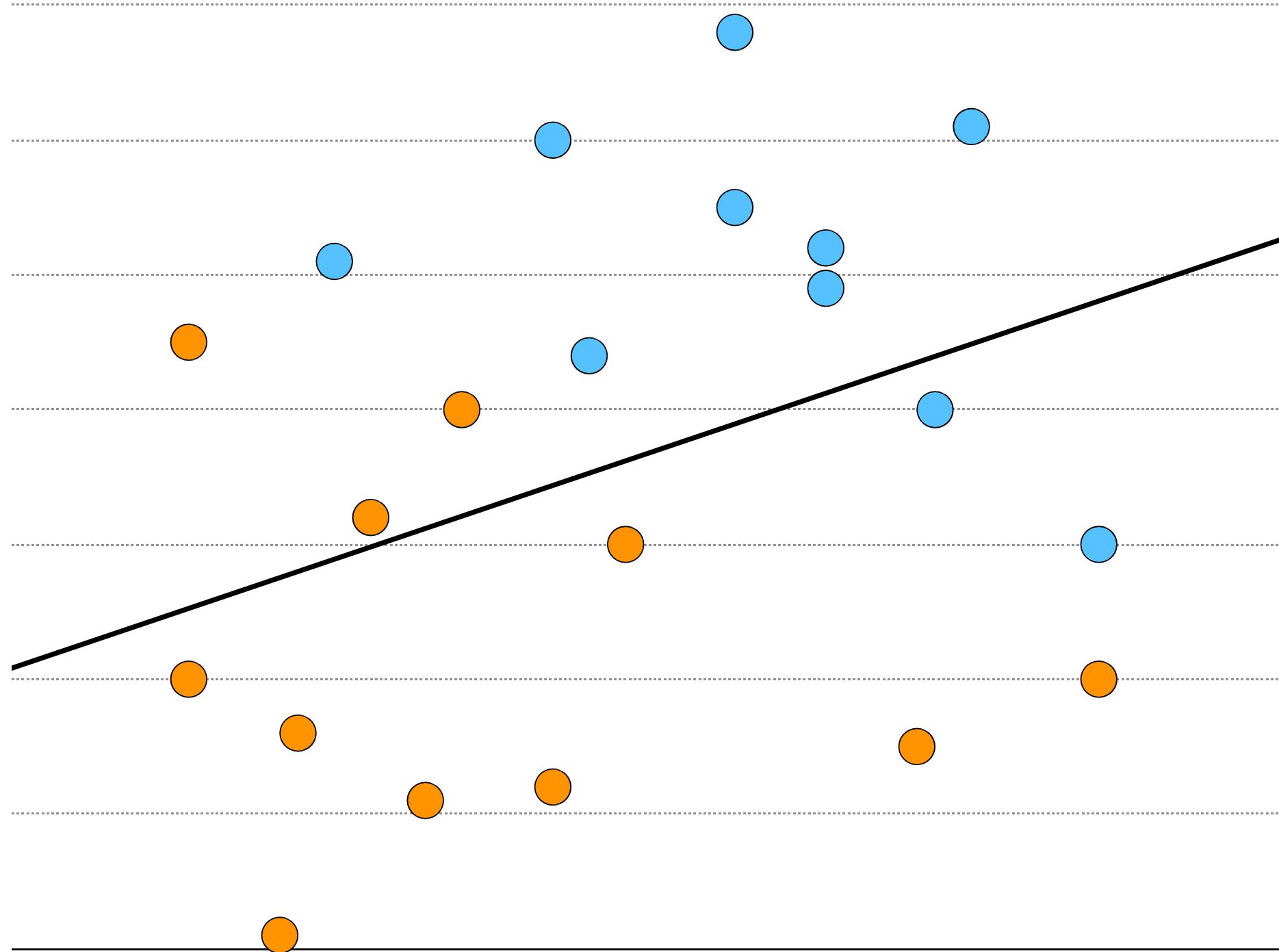


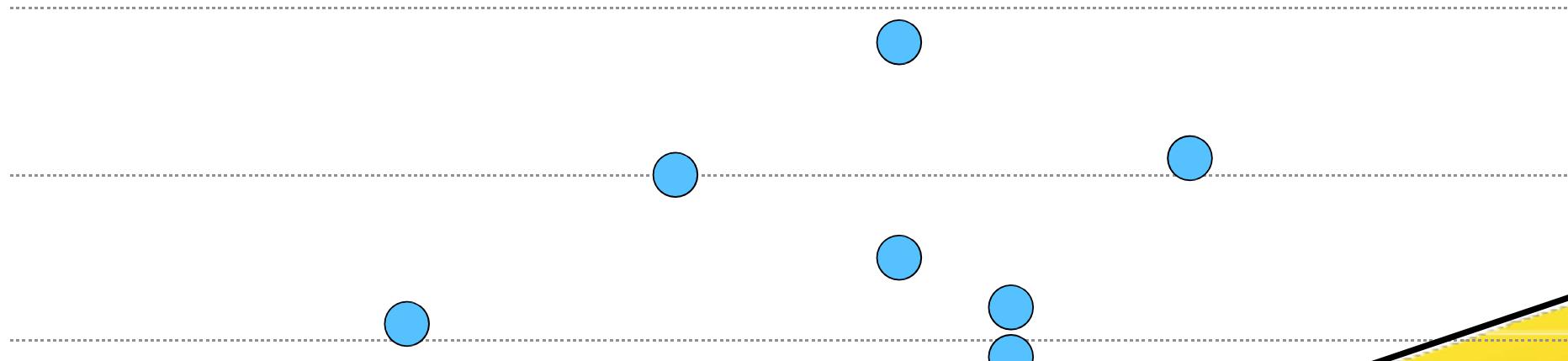
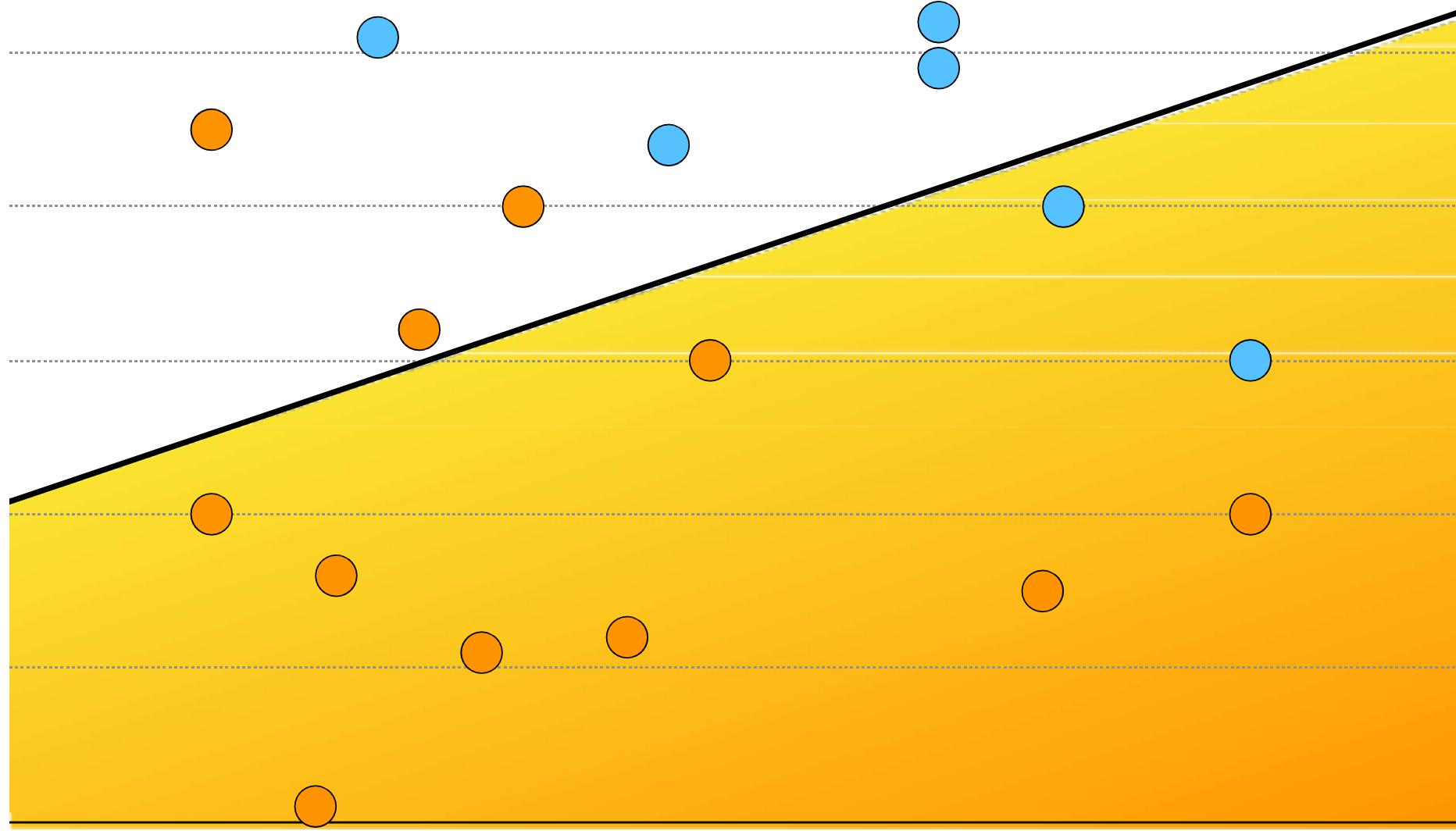
Classification

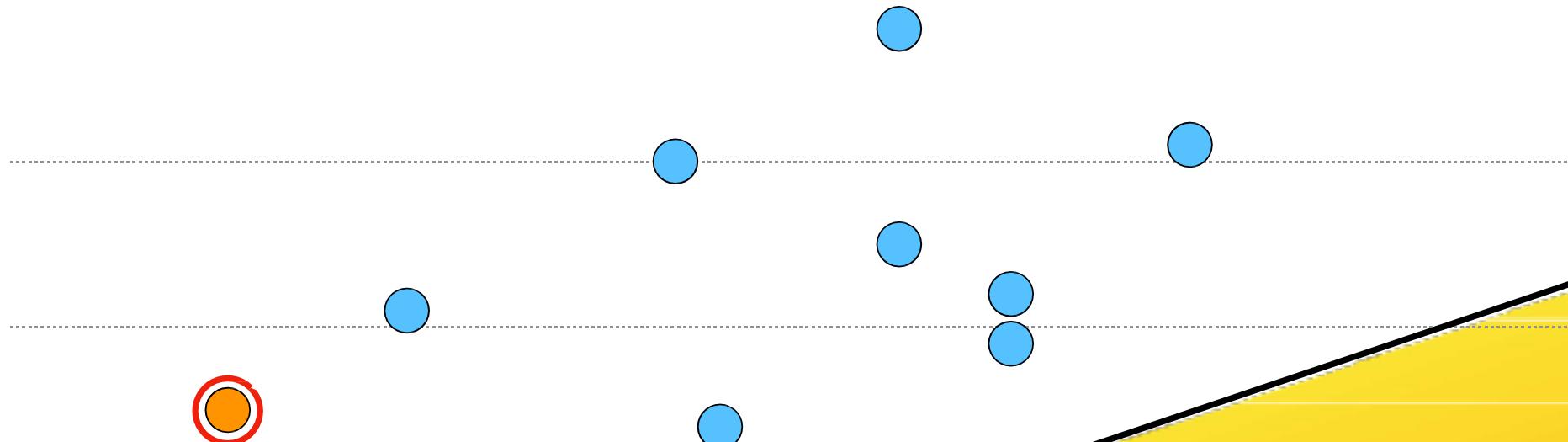
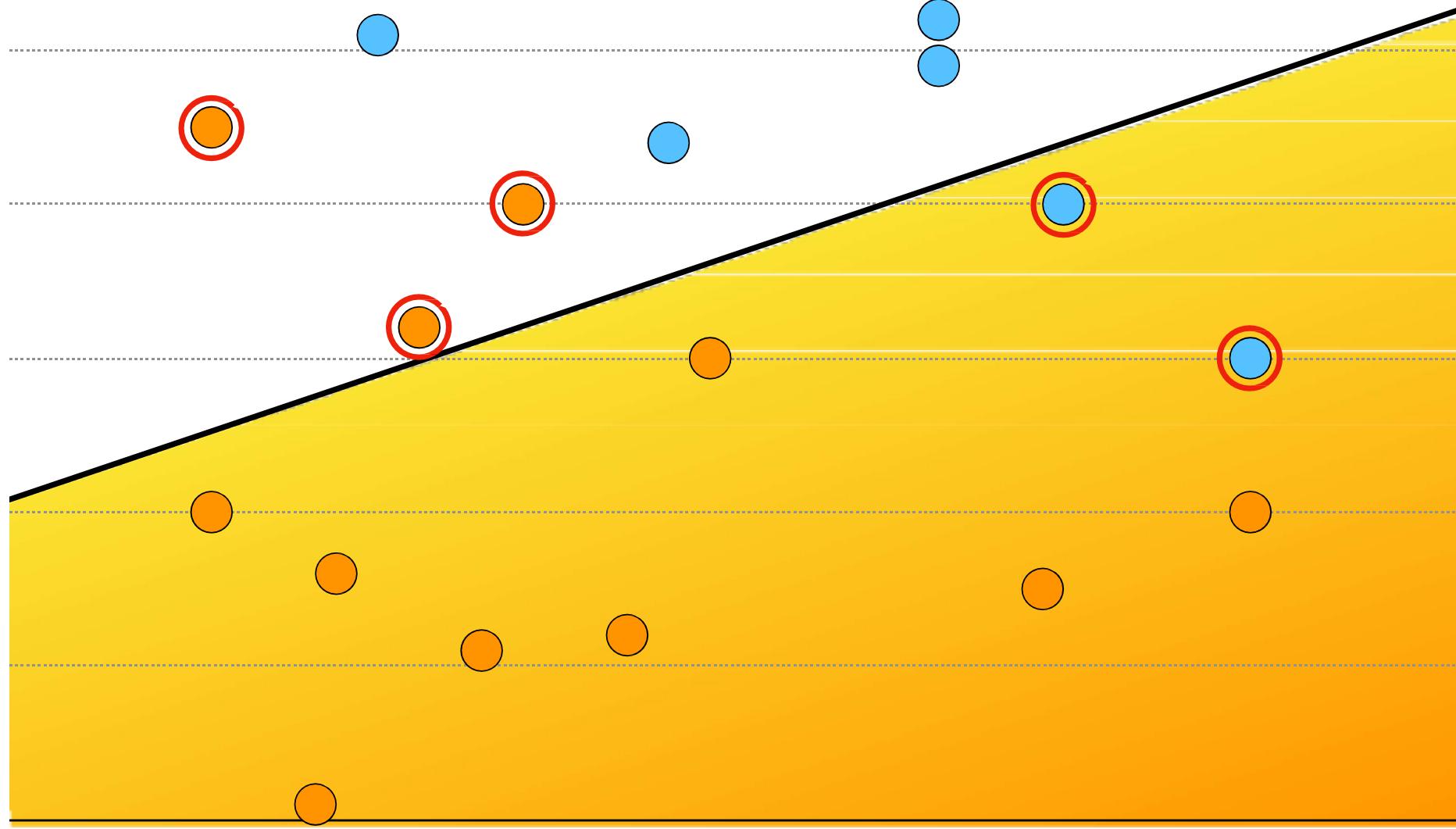
What is a classifier?



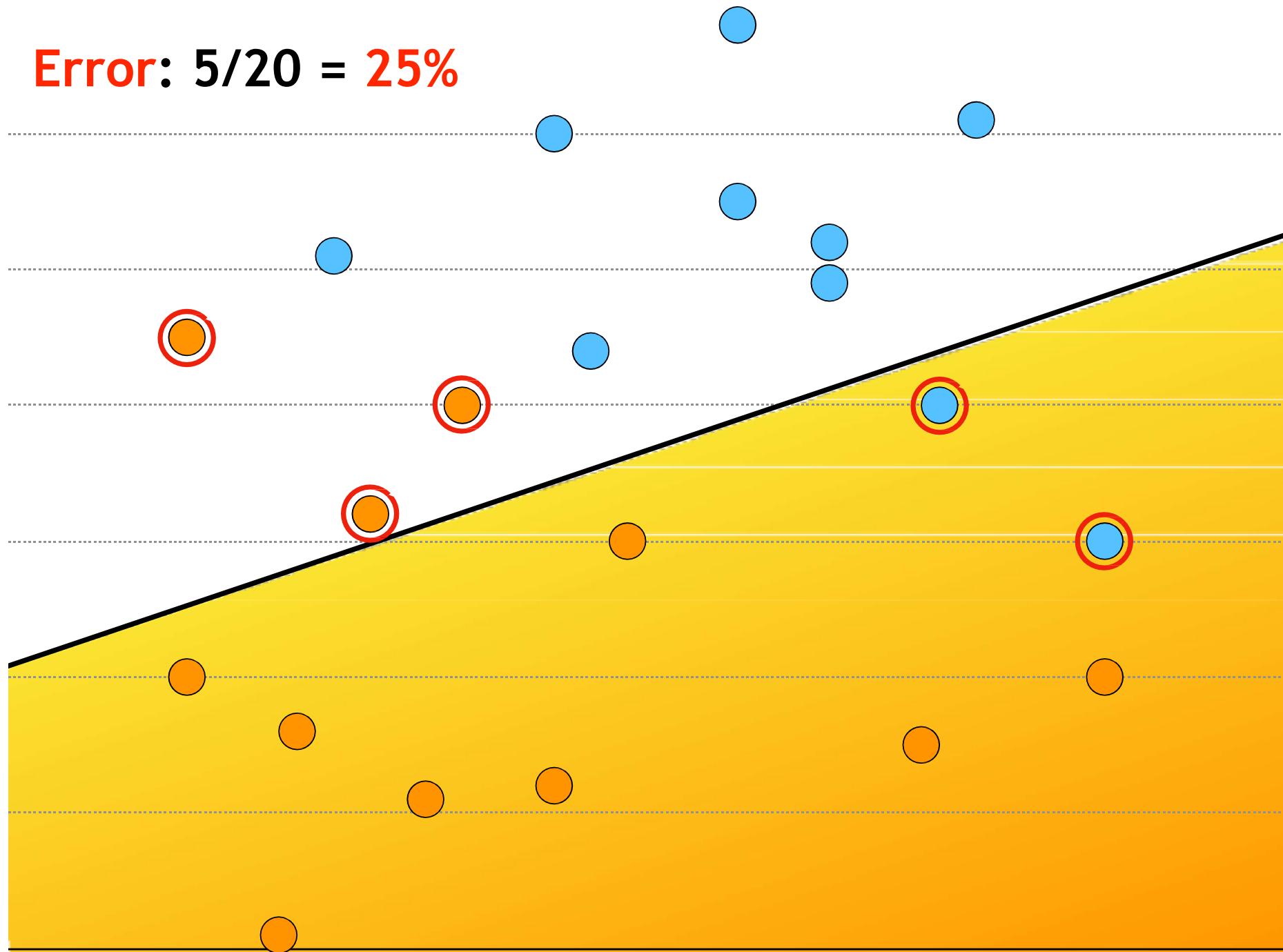


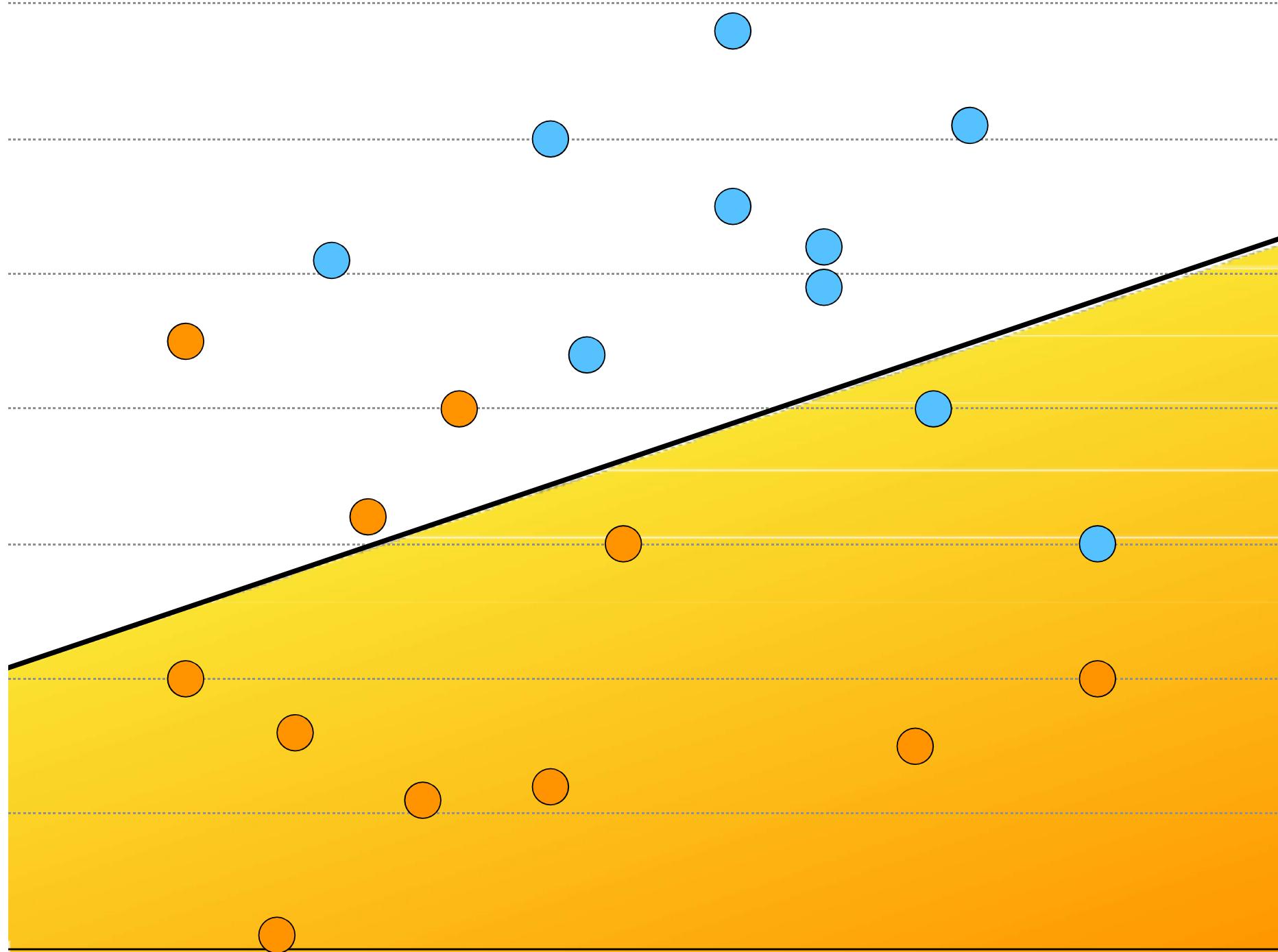


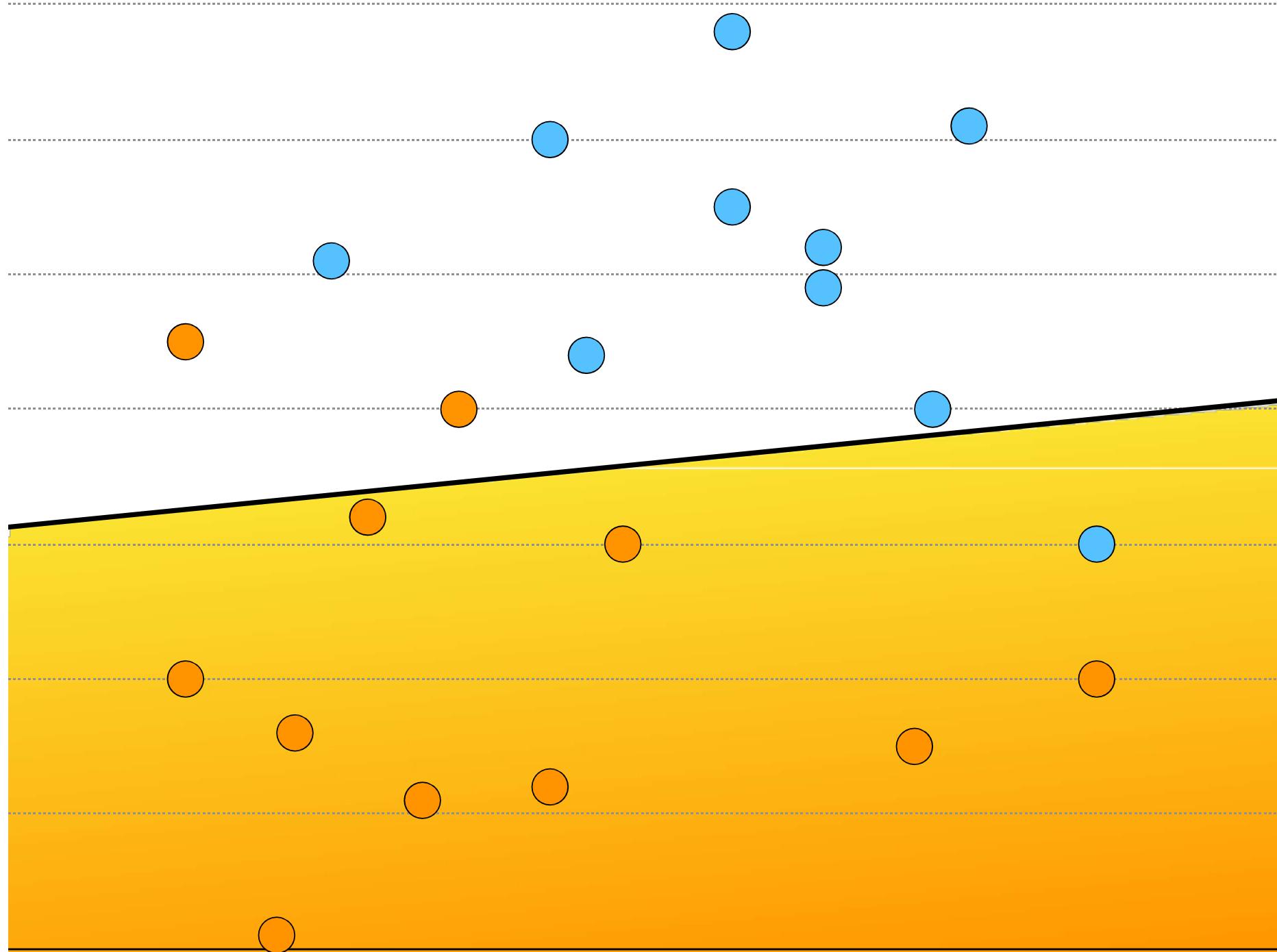


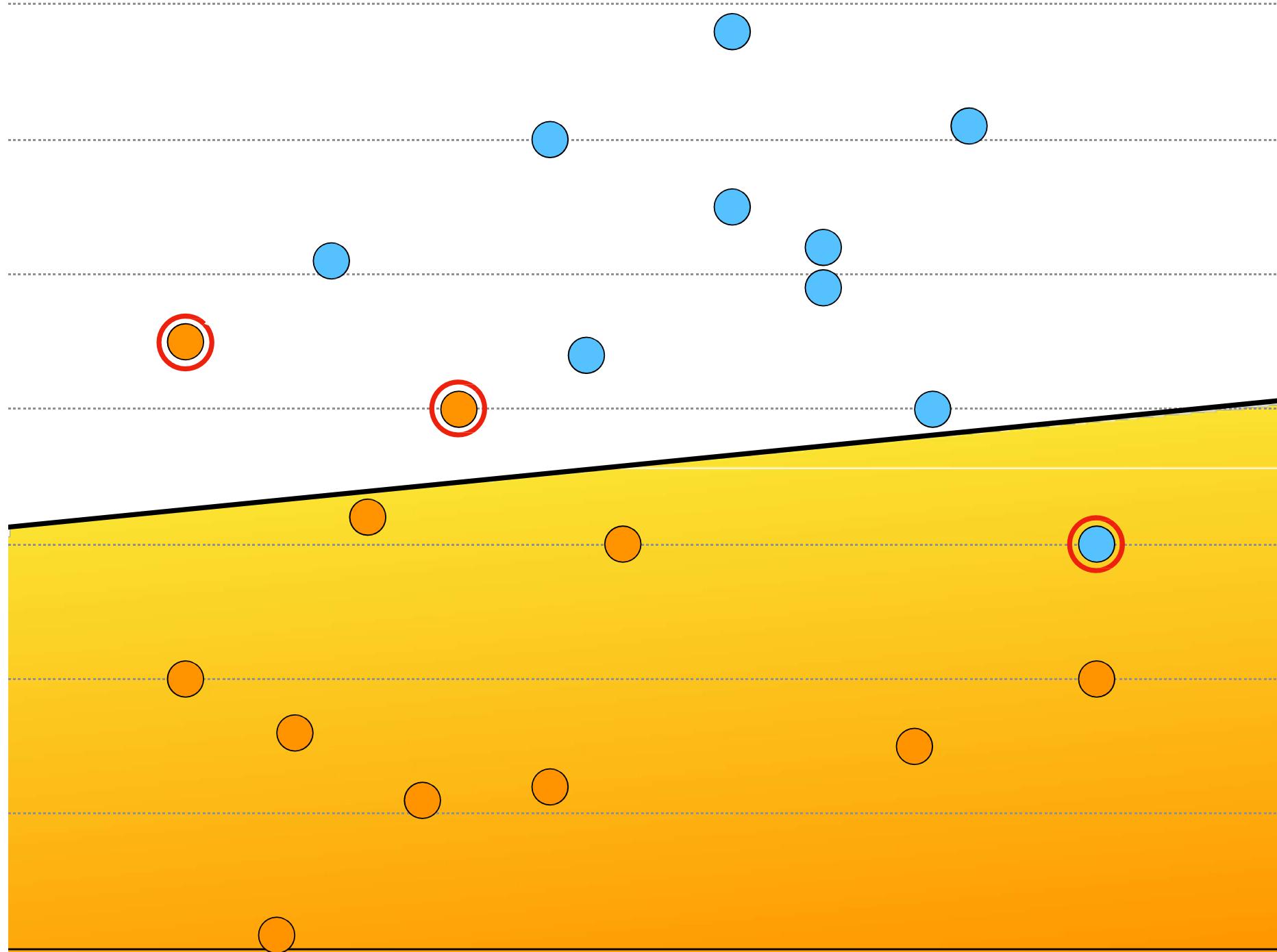


Error: $5/20 = 25\%$

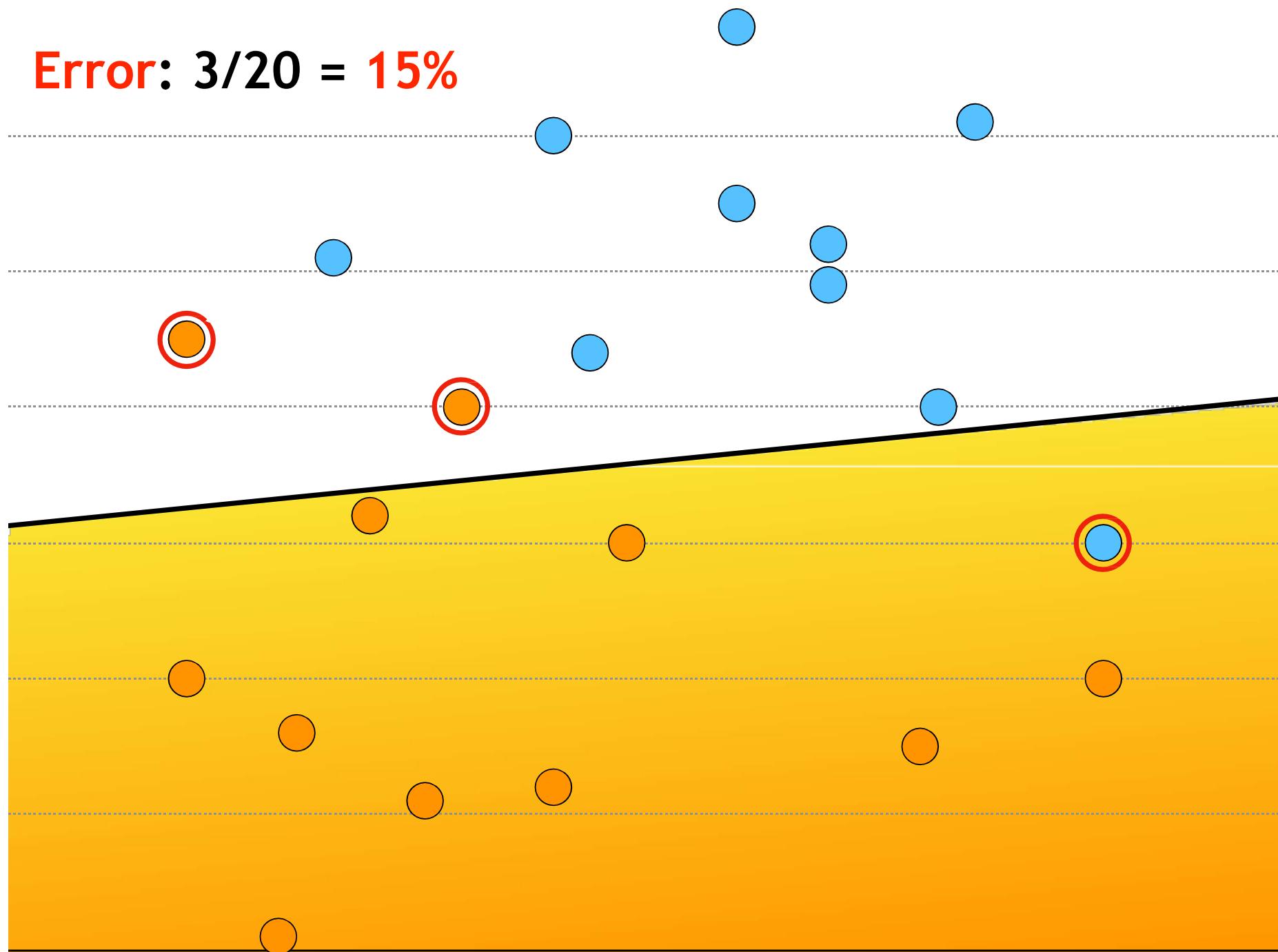


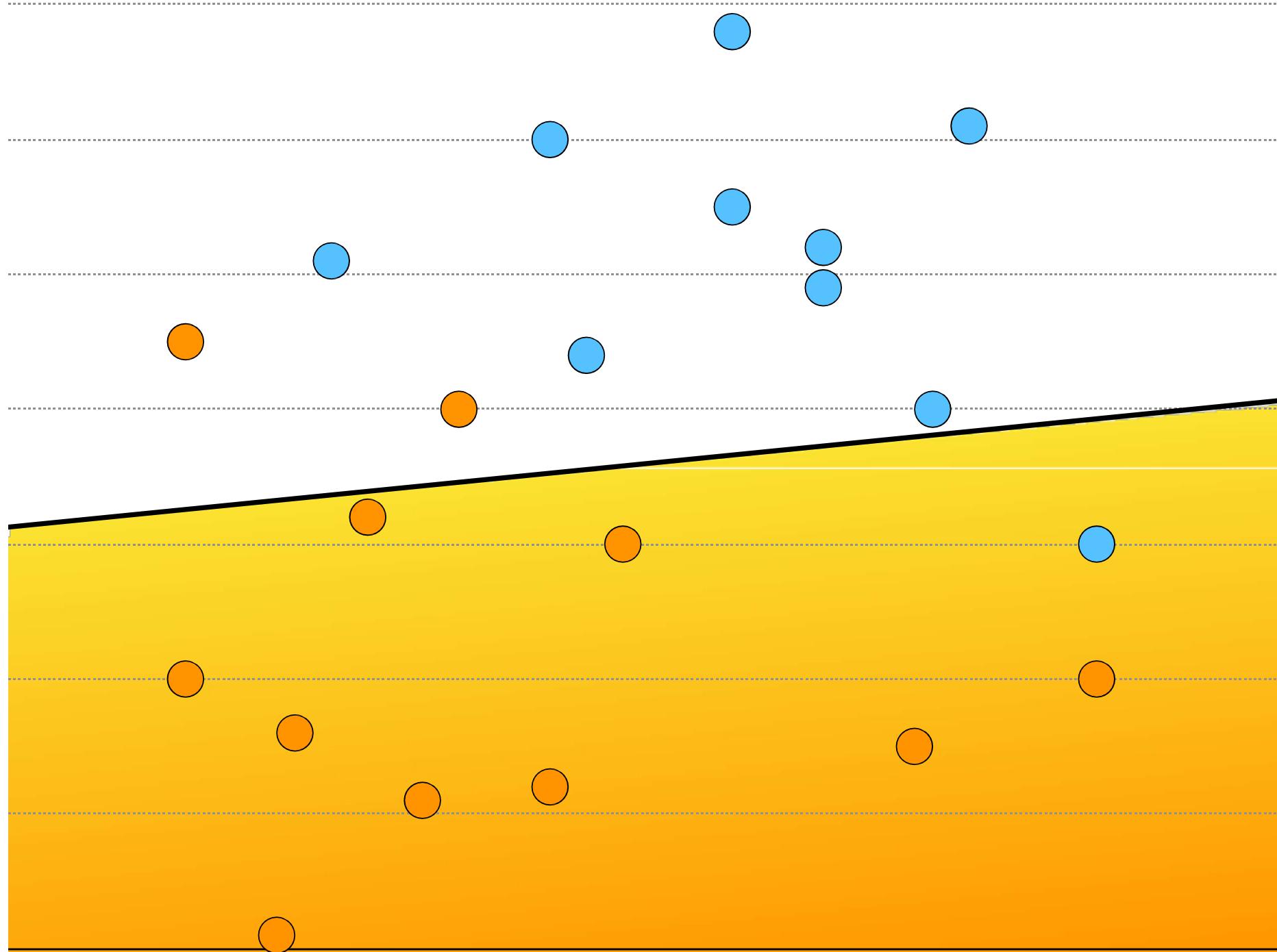


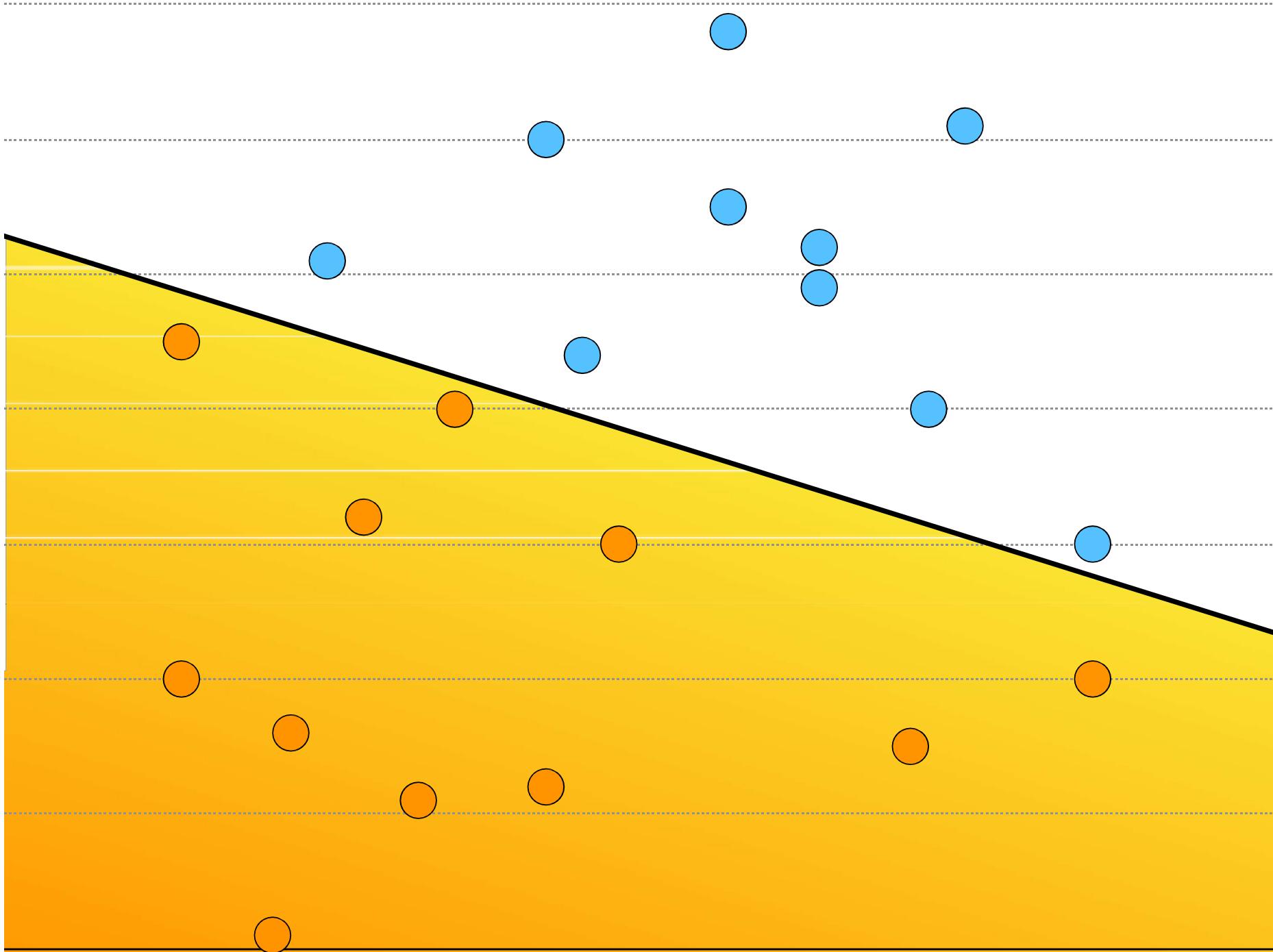




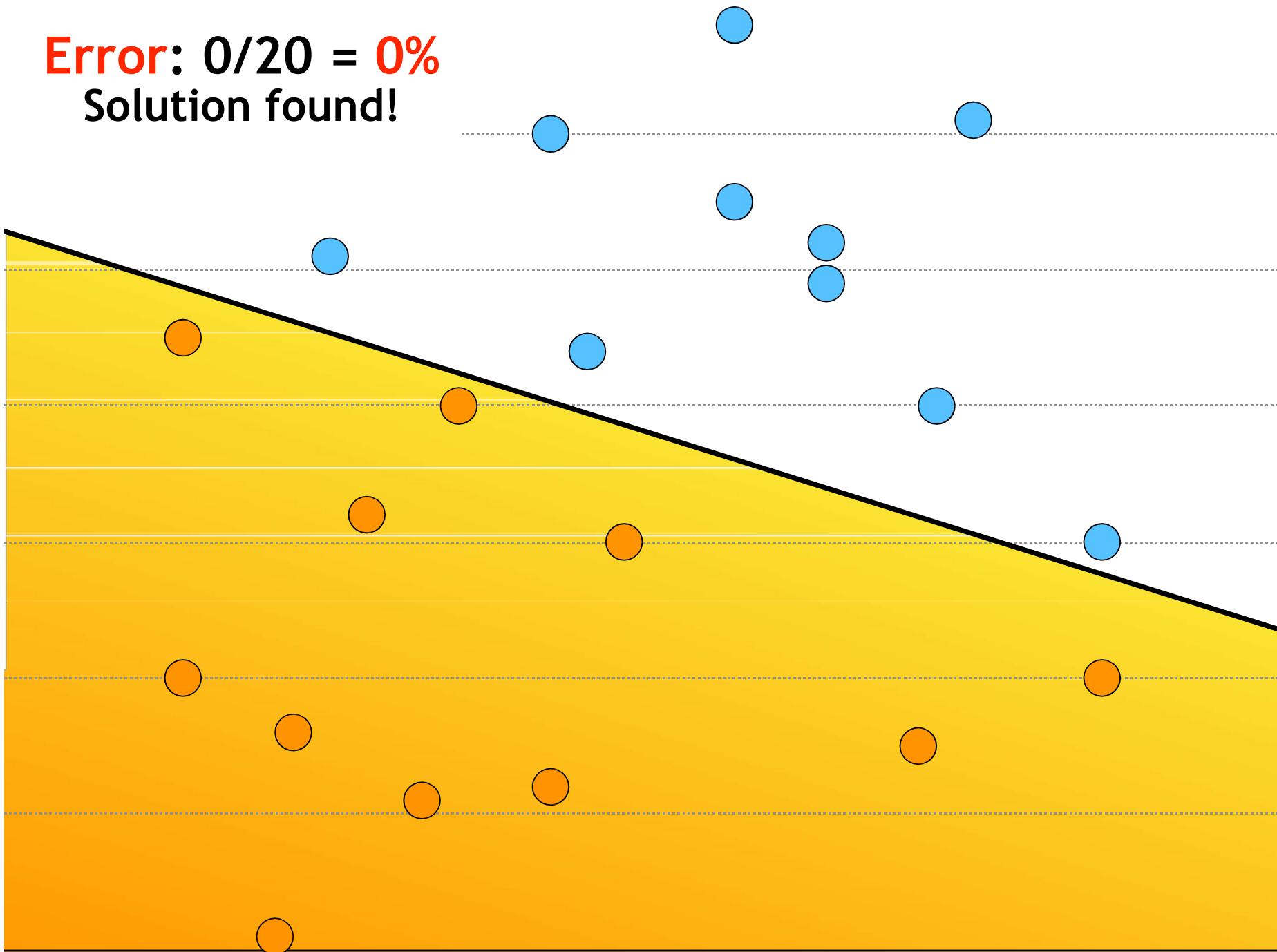
Error: $3/20 = 15\%$







Error: 0/20 = 0%
Solution found!



Logistic Regression

Logistic Regression

Binary Classifier

Logistic Regression

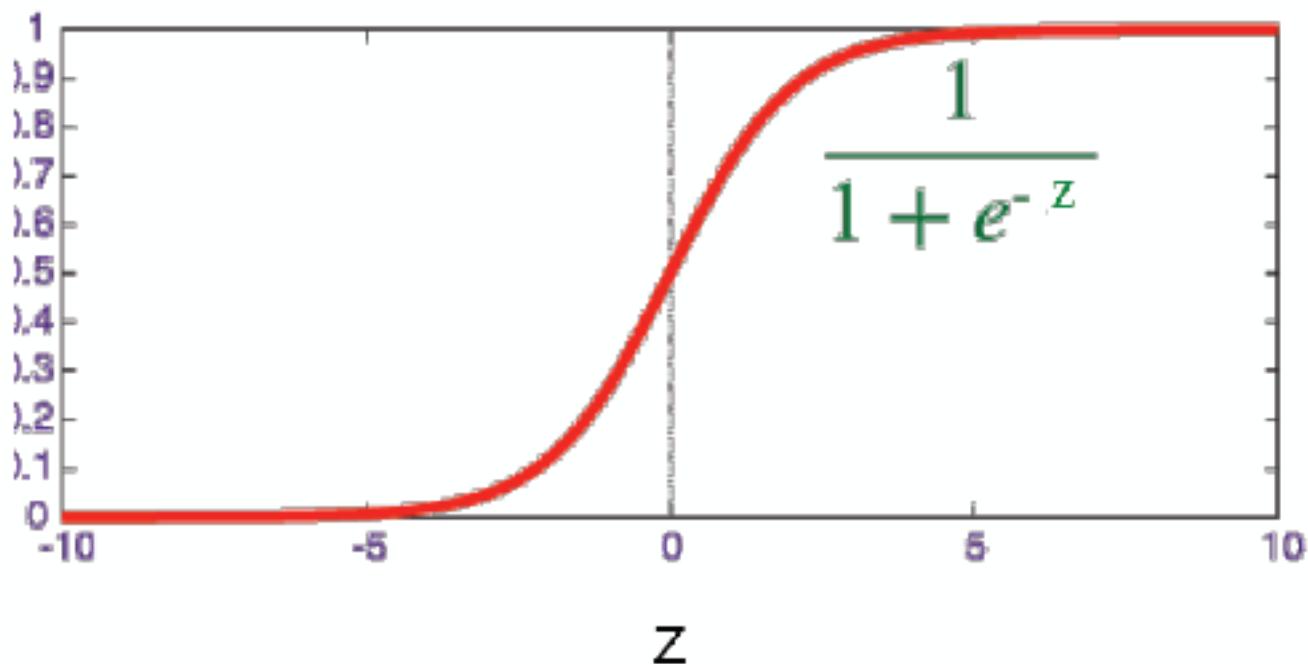
$$\beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \beta_3 \mathbf{x}_3$$

Logistic Regression

intermediate step

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

Logistic (Sigmoid) Function



Logistic Regression

$$\hat{f}(X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}}$$

Logistic Regression

$$\hat{f}(X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}}$$

Logistic Regression

$$\hat{f}(X) = \frac{1}{1 + \text{a very large number}}$$

Logistic Regression

$$\hat{f}(X) = \frac{1}{1 + \text{a very large number}}$$

approaches 0

Logistic Regression

$$\hat{f}(X) = \frac{1}{1 + \text{a very small number}}$$

Logistic Regression

$$\hat{f}(X) = \frac{1}{1 + \text{a very small number}}$$

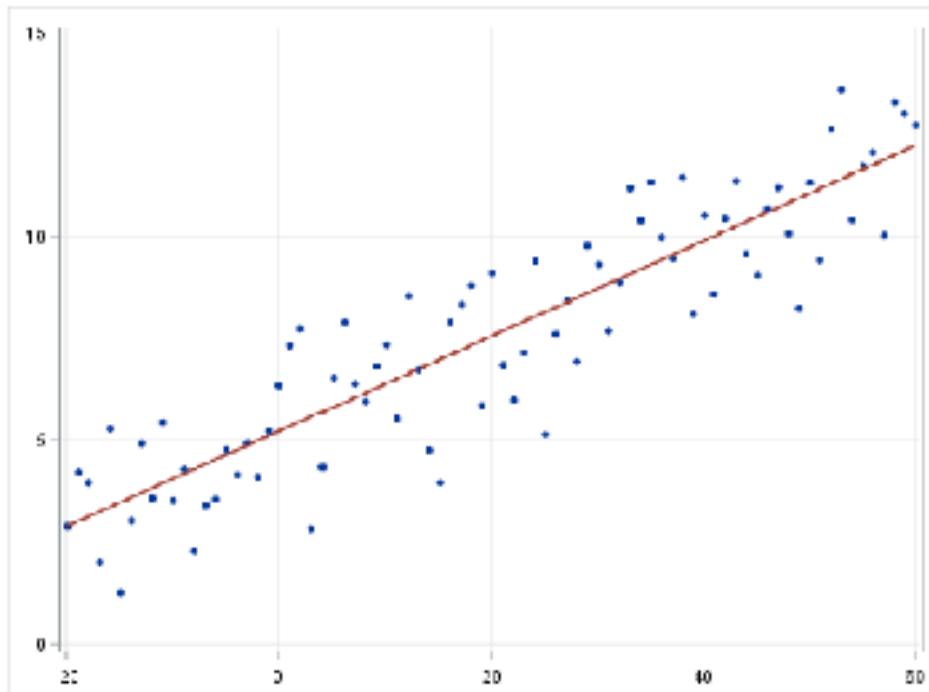
approaches 1

Logistic Regression

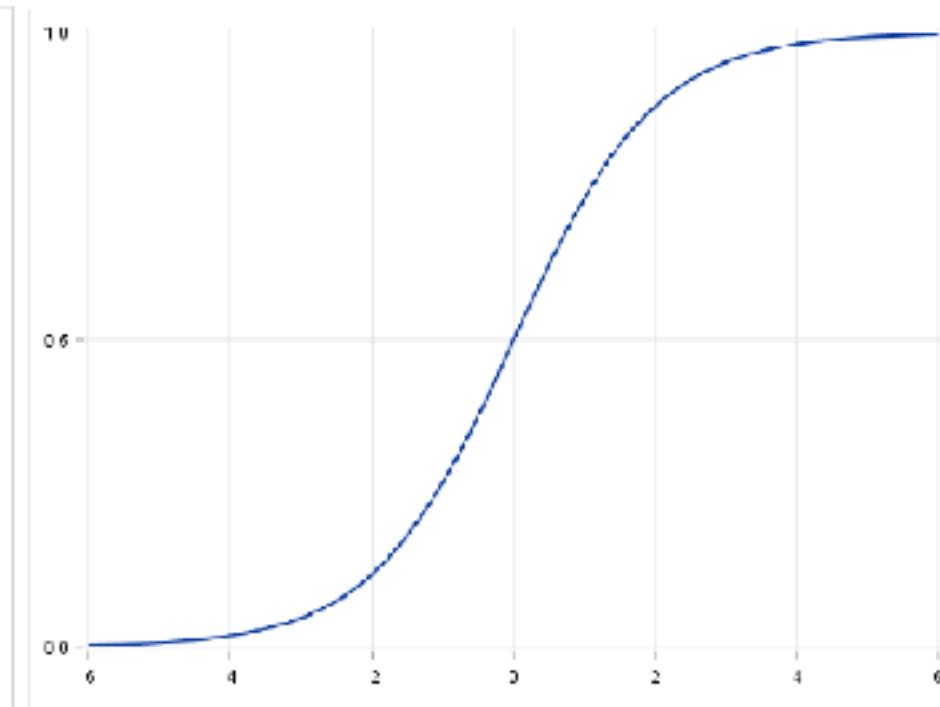
$$\hat{f}(X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}}$$

Gives us probabilities
(values between 0 and 1)

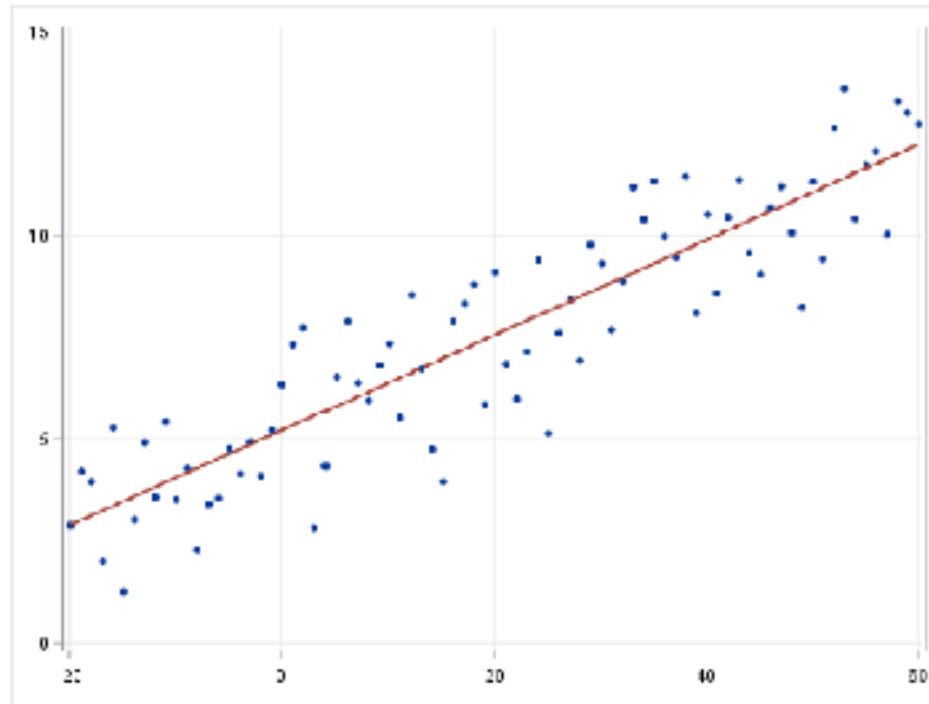
Linear Regression



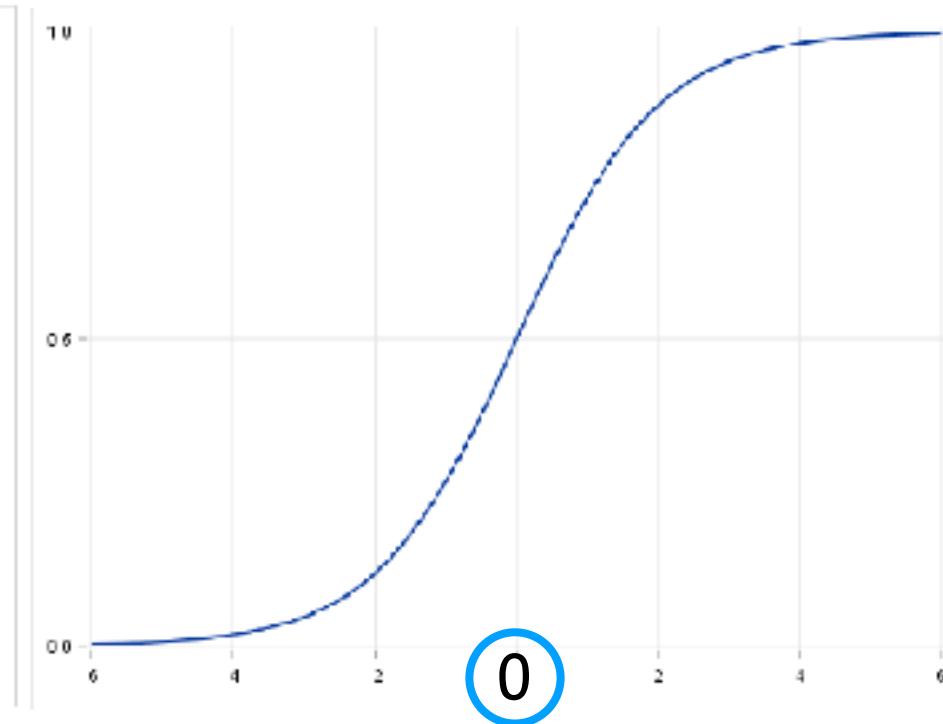
Logistic Regression



Linear Regression

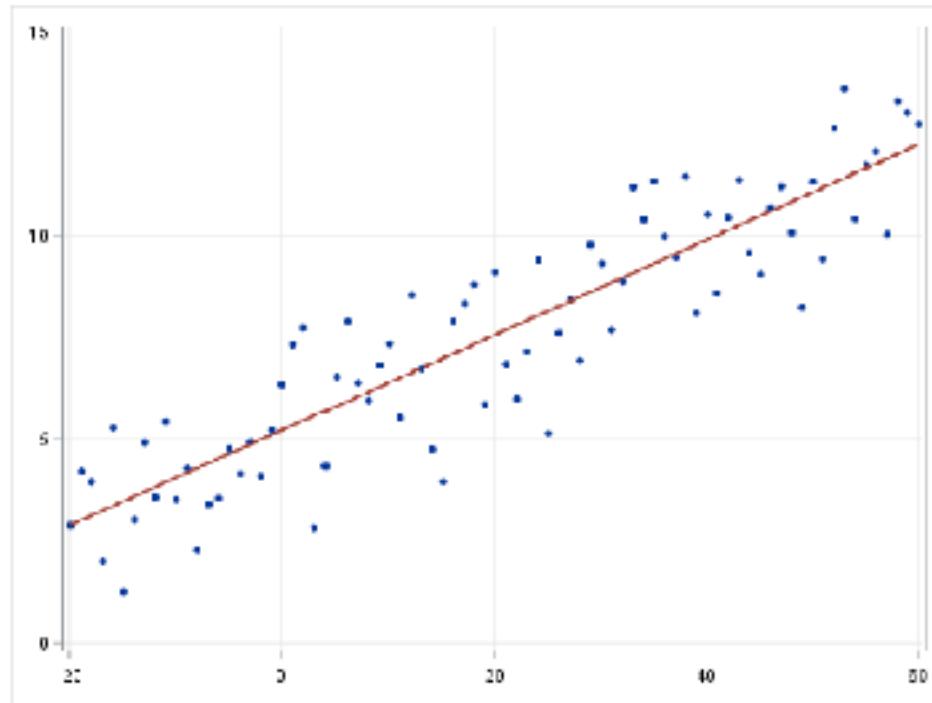


Logistic Regression

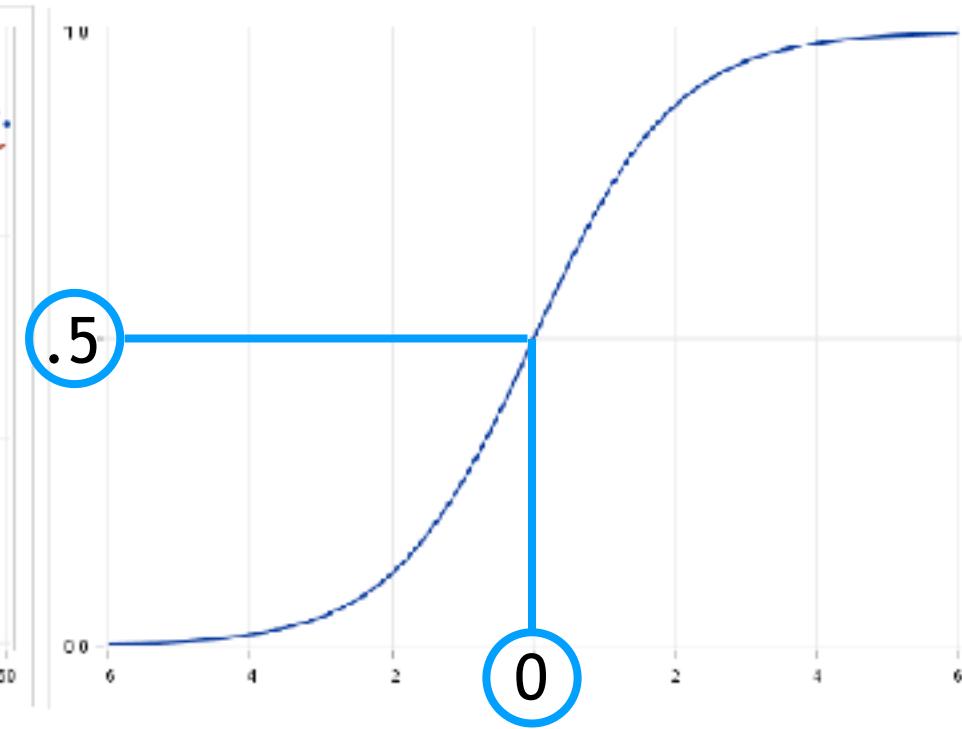


decision boundary

Linear Regression



Logistic Regression



decision boundary

Logistic Regression

positive class = 1

negative class = 0

predict 1 when $\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 \geq 0$

predict 0 when $\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 < 0$

The New York Times

The New York Times

*Who's Wearing a Mask? Women,
Democrats and City Dwellers*



Data

Race 1	Race 2	Race 3	Race 4	Gender	Age	City Dwell	Mask
1	0	0	0	0	49	1	1
0	0	0	0	1	43	0	0
0	0	1	0	0	37	0	0
0	1	0	0	0	28	1	1
0	1	0	0	0	25	1	0
0	0	0	1	0	23	0	0
0	0	1	0	0	48	0	1
0	0	0	0	1	33	0	1
1	0	0	0	0	39	1	1
0	0	0	1	0	47	1	0

Logistic Regression

Non-Mask Wearer = $\beta_0 + \beta_1 x_1 + \beta_2 x_2$

Decision Boundary

Non-Mask Wearer = $-6 + 1x_1 + 1x_2$

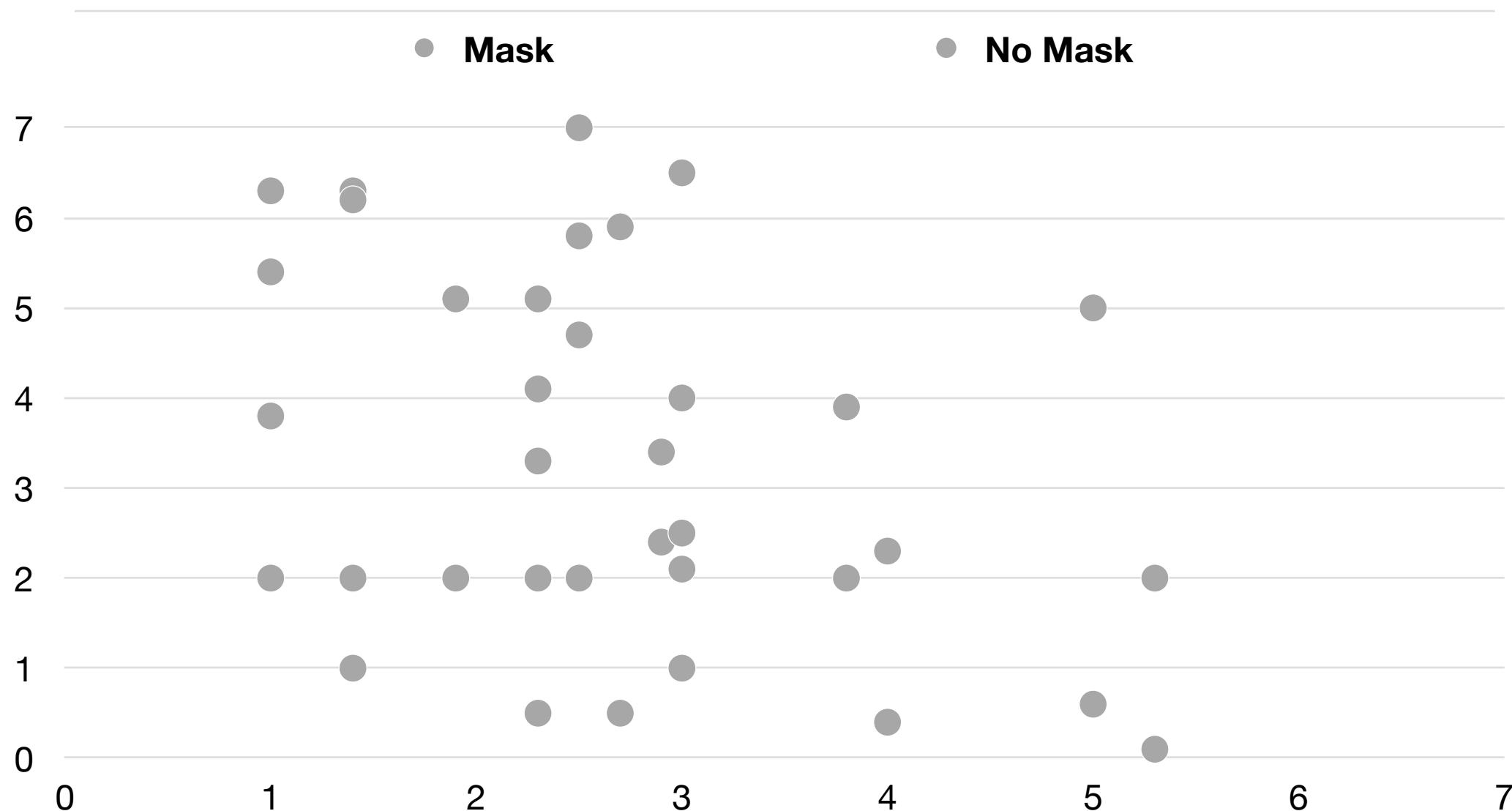
Decision Boundary

Non-Mask Wearer = $-6 + 1x_1 + 1x_2$

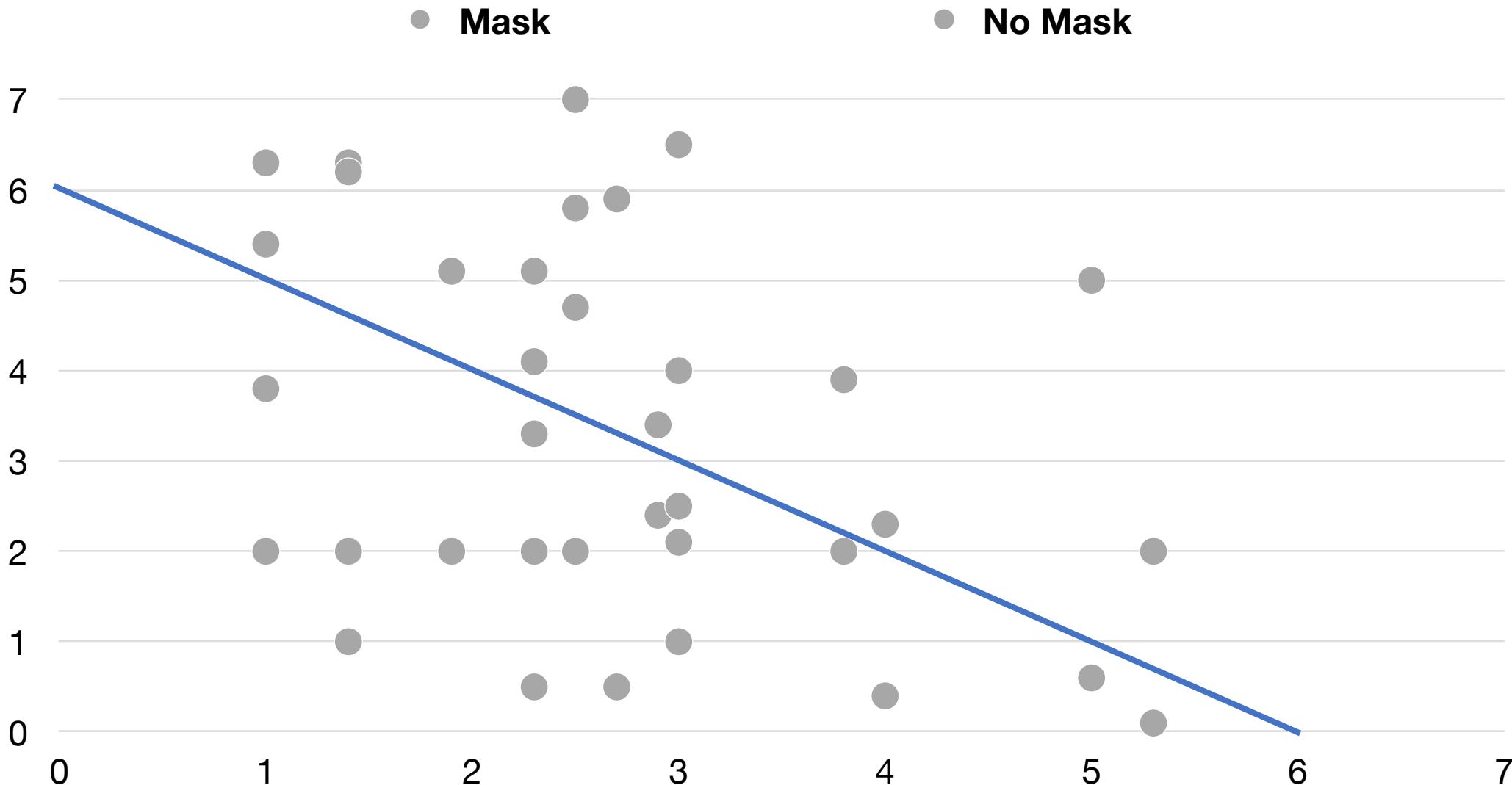
predict 1 when $-6 + 1x_1 + 1x_2 \geq 0$

predict 0 when $-6 + 1x_1 + 1x_2 < 0$

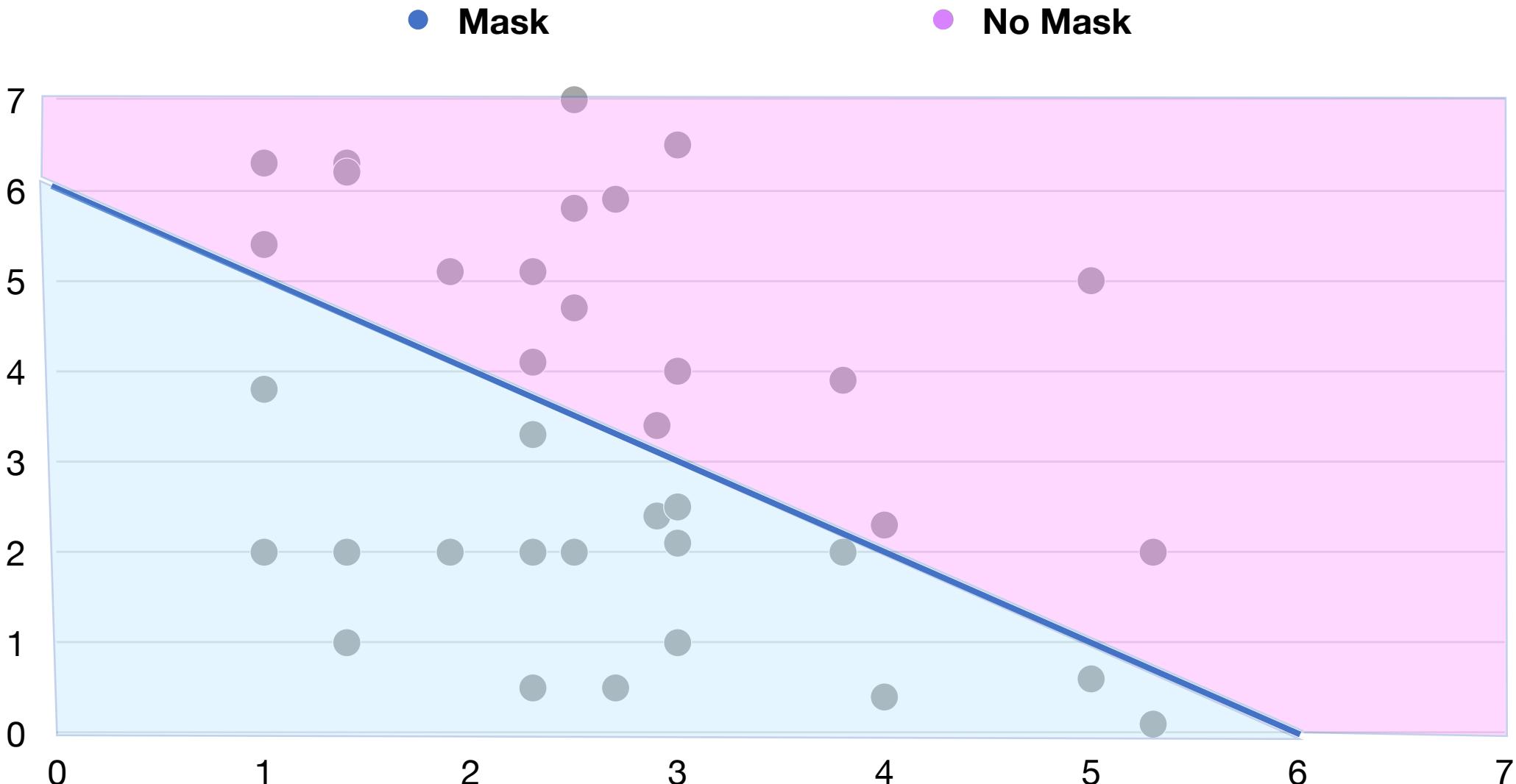
Mask Wearer Decision Boundary



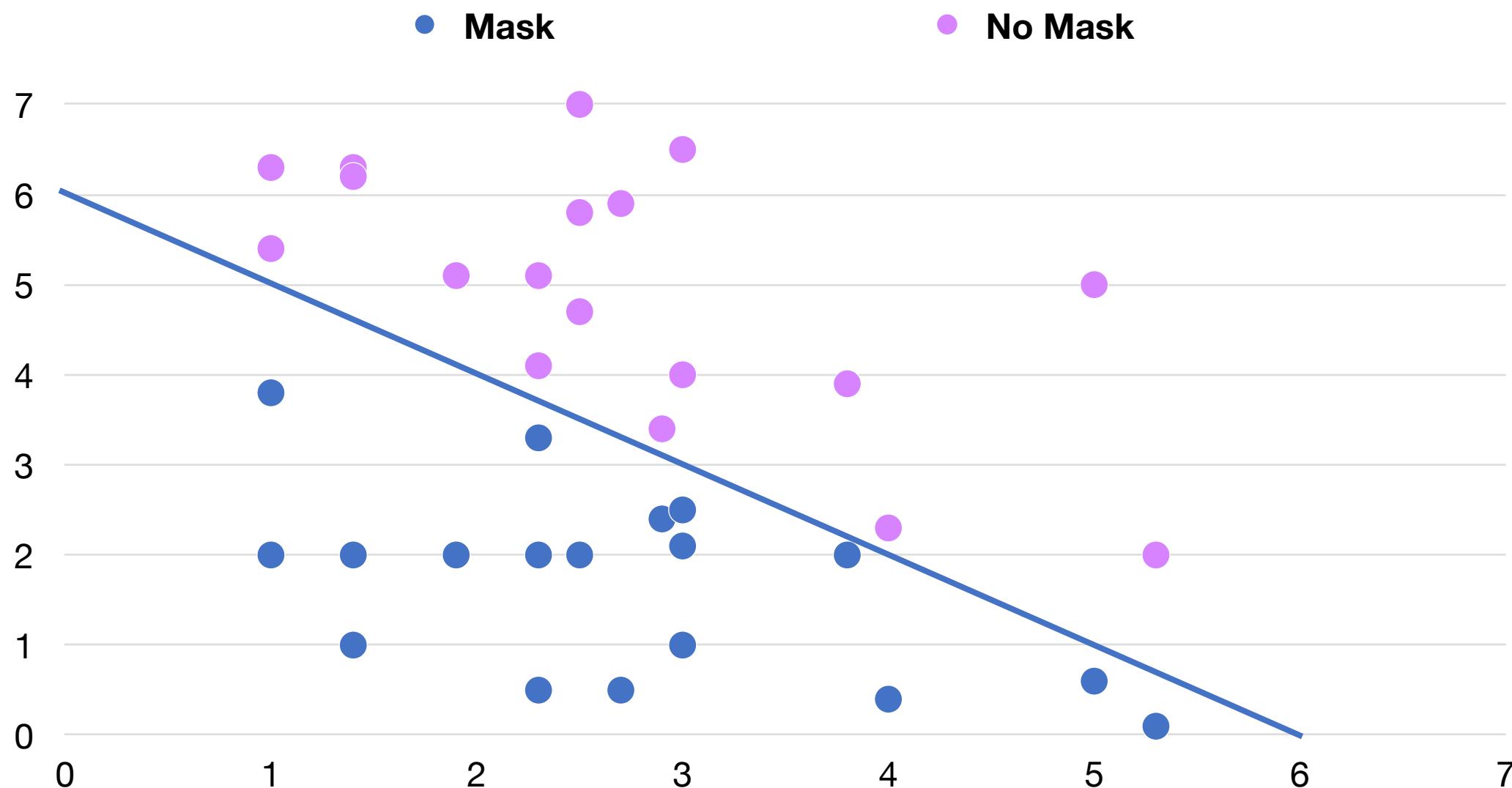
$$\hat{f}(X) = \frac{1}{1 + e^{(-6 + 1x_1 + 1x_2)}}$$



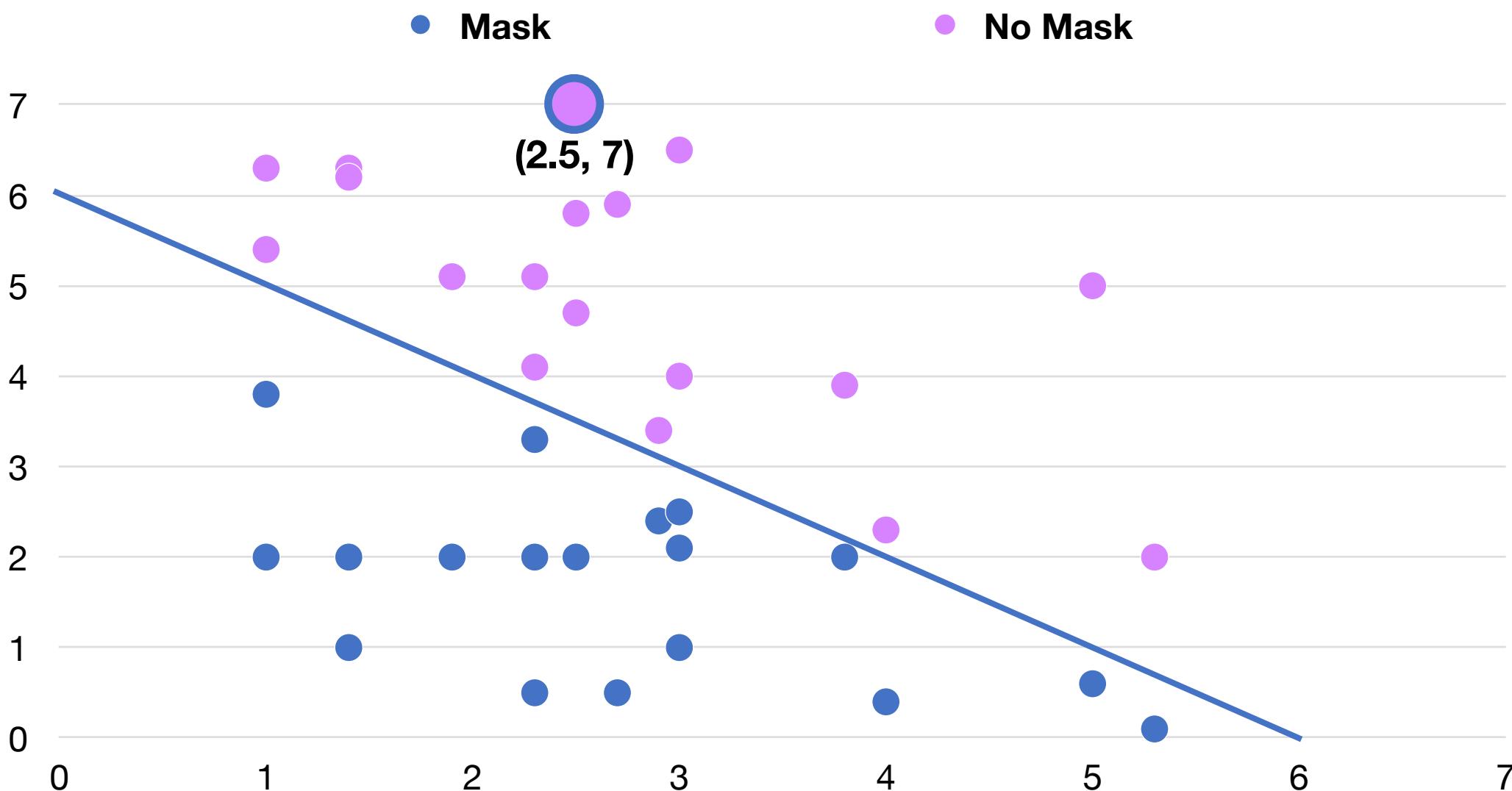
$$\hat{f}(X) = \frac{1}{1 + e^{(-6 + 1x_1 + 1x_2)}}$$



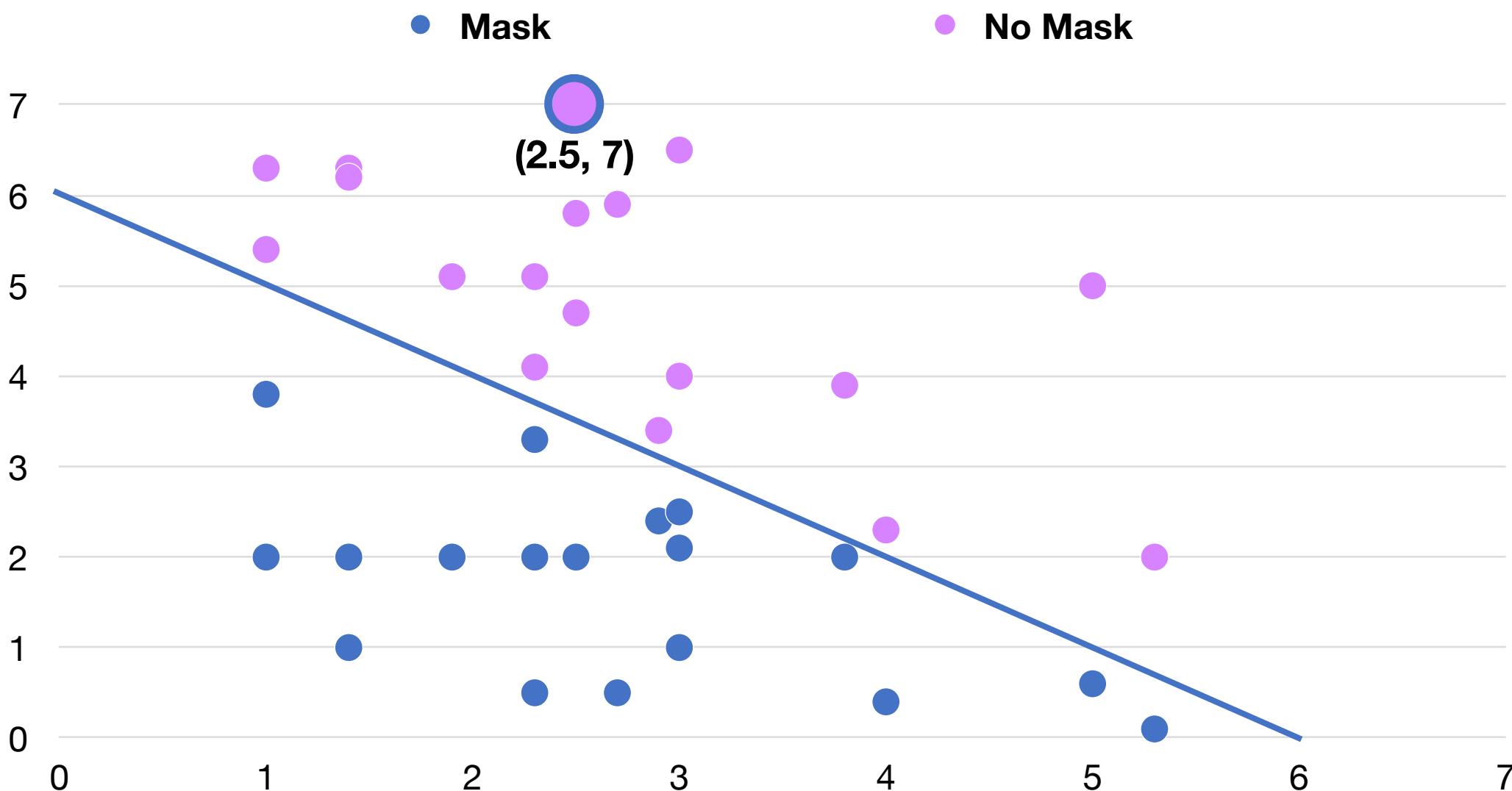
$$\hat{f}(X) = \frac{1}{1 + e^{(-6 + 1x_1 + 1x_2)}}$$



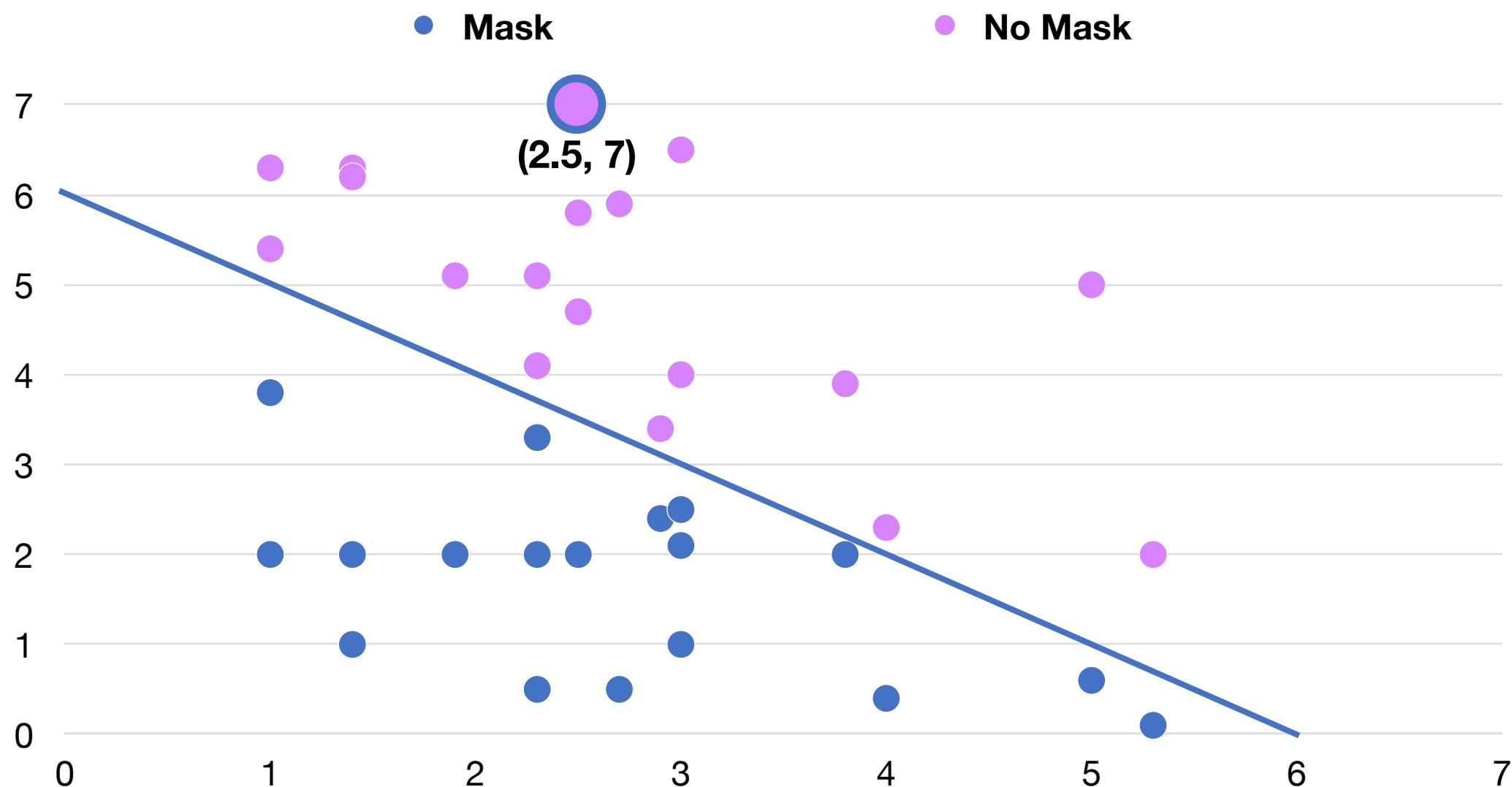
$$\hat{f}(X) = \frac{1}{1 + e^{(-6 + 1x_1 + 1x_2)}}$$



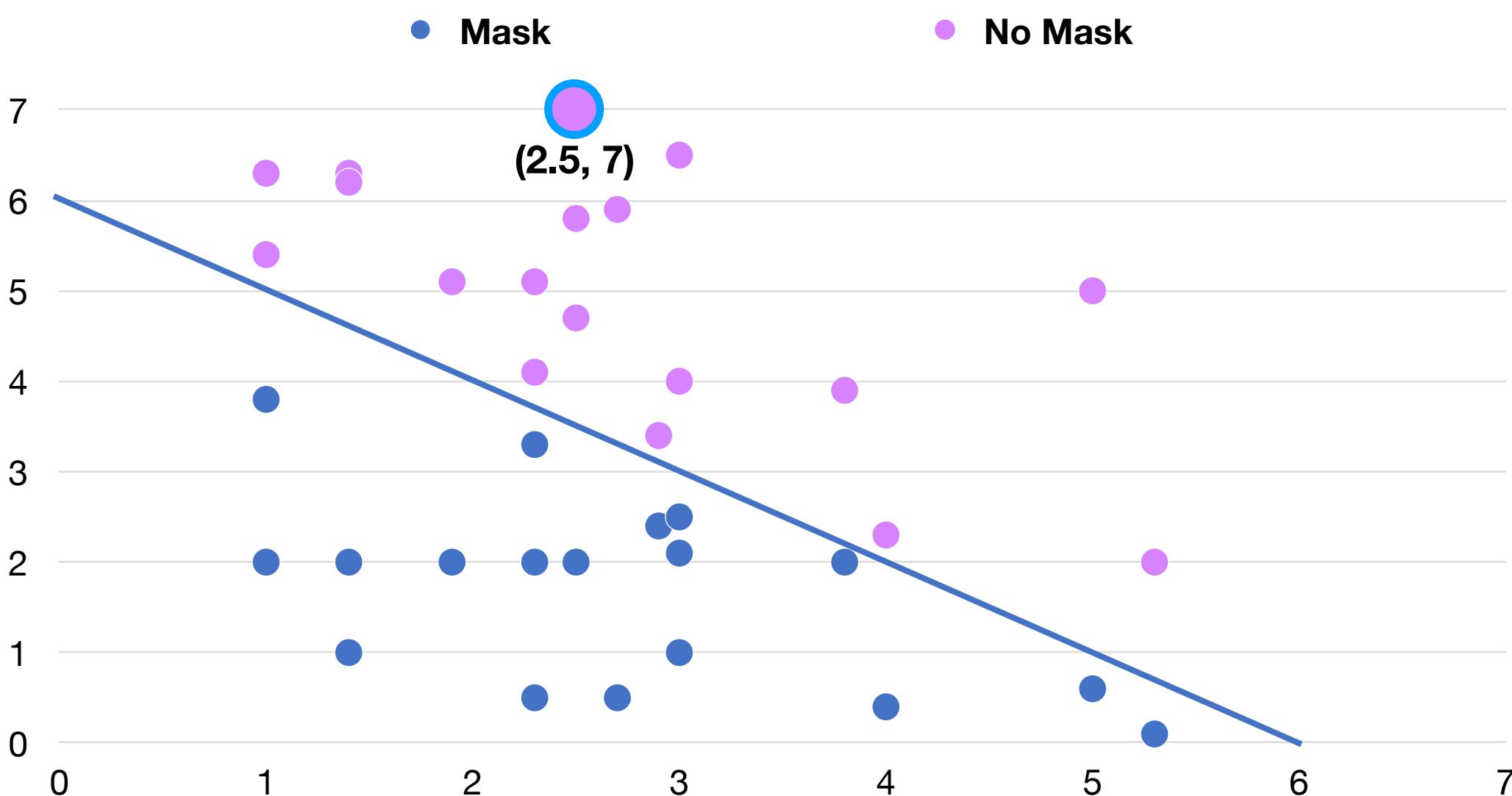
$$\hat{f}(X) = \frac{1}{1 + e^{(-6 + 2.5 + 7)}}$$



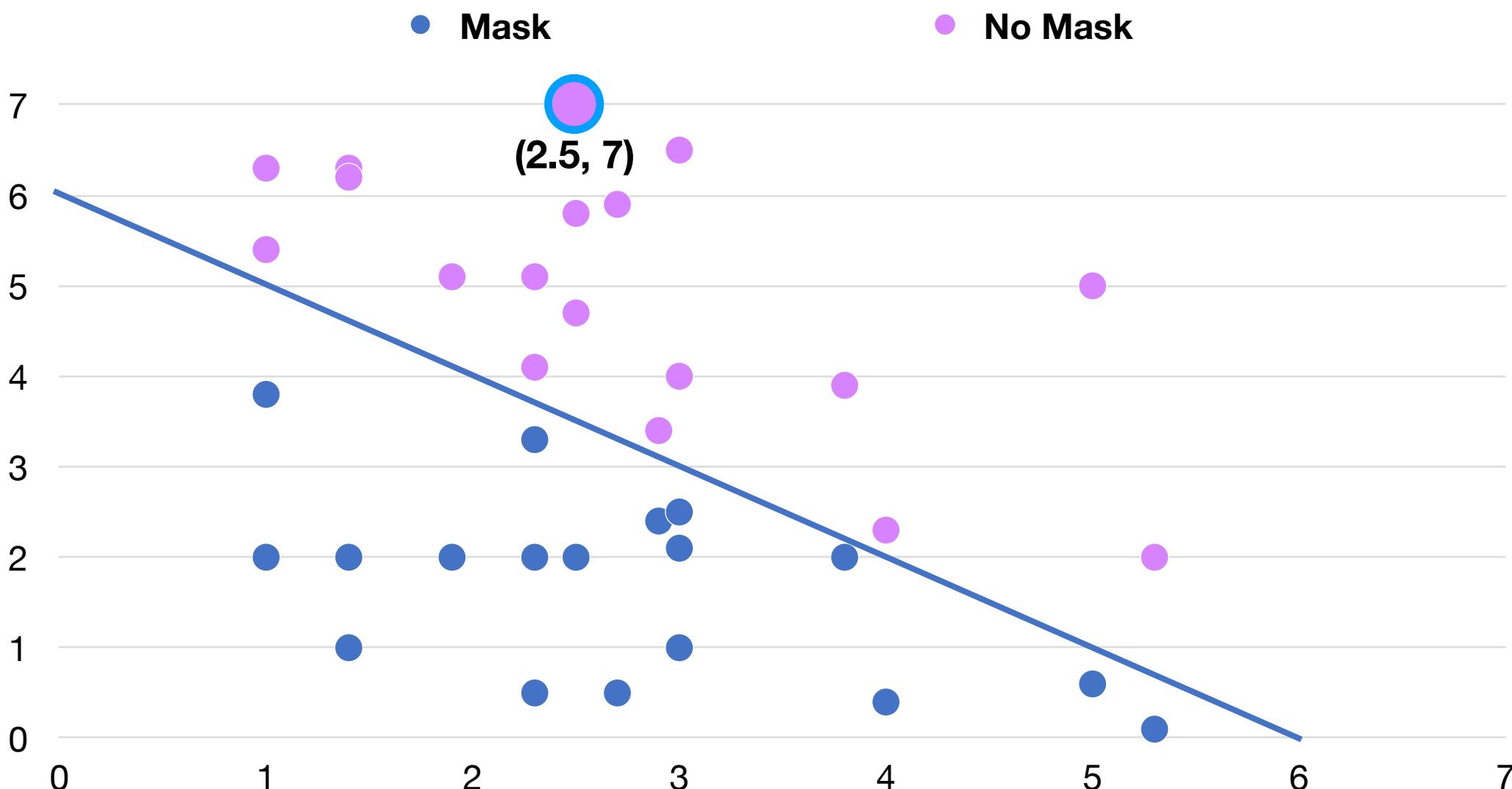
$$\hat{f}(X) = \frac{1}{1 + e^{(-6 + 9.5)}}$$



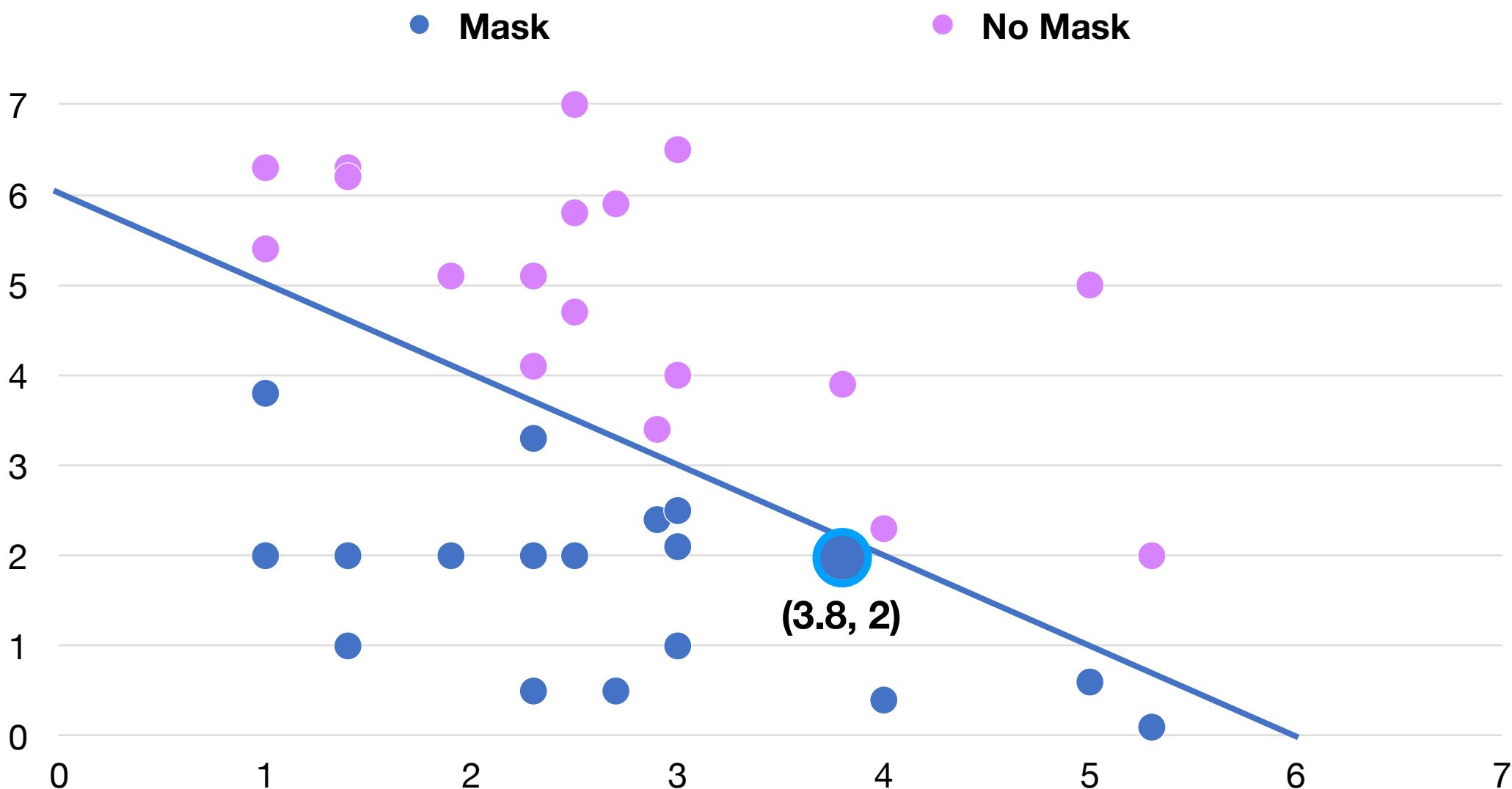
$$\hat{f}(X) = \frac{1}{1 + e^{-(3.5)}}$$



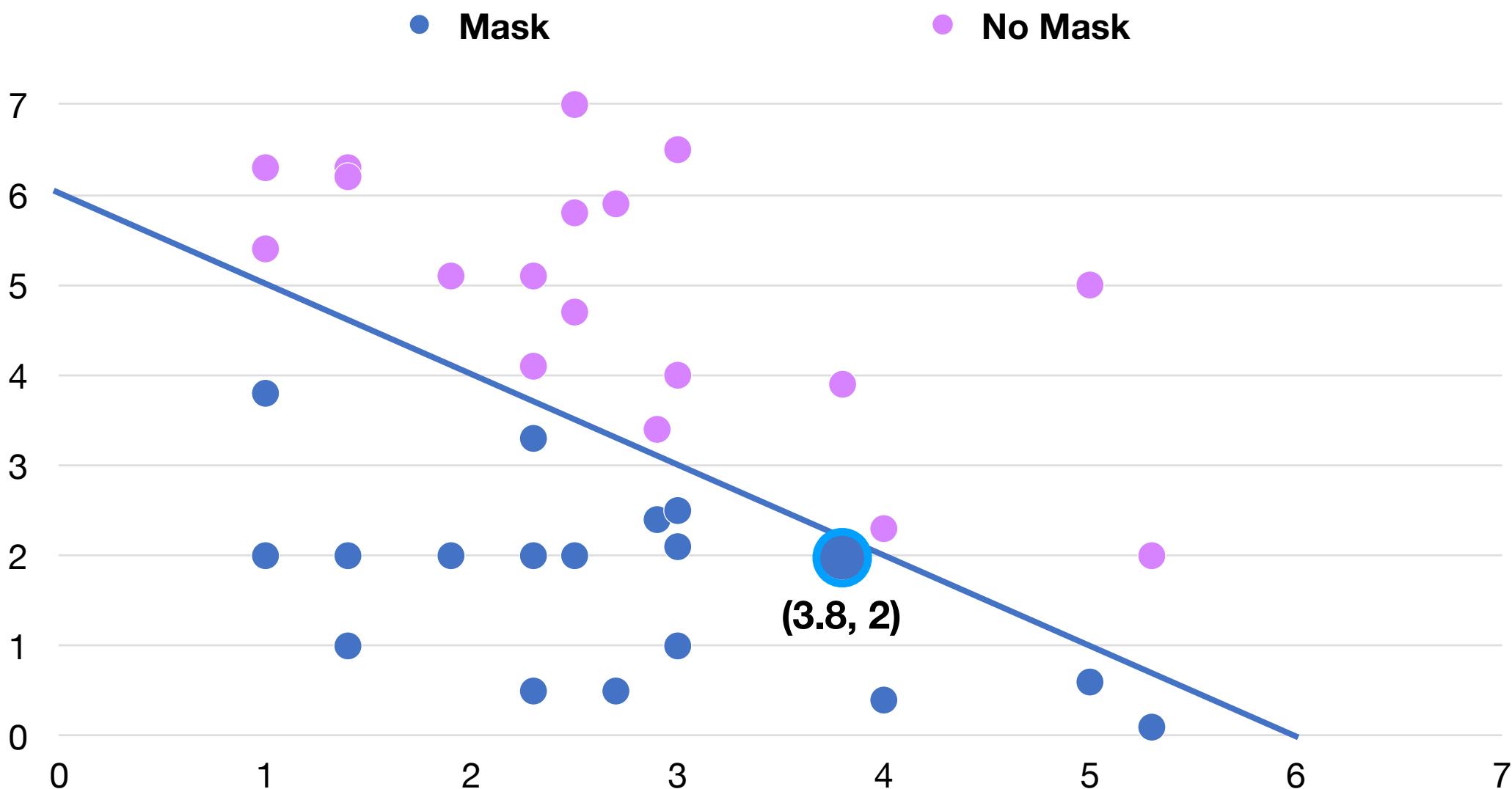
$$\hat{f}(X) = \frac{1}{1 + e^{-(3.5)}} = .97 \text{ (probability)}$$



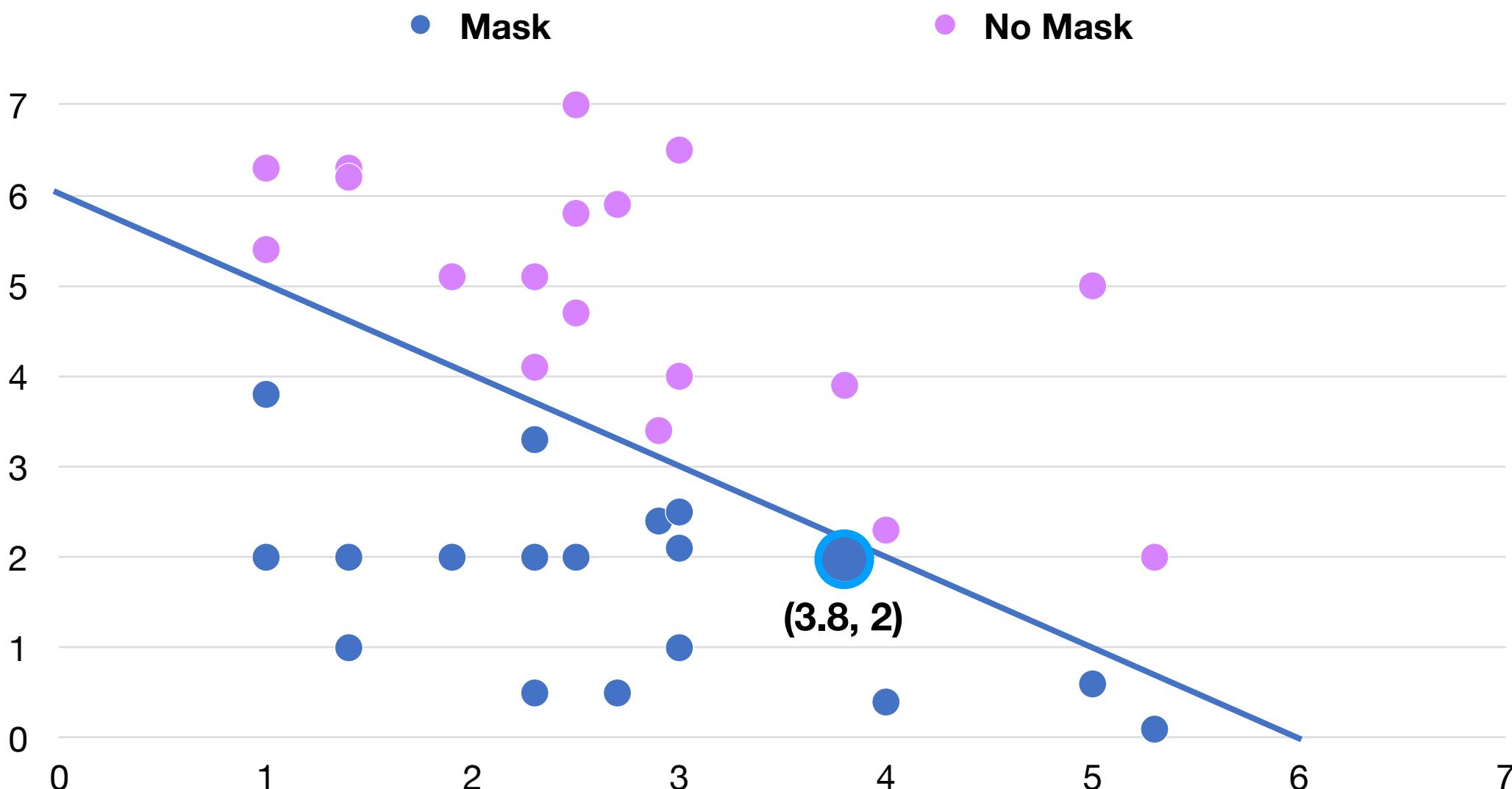
$$\hat{f}(X) = \frac{1}{1 + e^{(-6 + 3.8 + 2)}}$$



$$\hat{f}(X) = \frac{1}{1 + e^{(-6 + 5.8)}}$$



$$\hat{f}(X) = \frac{1}{1 + e^{-(0.2)}} = .45 \text{ (probability)}$$



Support Vector Machine

Support Vector Machine

$$\beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \beta_3 \mathbf{x}_3$$

Support Vector Machine

predict 1 when $\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 \geq 0$

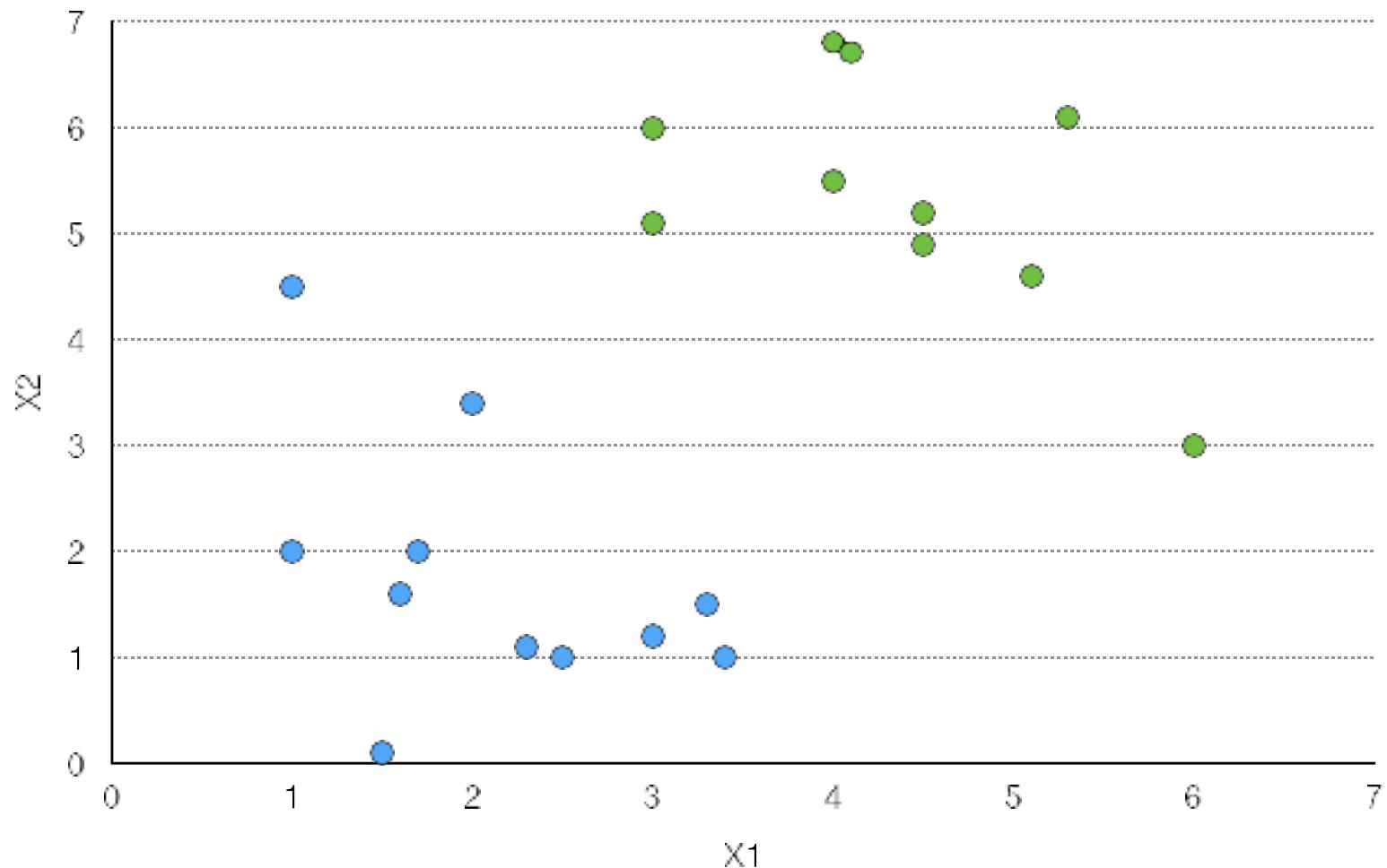
predict 0 when $\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 < 0$

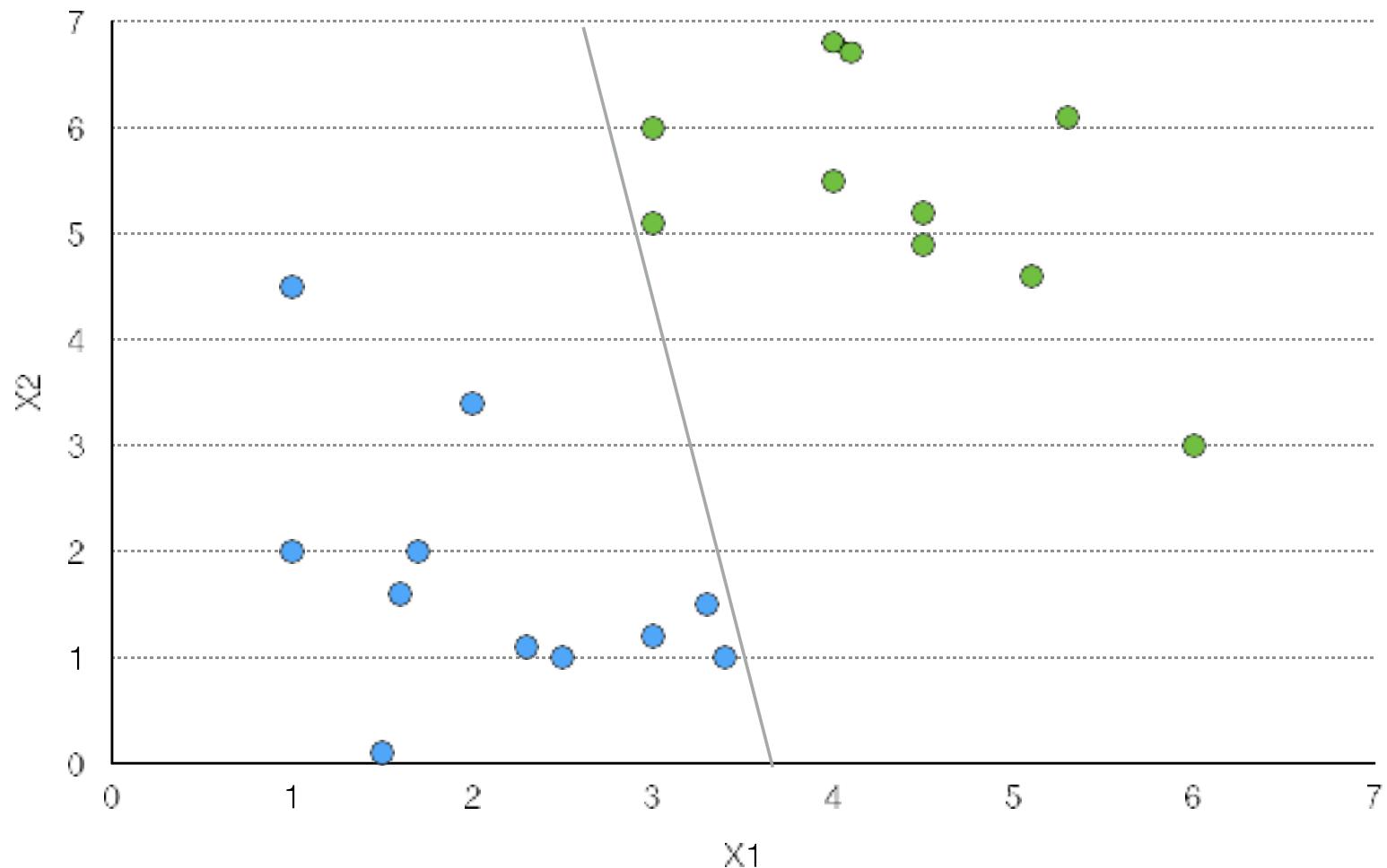
Support Vector Machine

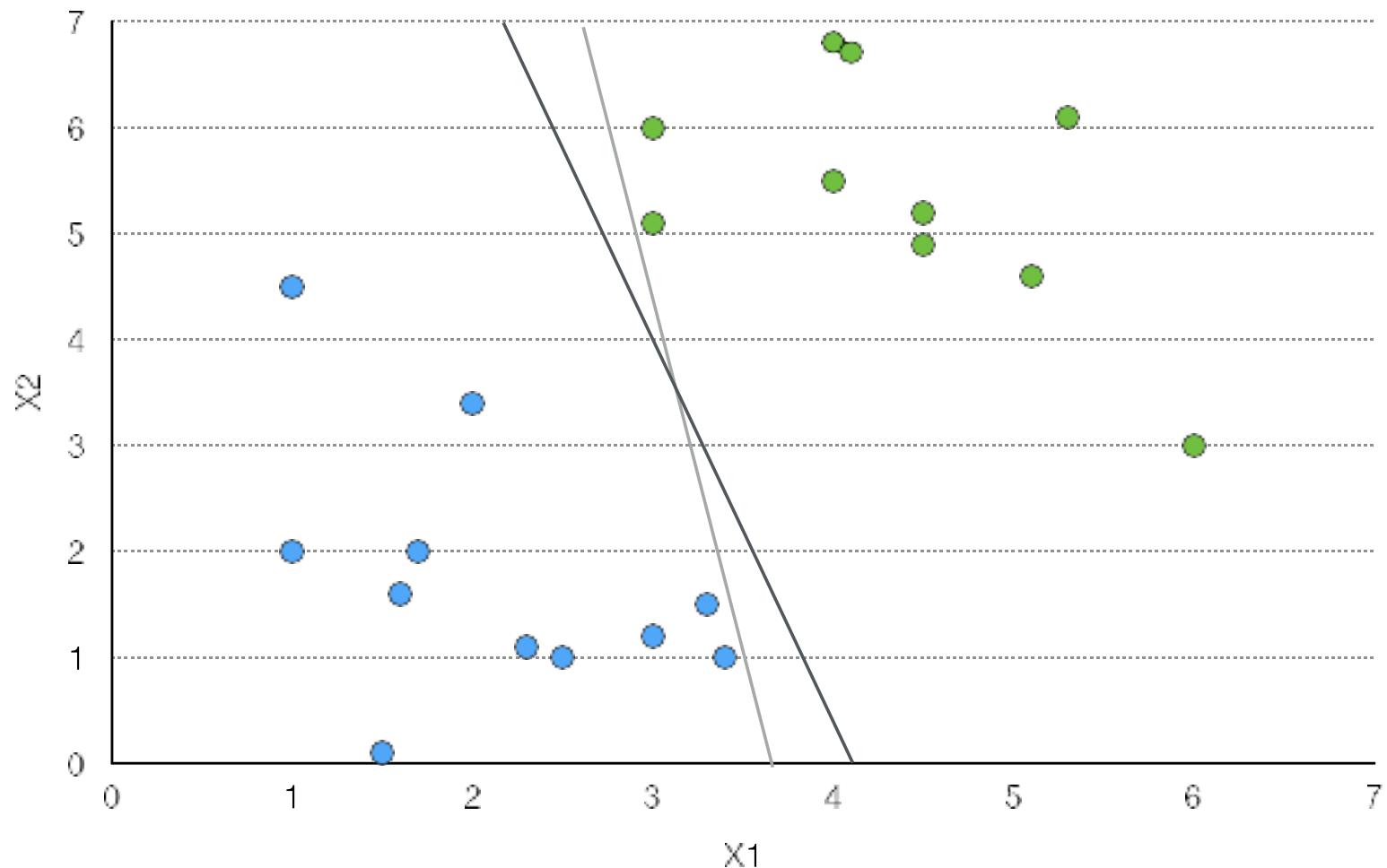
Large Margin Classifier

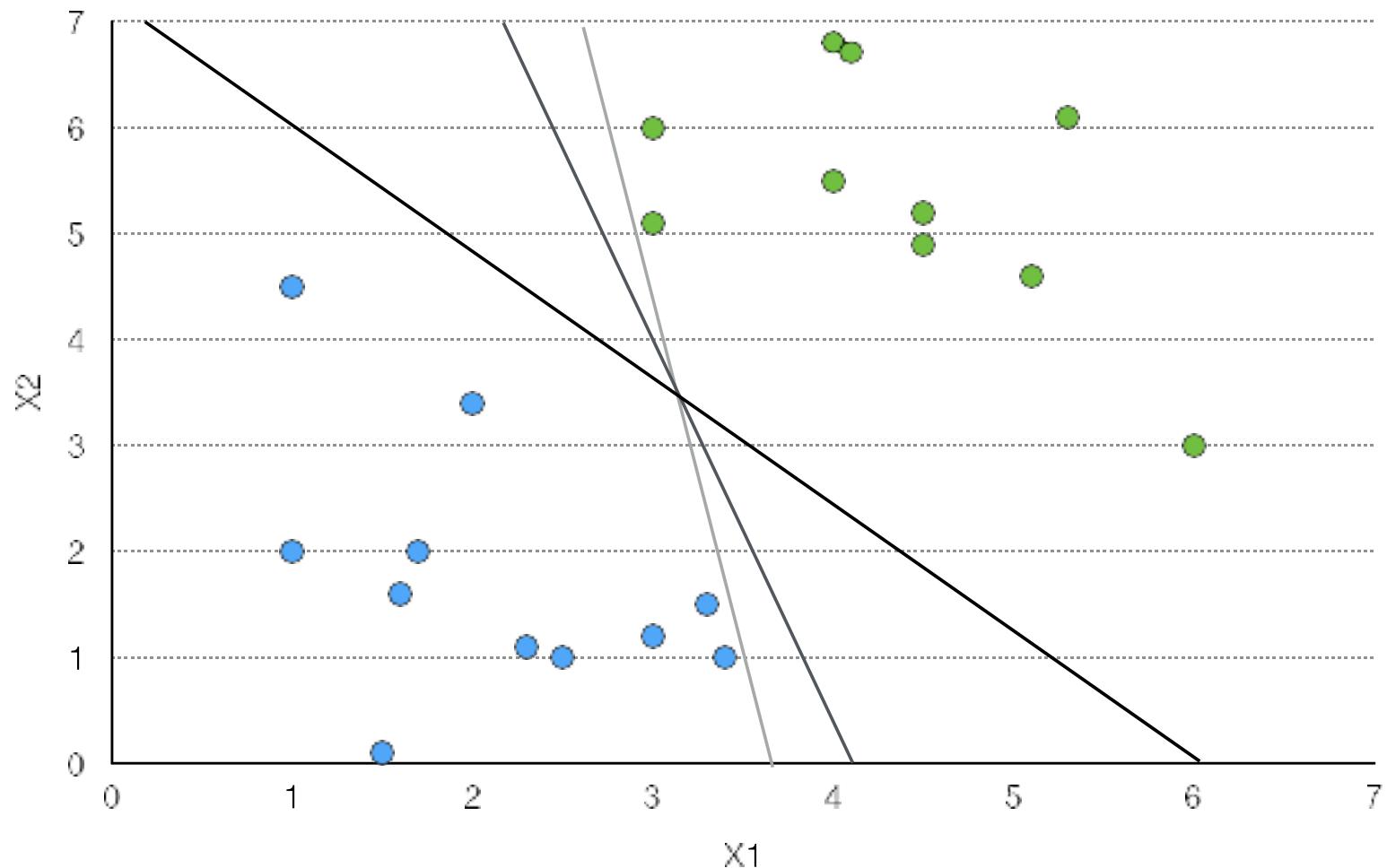
predict 1 when $\beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \beta_3 \mathbf{x}_3 \geq 1$

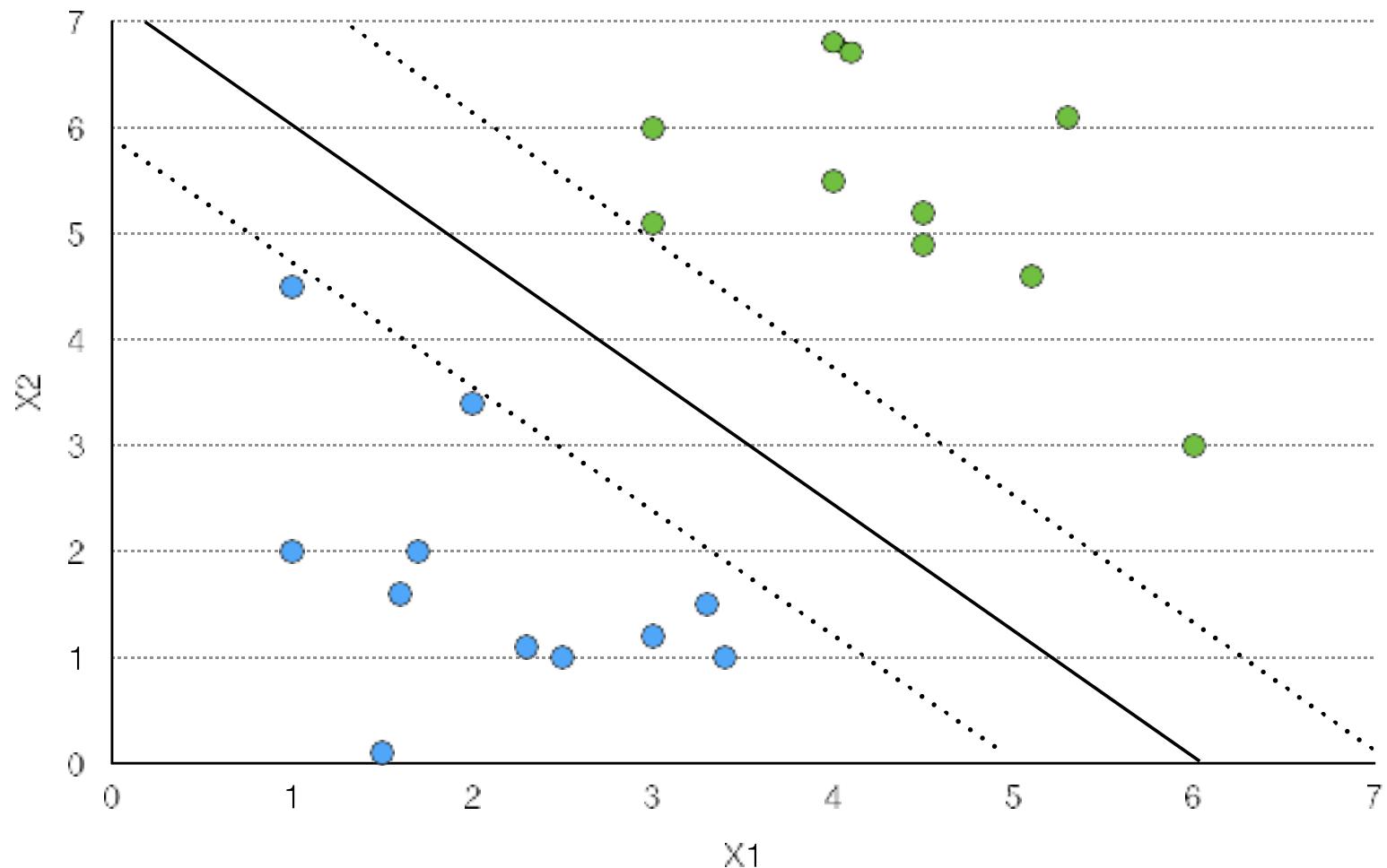
predict 0 when $\beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \beta_3 \mathbf{x}_3 < -1$













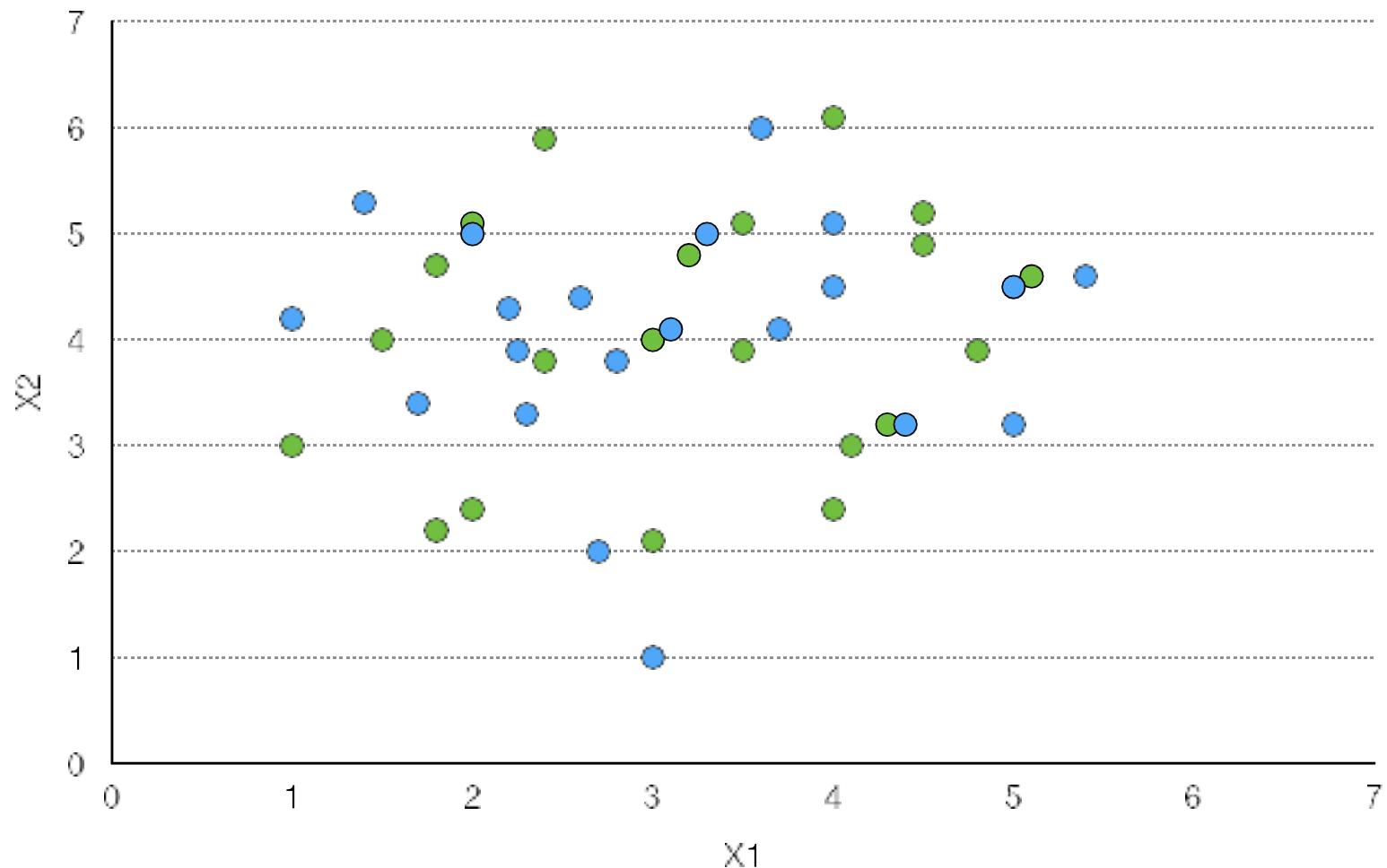




40 YARD DASH



40-yard dash	Weight	Height	Drafted
5.10	290	74	1
4.92	275	75.5	1
4.43	178	69	0
4.62	221	74.5	1
4.91	248	75	0
5.53	303	77	0
4.47	189	71	1
4.56	205	71	1
4.75	267	73	0
4.84	261	74	1



40-yard dash	Weight	Height	Drafted
5.10	290	74	1
4.92	275	75.5	1
4.43	178	69	0
4.62	221	74.5	1
4.91	248	75	0
5.53	303	77	0
4.47	189	71	1
4.56	205	71	1
4.75	267	73	0
4.84	261	74	1

Feature Engineering

40-yard dash	BMI (wt/ht ²)	Drafted
5.10	37.2	1
4.92	33.9	1
4.43	26.3	0
4.62	28	1
4.91	31	0
5.53	35.9	0
4.47	26.4	1
4.56	28.6	1
4.75	35.2	0
4.84	33.5	1

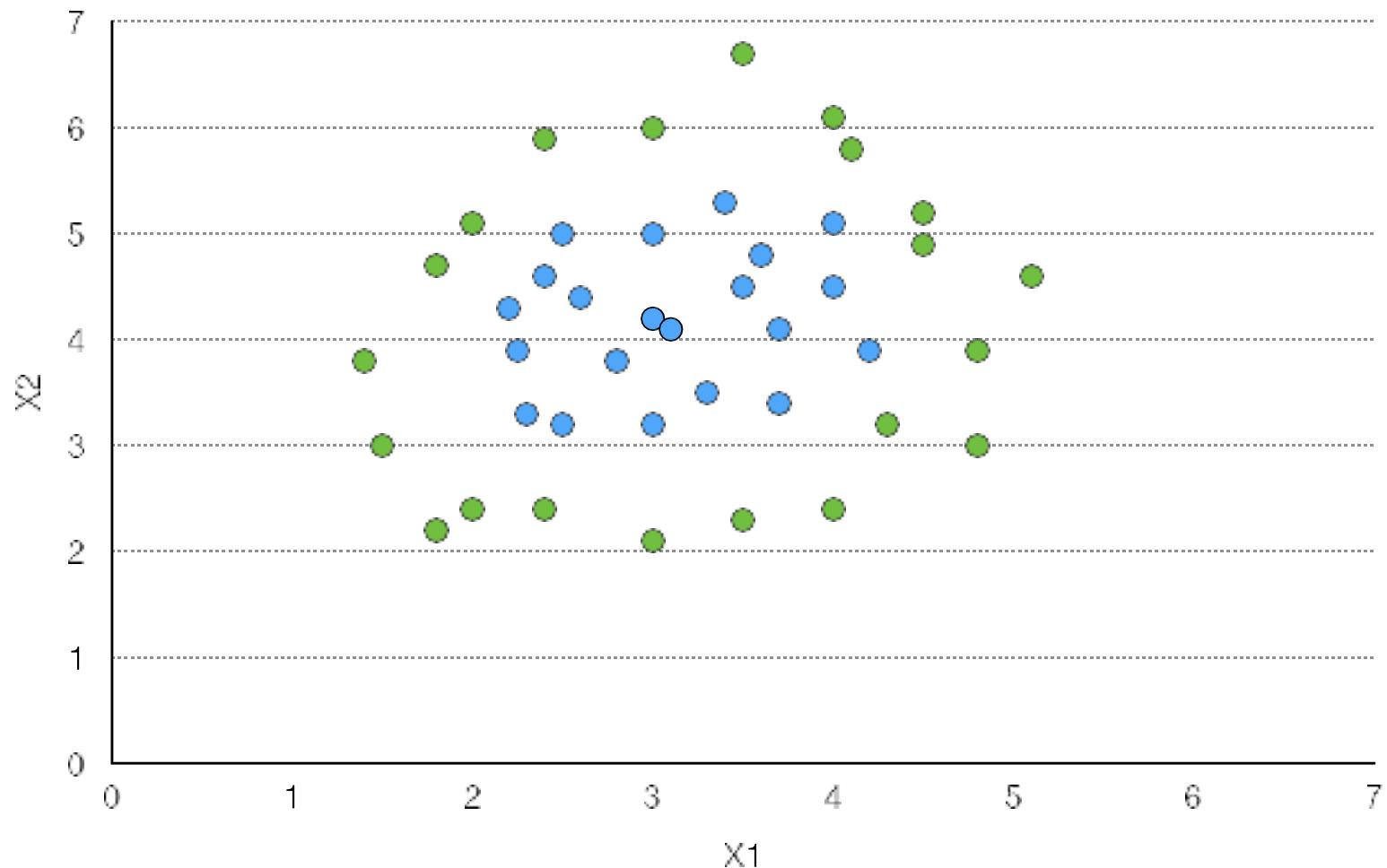


Feature Engineering

40-yard dash	BMI (wt/ht ²)	Drafted
5.10	37.2	1
4.92	33.9	1
4.43	26.3	0
4.62	28	1
4.91	31	0
5.53	35.9	0
4.47	26.4	1
4.56	28.6	1
4.75	35.2	0
4.84	33.5	1

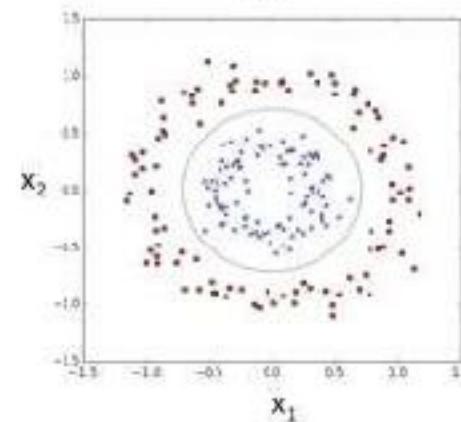
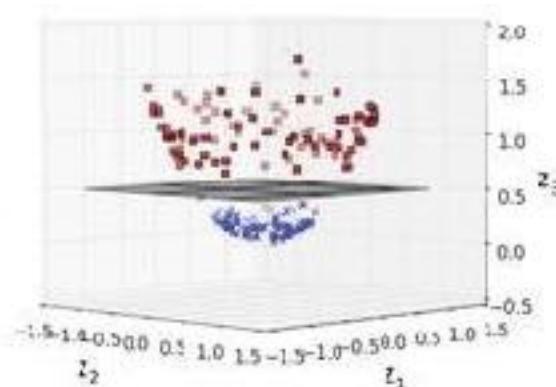
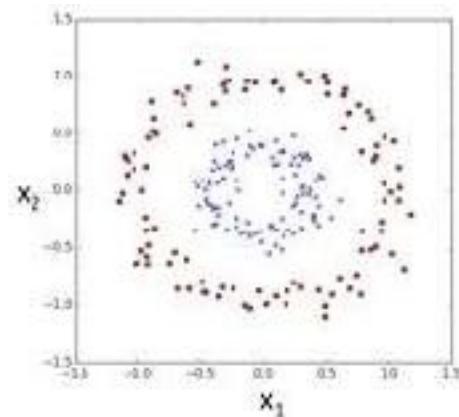
Feature Engineering

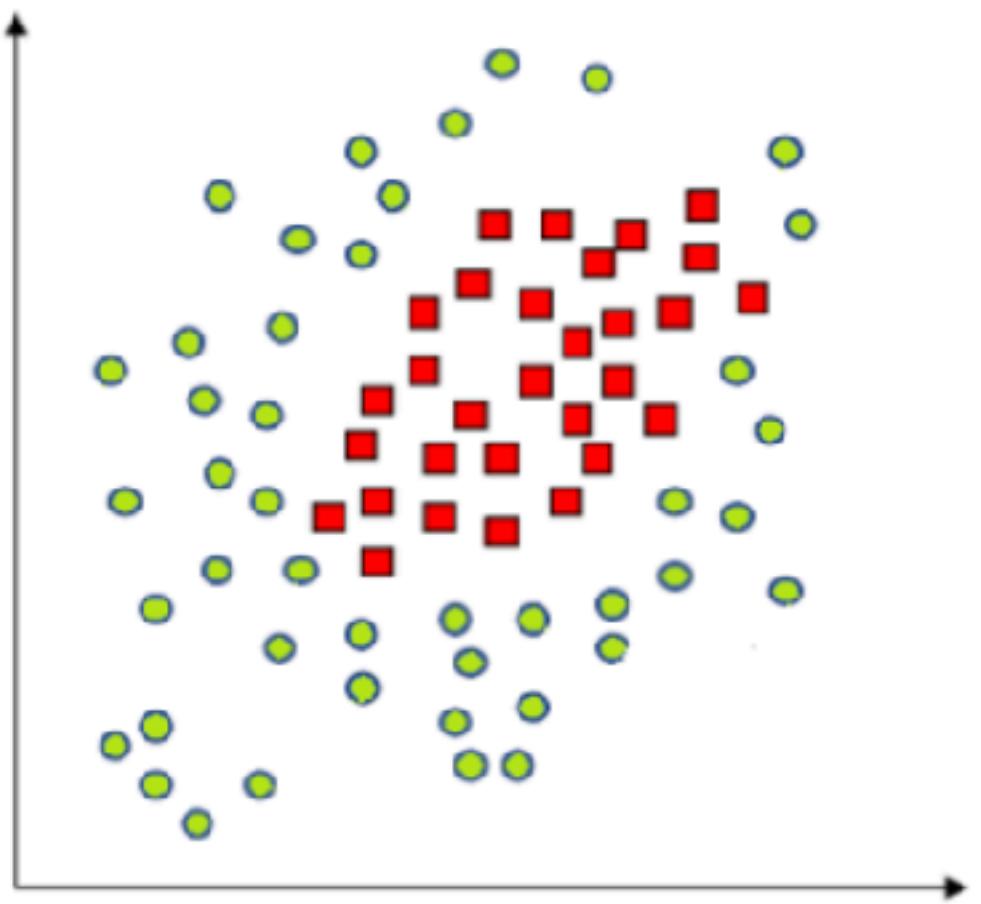
Speed-to-Size (40yd/bsa)	BMI (wt/ht ²)	Drafted
2.16	37.2	1
2.06	33.9	1
2.02	26.3	0
1.97	28	1
2.23	31	0
2.00	35.9	0
2.03	26.4	1
1.99	28.6	1
1.85	35.2	0
2.03	33.5	1



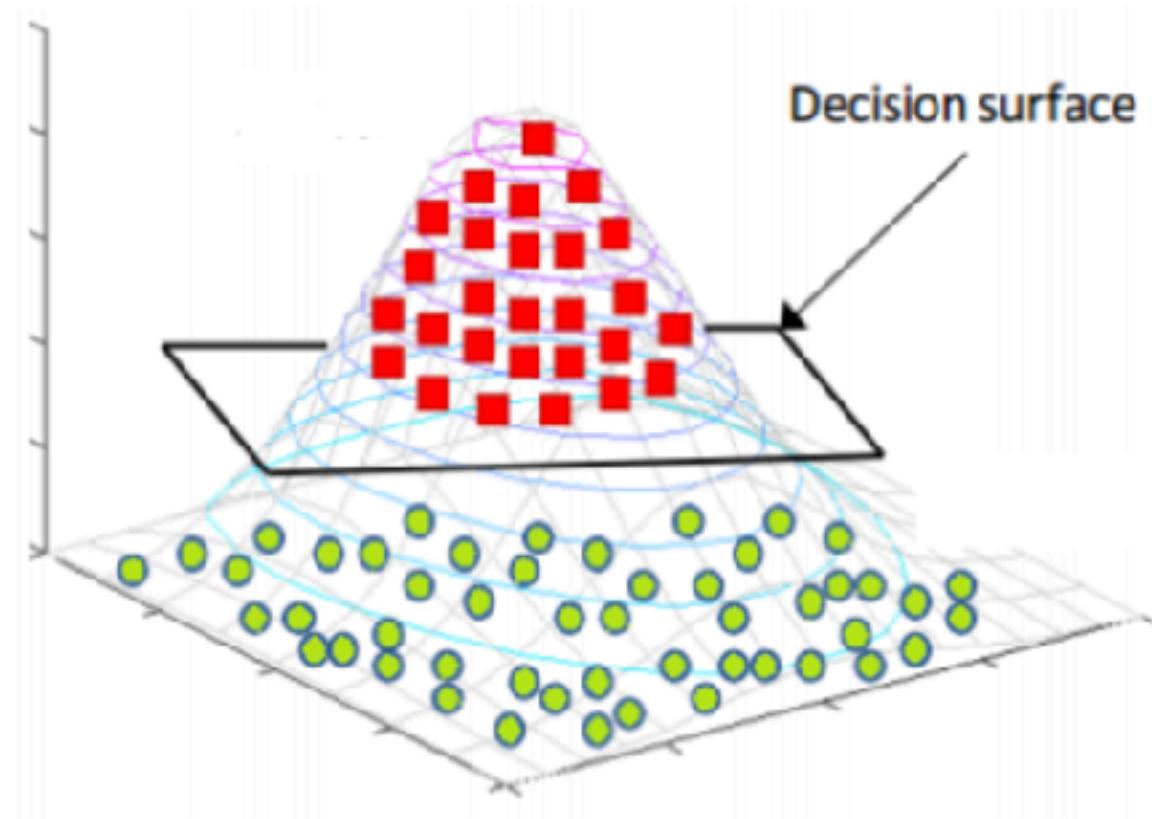
Kernel

Non-linear Classification

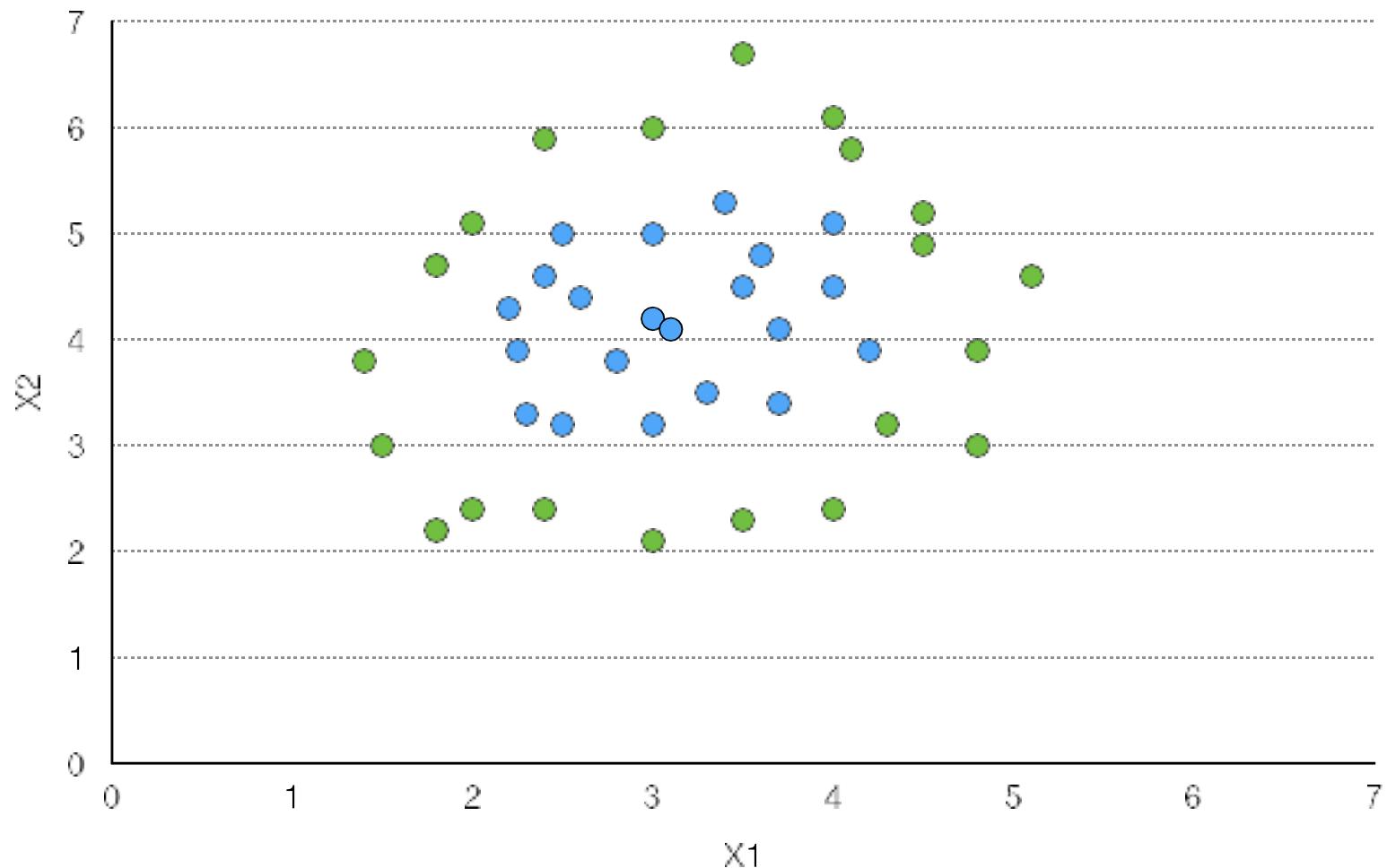


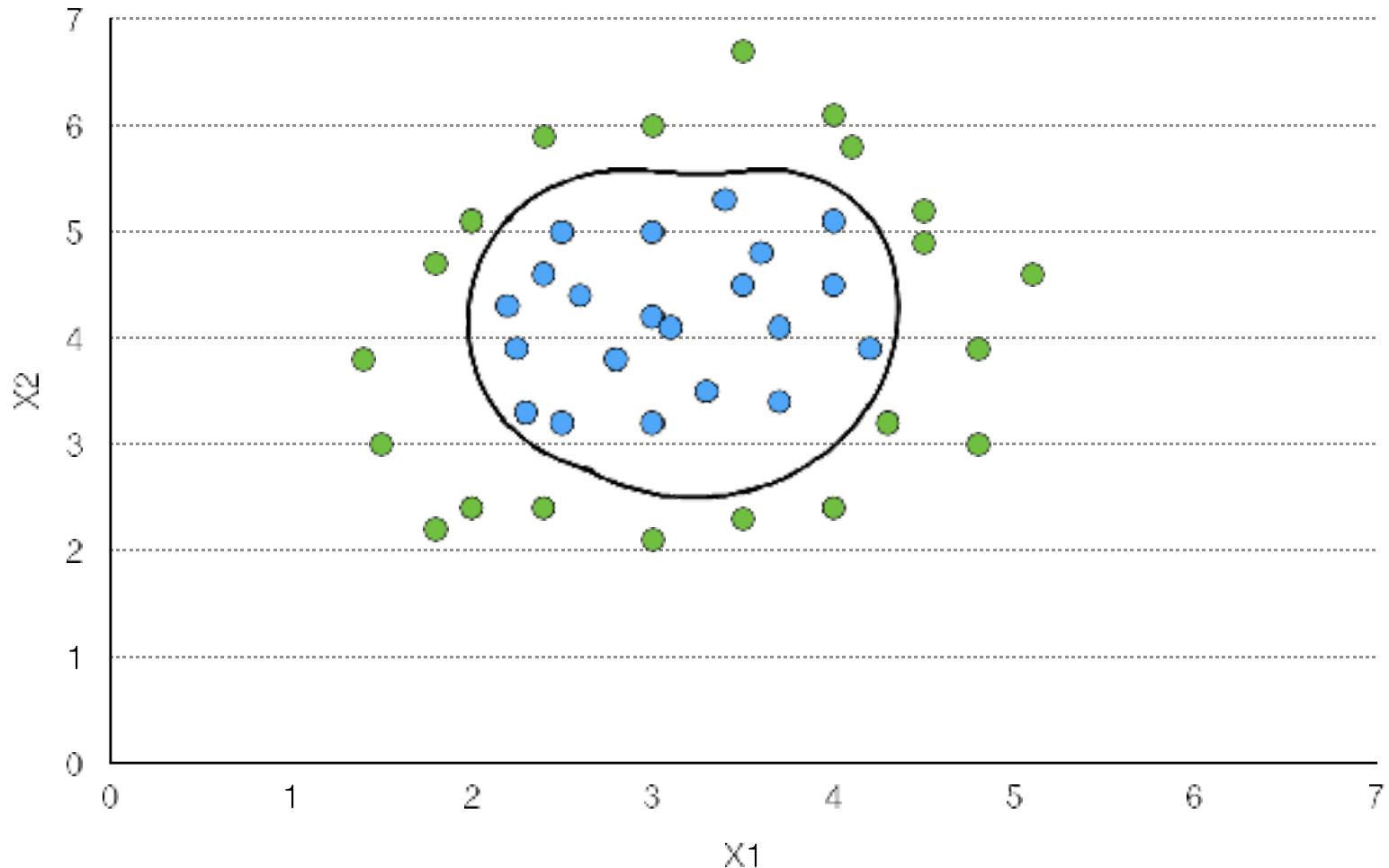


kernel



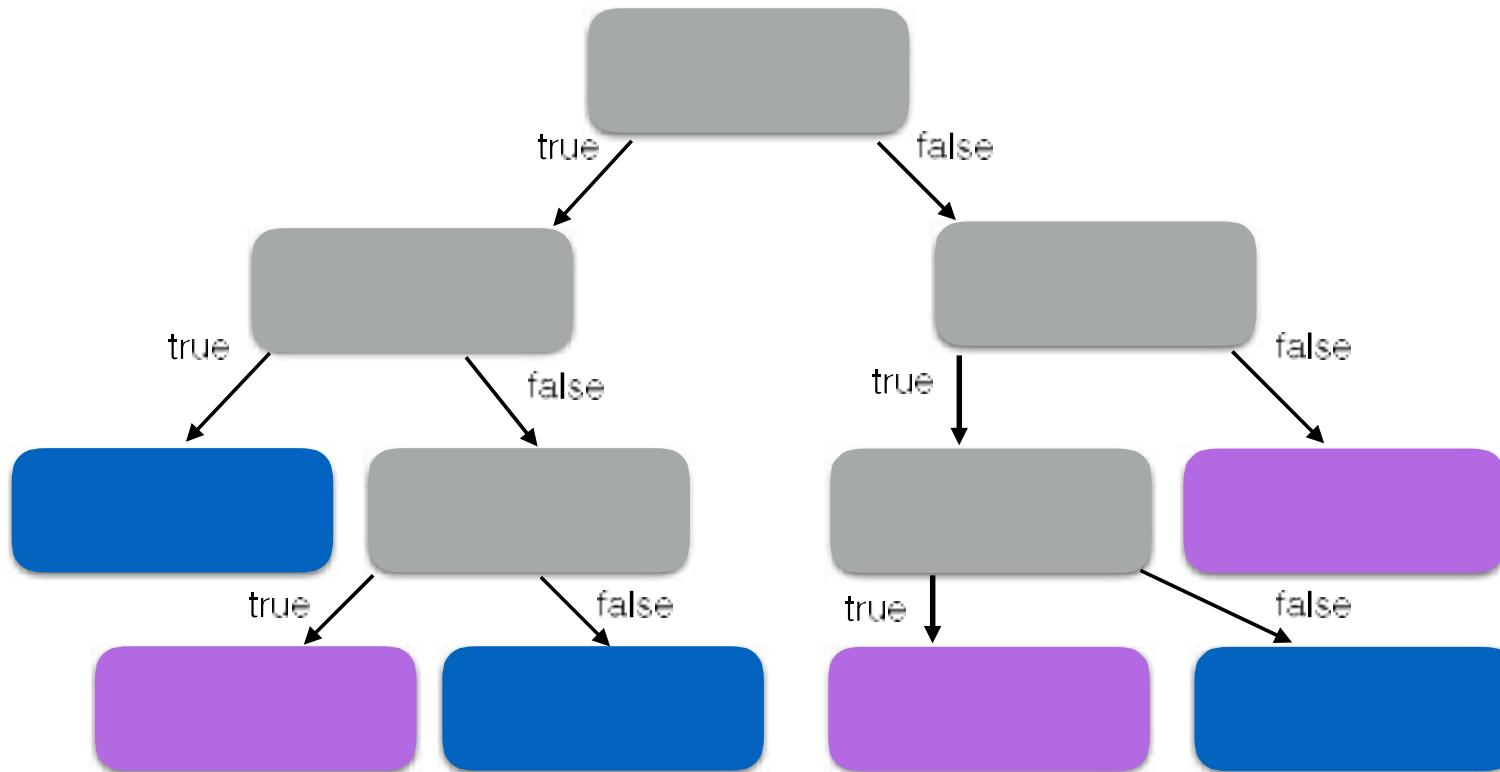
Kernel Trick





Decision Tree

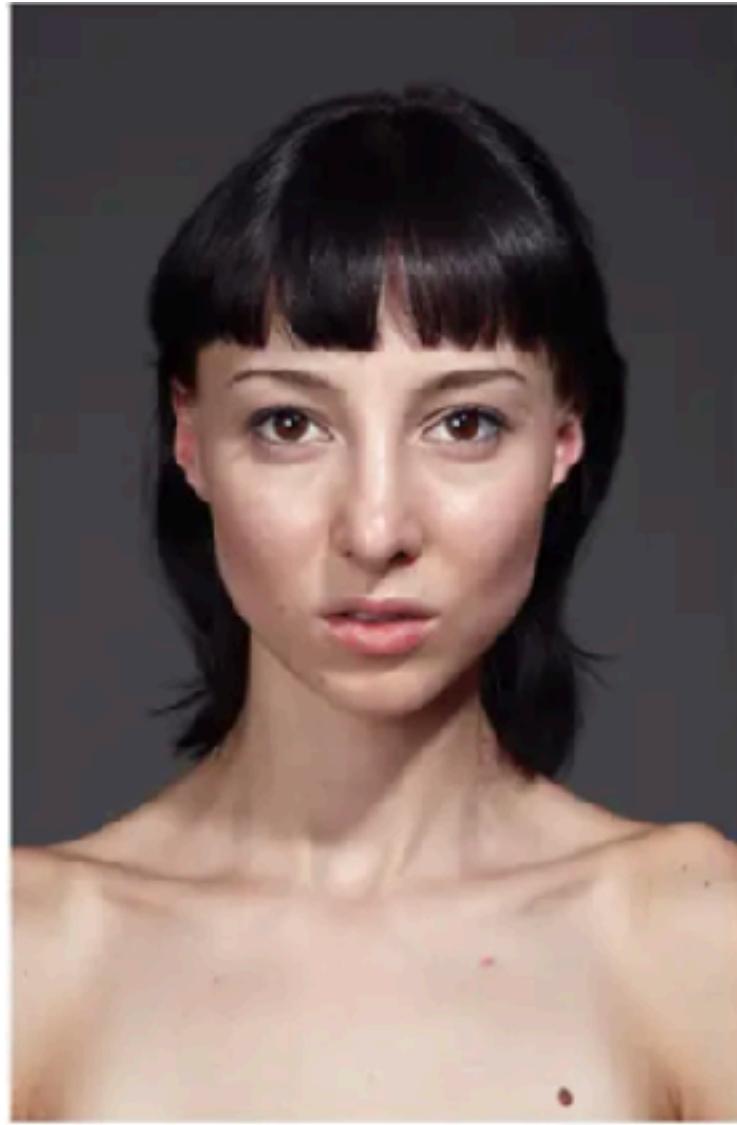
Decision Tree

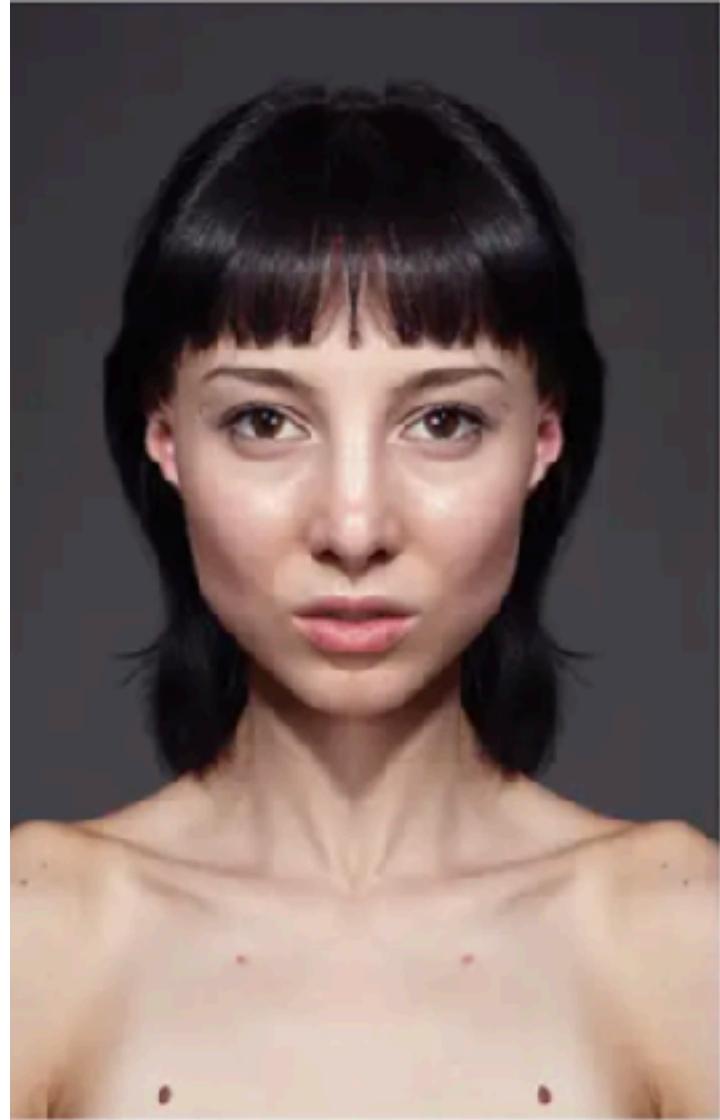
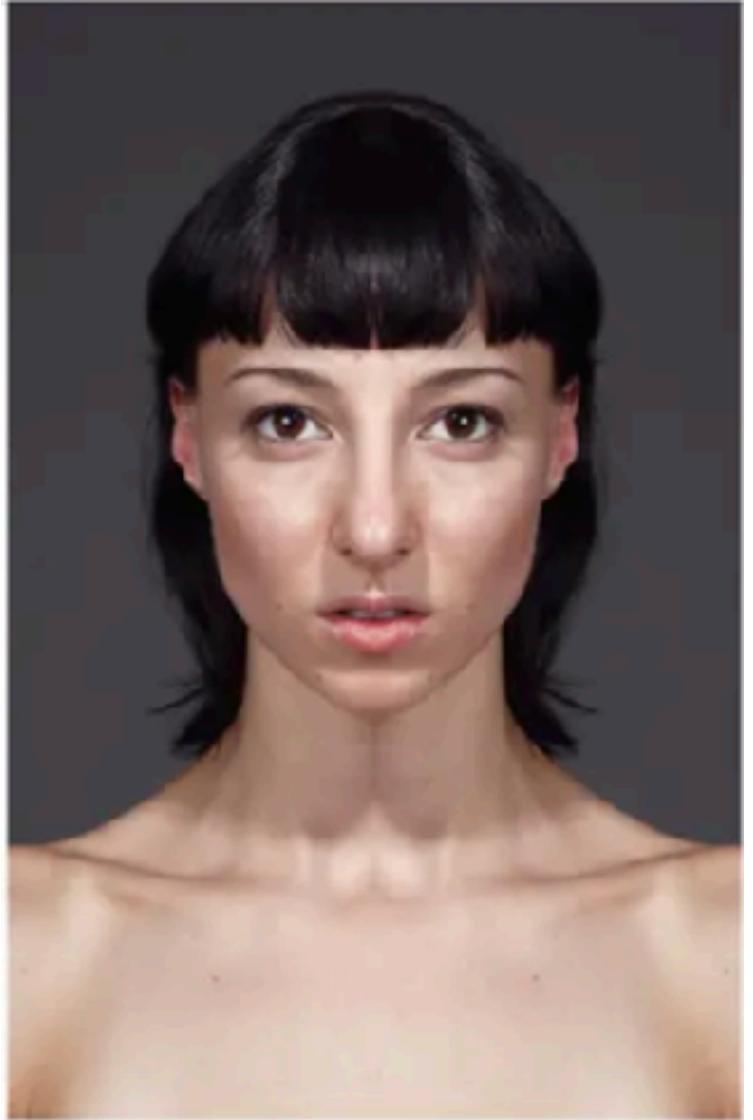














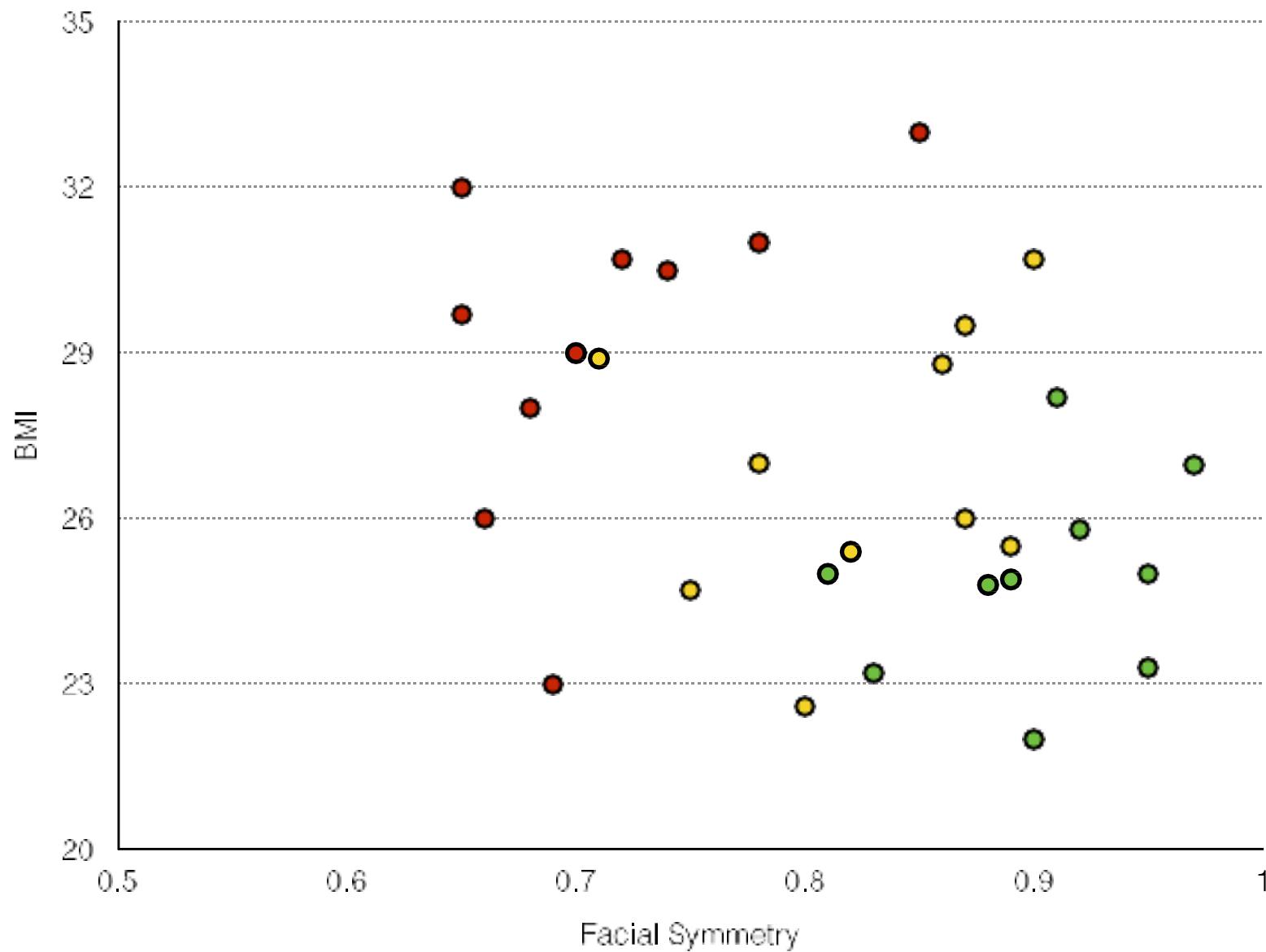


Short-term Attractiveness



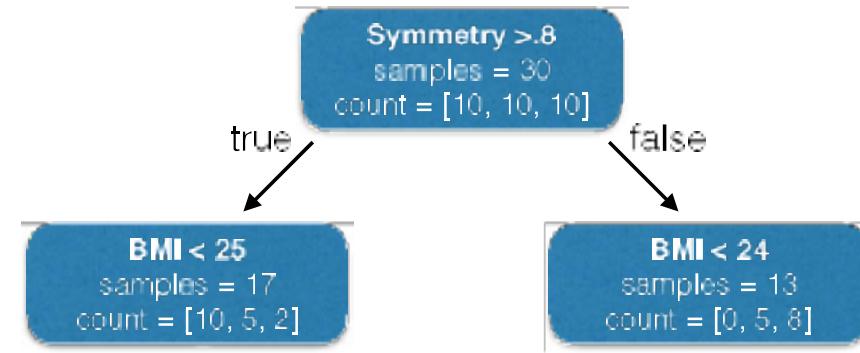
Short-term Attractiveness

Facial Symmetry	BMI	Waist-to-Hip	Well-Groomed
0.9	23.4	0.93	1
0.85	27.9	0.87	0
0.65	27.1	0.79	1
0.85	22.6	0.91	1
0.9	30.3	0.82	0
0.75	29.0	0.82	0
0.85	22.3	0.89	1
0.7	37.6	0.73	0
0.85	24.2	0.85	0

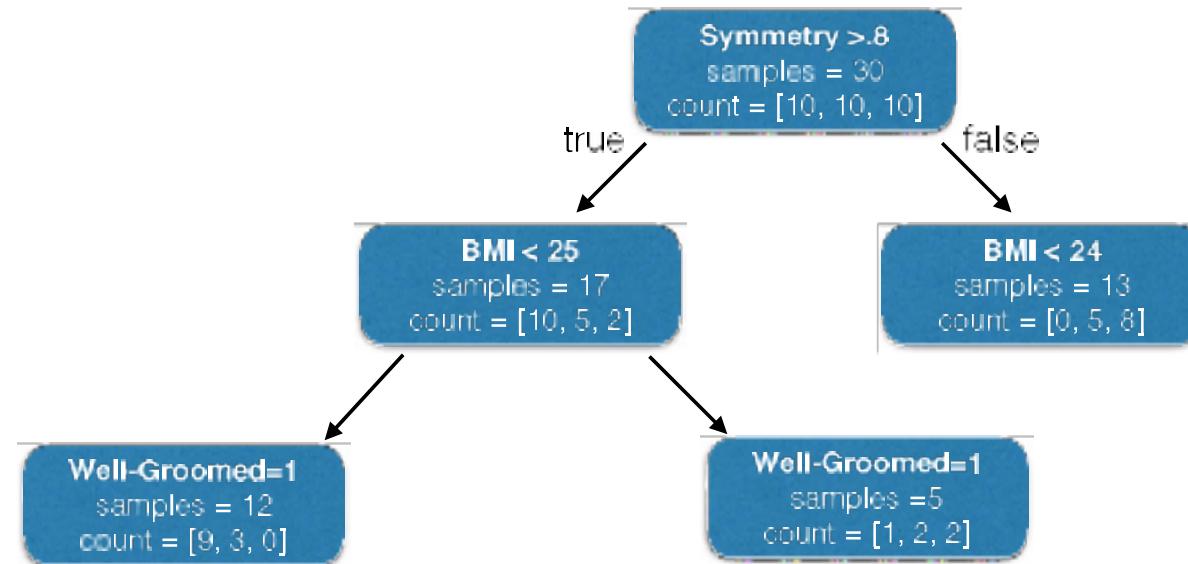


Symmetry >.8
samples = 30
count = [10, 10, 10]

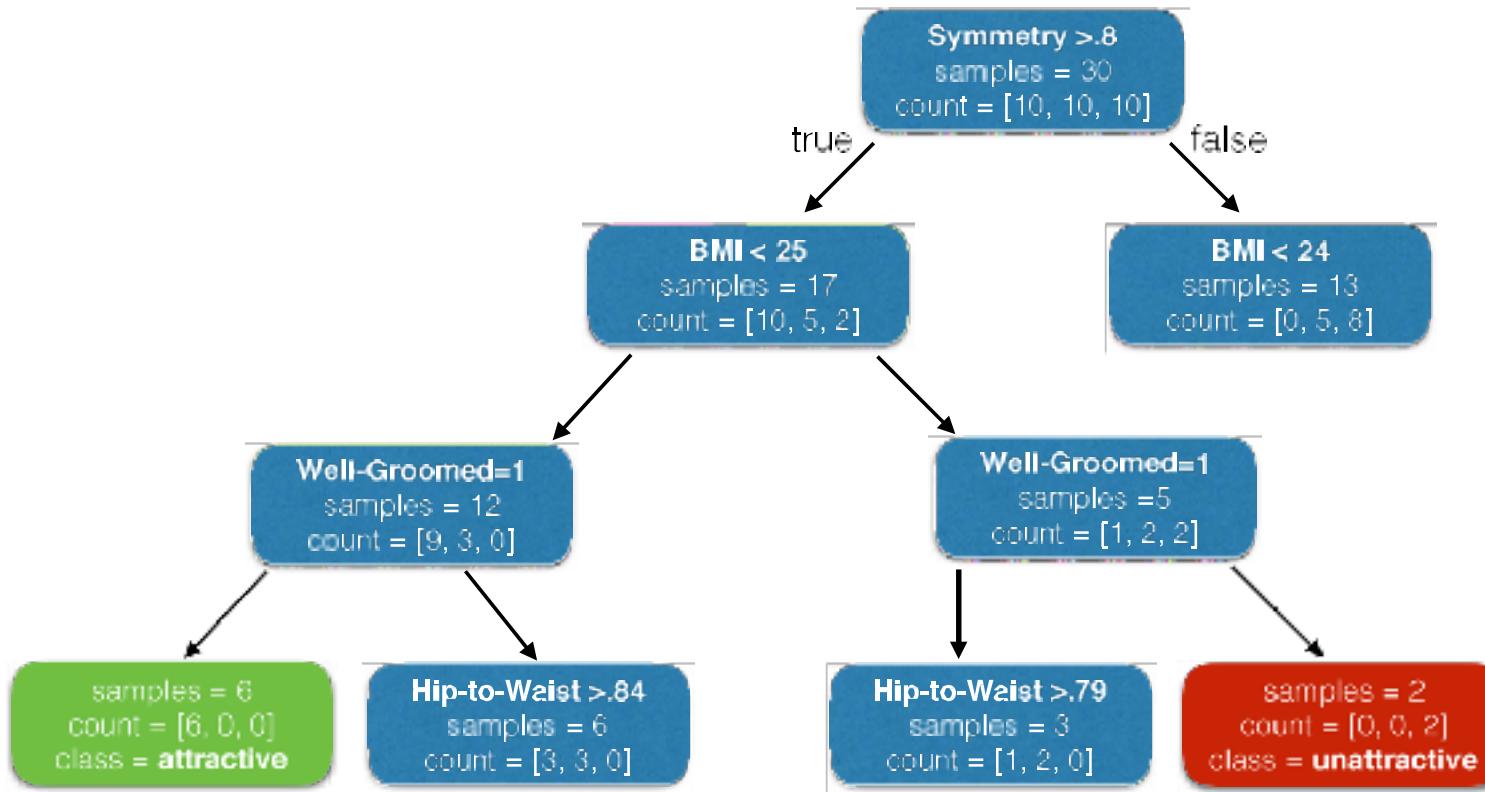
[att, ave, un]



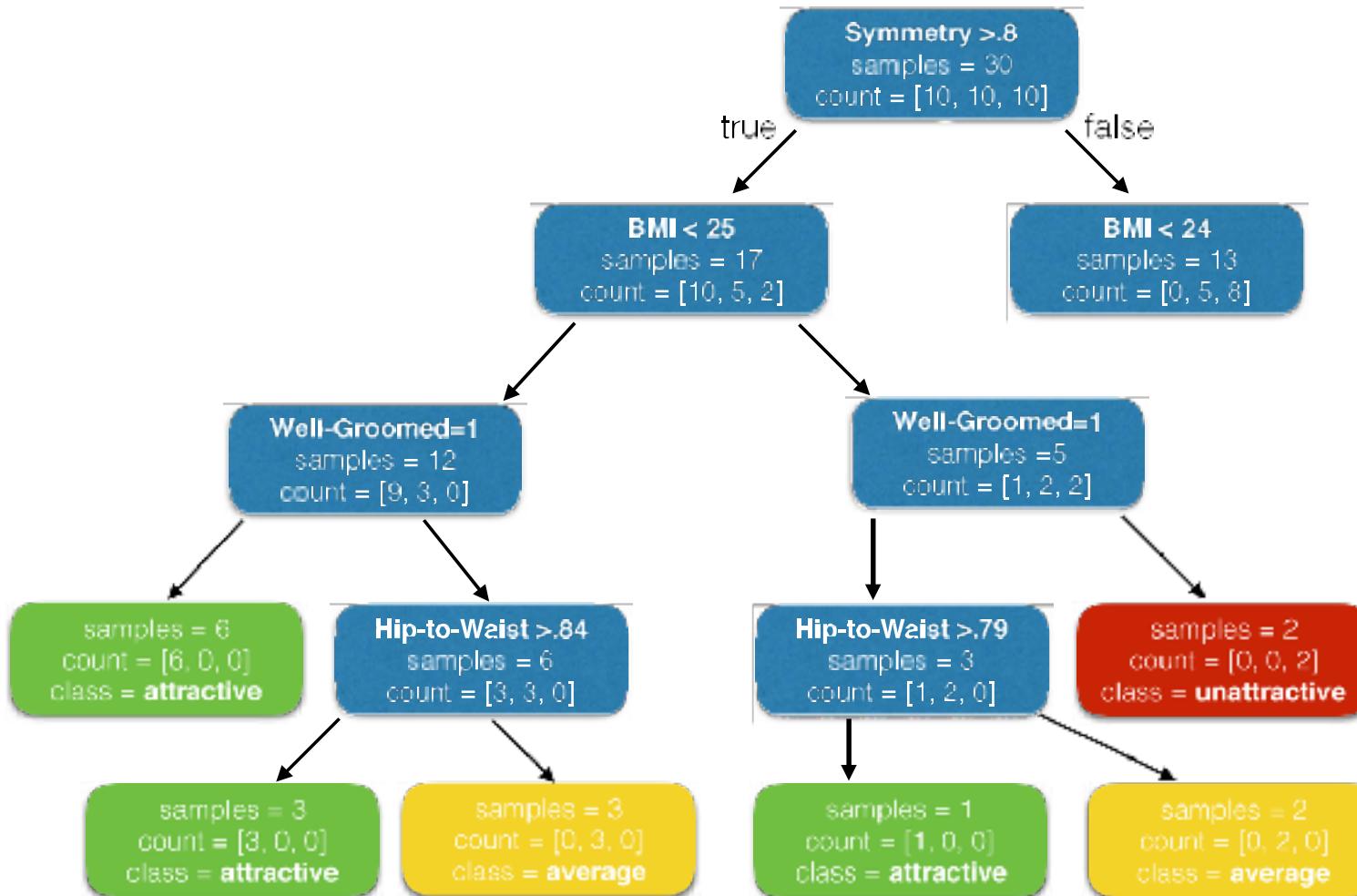
[att, ave, un]



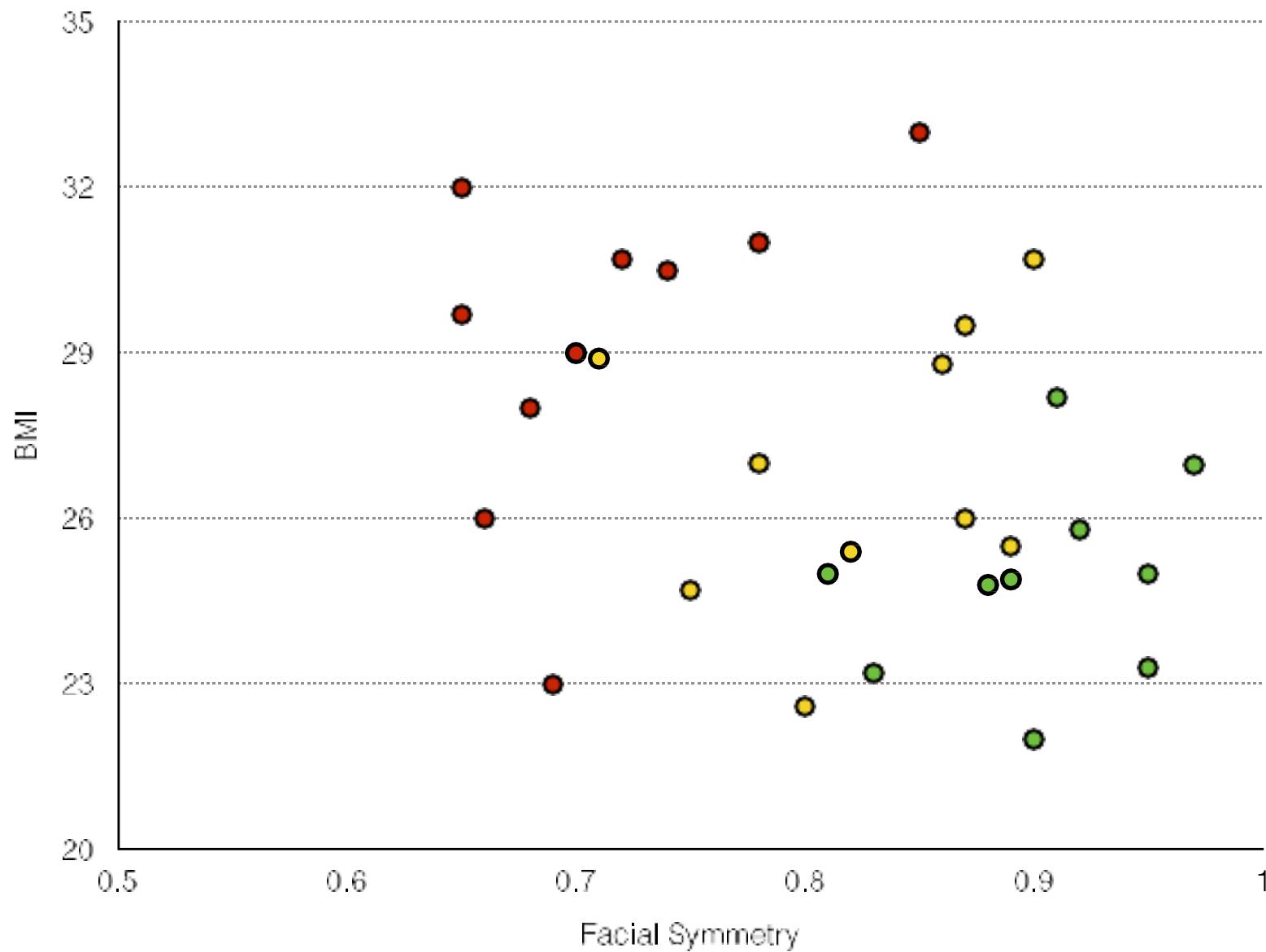
[att, ave, un]

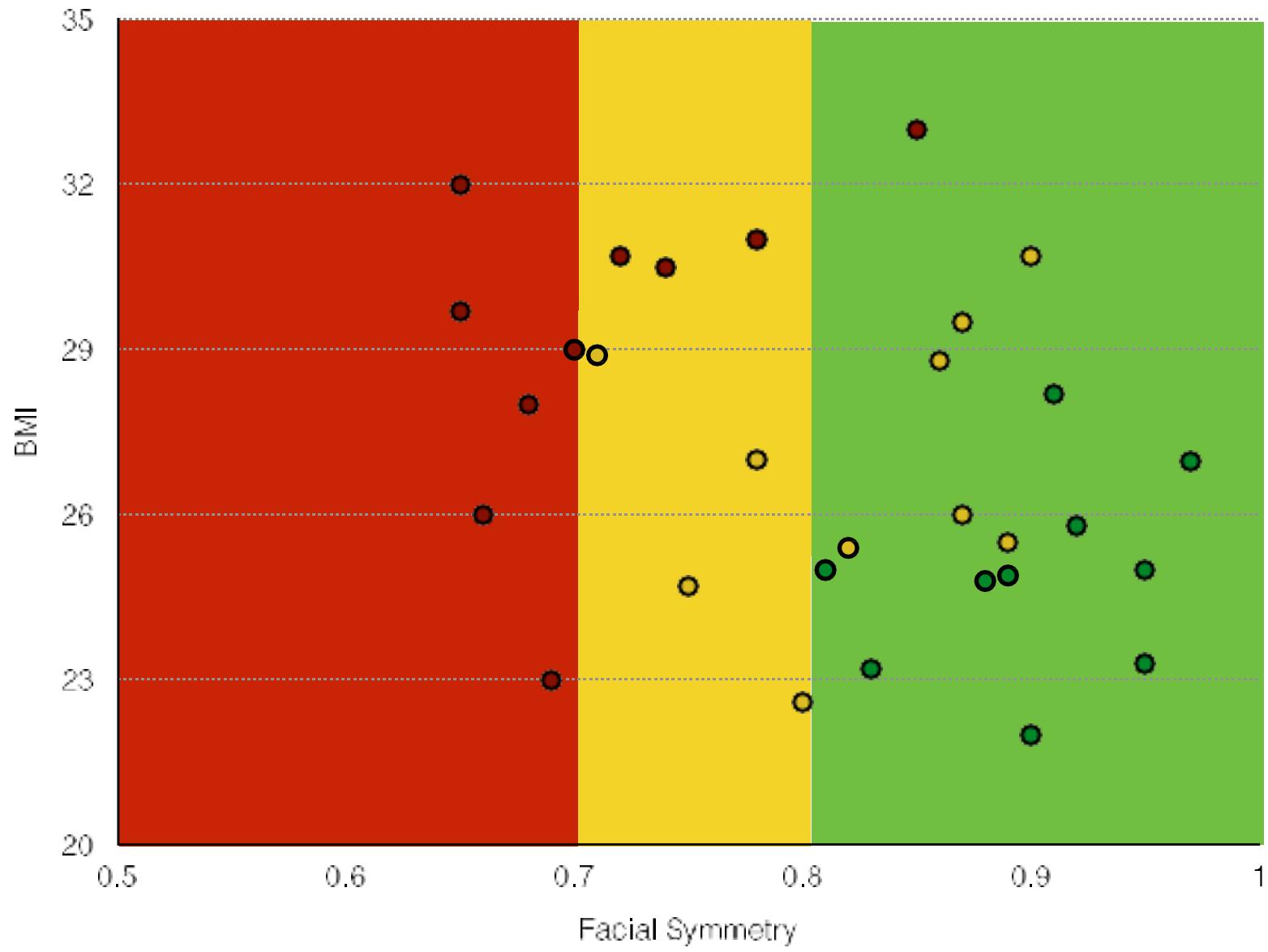


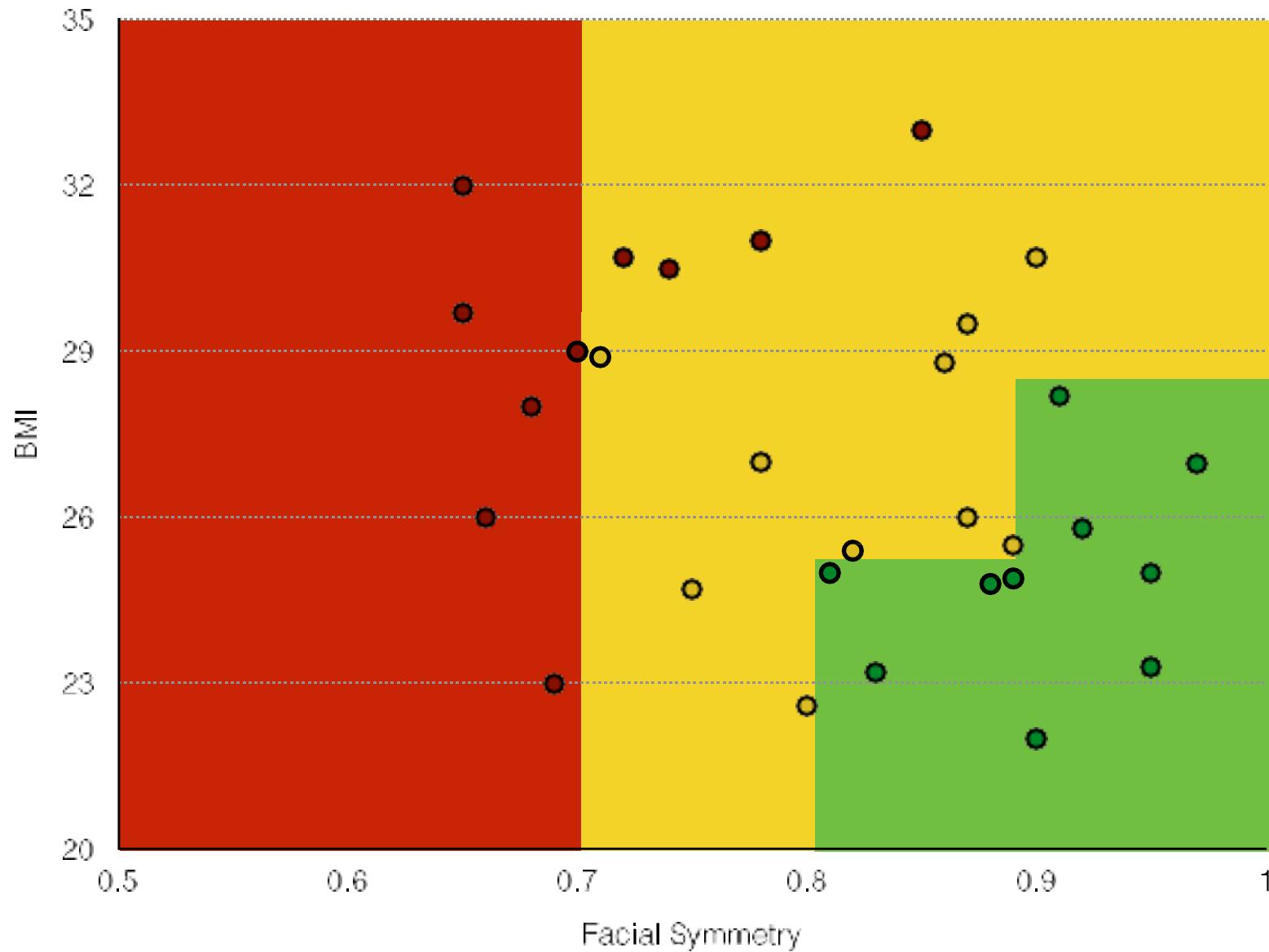
[att, ave, un]

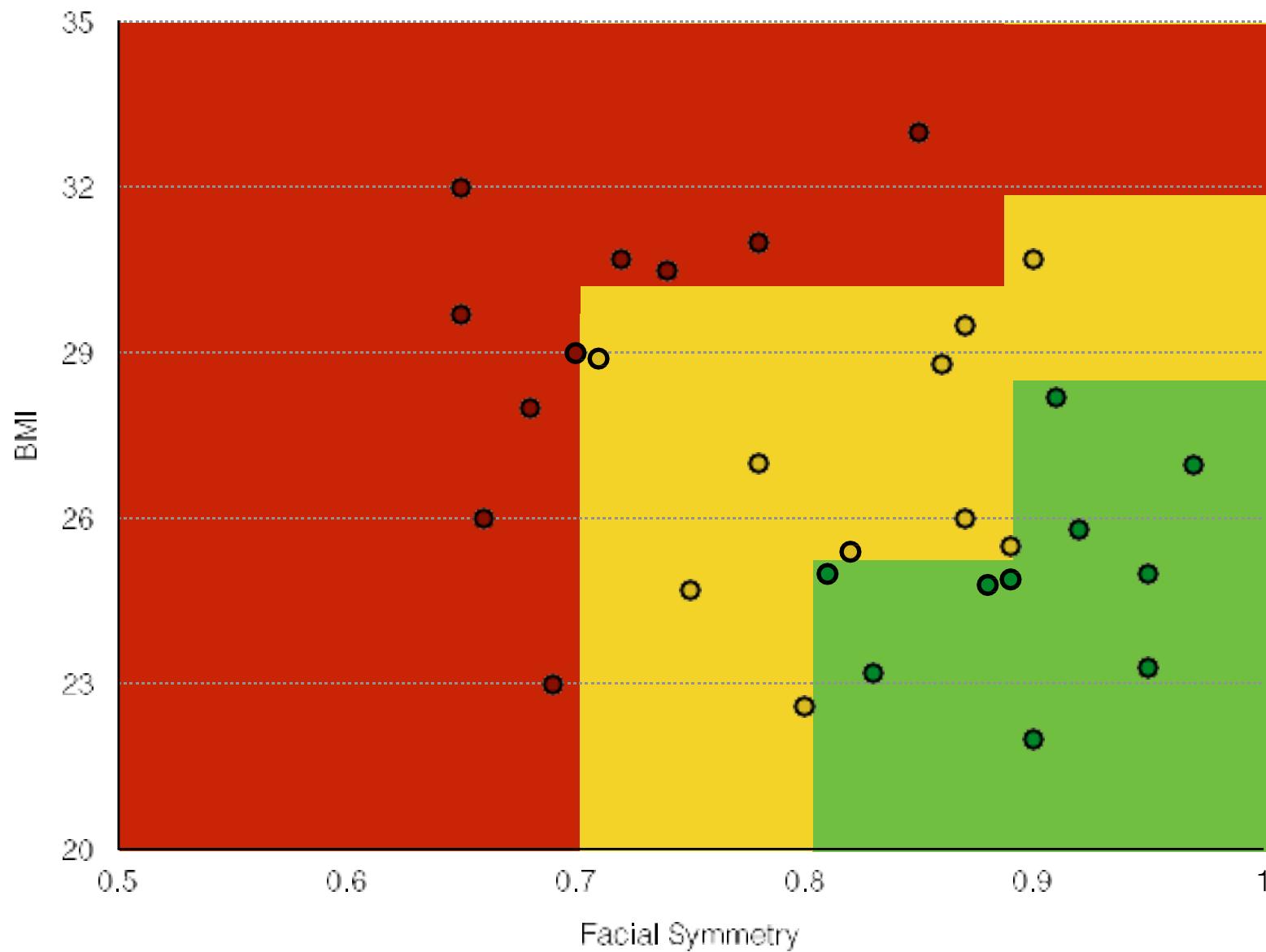


[att, ave, un]



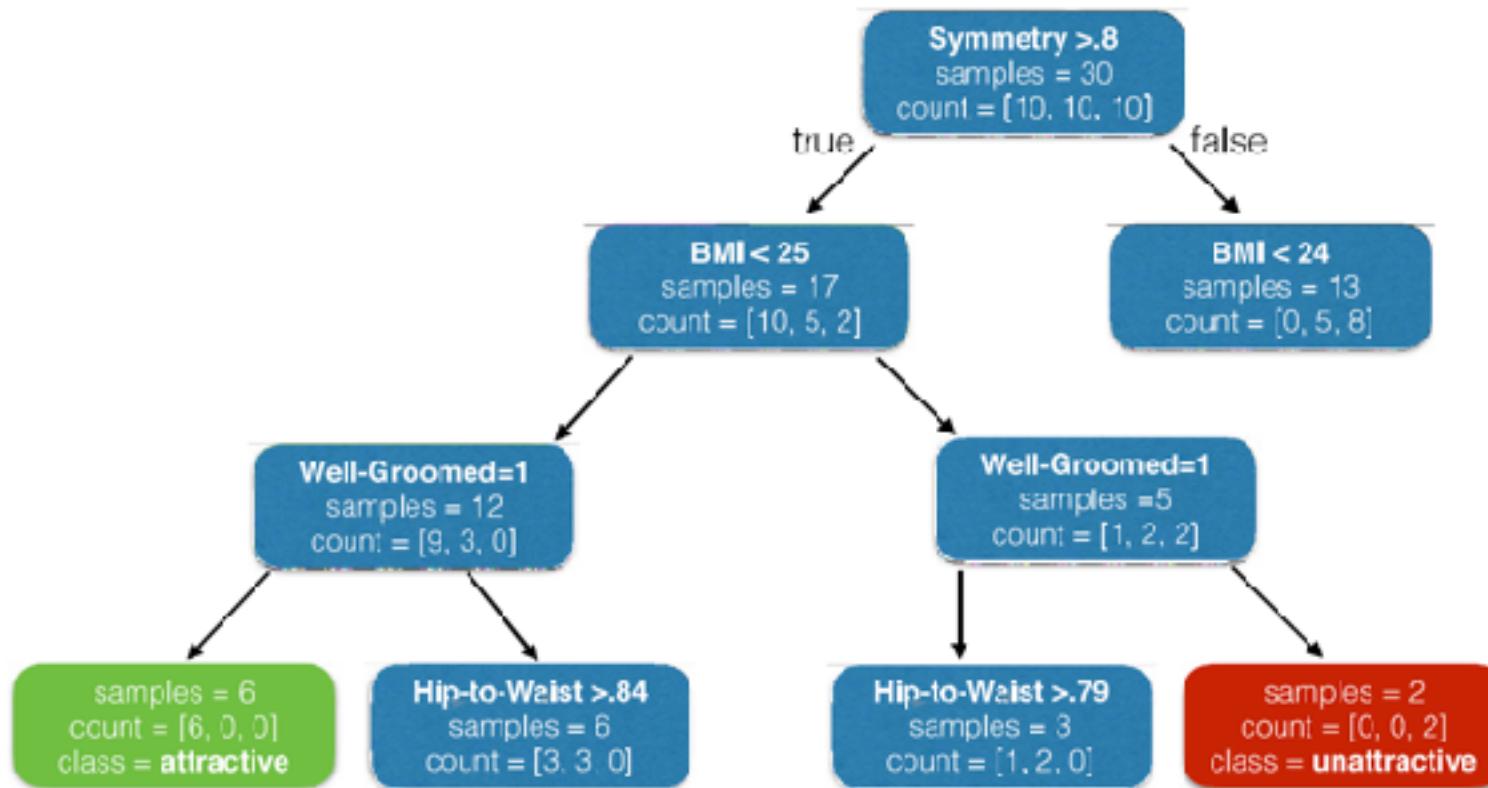




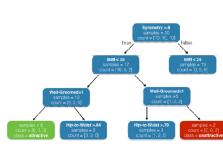
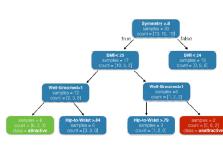
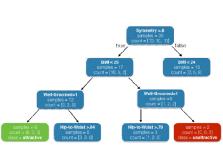


Random Forest

Random Forest



Random Forest



K-nearest Neighbors

Resumes

CV template for a sales executive. The resume is divided into several sections: Profile, Experience, Skills, Languages, Education, and Other Credentials.

Profile: Sales Executive (Manager of Sales)

Experience:

- Progressive Technologies Inc.** July 2012 to Present
Sales Executive
 - Increased territory sales from less than \$4 million to \$8.2 million within two years, exceeding quota by 12% in 2013 and 15% in 2014.
 - Banked as #1 sales manager (out of 12) in 2013 and 2014.
 - Developed a robust, sustainable network of buyers from London to Cardiff, leveraging strong listening, presentation and closing skills to optimize client results despite previously dominant competitor advantage.
 - Created and marketed multimedia, computer-based training, graphics and VHS design, photoimaging, sound editing, and web page development tools to key Fortune 500 and 1000 accounts.
- Niche Software Ltd.** January 2010 to June 2012
Junior Sales Executive
 - Managed daily operations of the IT service department generating \$12.5 million annually. Provided floor sales leadership and supervised eight associates. Rapidly promoted from initial junior sales executive position.
 - Tracked sales using relevant software to provide accurate reports and monitored competitor activities closely to identify any business threats.
 - Gathered customer satisfaction surveys and regularly updated contact database.
 - Contributed to identifying, uncovering product opportunities – supported product development to develop new products.

Skills:

- Microsoft Office Suite
- Financial Forecasting
- Single team supervision
- Presentation & Proposals
- Relationship management
- Effectively meet deadlines
- Achieve targets and work under pressure
- Accounting-related computer literacy
- Excellent communication skills
- Proactive and intuitive leader

Languages:

- Mandarin Chinese (Proficiency), English (Professional proficiency)

Education:

University of Arizona 2009 to 2010
MBA, Master's in Business Administration

- Graduated summa cum laude
- Won the "Next Gen Entrepreneur Award" for a sales and marketing plan development for a software start-up.

Other Credentials:

- Dale Carnegie Sales Training
- Leadership Through Quality
- Account Management-Selling System
- WFIN Sales and Marketing Programs, New York
- Certification Program in Sales Management (CPSM)

	Phone Screen	Experience
Candidate A	3	7
Candidate B	6	11
Candidate C	8	9
Candidate D	7	10

Euclidean Distance

candidate a = [a₁, a₂]

candidate b = [b₁, b₂]

Euclidean Distance

candidate a = [a₁, a₂]

candidate b = [b₁, b₂]

Two dimensions (features)

$$\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$$

Euclidean Distance

Feature Vector

$$a = [3, 7]$$

$$b = [6, 11]$$

Euclidean Distance

$$\sqrt{(3-6)^2 + (7-11)^2}$$

Feature Vector

$$a = [3, 7]$$

$$b = [6, 11]$$

Euclidean Distance

$$\sqrt{(3-6)^2 + (7-11)^2}$$

Feature Vector

$$a = [3, 7]$$

$$b = [6, 11]$$

$$\sqrt{(-3)^2 + (-4)^2}$$

Euclidean Distance

$$\sqrt{(3-6)^2 + (7-11)^2}$$

Feature Vector

$$a = [3, 7]$$

$$b = [6, 11]$$

$$\sqrt{(-3)^2 + (-4)^2}$$

$$\sqrt{9+16}$$

Euclidean Distance

$$\sqrt{(3-6)^2 + (7-11)^2}$$

Feature Vector

$$a = [3, 7]$$

$$b = [6, 11]$$

$$\sqrt{(-3)^2 + (-4)^2}$$

$$\sqrt{9+16}$$

Distance between points (vectors) a and b:

$$\sqrt{25}$$

Euclidean Distance

$$\sqrt{(3-6)^2 + (7-11)^2}$$

Feature Vector

$$a = [3, 7]$$

$$b = [6, 11]$$

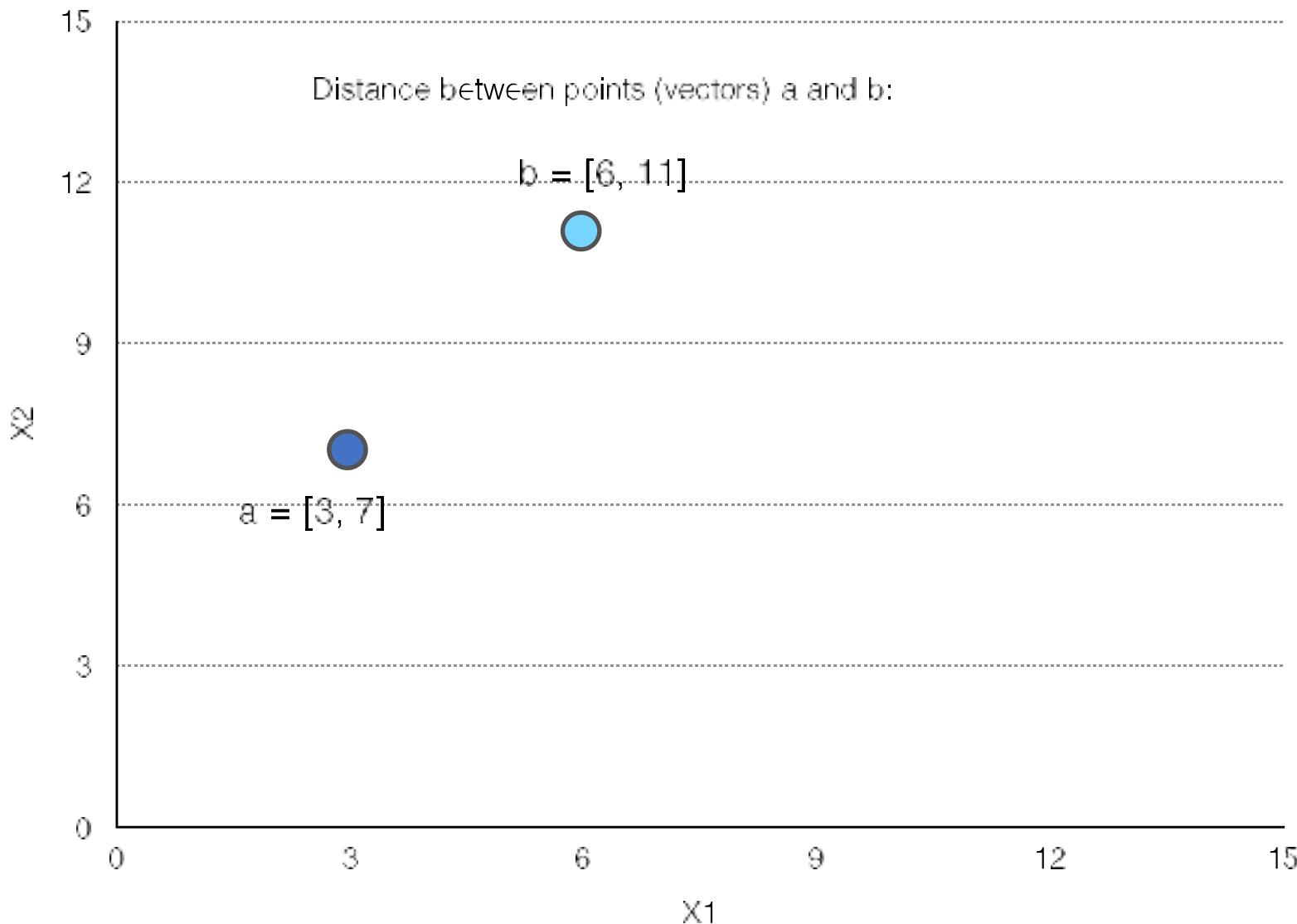
$$\sqrt{(-3)^2 + (-4)^2}$$

$$\sqrt{9+16}$$

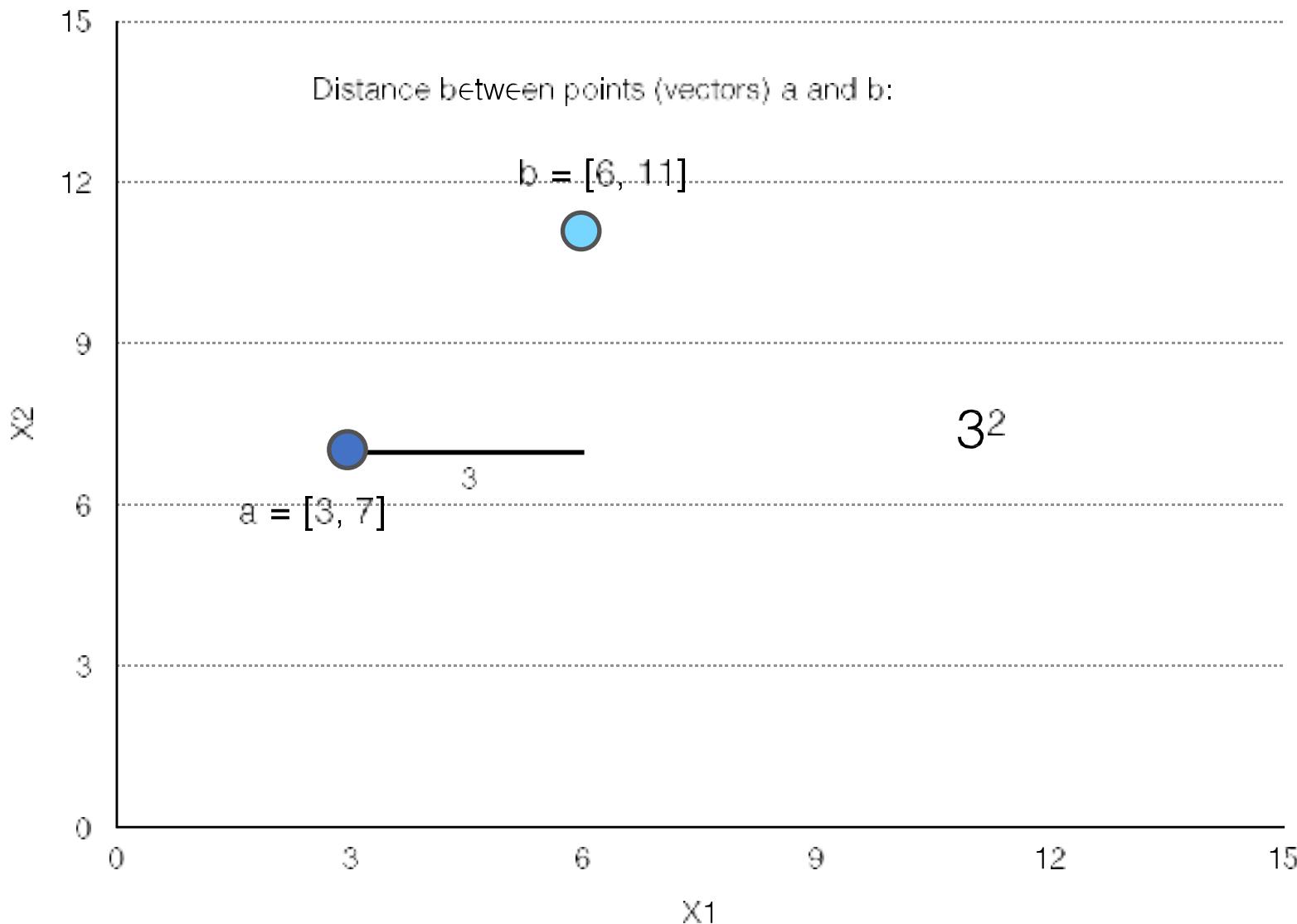
Distance between points (vectors) a and b:

$$\sqrt{25} = 5$$

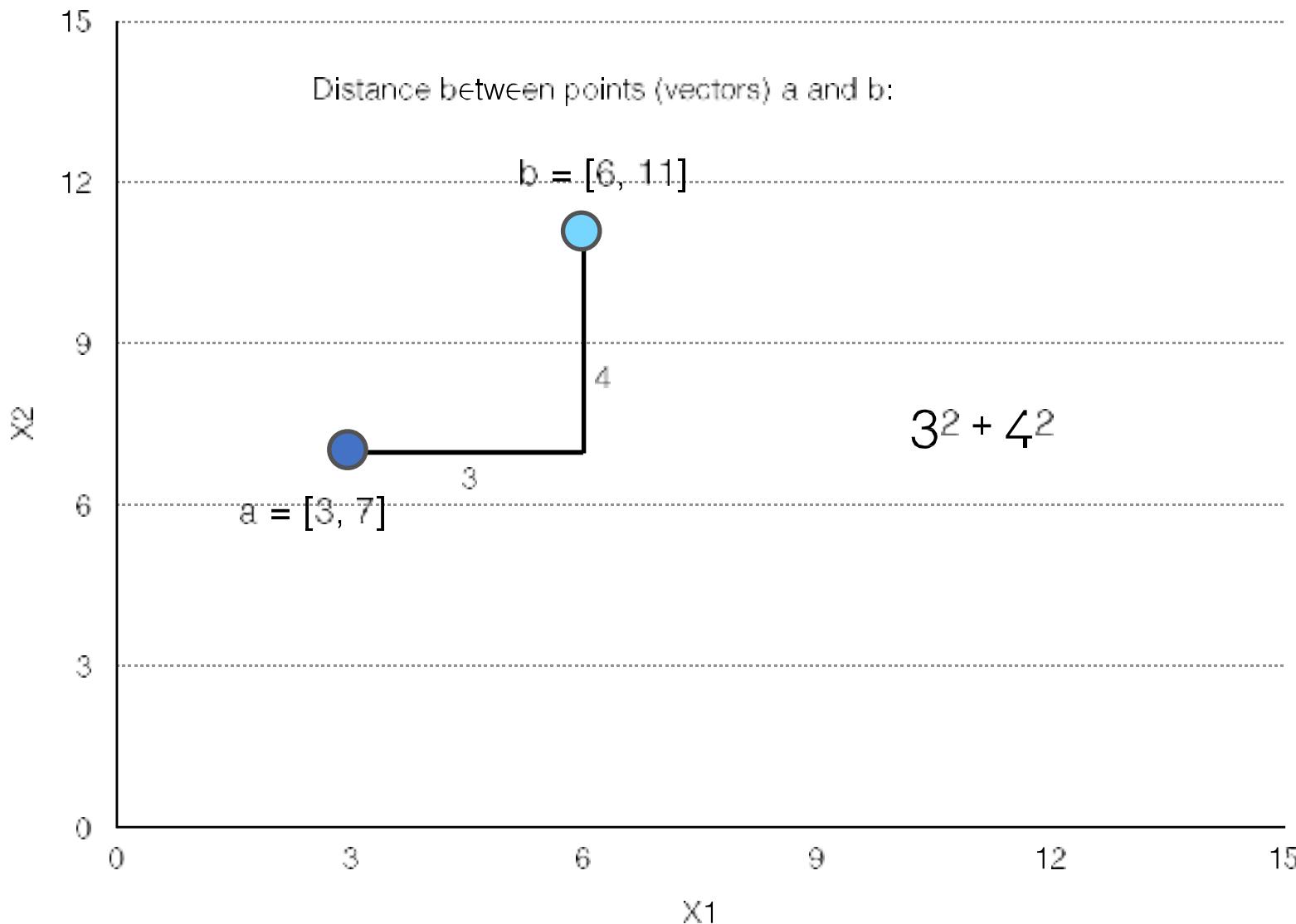
Euclidean Distance



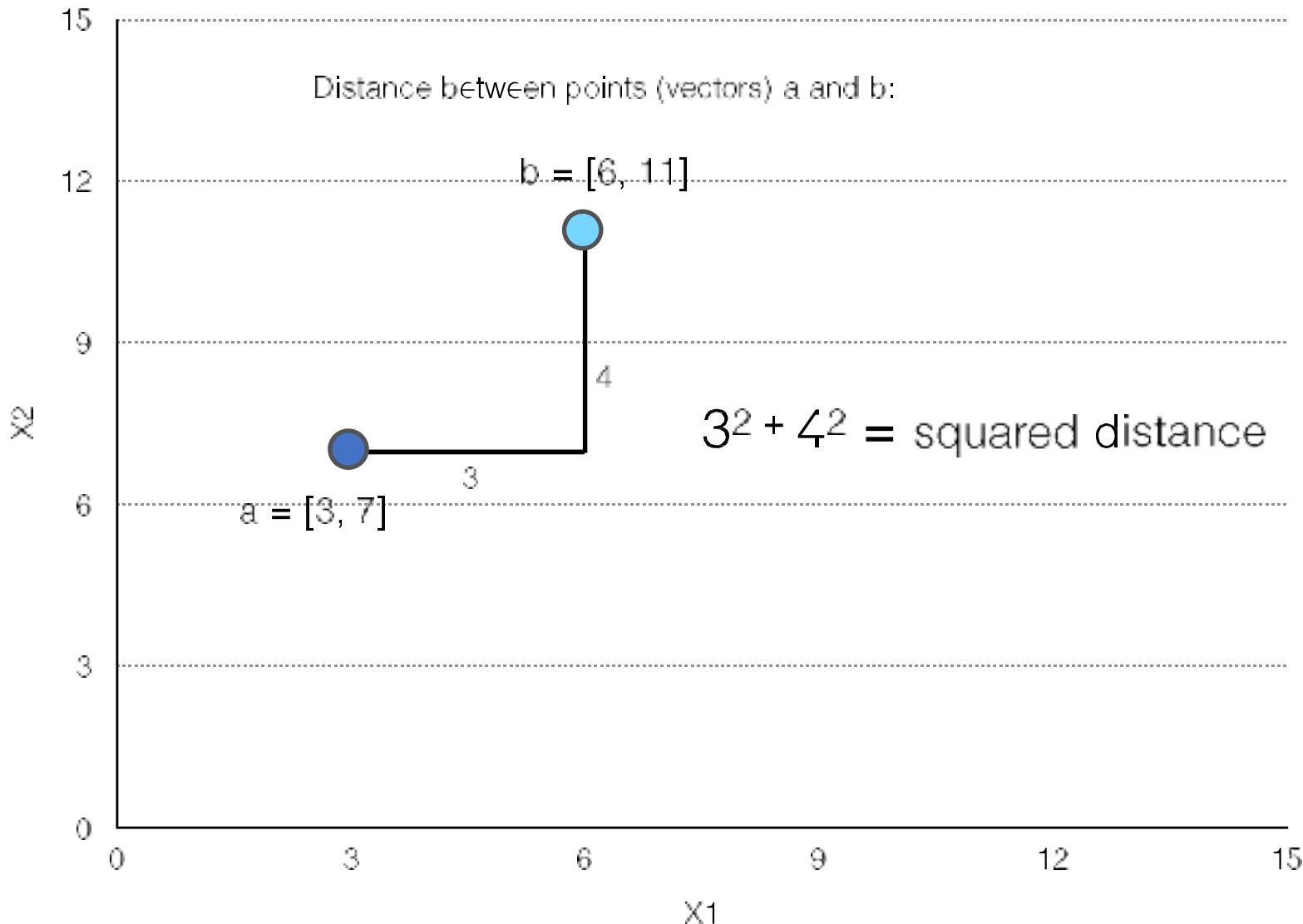
Euclidean Distance



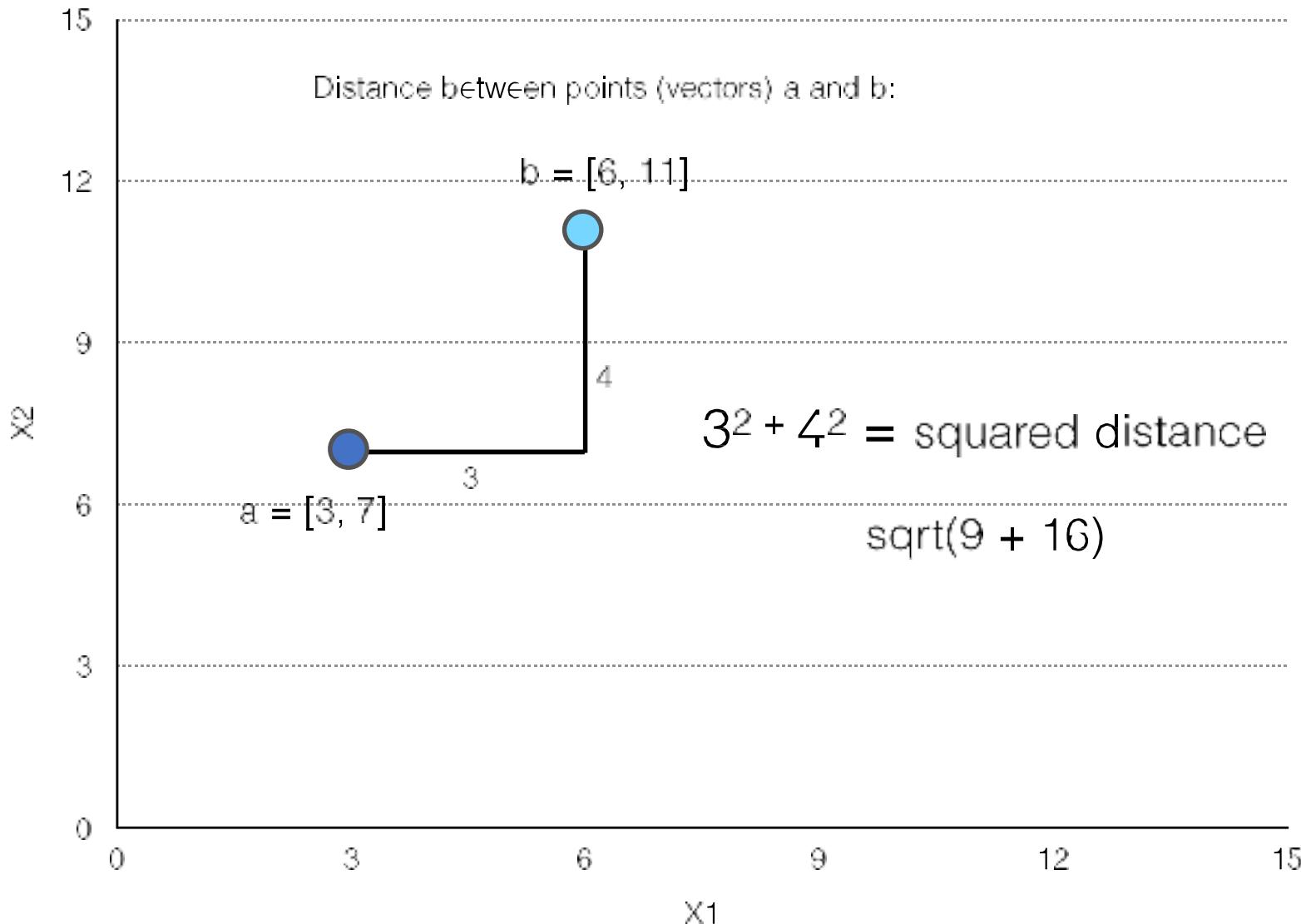
Euclidean Distance



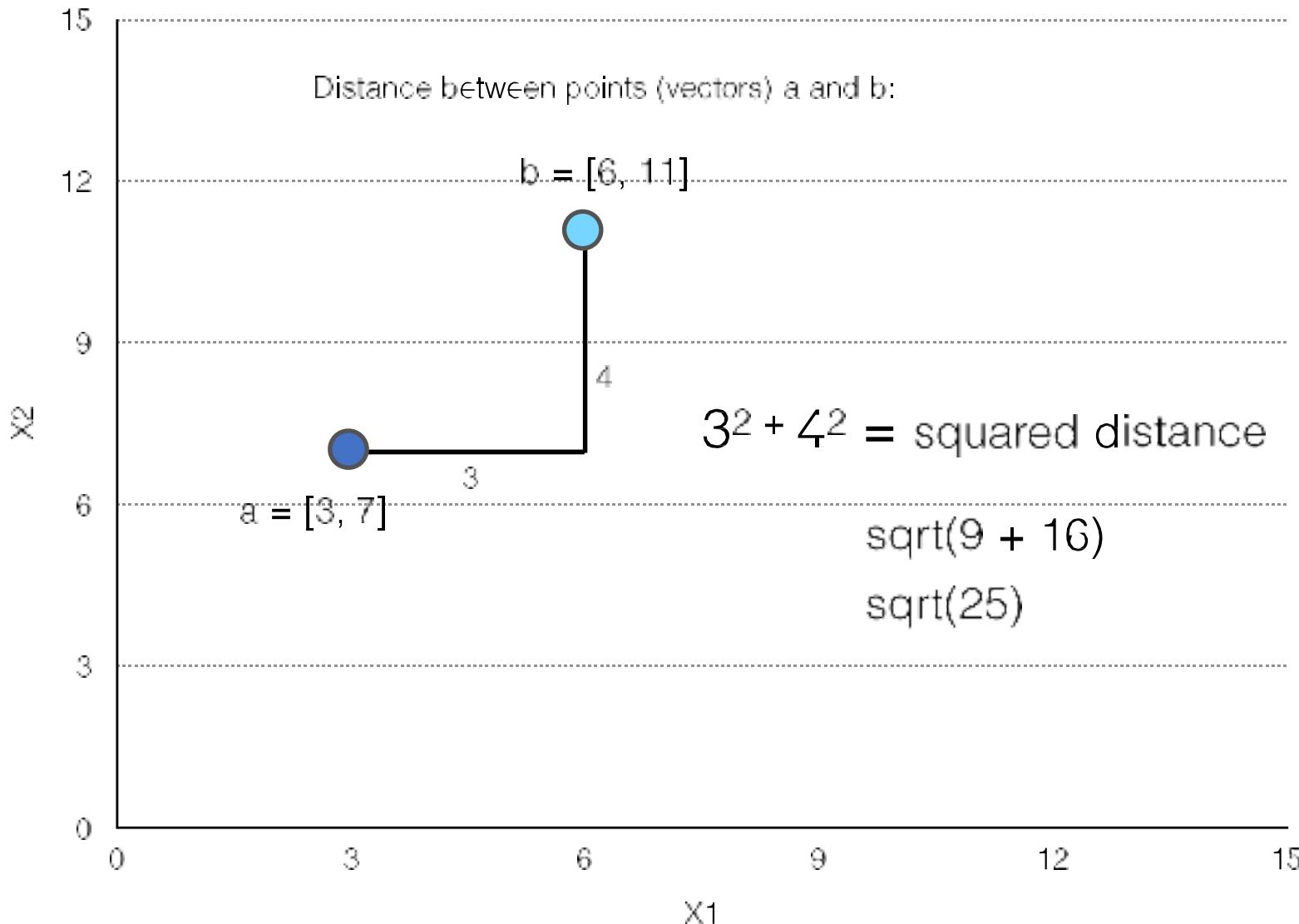
Euclidean Distance



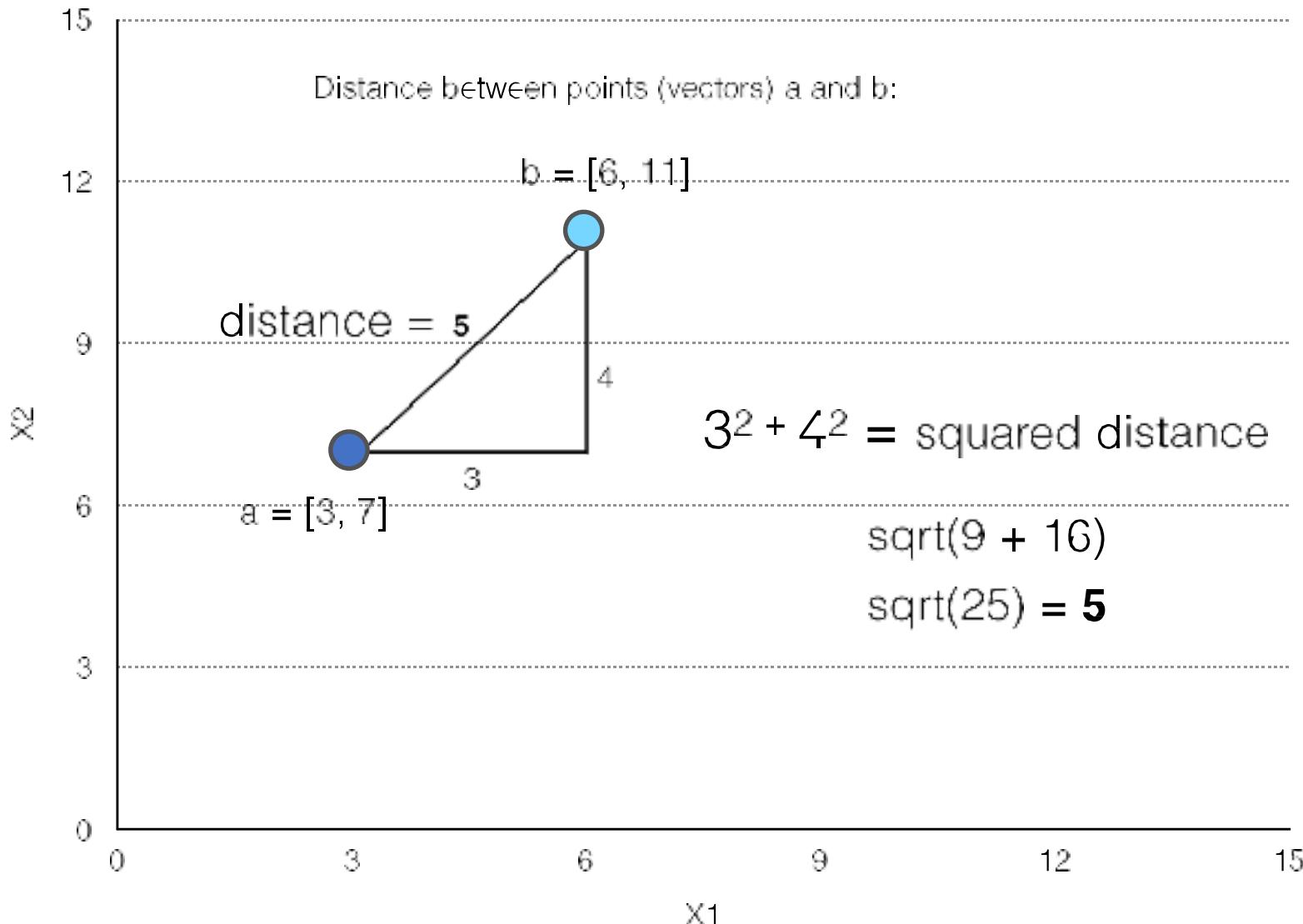
Euclidean Distance



Euclidean Distance



Euclidean Distance



Resumes

CV template for a sales executive. The resume is divided into several sections: Profile, Experience, Skills, Languages, Education, and Other Credentials.

Profile: Sales Executive (Manager of Sales)

Experience:

- Progressive Technologies Inc.** July 2012 to Present
Sales Executive
 - Increased territory sales from less than \$4 million to \$8.2 million within two years, exceeding quota by 12% in 2013 and 15% in 2014.
 - Banked as #1 sales manager (out of 12) in 2013 and 2014.
 - Developed a robust, sustainable network of buyers from London to Cardiff, leveraging strong listening, presentation and closing skills to optimize client results despite previously dominant competitor advantage.
 - Created and marketed multimedia, computer-based training, graphics and VHS design, photoimaging, sound editing, and web page development tools to key Fortune 500 and 1000 accounts.
- Niche Software Ltd.** January 2010 to June 2012
Junior Sales Executive
 - Managed daily operations of the IT service department generating \$12.5 million annually. Provided floor sales leadership and supervised eight associates. Rapidly promoted from initial junior sales executive position.
 - Tracked sales using relevant software to provide accurate reports and monitored competitor activities closely to identify any business threats.
 - Gathered customer satisfaction surveys and regularly updated contact database.
 - Contributed to identifying, uncovering product opportunities – supported product development to develop new products.

Skills:

- Microsoft Office Suite
- Financial Forecasting
- Single team supervision
- Presentation & Proposals
- Relationship management
- Effectively meet deadlines
- Achieve targets and work under pressure
- Accounting-related computer literacy
- Excellent communication skills
- Proactive and intuitive leader

Languages:

- Mandarin Chinese (Proficiency), English (Professional proficiency)

Education:

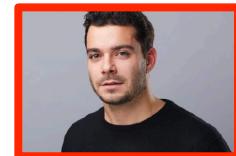
University of Arizona 2009 to 2010
MBA, Master's in Business Administration

- Graduated summa cum laude
- Won the "Next Gen Entrepreneur Award" for a sales and marketing plan development for a software start-up.

Other Credentials:

- Dale Carnegie Sales Training
- Leadership Through Quality
- Account Management-Selling System
- WFIN Sales and Marketing Programs, New York
- Certification Program in Sales Management (CPSM)

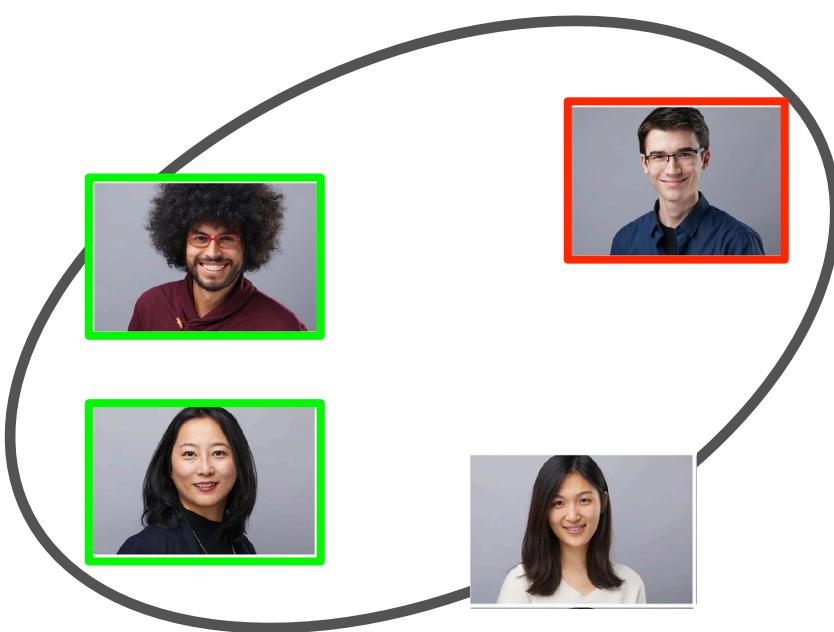




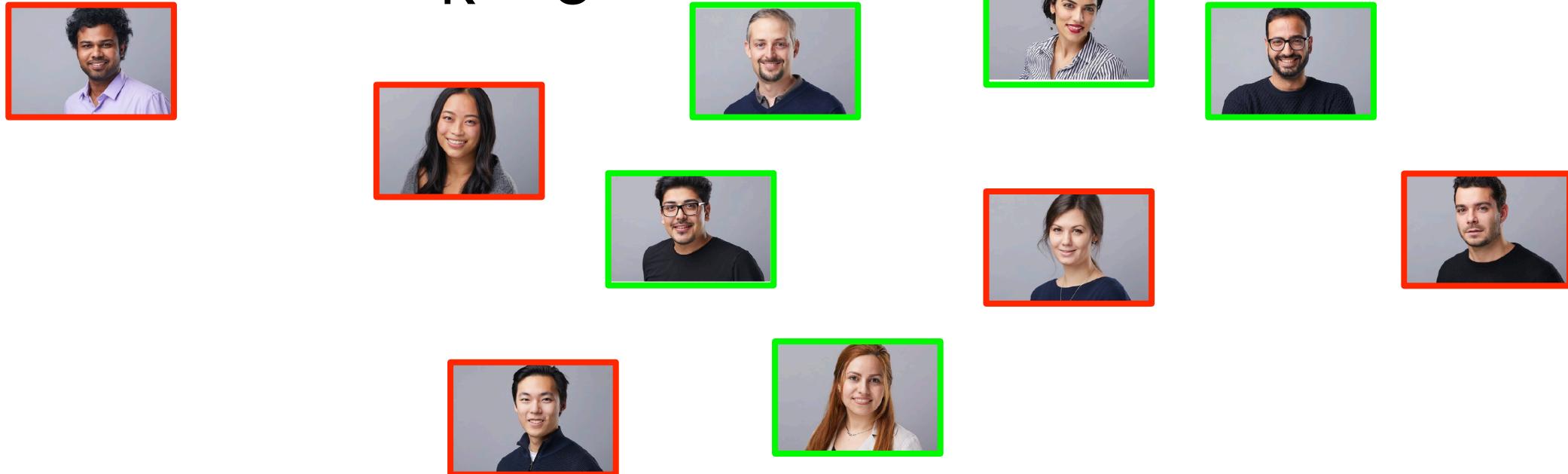


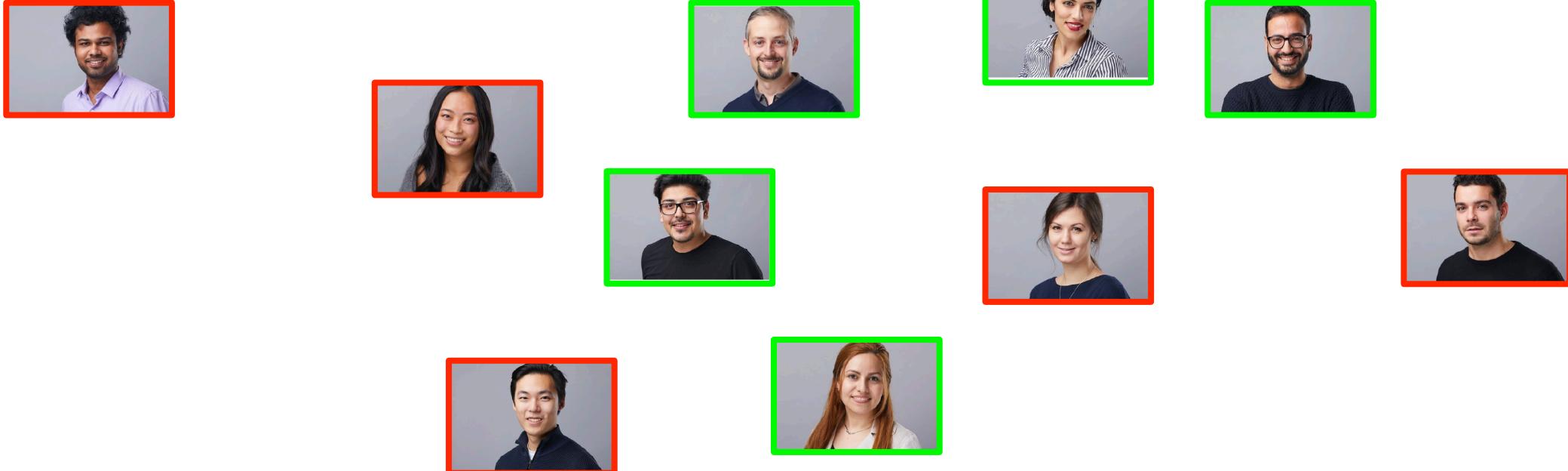
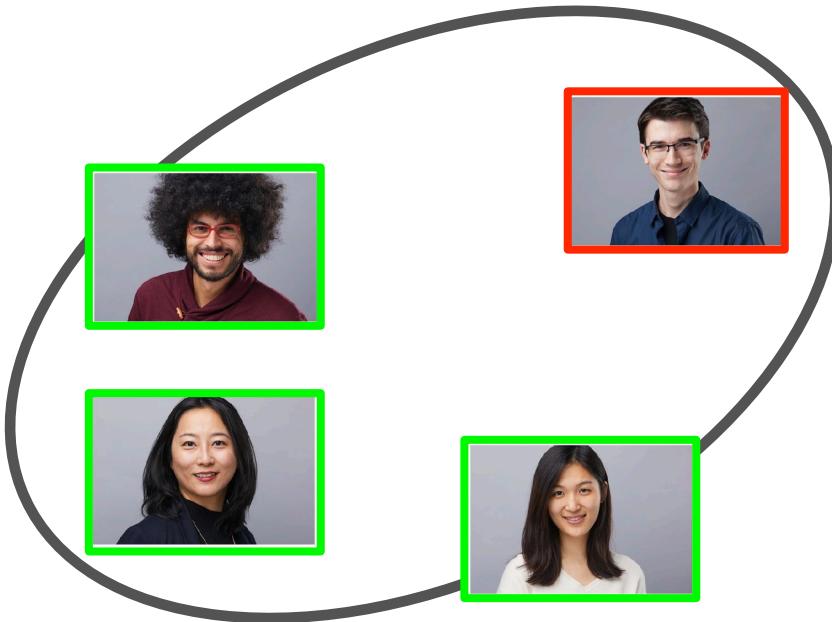
$k = 3$





$k = 3$



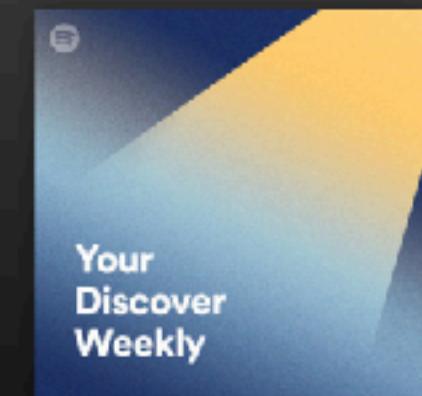




Recommendation System

Collaborative Filtering

Model



MADE FOR HISBIZ

Discover Weekly

Your weekly mixtape of fresh music. Enjoy new discoveries and deep cuts chosen just for you. Updated every Monday, so save your favorites!

Made for hisbiz by Spotify • 30 songs, 2 hr 24 min

PLAY

FOLLOWER
1

Q Filter

Download

TITLE	ARTIST	ALBUM	
Land	Hidden Spheres	Well Well	7 days ago
Where You Are	Ambient Solie	Where You Are	7 days ago
falling	Galavant Cross	falling	7 days ago
XOXO	Emmit Fenn	Before We Begin	7 days ago
Orissa	Wayward	Orissa EP	7 days ago
What I Said	Fuzhou	What I Said	7 days ago
A Beautiful Life	Broke For Free	Petal	7 days ago
Strangers	Giyo	Strangers	7 days ago
Trust - Christian Löffler Revision	Max Cooper, To...	Emergence Remi...	7 days ago
Leave	LycorisCoris	Return Flight	7 days ago

Collaborative Filtering

	My Time	Break	Slow Beats	Just Friends
Man 1		✓	✓	✓
Woman 1	✓	✓	✓	
Man 2		✓		
Woman 2	✓			

Collaborative Filtering

	My Time	Break	Slow Beats	Just Friends
		✓	✓	✓
	✓	✓	✓	
		✓		
	✓			

Collaborative Filtering

	My Time	Break	Slow Beats	Just Friends
		✓	✓	✓
	✓	✓	✓	
		✓		
	✓			

Collaborative Filtering

	My Time	Break	Slow Beats	Just Friends
		✓	✓	✓
	✓	✓	✓	
		✓		
	✓			

Collaborative Filtering

	My Time	Break	Slow Beats	Just Friends
		✓	✓	✓
	✓	✓	✓	✓
		✓		
	✓			

Collaborative Filtering

	My Time	Break	Slow Beats	Just Friends
		✓	✓	✓
	✓	✓	✓	✓
		✓		
	✓			

Collaborative Filtering

	My Time	Break	Slow Beats	Just Friends
	✓	✓	✓	✓
	✓	✓	✓	✓
		✓		
	✓			

Collaborative Filtering

	My Time	Break	Slow Beats	Just Friends
Man 1	✓	✓	✓	✓
Woman 1	✓	✓	✓	✓
Man 2		✓		
Woman 2	✓			

Ensemble Methods

Ensemble Methods

Bagging

Boosting

Stacking

Bagging

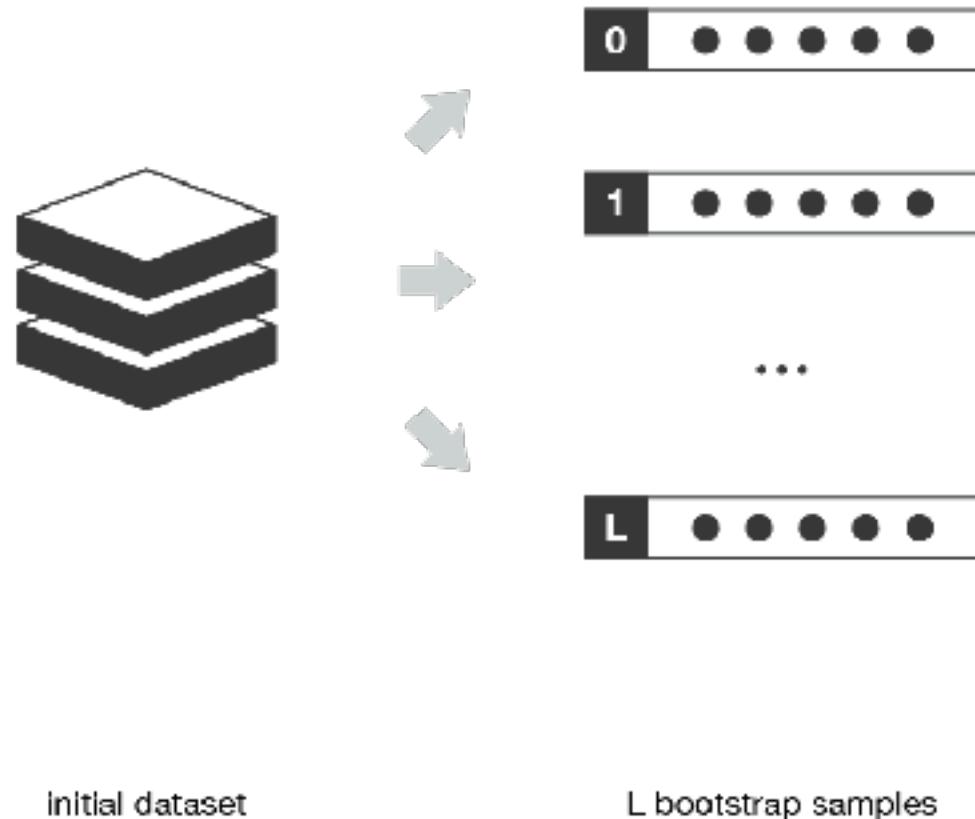
Bootstrap Aggregating



Initial dataset

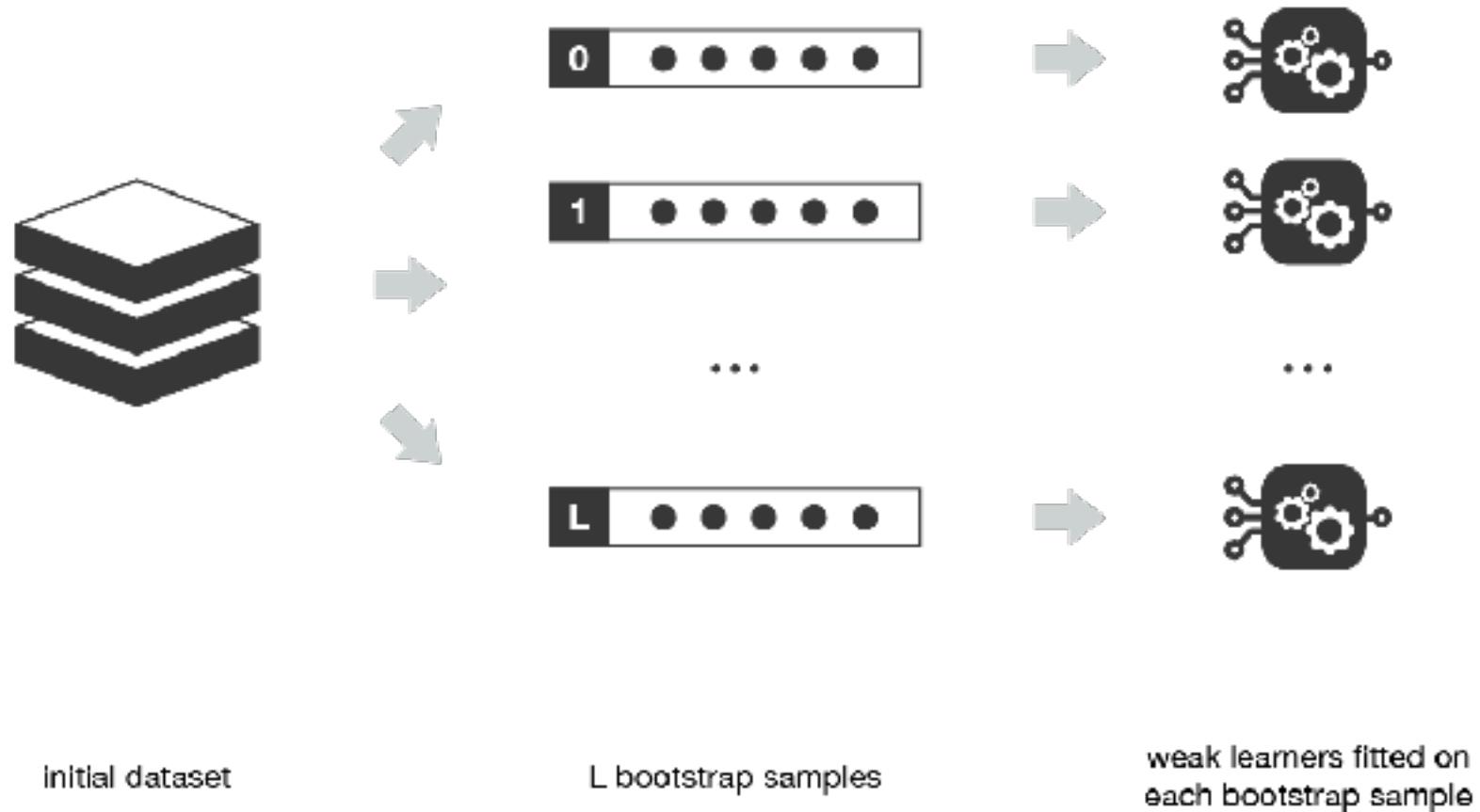
Bagging

Bootstrap Aggregating



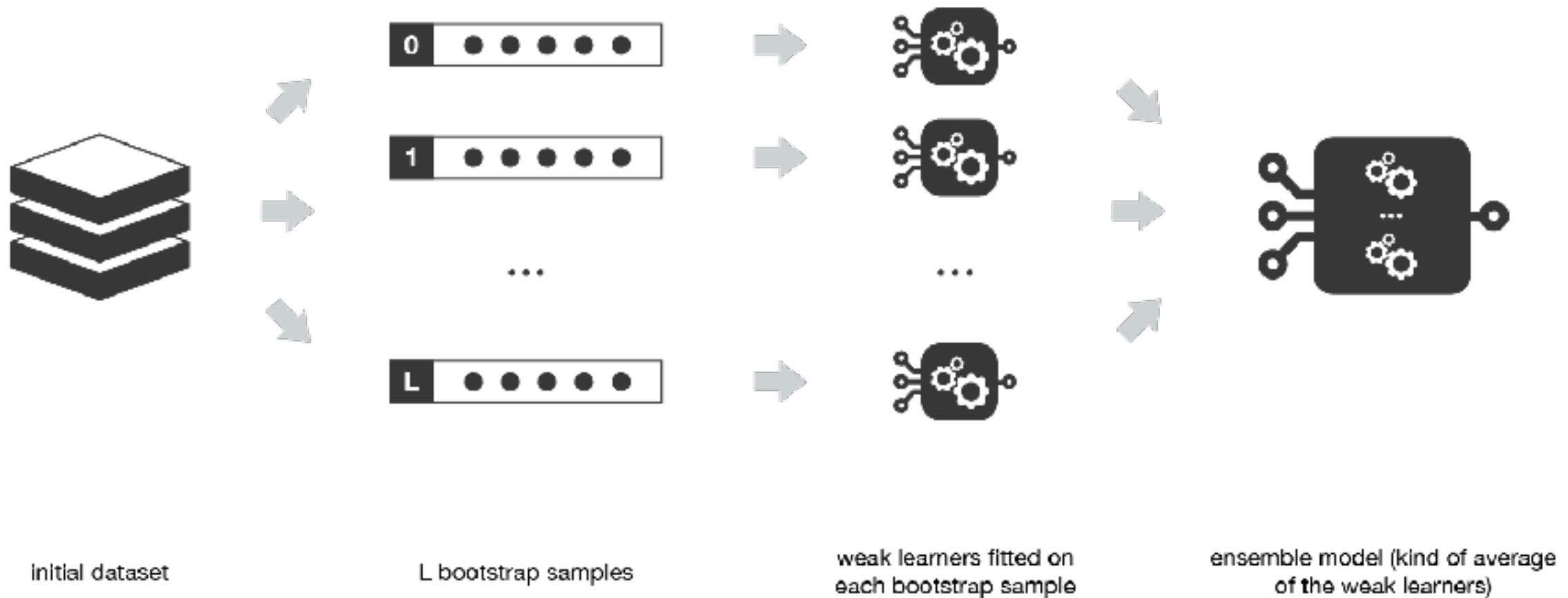
Bagging

Bootstrap Aggregating

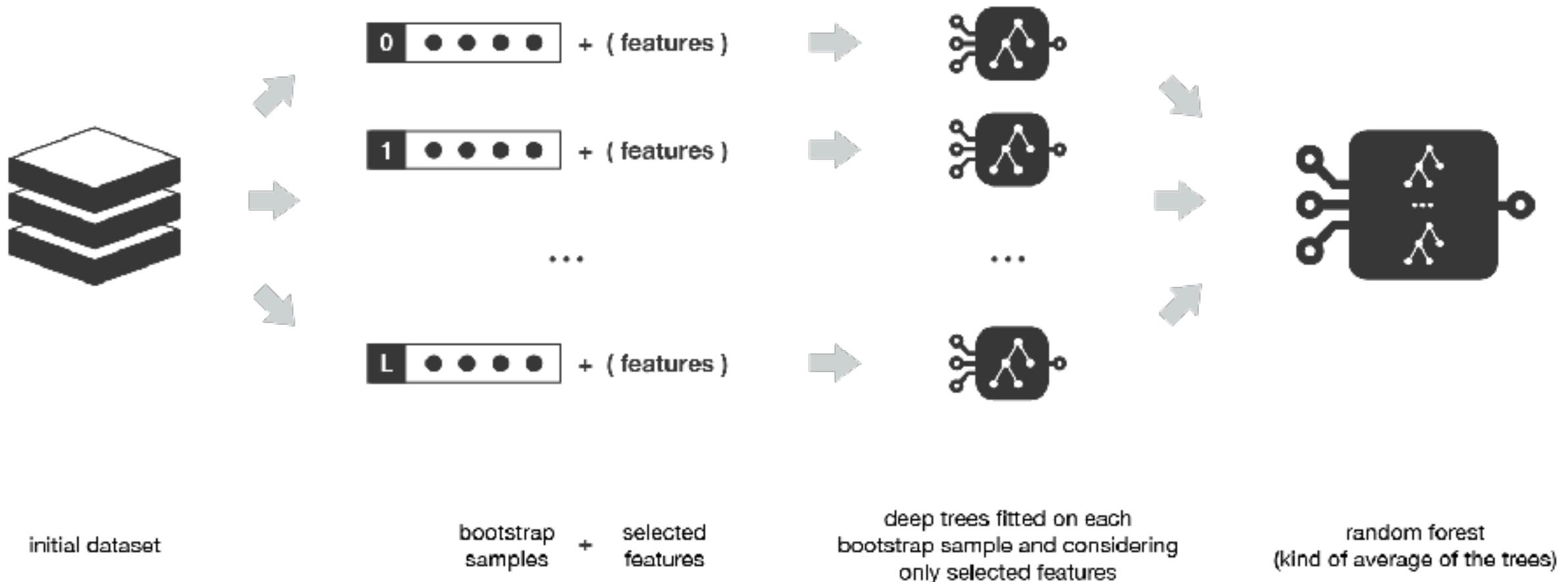


Bagging

Bootstrap Aggregating

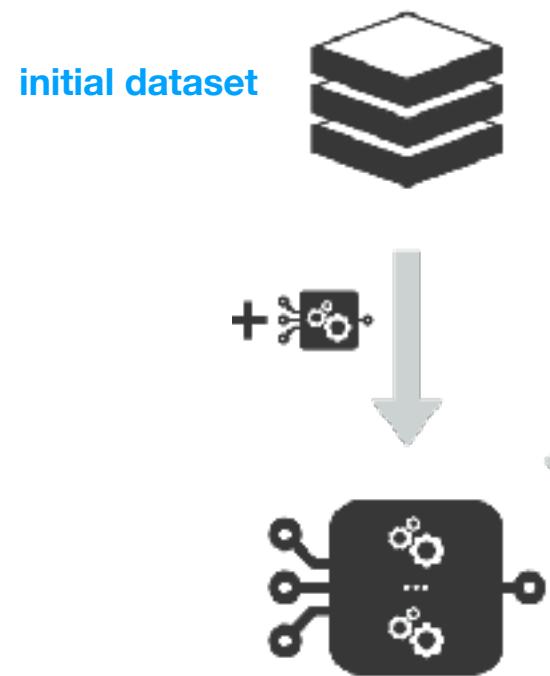


Random Forest

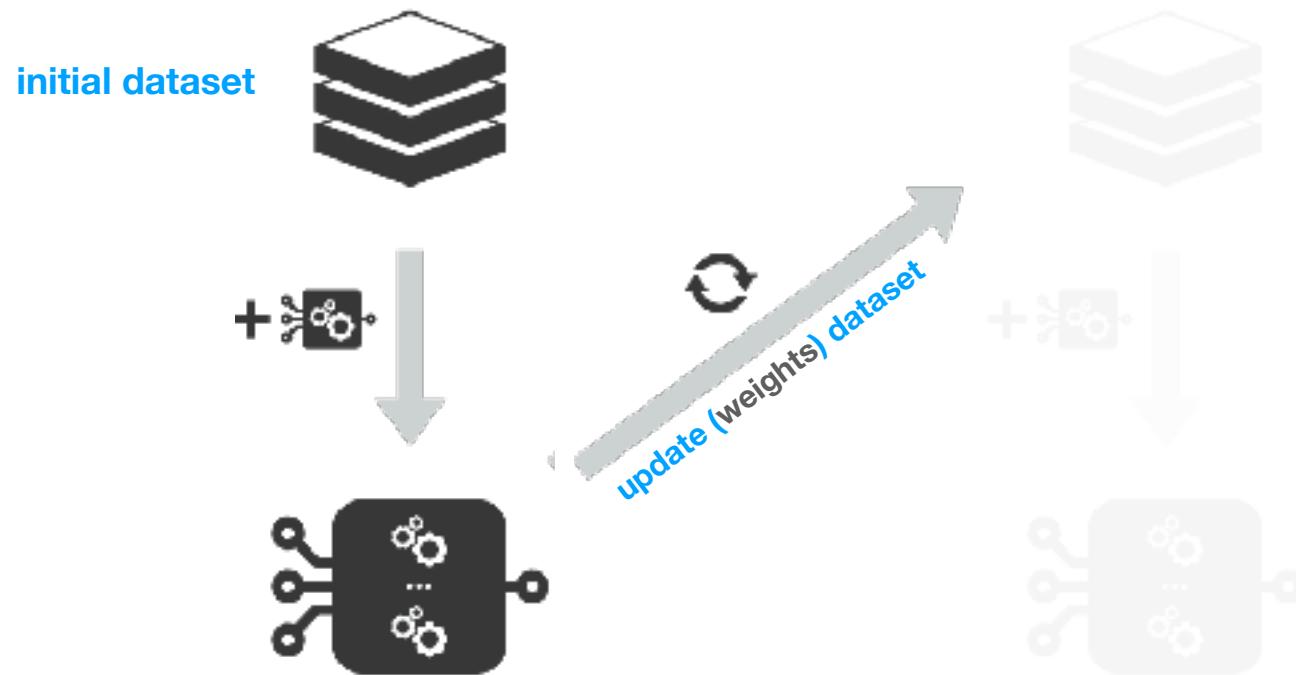


Boosting

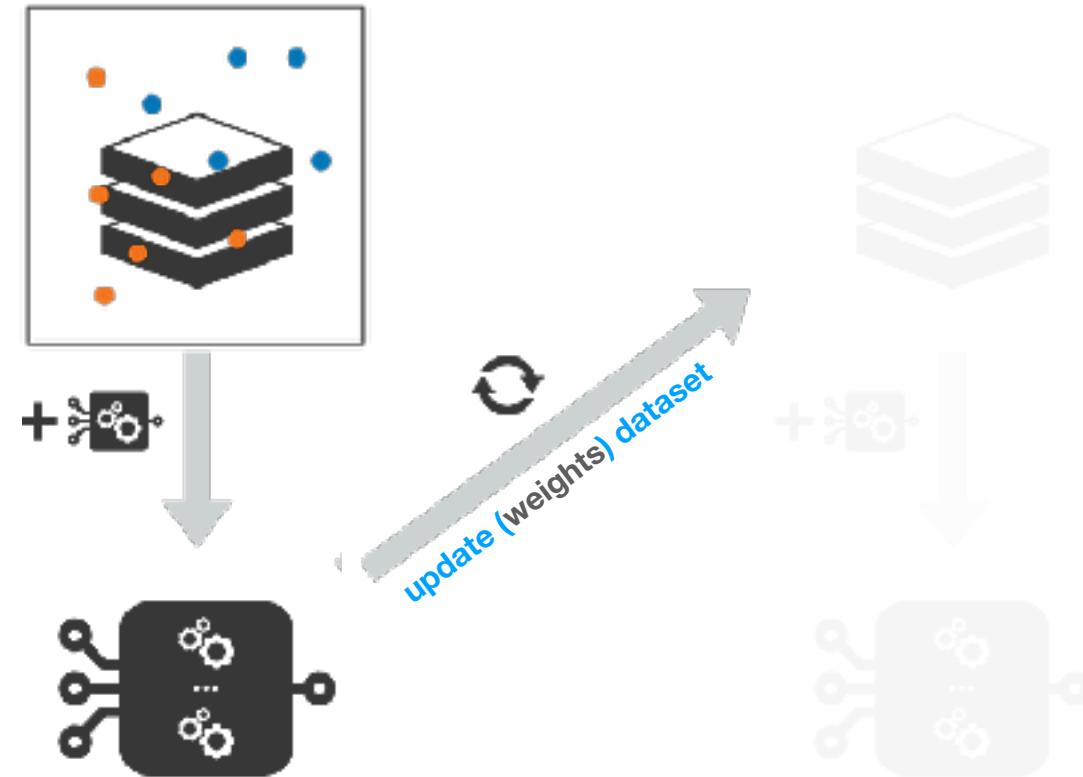
Boosting



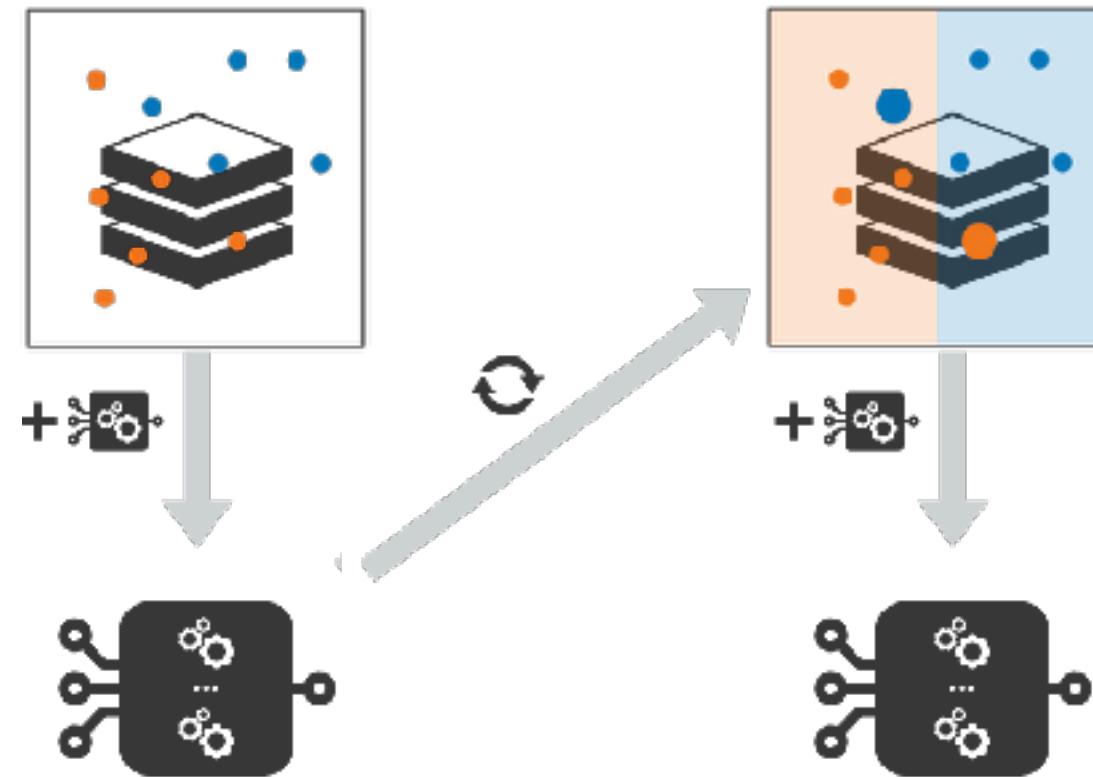
Boosting



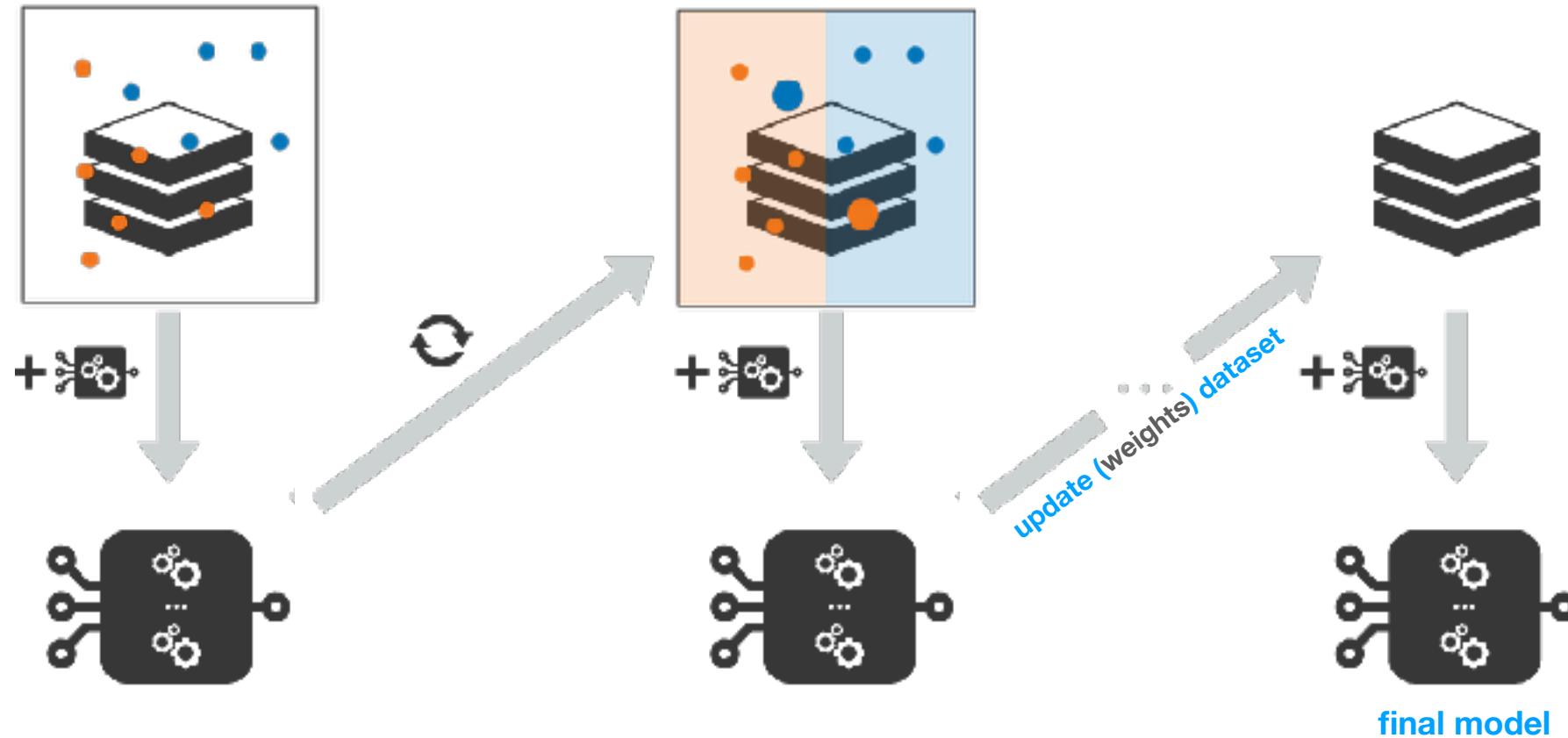
Boosting



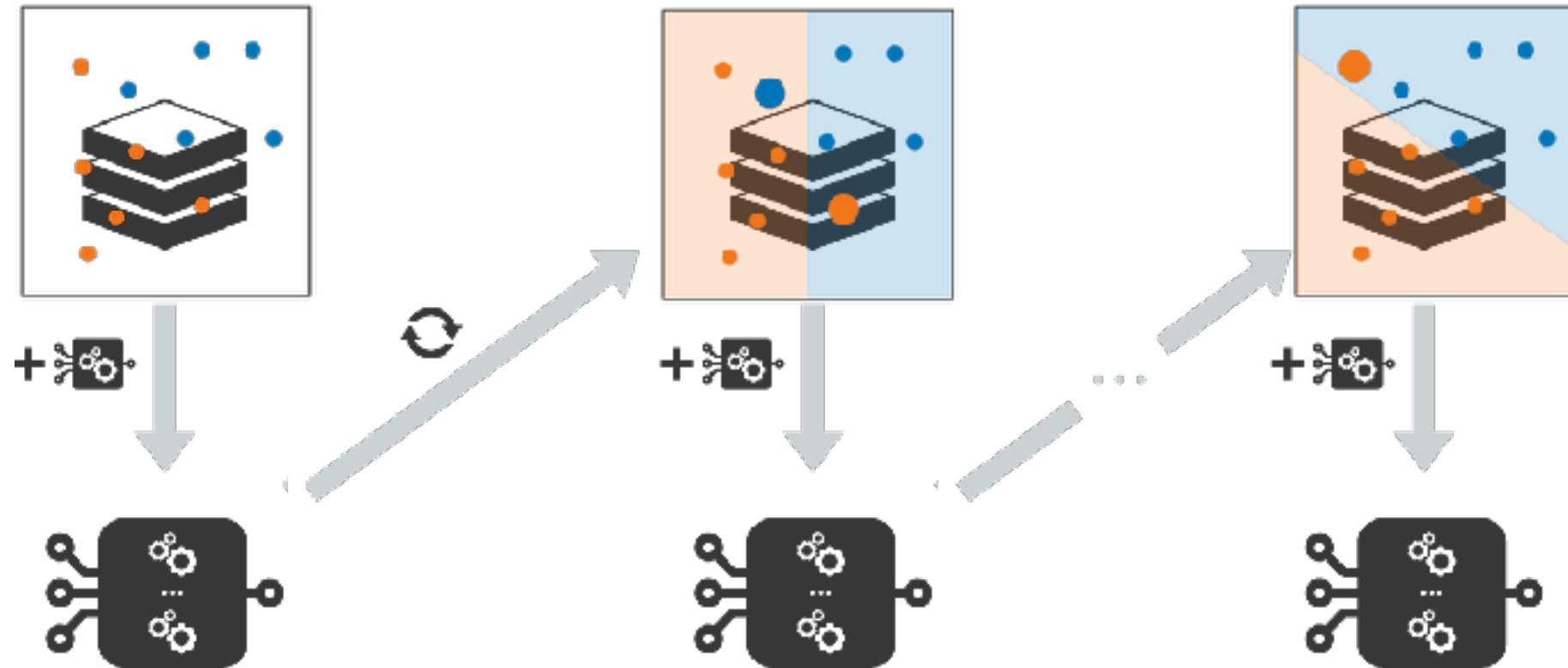
Boosting



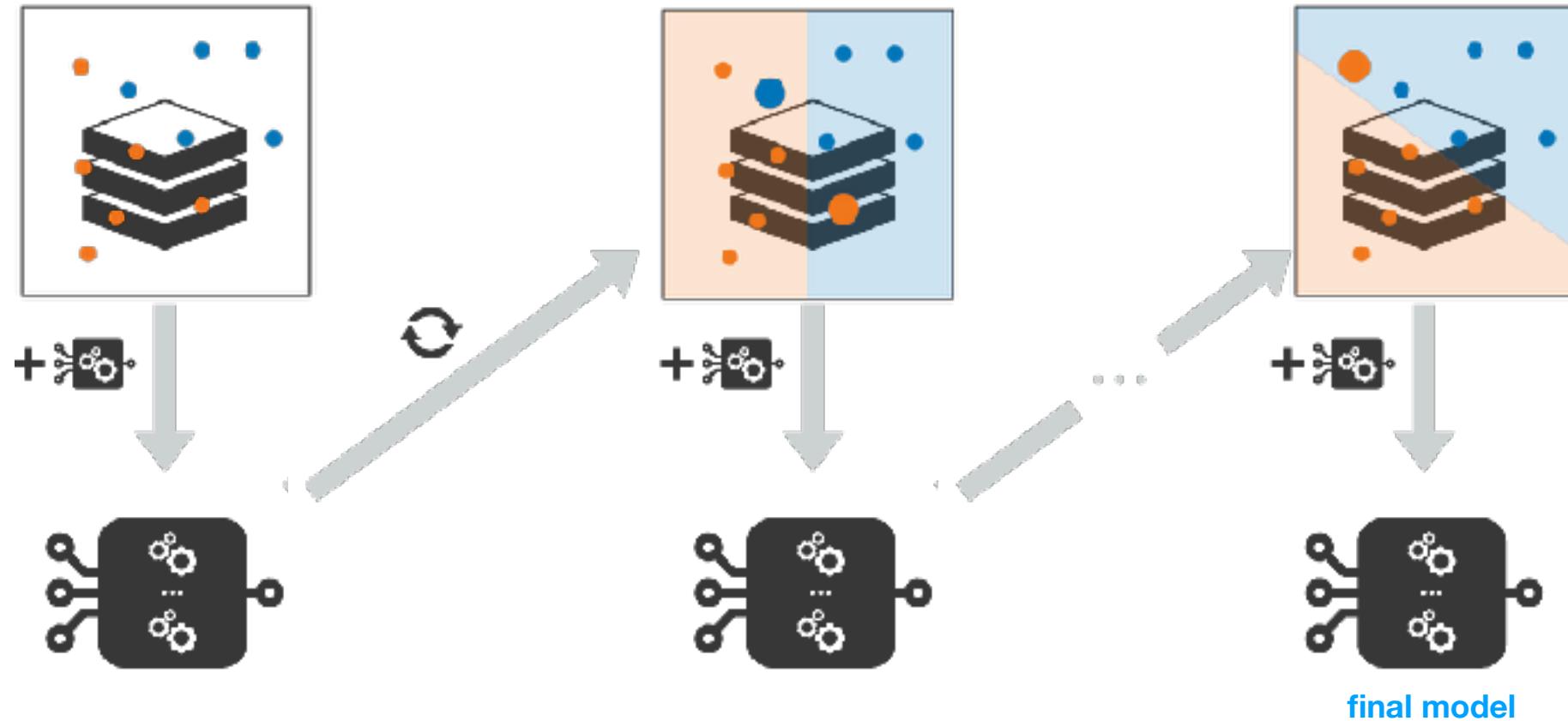
Boosting



Boosting

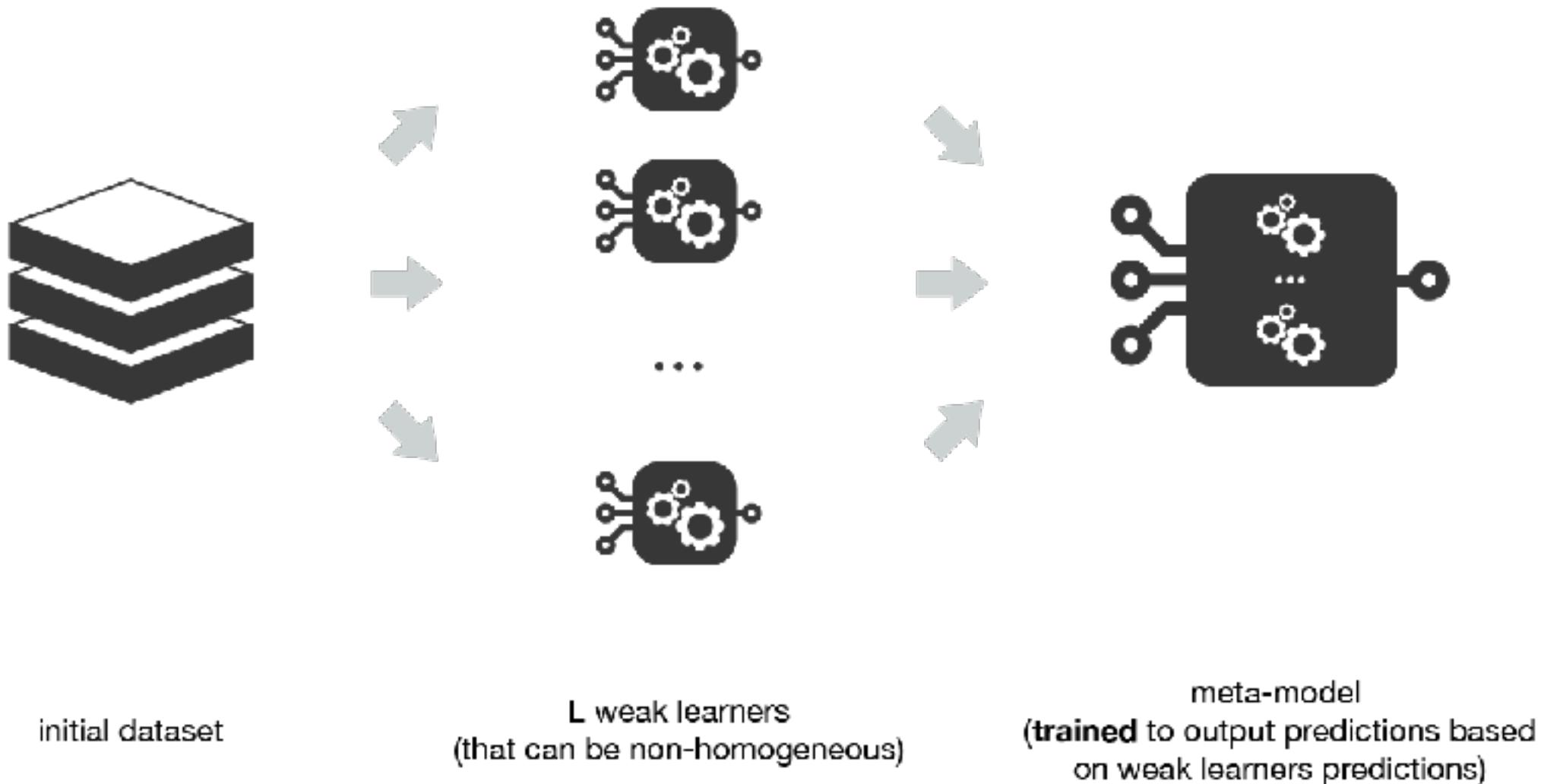


Boosting

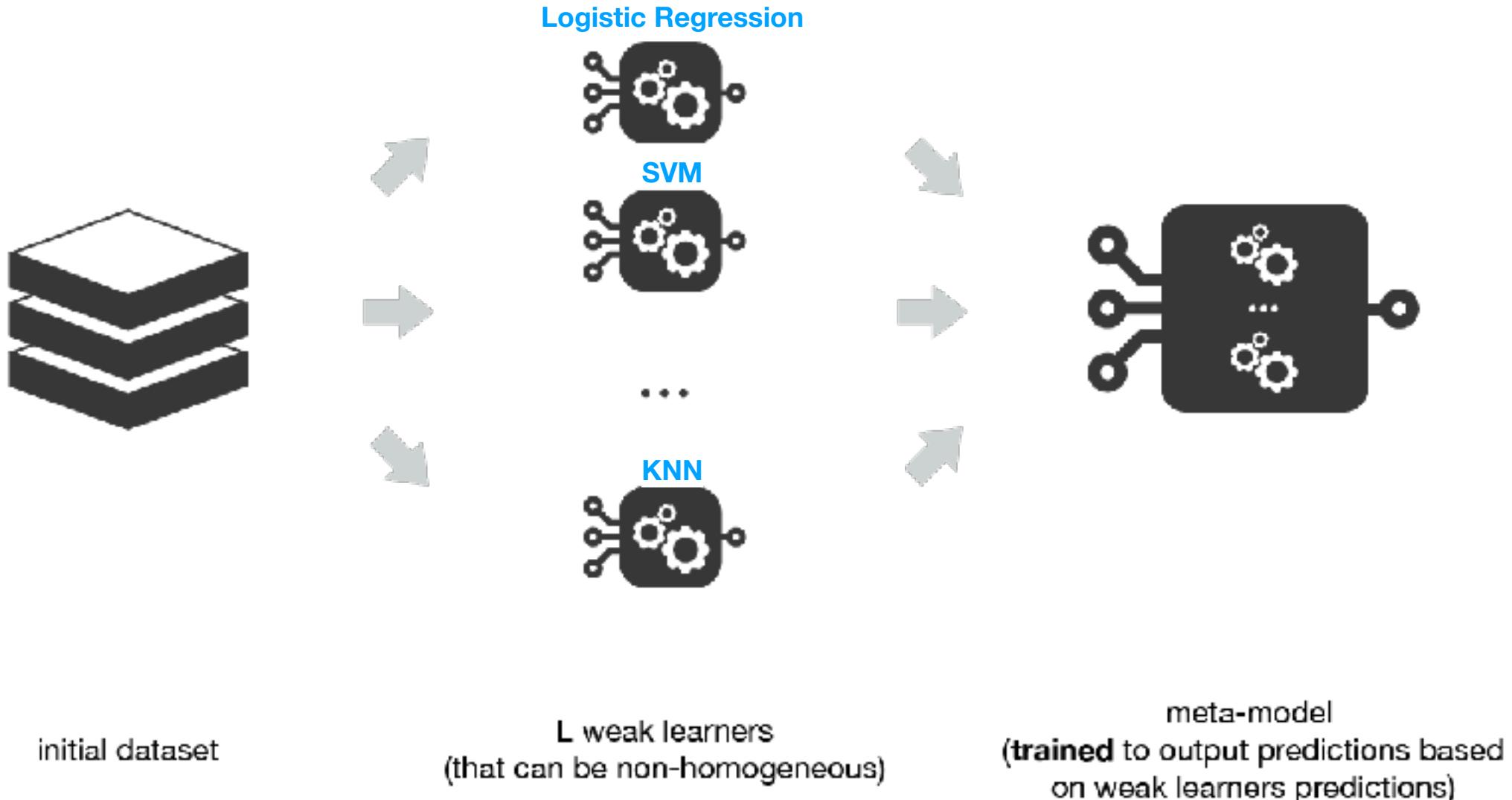


Stacking

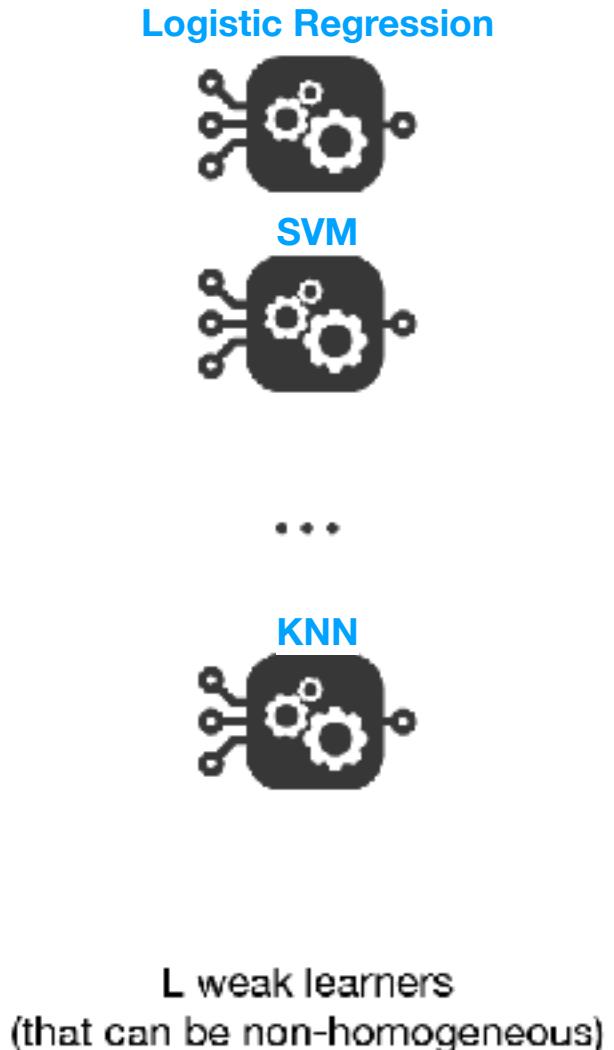
Stacking



Stacking

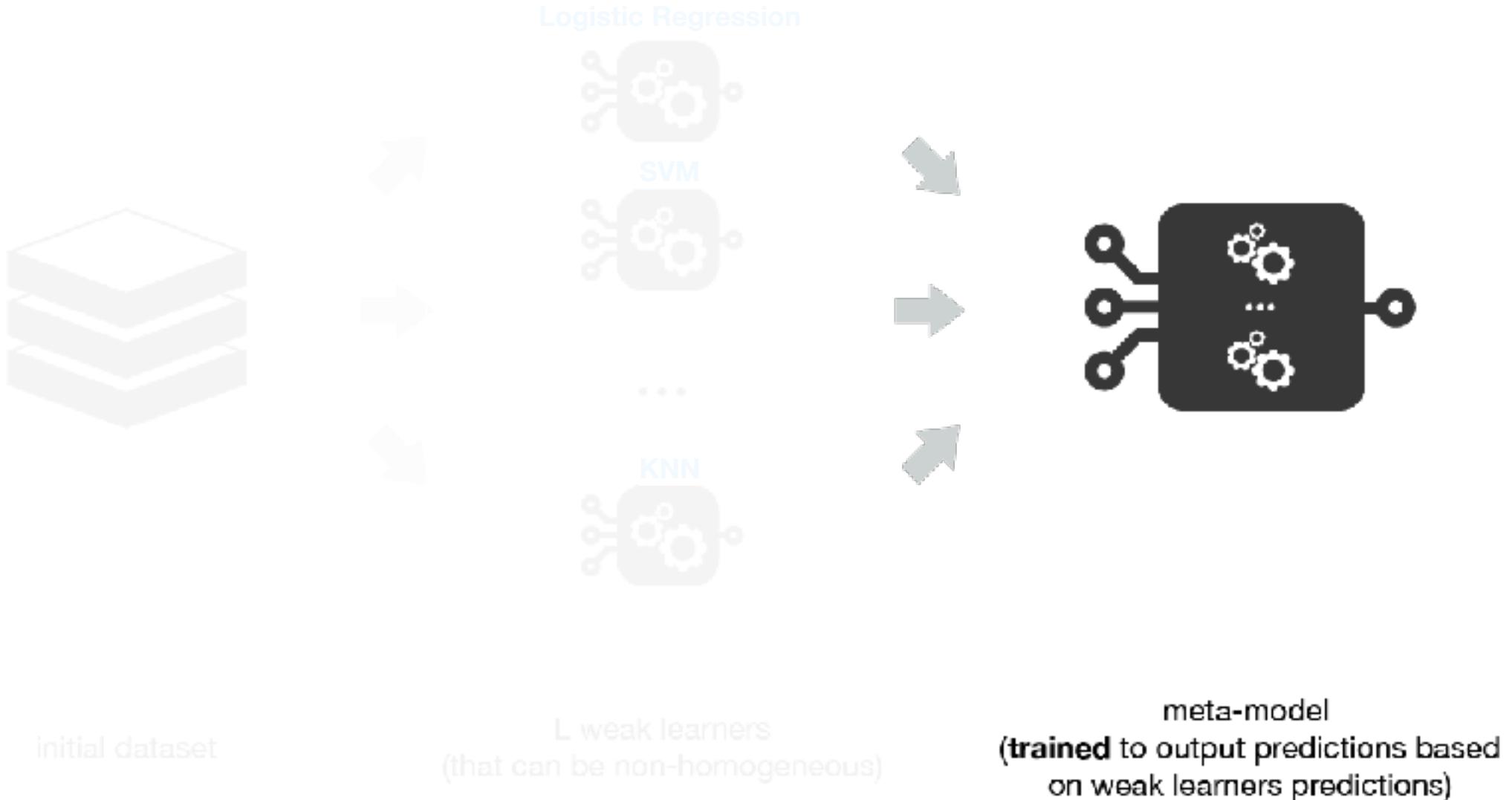


Stacking



meta-model
(trained to output predictions based
on weak learners predictions)

Stacking



Stacking

