Machine Learning Algorithms

Common ML Algorithms

Linear Regression

Logistic Regression

Support Vector Machine

Decision Tree

K-Nearest Neighbor

Neural Network

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

The Prediction

output $\hat{y}$ = $\hat{f}(X)$ 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

$$y = mx + b$$

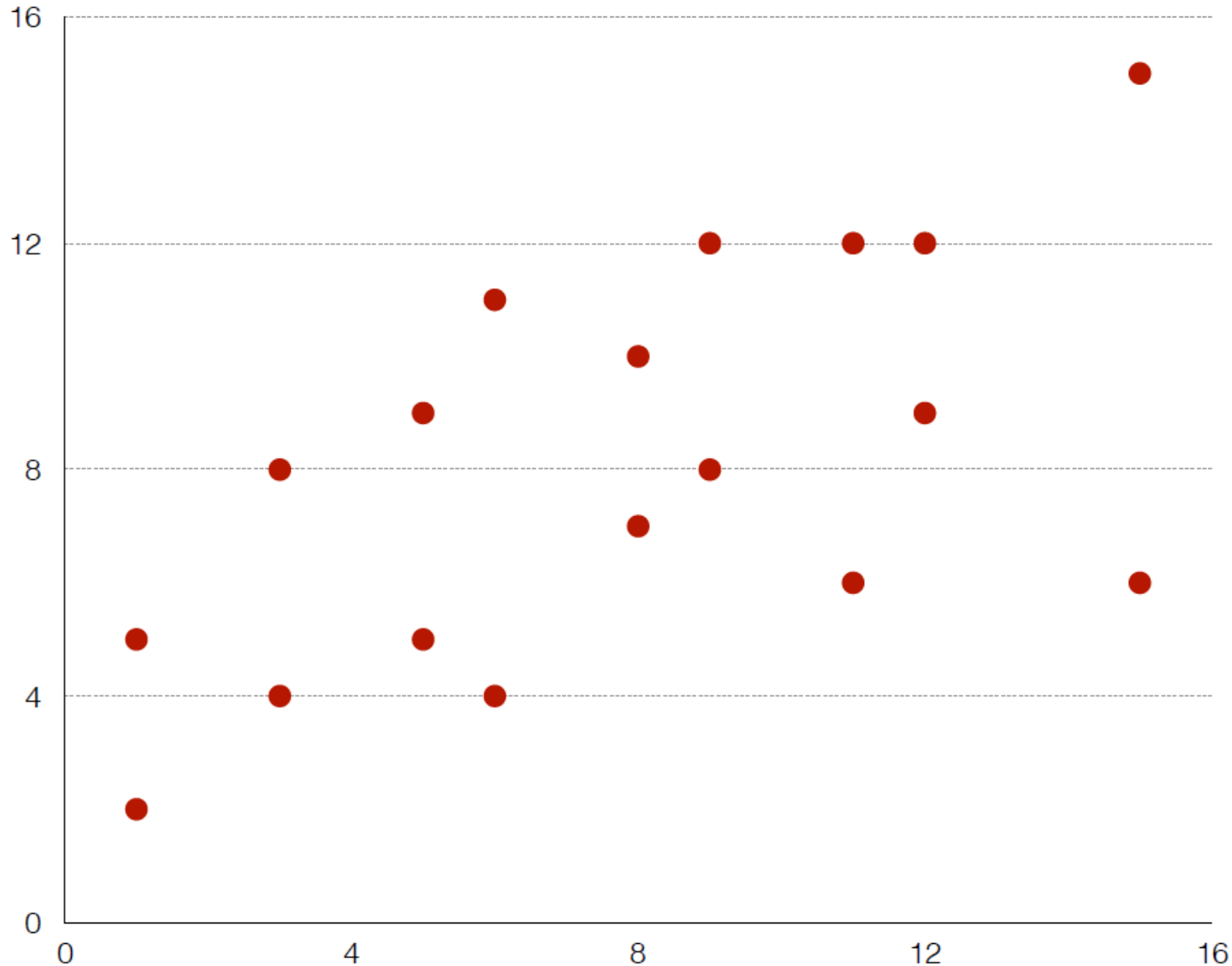
linear regression

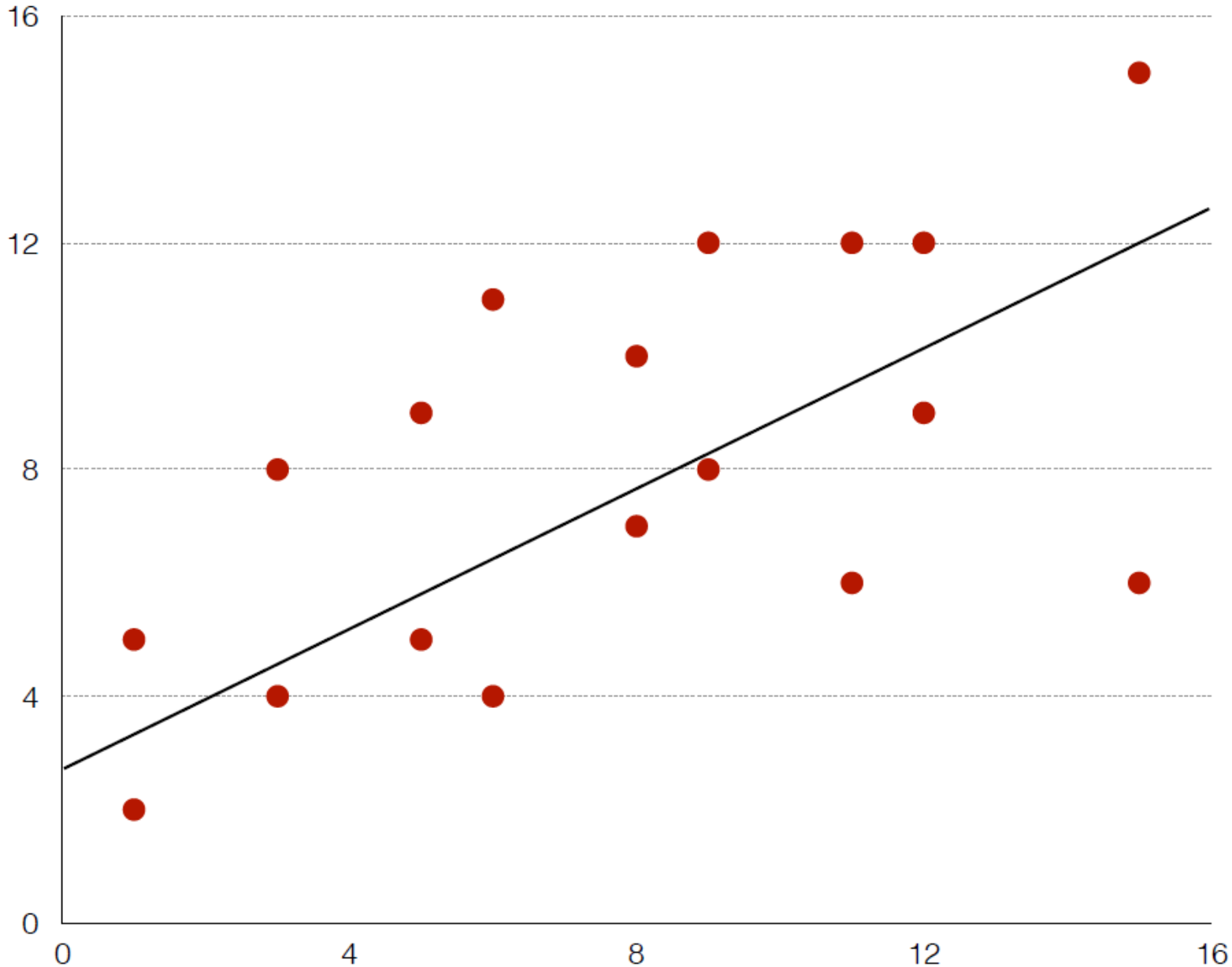
$$y = \beta_0 + \beta_1 x$$

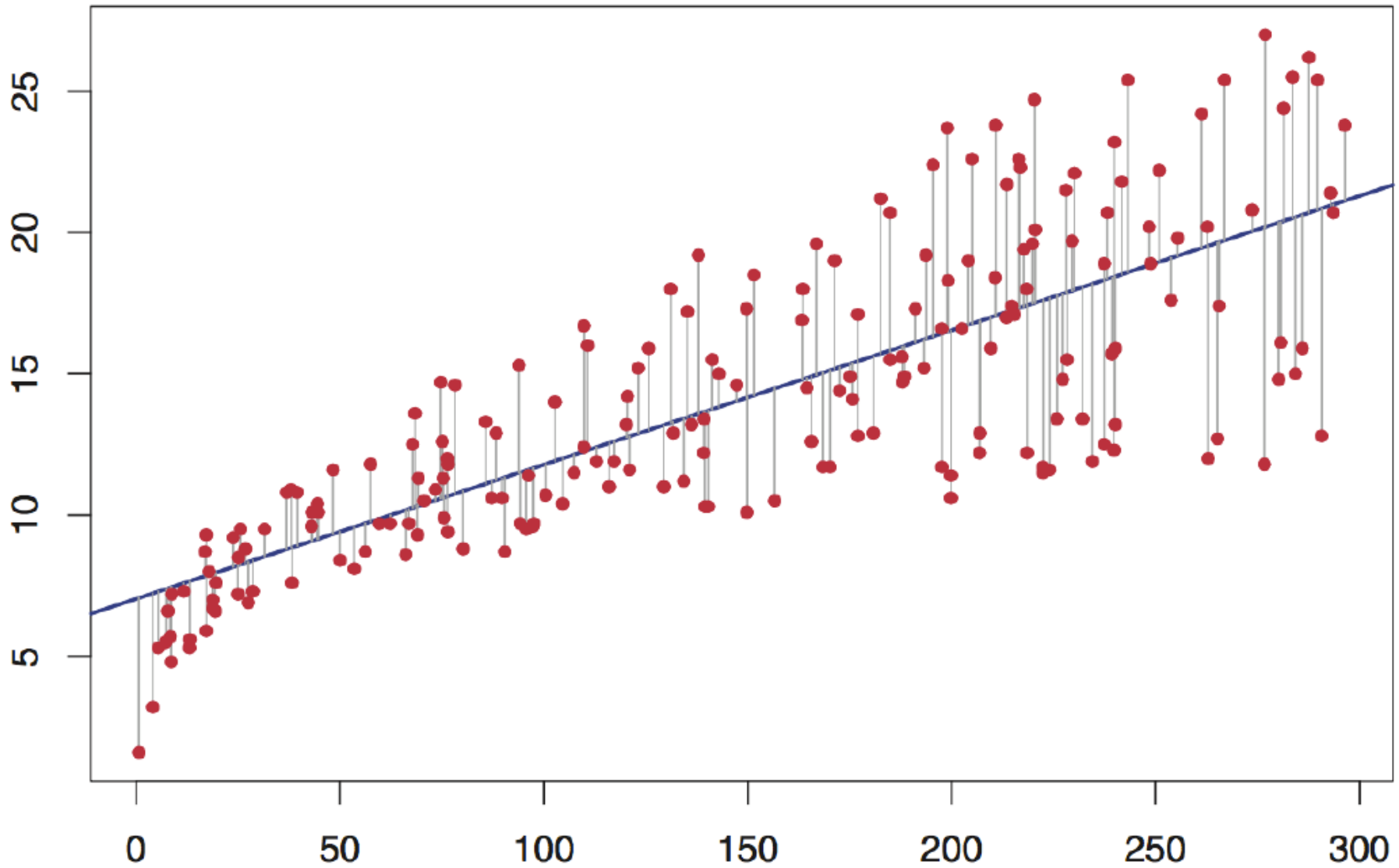
Simple Linear Regression

Input	learned coefficients (weights)	output
x_1	β_0, β_1	y

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


16
12
8
4
0
0
4
8
12
16


16
12
8
4
0
0
4
8
12
16


25
20
15
10
5
0
50
100
150
200
250
300

Multiple Linear Regression

$$\mathbf{LOC} = 227.63 + 9.51\mathbf{x_1} + 2.7\mathbf{x_2} - 7.08\mathbf{x_3}$$

$\mathbf{x_1}$ = hour of pair programming

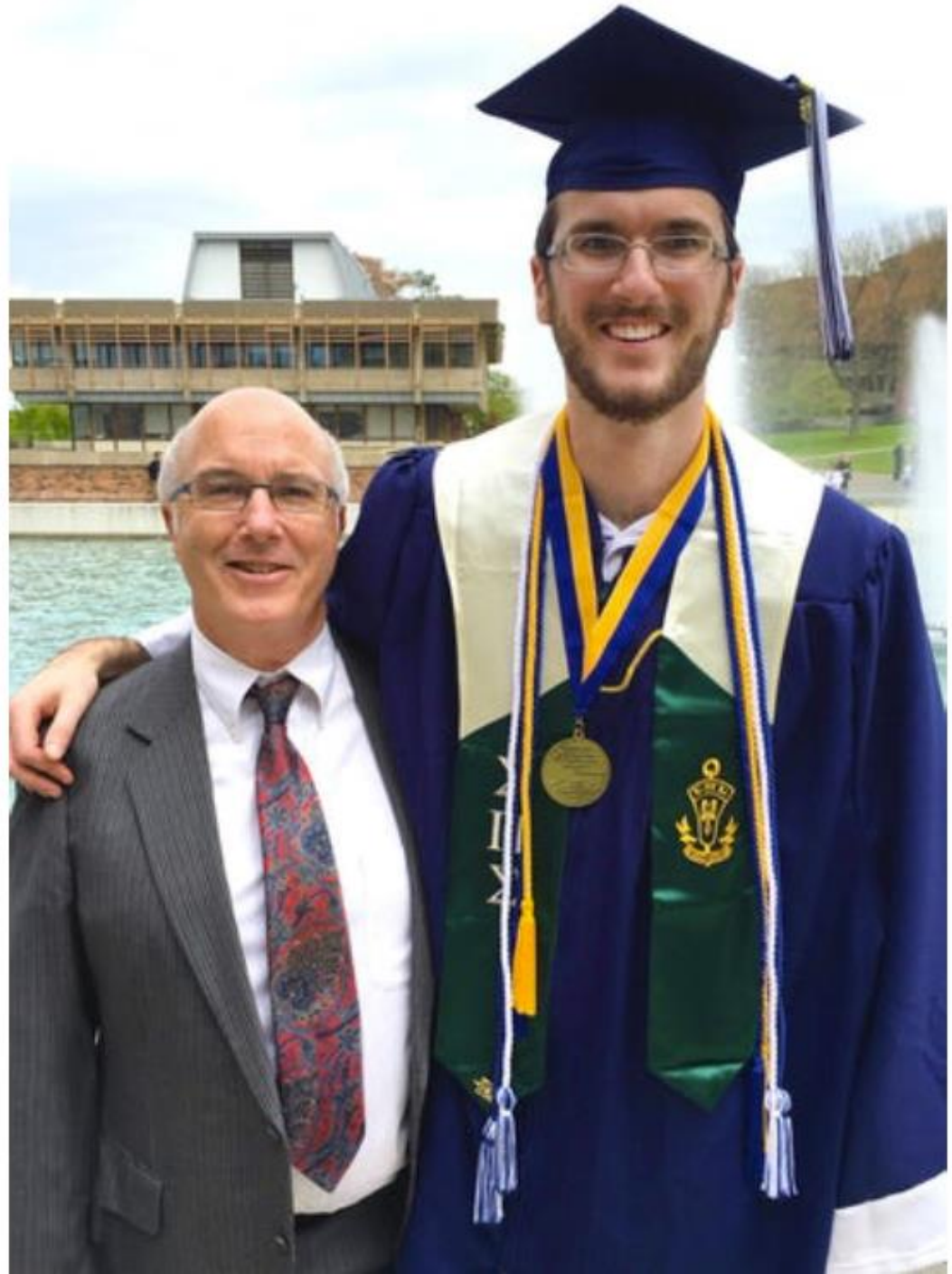
$\mathbf{x_2}$ = gender (m=0, female=1)

$\mathbf{x_3}$ = number of social accounts

Multiple Linear Regression

$$\textbf{Apps Sold} = 46.55 + 35.03x_1 + 7.11x_2 + 52.48x_3$$

x_1 = per $100 of advertising
x_2 = public talks
x_3 = targeted podcasts



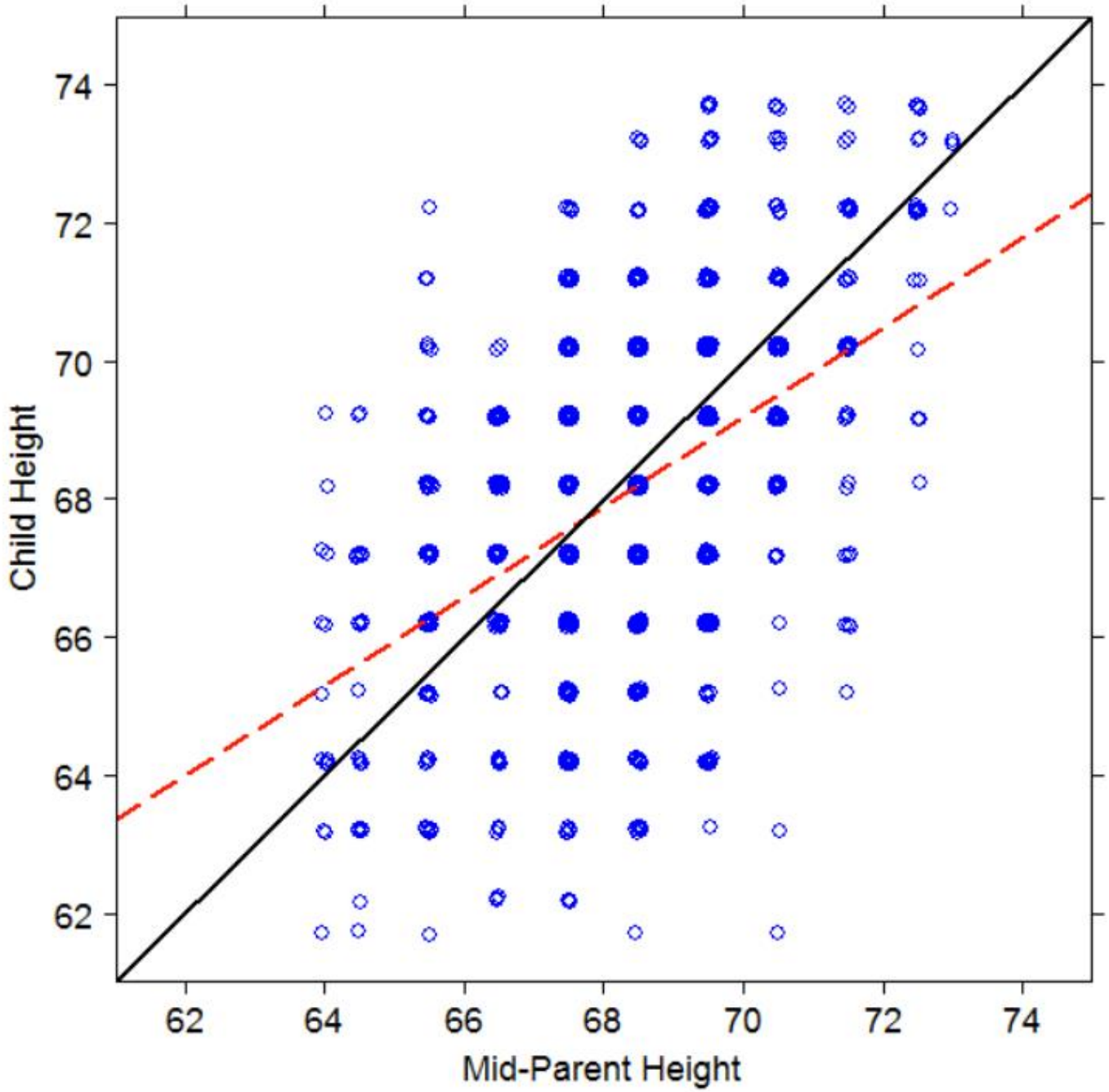
HEIGHT

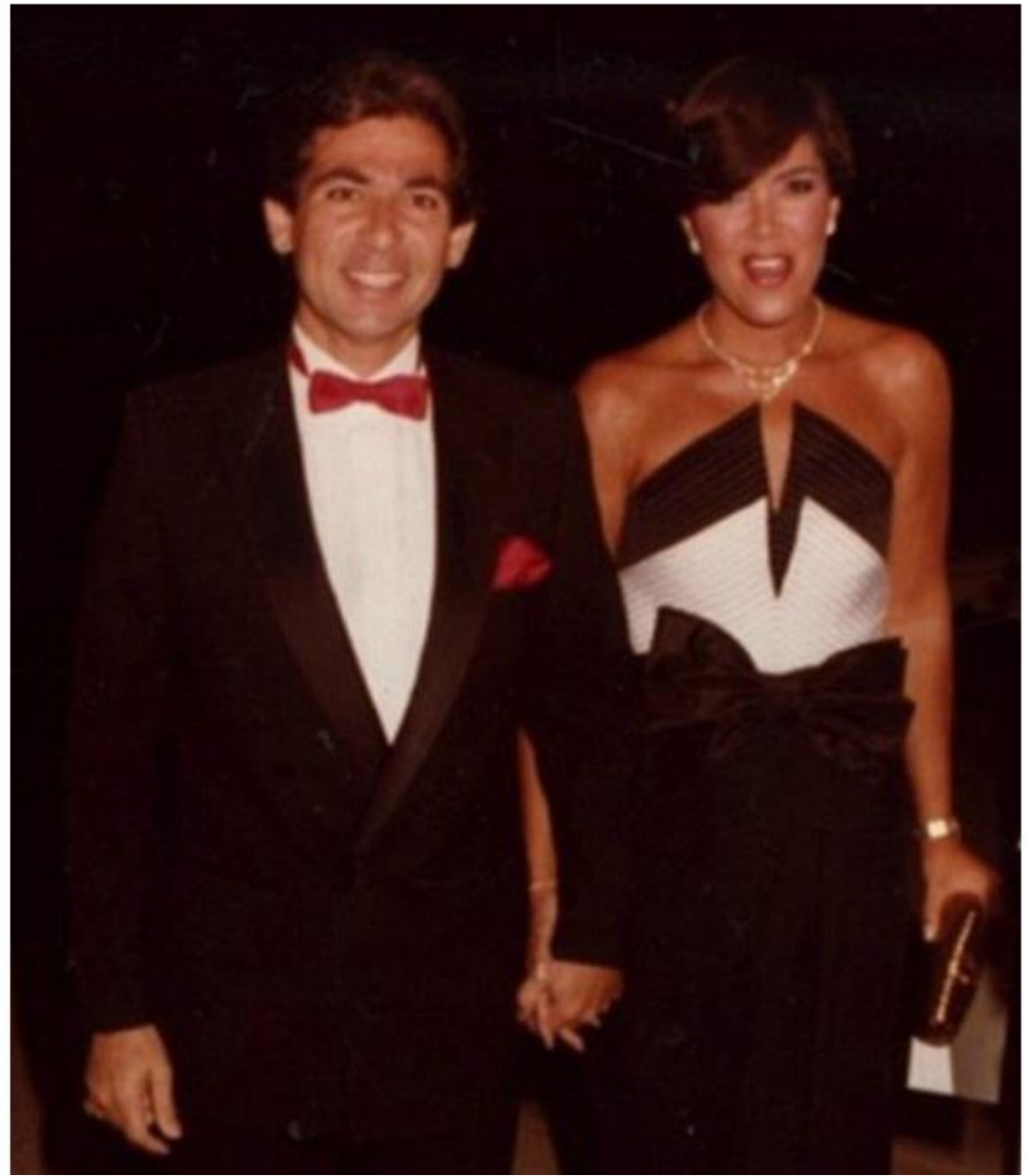
A MALE WILL GROW TO BE ABOUT THE HEIGHT OF

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

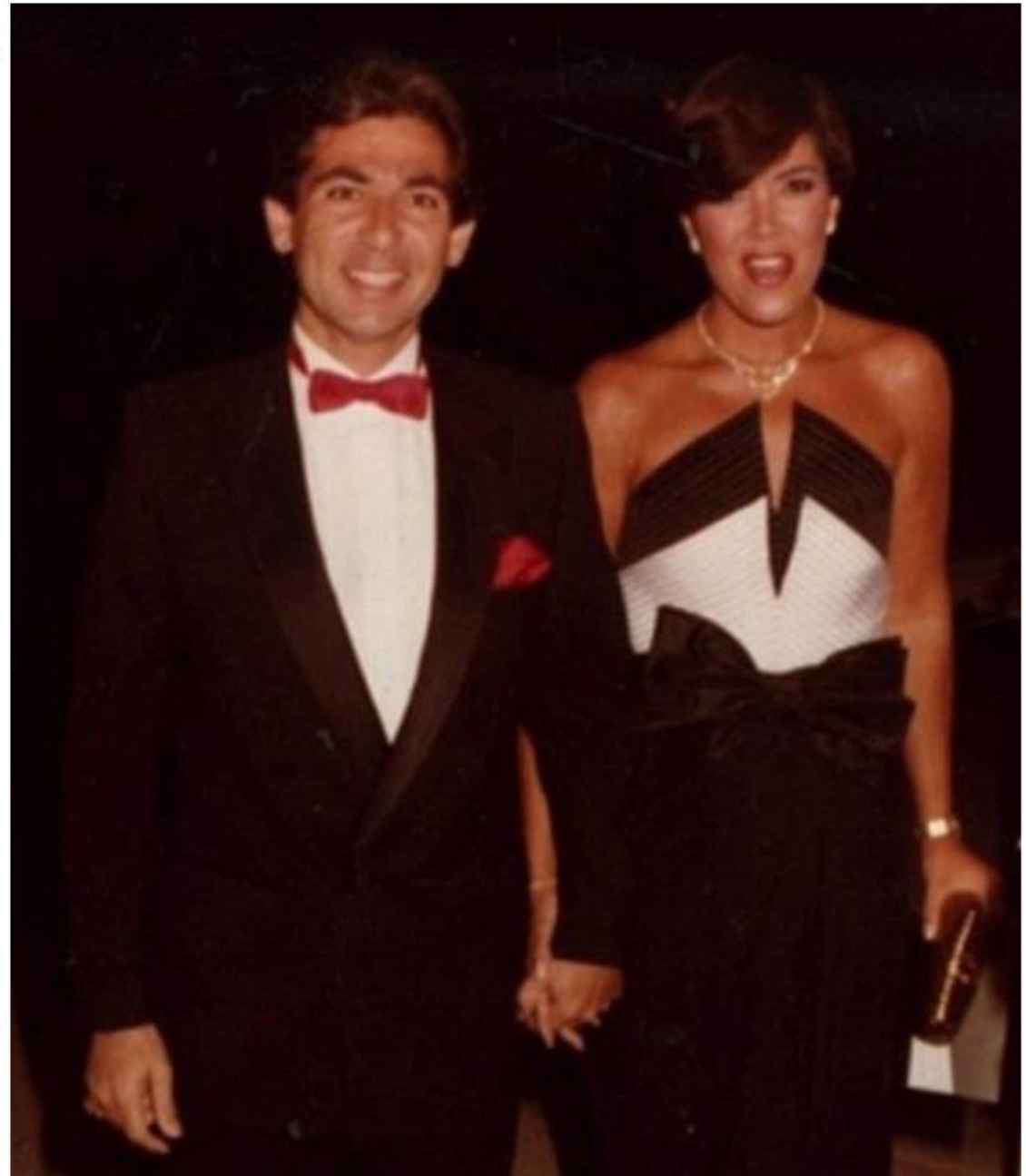
A FEMALE WILL GROW TO BE ABOUT THE HEIGHT OF

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

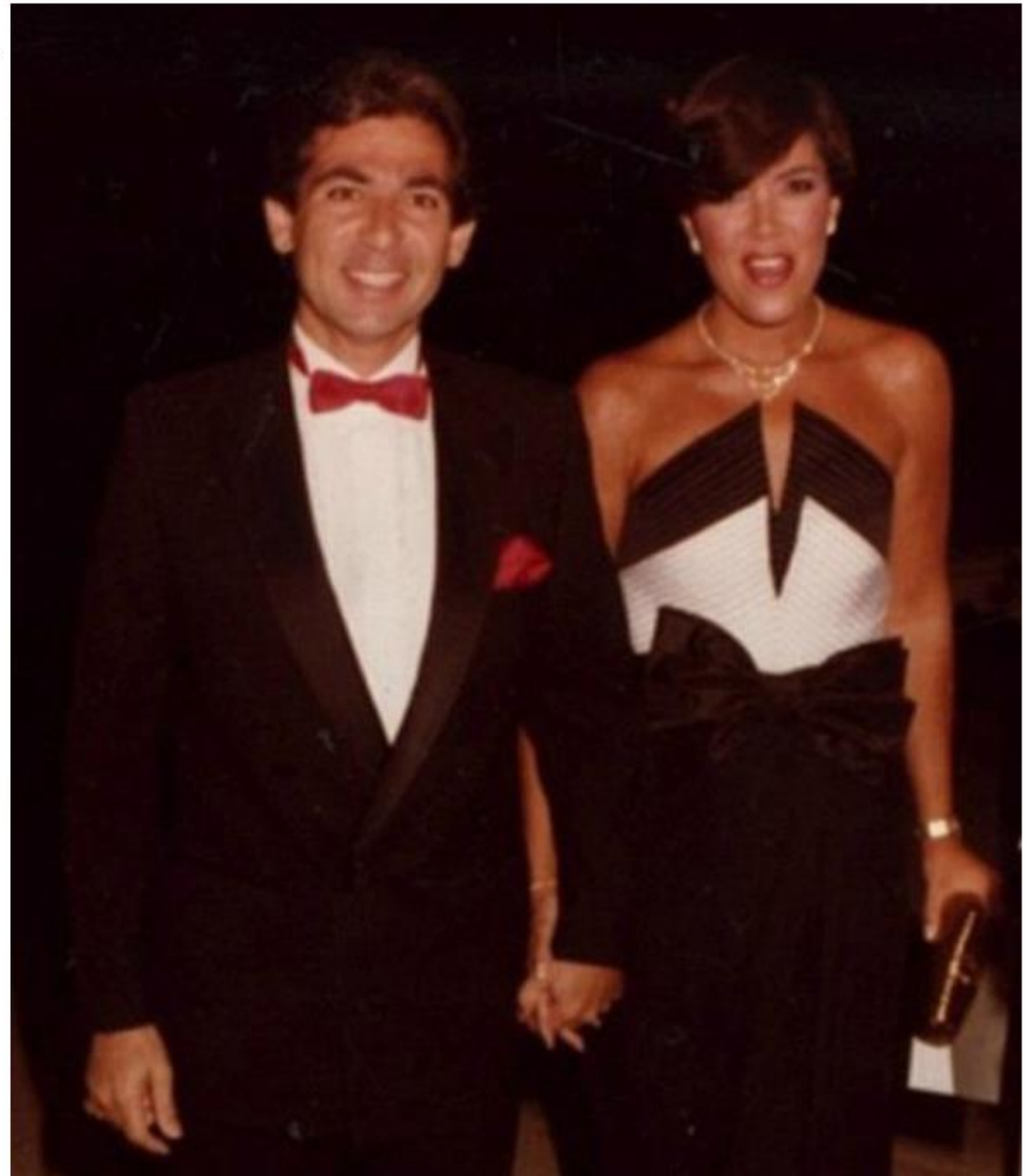

Child Height
Mid-Parent Height
62
64
66
68
70
72
74



5’7



5’7

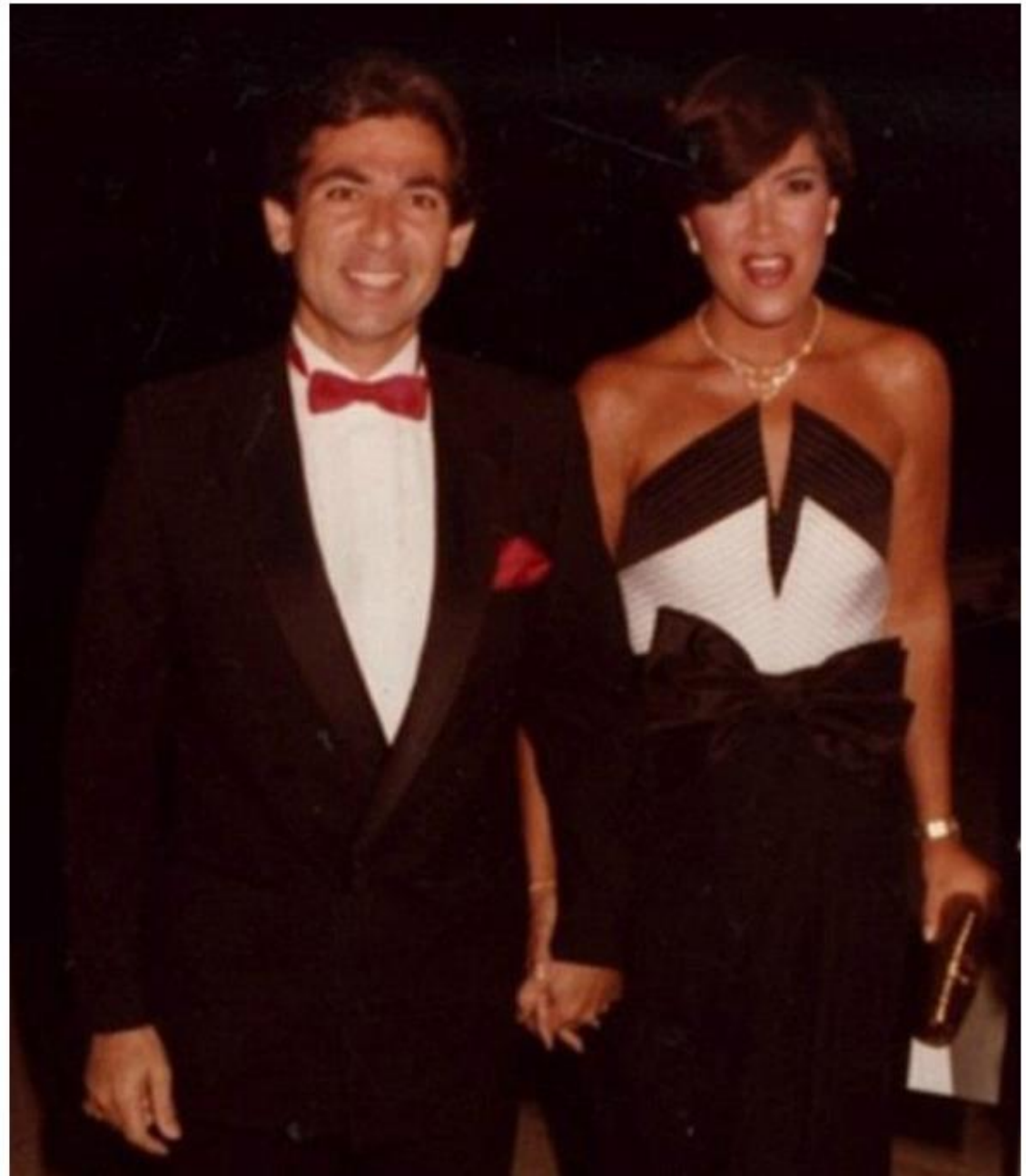


5’6

Mid-parent Height

5’6.5

5’7



5’6

5’0

5’3

5’10

5’0

5’3

5’10

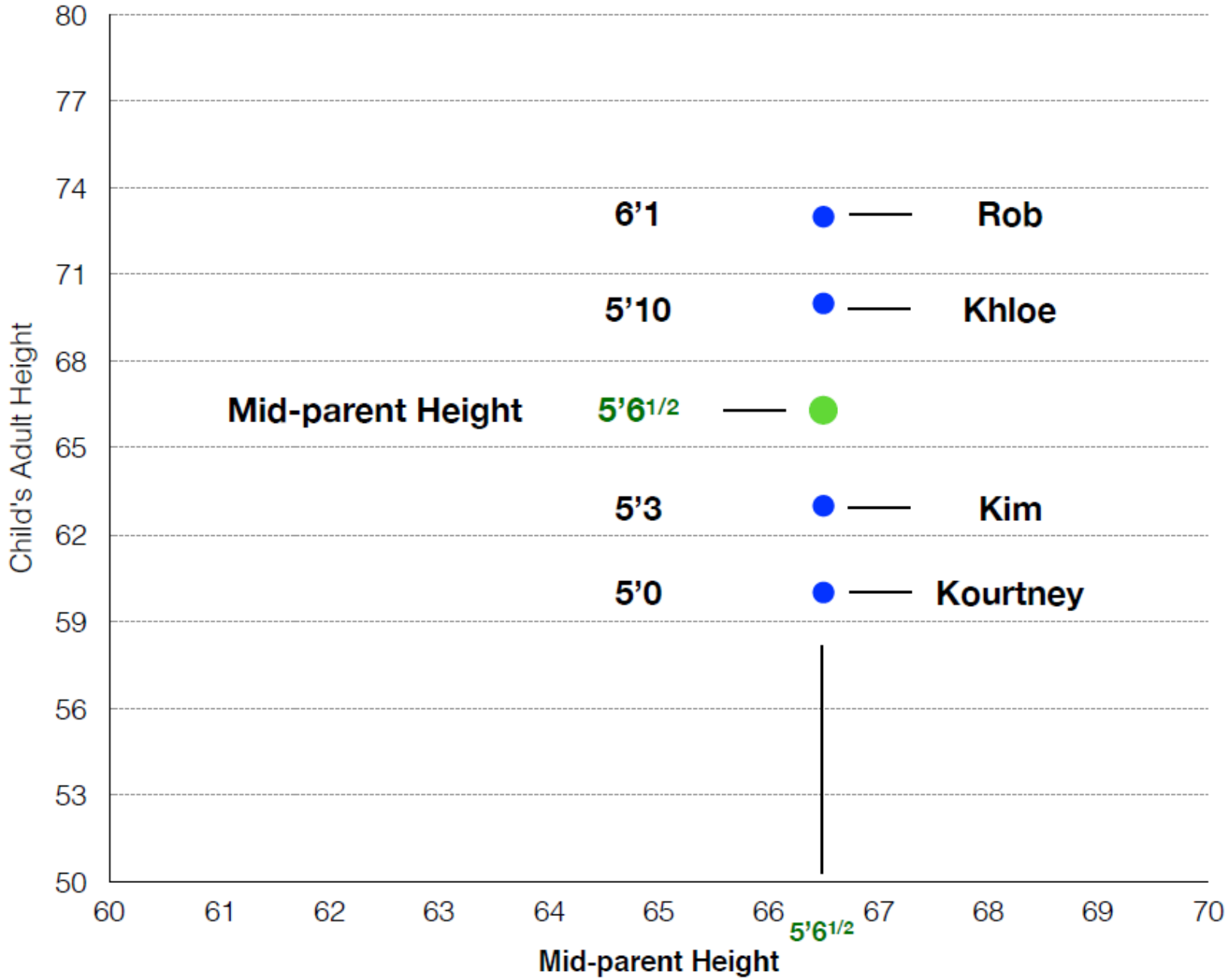
Average

5’4 1/3


6’1
5’3
5’0
5’10
KARDASHIAN KHAOS
LAS VEGAS
Mirage
Vegas Starts Here

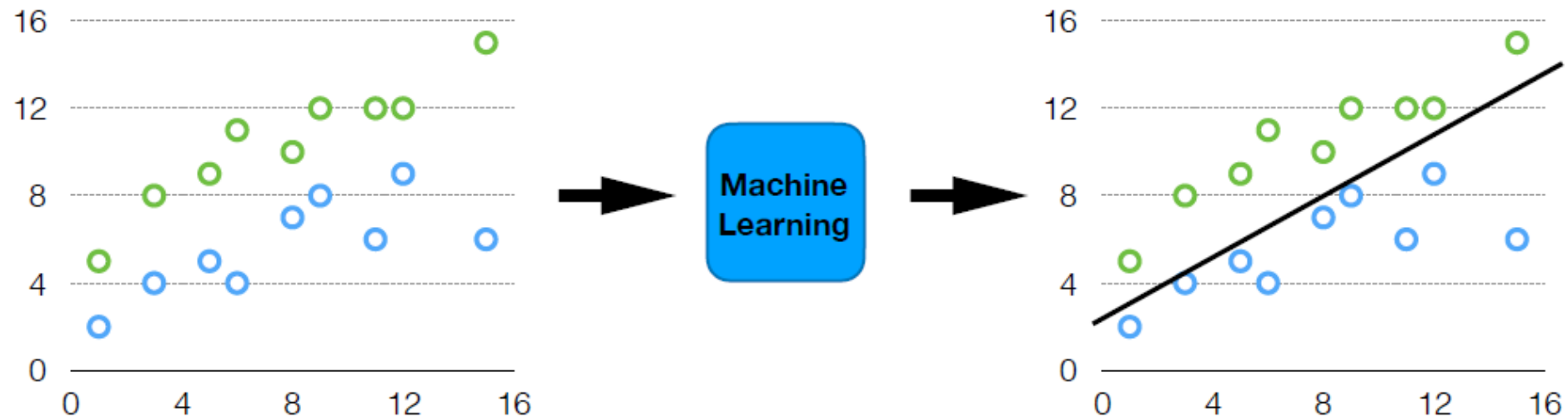
Mid-parent Height

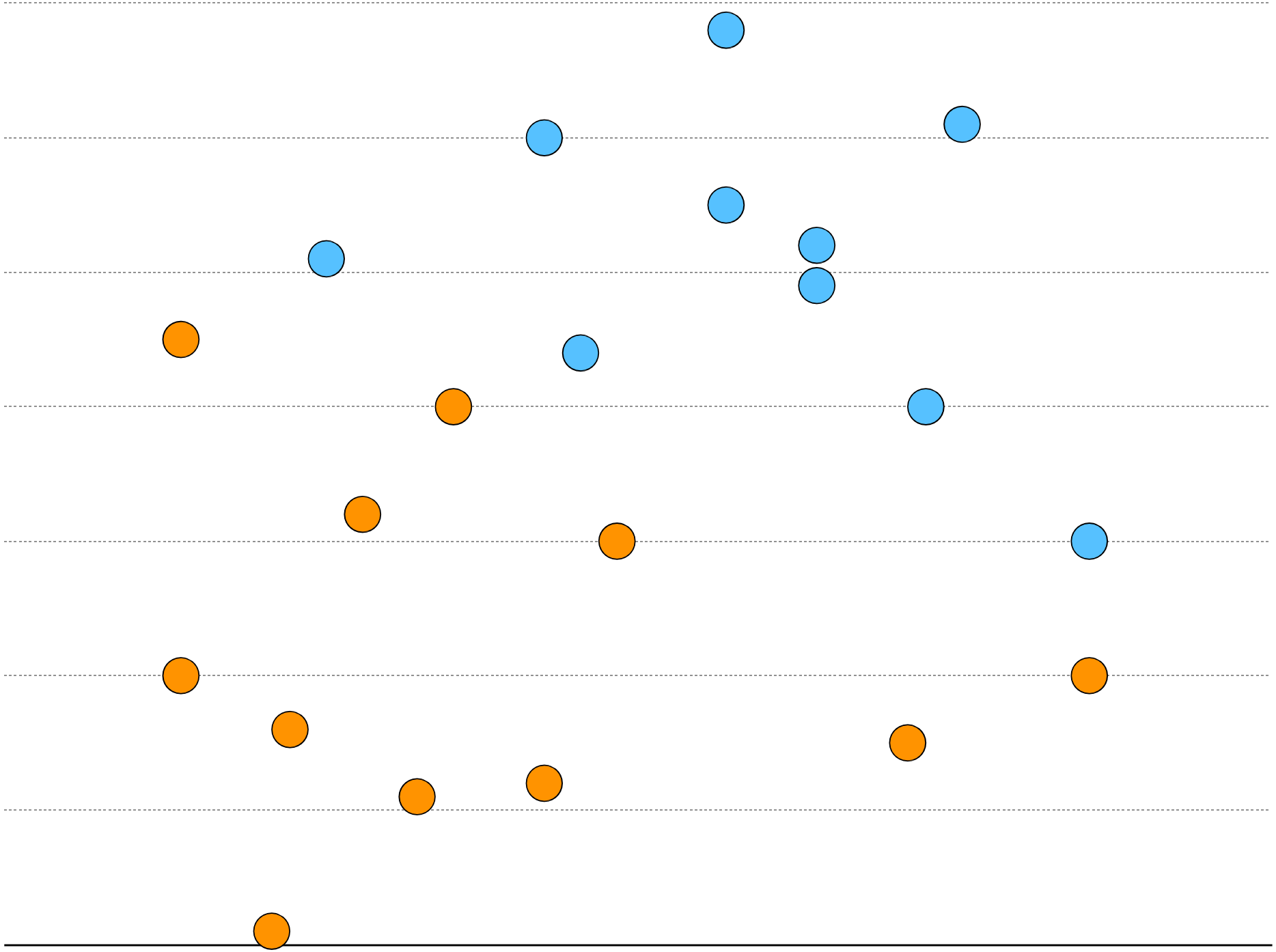
5’6.5

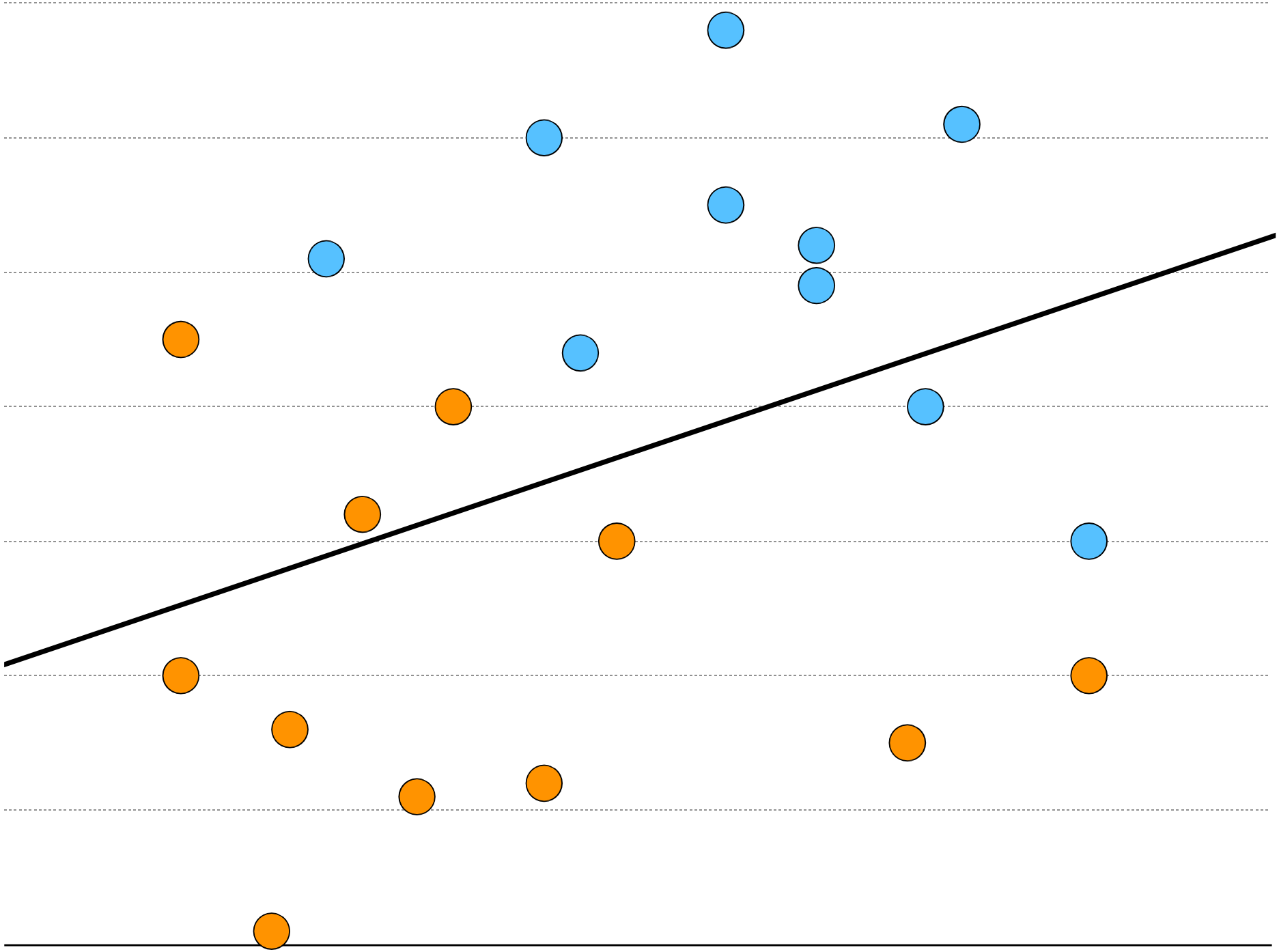

80
77
74
71
68
65
62
59
56
53
50
Child's Adult Height
6'1
Rob
5'10
Khloe
Mid-parent Height
5'6 1/2
5'3
Kim
5'0
Kourtney
60
61
62
63
64
65
66
67
68
69
70
5'6 1/2
Mid-parent Height

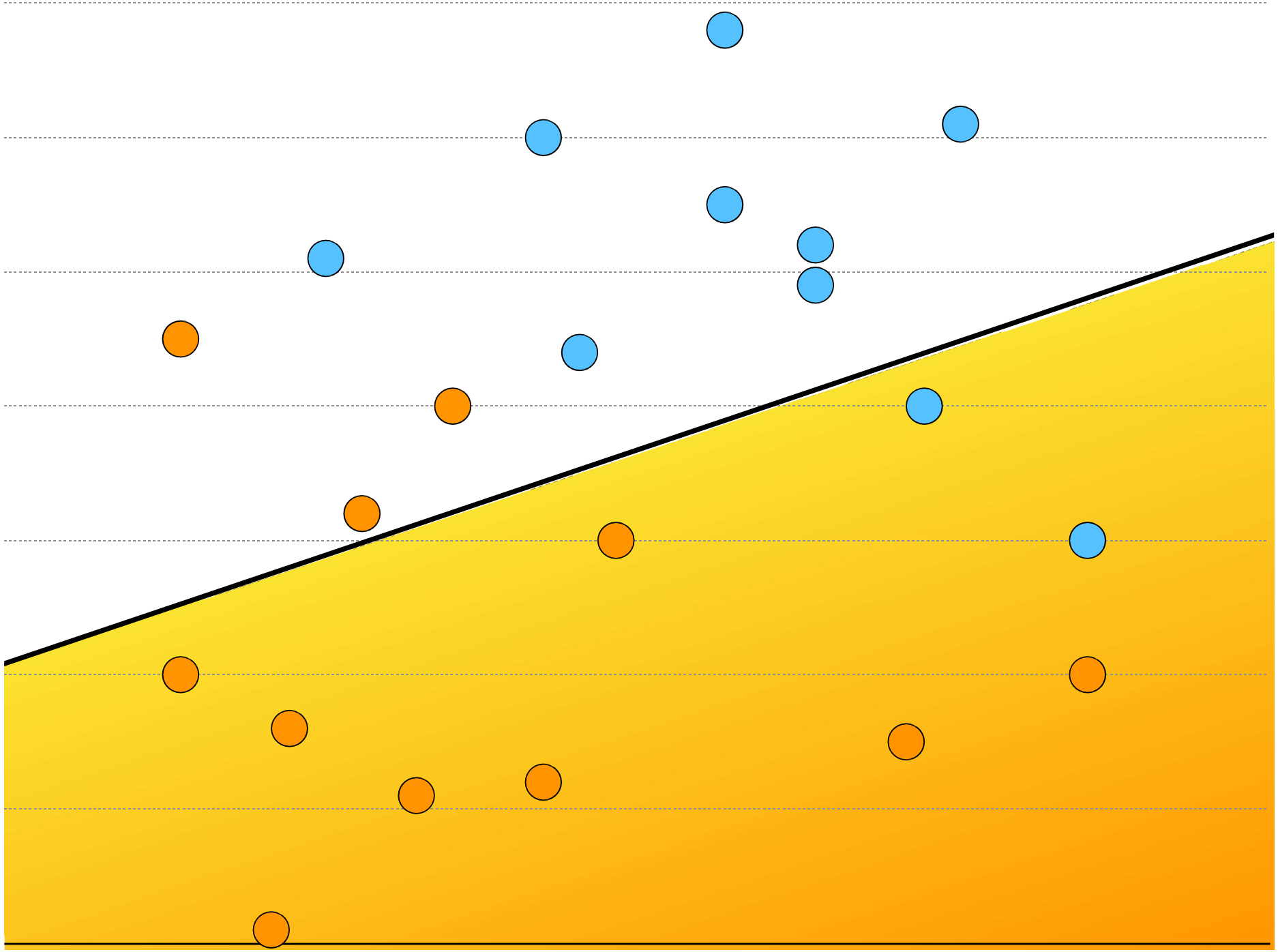
Logistic Regression

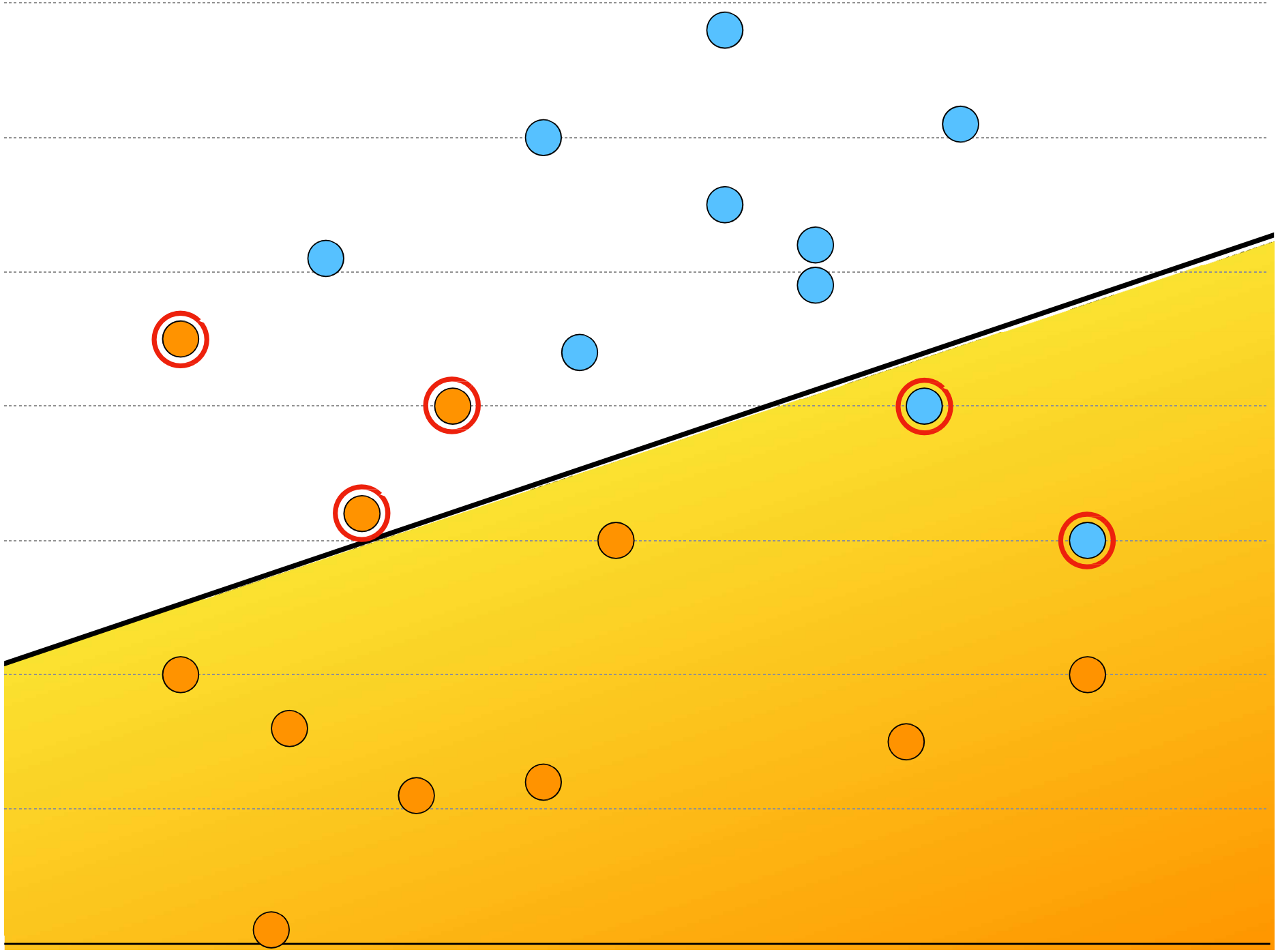
What is a classifier?

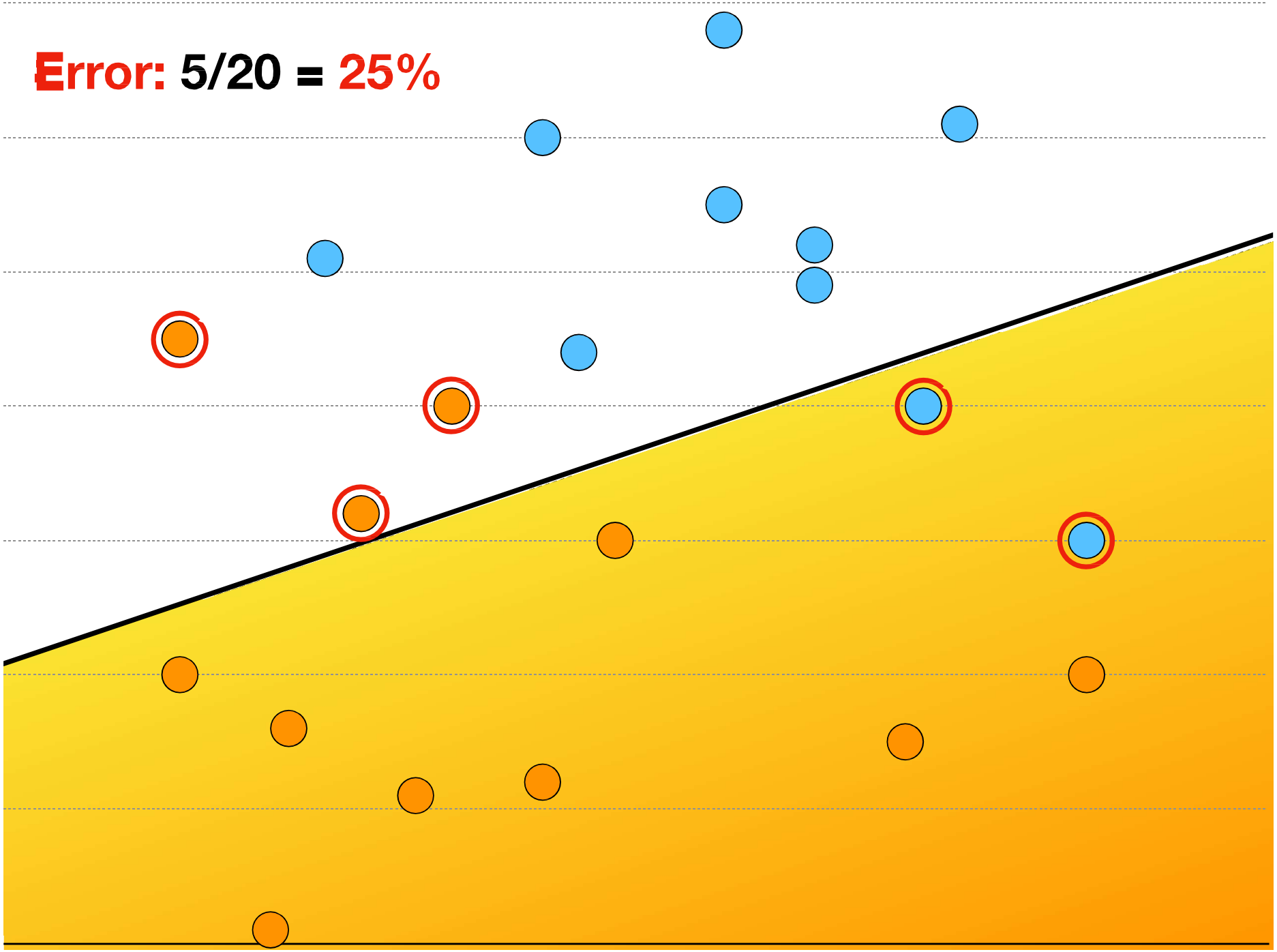


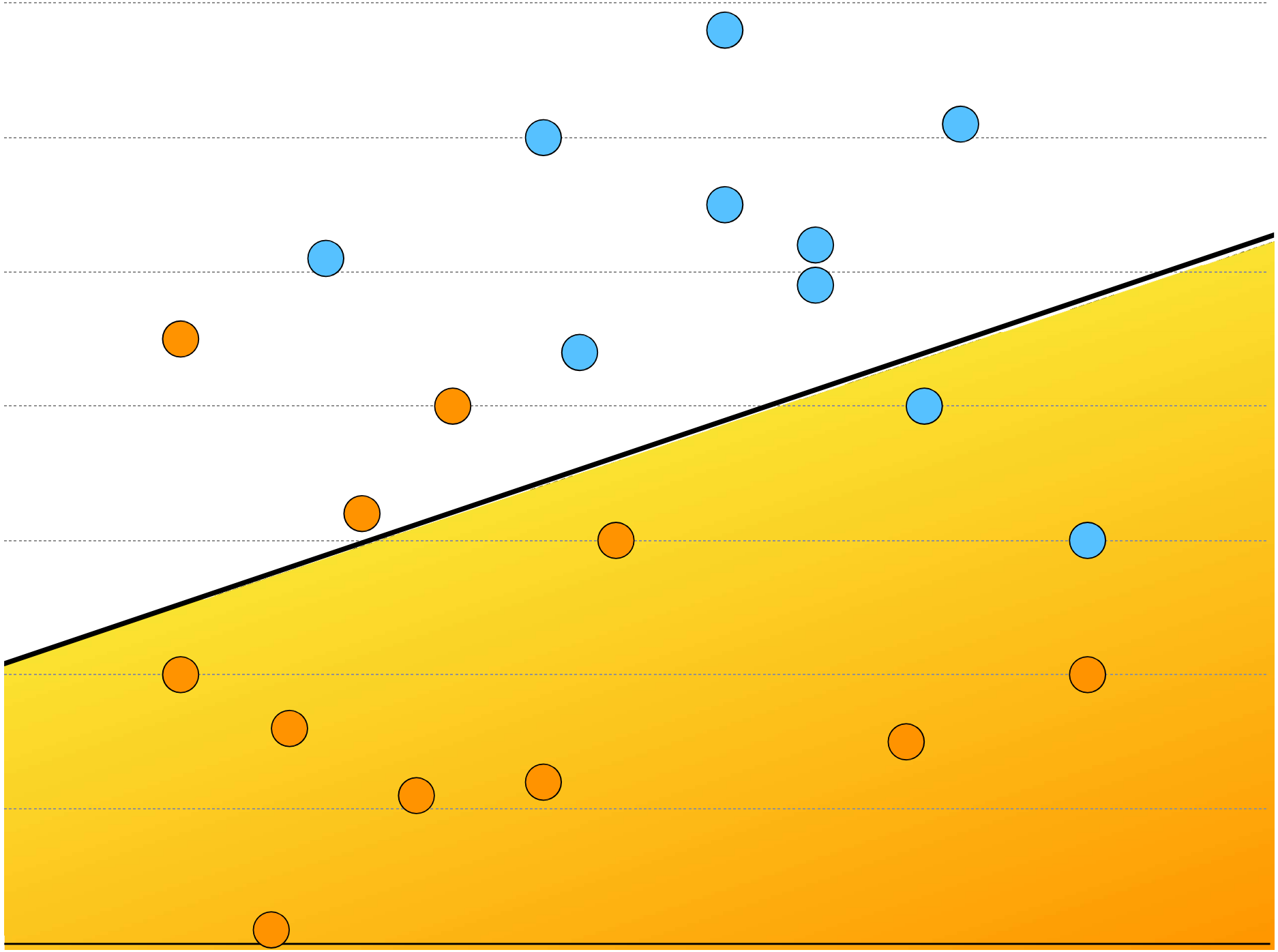


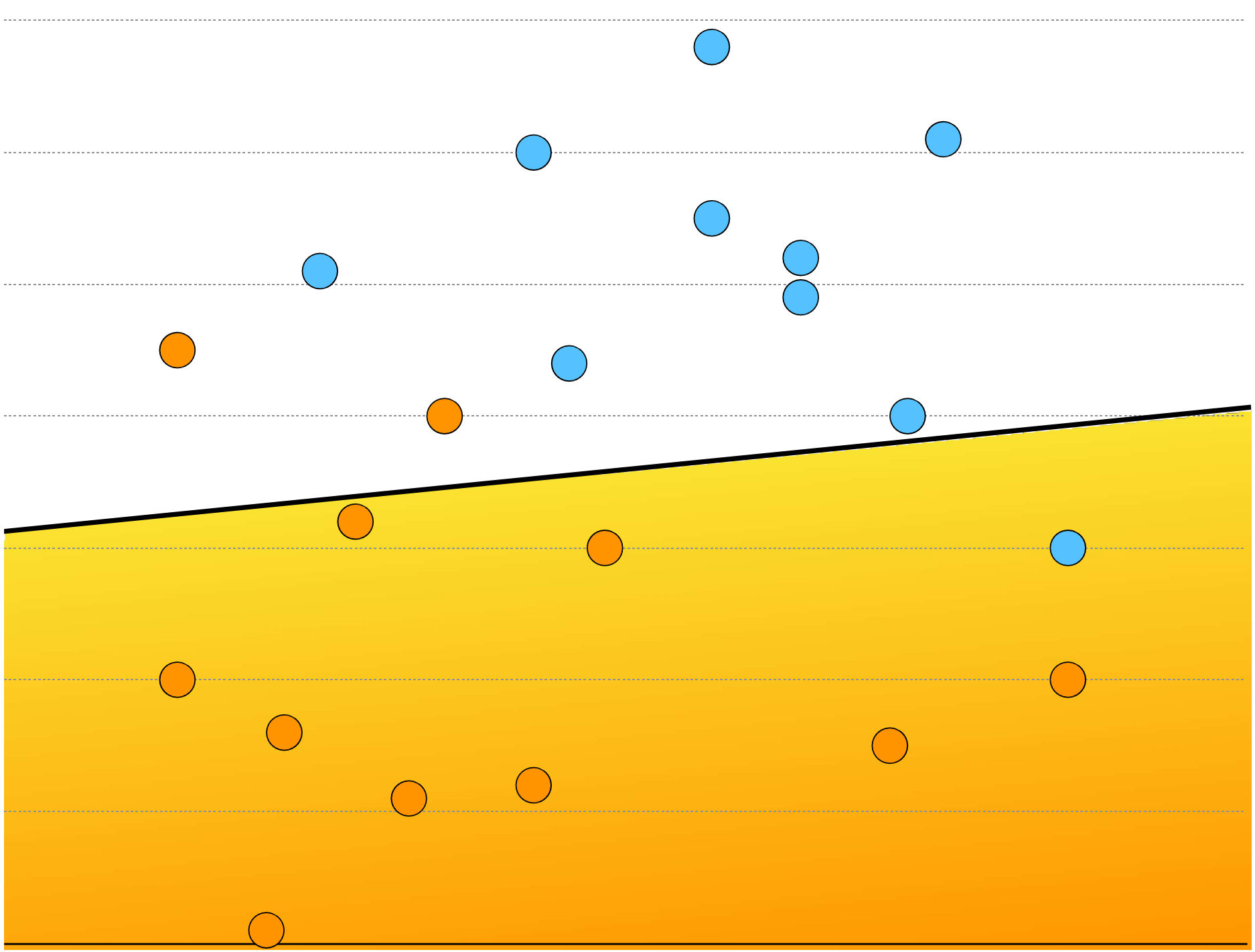


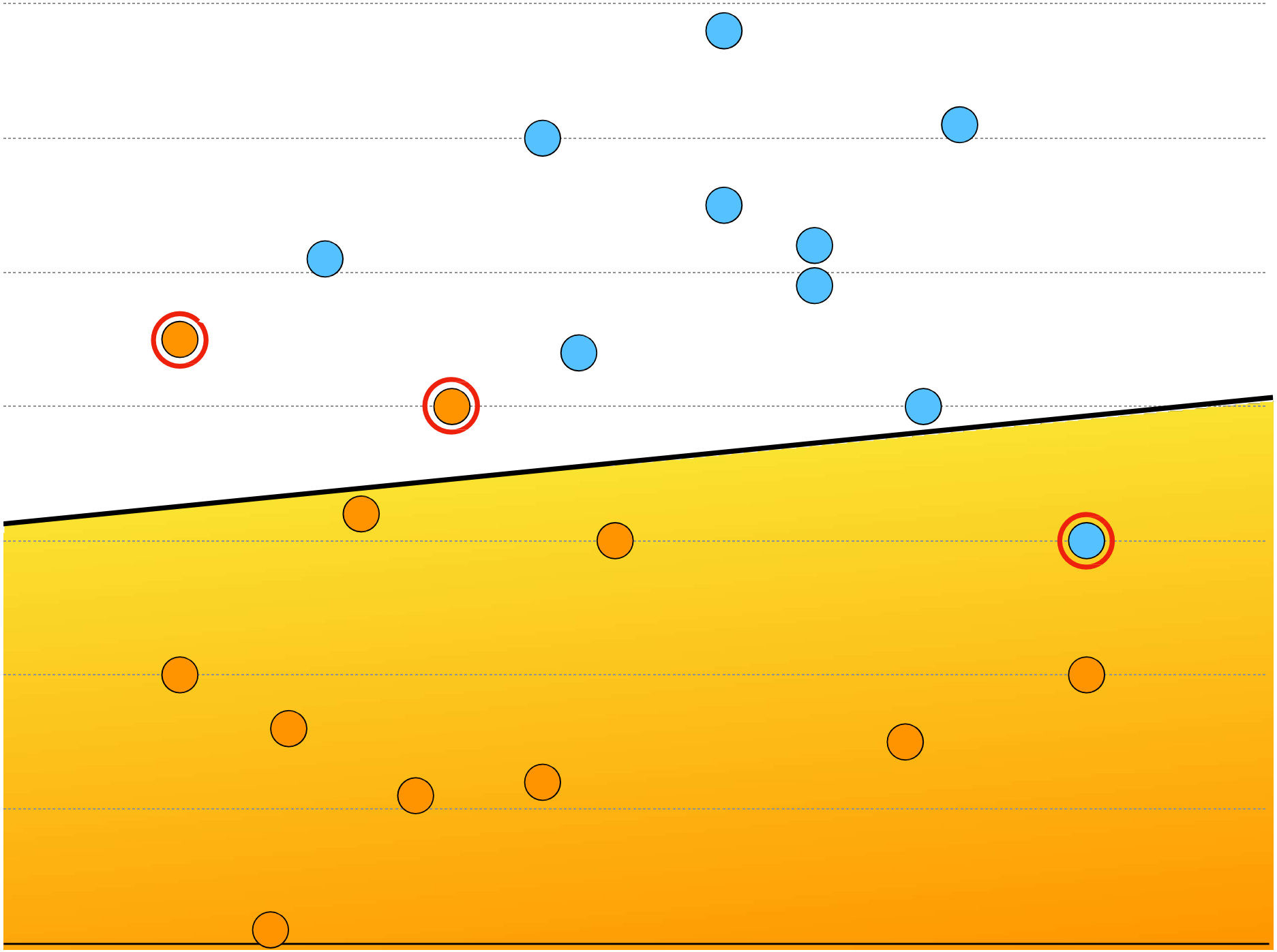


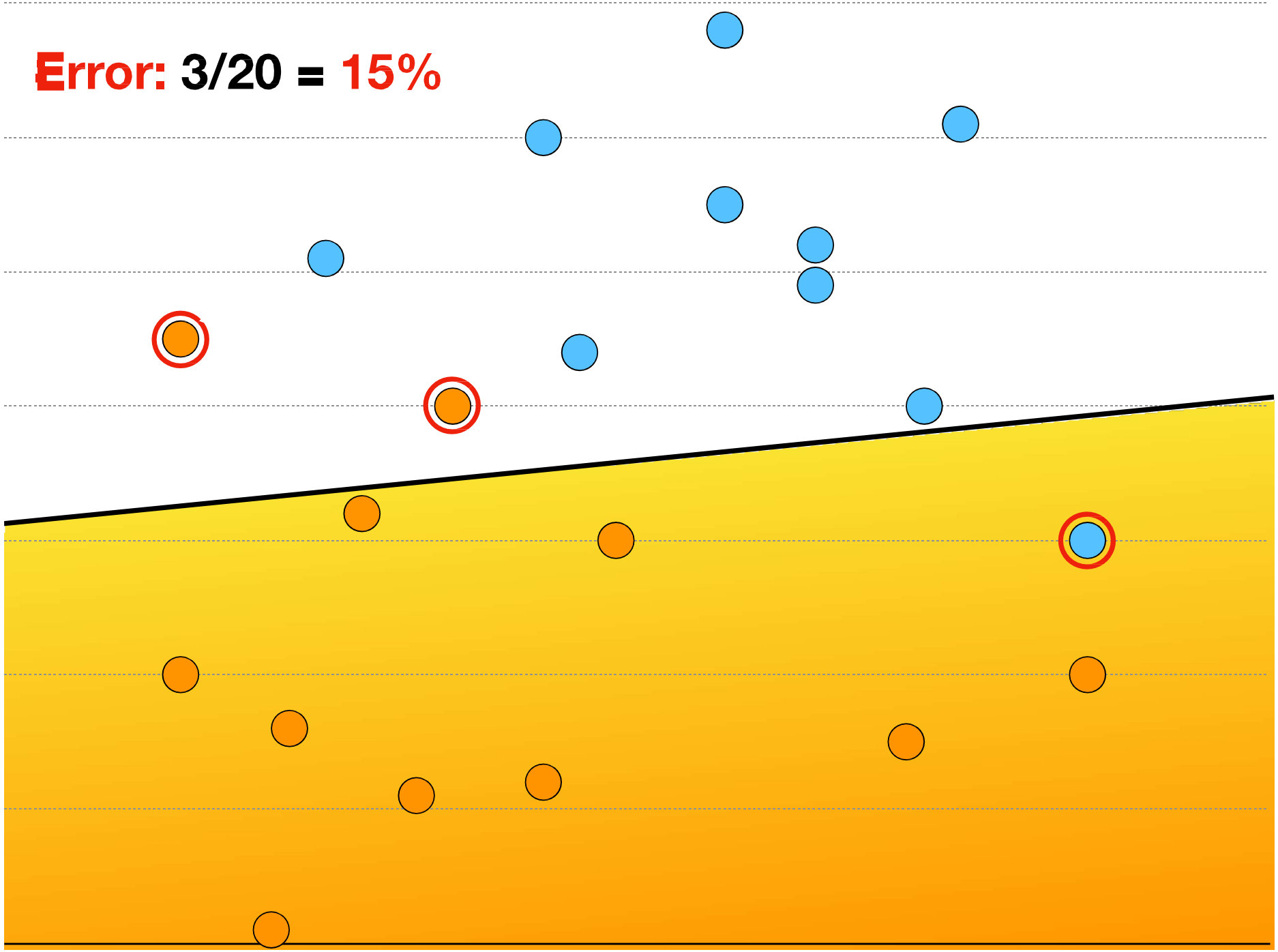


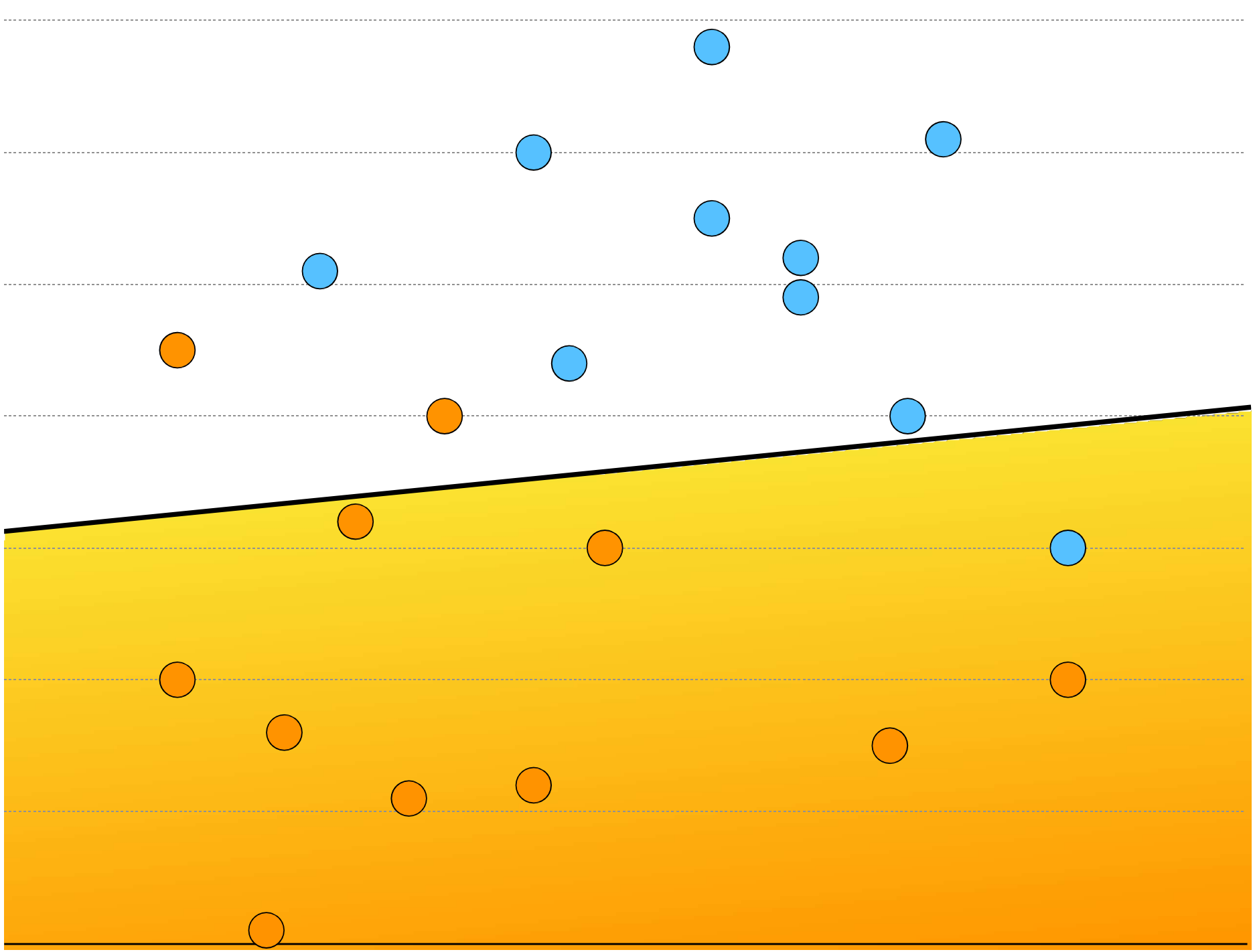

Error: 5/20 = 25%

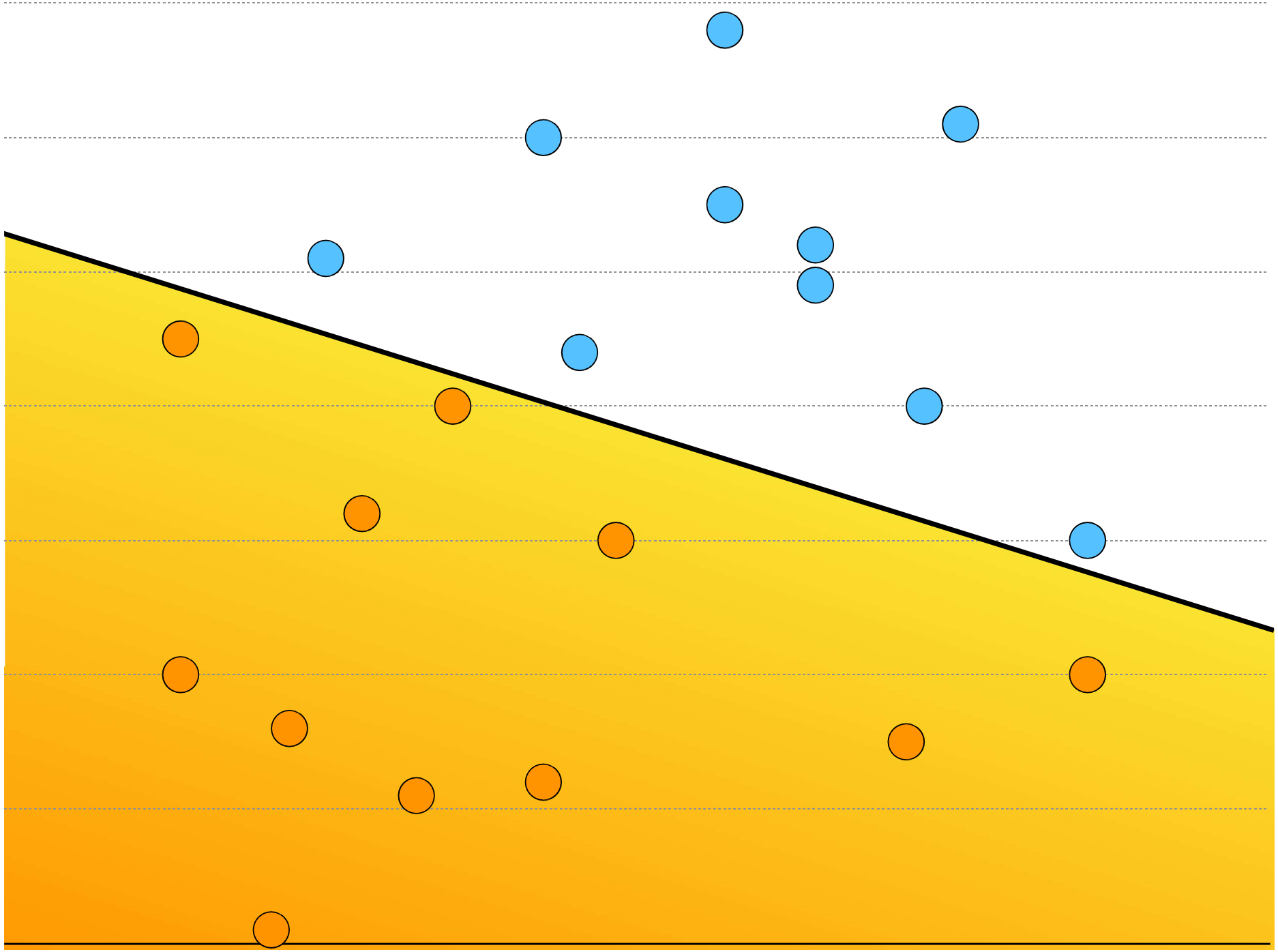


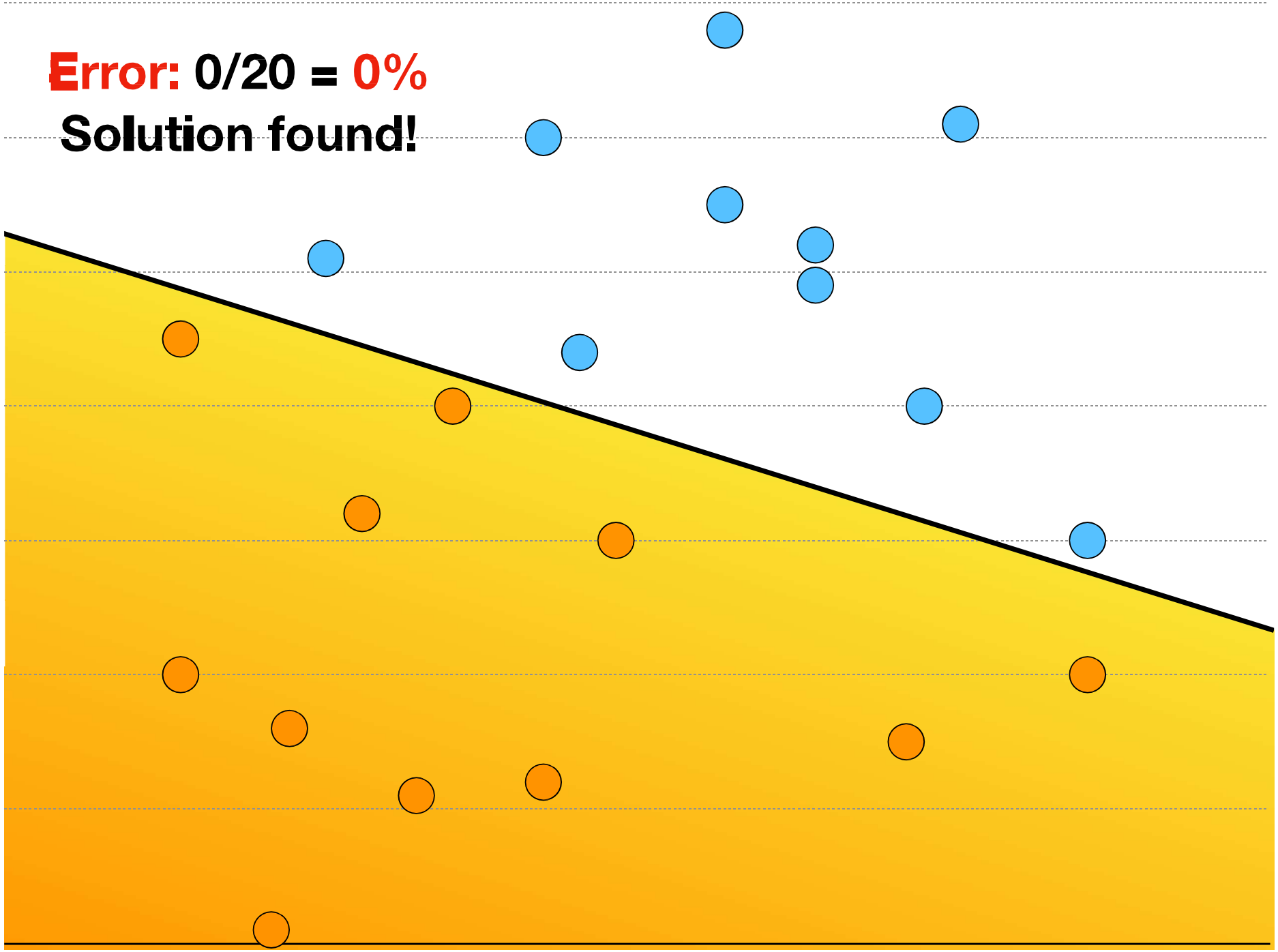





Error: 3/20 = 15%






Error: 0/20 = 0%
Solution found!

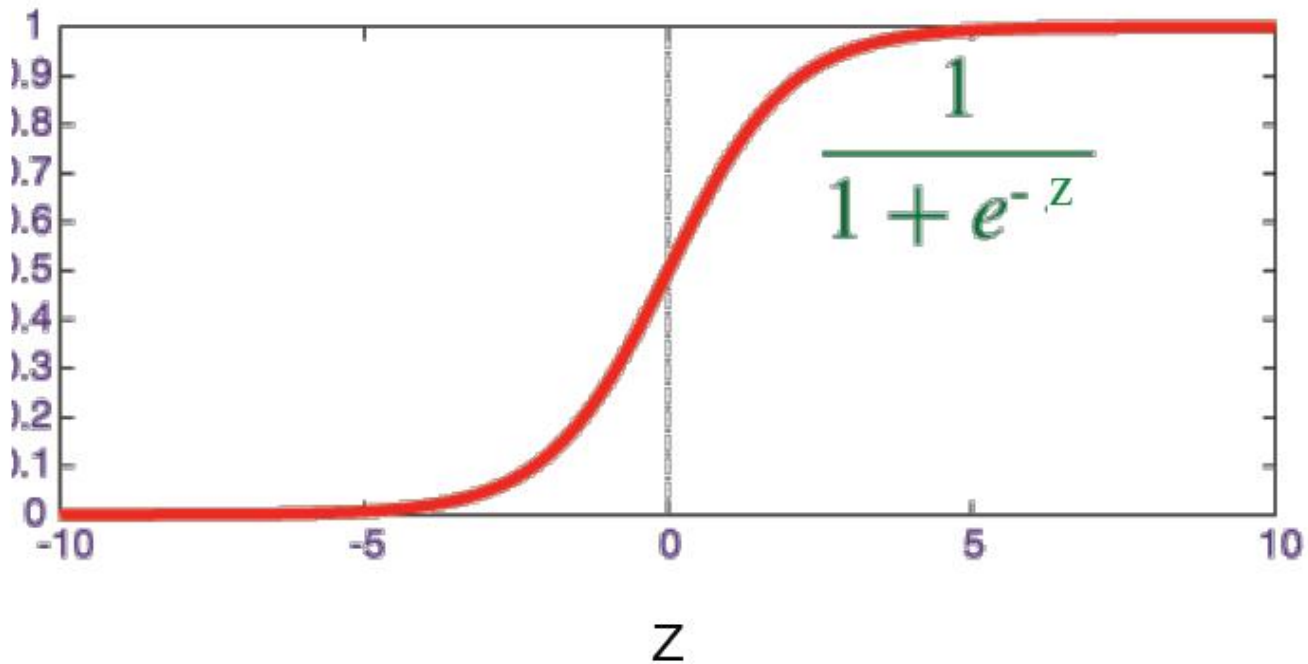
Logistic Regression

Logistic Regression

intermediate step

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

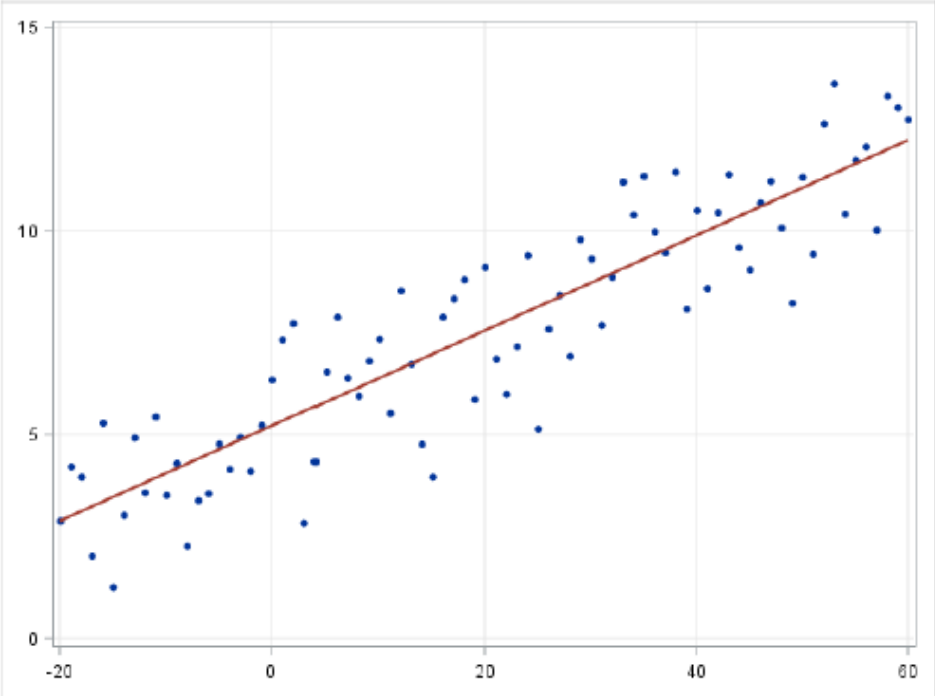
Logistic (Sigmoid) Function



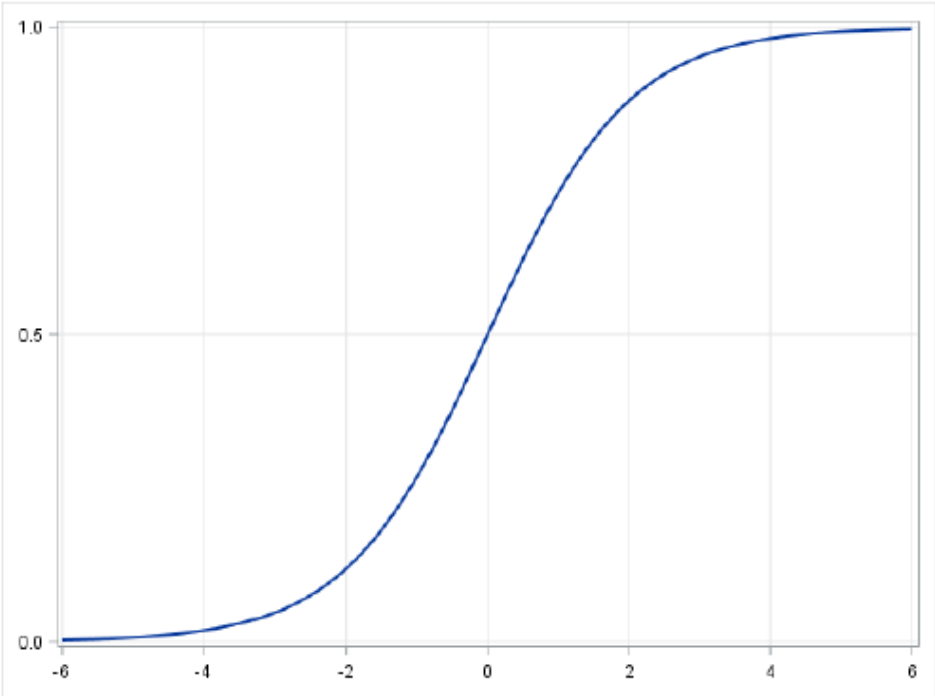
Logistic Regression

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

Linear Regression



Logistic Regression



positive class = 1

negative class = 0

predict 1 when $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 >= 0$

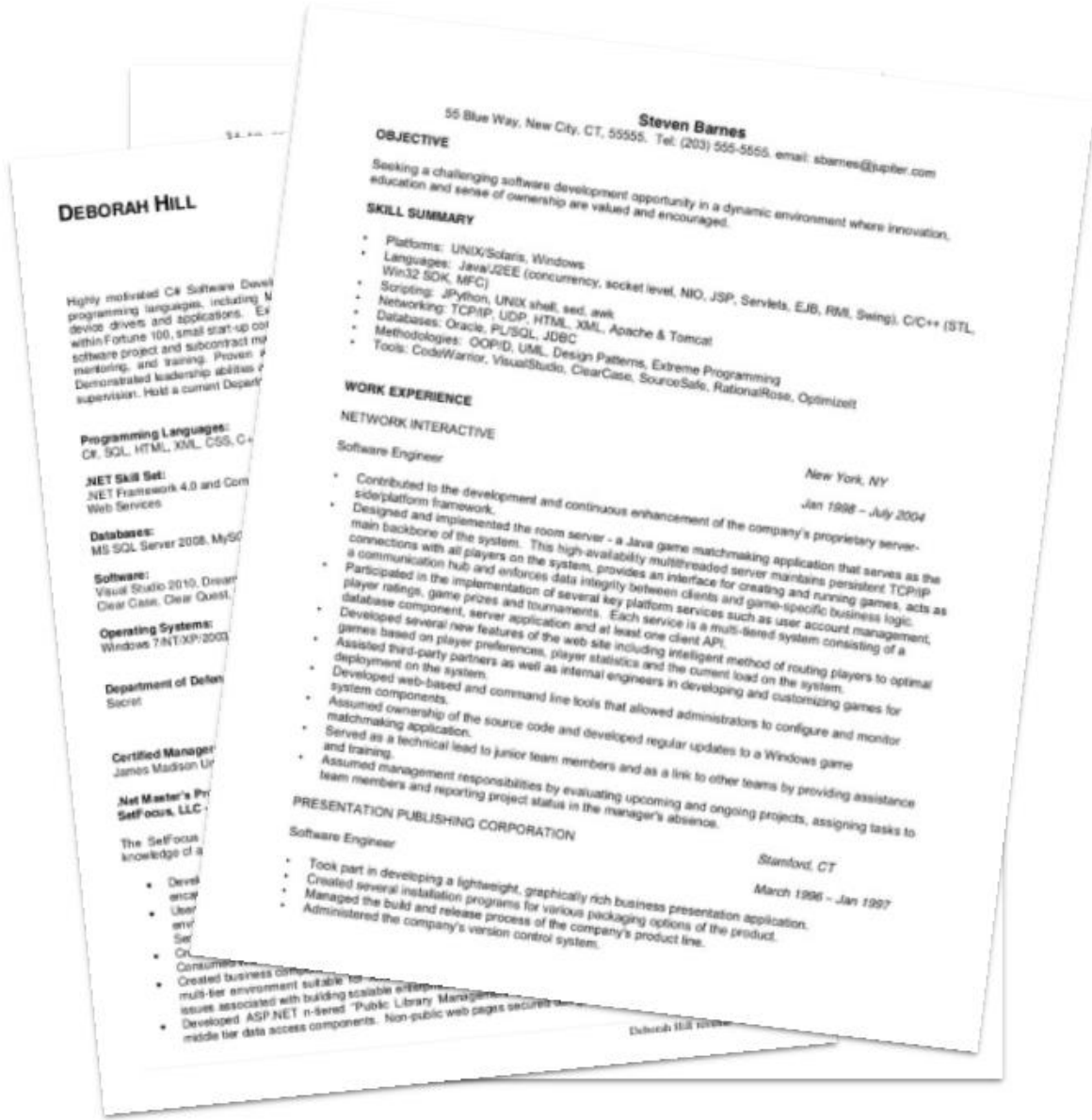
predict 0 when $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 < 0$

positive class = 1

negative class = 0

if probability >= 0.5: predict 1

if probability < 0.5: predict 0



DEBORAH HILL

Highly motivated C# Software Develo
programming languages, including M
device drivers and applications. Ex
within Fortune 100, small start-up co
software project and subcontract m
mentoring, and training. Proven a
Demonstrated leadership abilities a
supervision. Hold a current Departm

Programming Languages:
C#, SQL, HTML, XML, CSS, C+

.NET Skill Set:
.NET Framework 4.0 and Com
Web Services

Databases:
MS SQL Server 2008, MySQ

Software:
Visual Studio 2010, Dream
Clear Case, Clear Quest,

Operating Systems:
Windows 7/NT/XP/2003

Department of Defen
Secret

Certified Manager
James Madison Un

.Net Master's Pro
SetFocus, LLC -

The SetFocus
knowledge of a

- Devel
encap
- User
envi
Set
- Cre
Consumes
- Created business com
multi-tier environment suitable for
issues associated with building scalable enterpri
- Developed ASP.NET n-tiered "Public Library Manageme
middle tier data access components. Non-public web pages secured

Deborah Hill Resume

Steven Barnes

55 Blue Way, New City, CT, 55555. Tel: (203) 555-5555, email: sbarnes@tupiter.com

OBJECTIVE

Seeking a challenging software development opportunity in a dynamic environment where innovation, education and sense of ownership are valued and encouraged.

SKILL SUMMARY

- Platforms: UNIX/Solaris, Windows
- Languages: Java/J2EE (concurrency, socket level, NIO, JSP, Servlets, EJB, RMI, Swing), C/C++ (STL, Win32 SDK, MFC)
- Scripting: JPython, UNIX shell, sed, awk
- Networking: TCP/IP, UDP, HTML, XML, Apache & Tomcat
- Databases: Oracle, PL/SQL, JDBC
- Methodologies: OOP/D, UML, Design Patterns, Extreme Programming
- Tools: CodeWarrior, VisualStudio, ClearCase, SourceSafe, RationalRose, Optimizelt

WORK EXPERIENCE

NETWORK INTERACTIVE

Software Engineer — *New York, NY*

Jan 1998 – July 2004

- Contributed to the development and continuous enhancement of the company's proprietary server-side/platform framework.
- Designed and implemented the room server - a Java game matchmaking application that serves as the main backbone of the system. This high-availability multithreaded server maintains persistent TCP/IP connections with all players on the system, provides an interface for creating and running games, acts as a communication hub and enforces data integrity between clients and game-specific business logic.
- Participated in the implementation of several key platform services such as user account management, player ratings, game prizes and tournaments. Each service is a multi-tiered system consisting of a database component, server application and at least one client API.
- Developed several new features of the web site including intelligent method of routing players to optimal games based on player preferences, player statistics and the current load on the system.
- Assisted third-party partners as well as internal engineers in developing and customizing games for deployment on the system.
- Developed web-based and command line tools that allowed administrators to configure and monitor system components.
- Assumed ownership of the source code and developed regular updates to a Windows game matchmaking application.
- Served as a technical lead to junior team members and as a link to other teams by providing assistance and training.
- Assumed management responsibilities by evaluating upcoming and ongoing projects, assigning tasks to team members and reporting project status in the manager's absence.

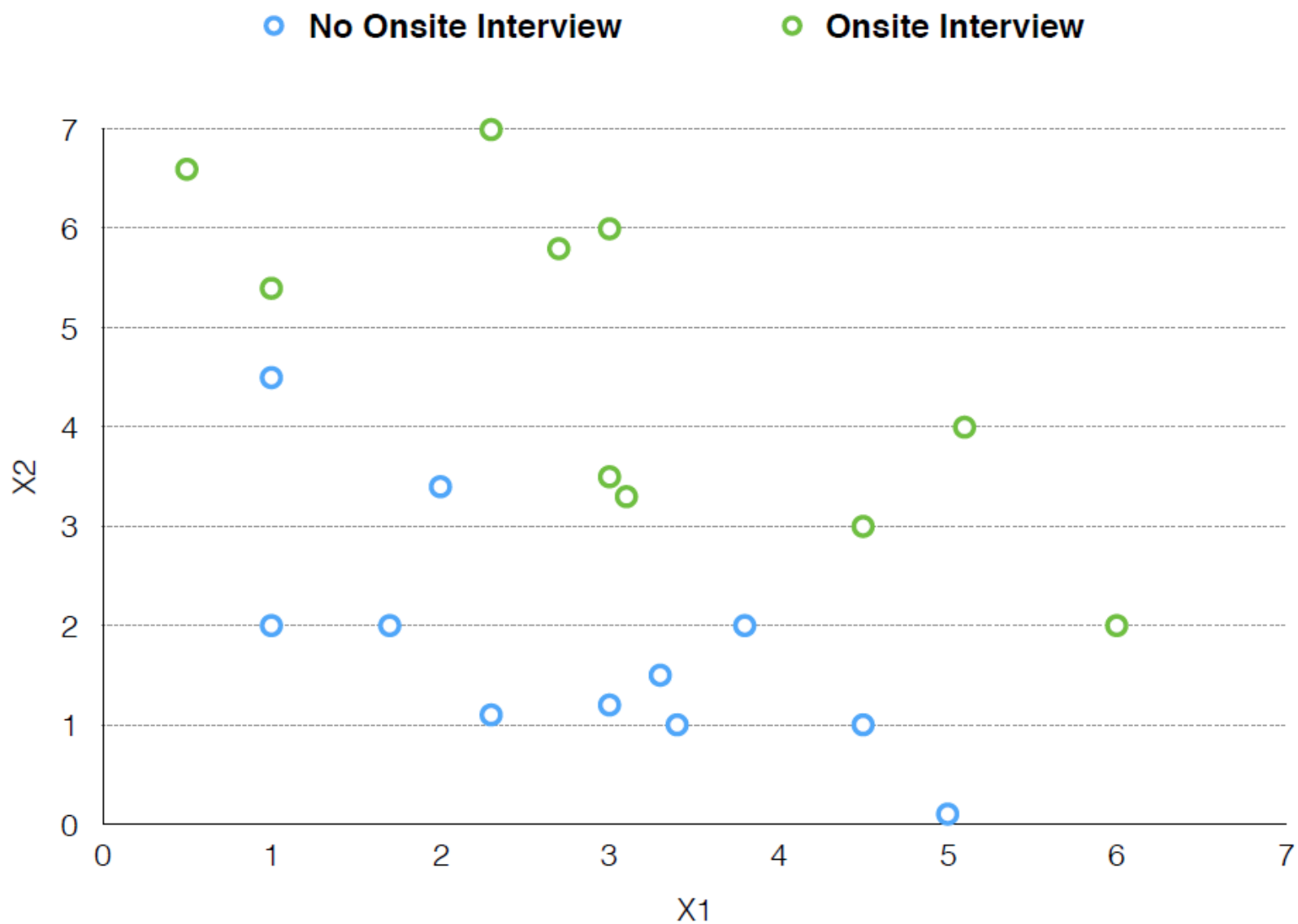
PRESENTATION PUBLISHING CORPORATION

Software Engineer — *Stamford, CT*

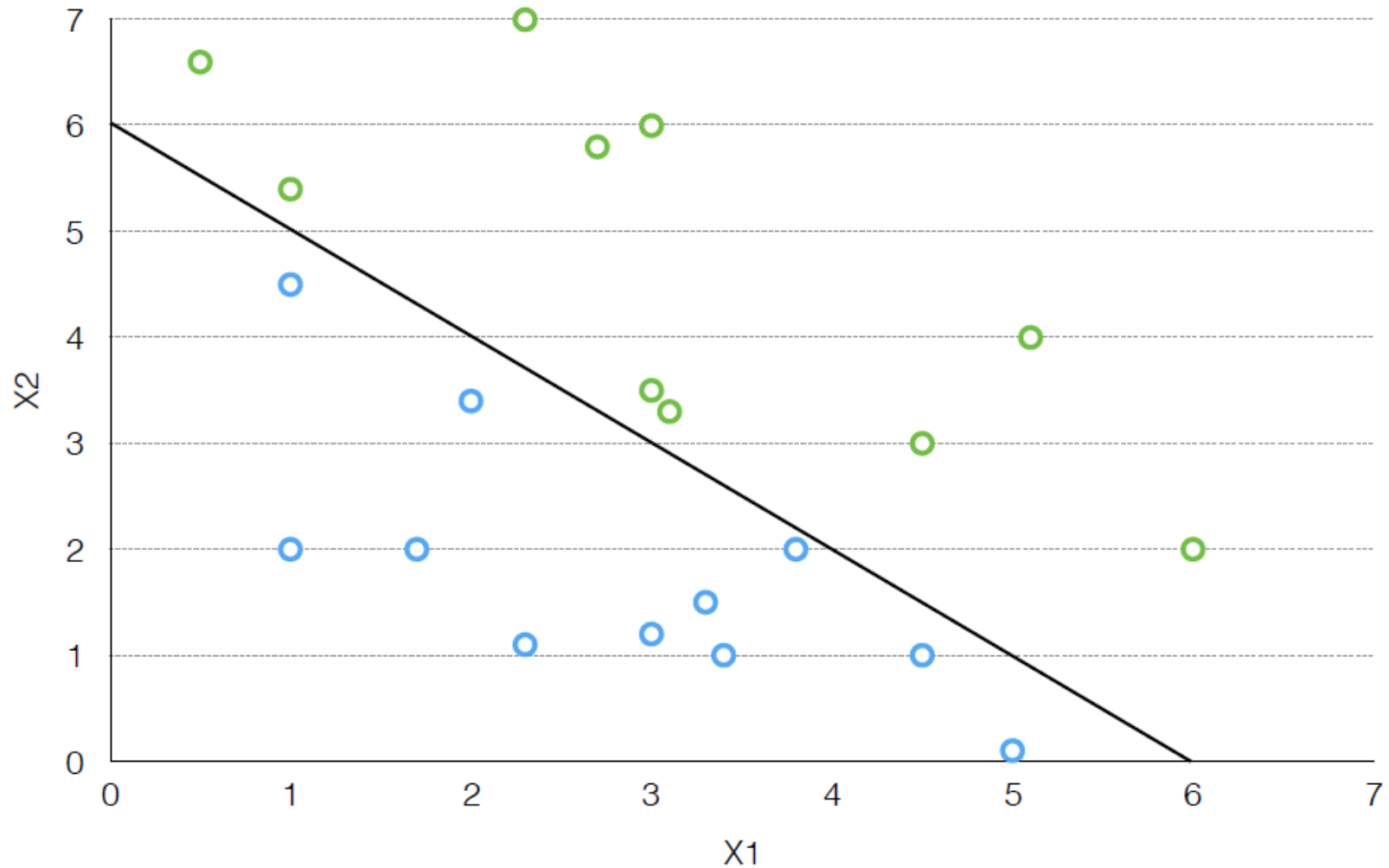
March 1996 – Jan 1997

- Took part in developing a lightweight, graphically rich business presentation application.
- Created several installation programs for various packaging options of the product.
- Managed the build and release process of the company's product line.
- Administered the company's version control system.

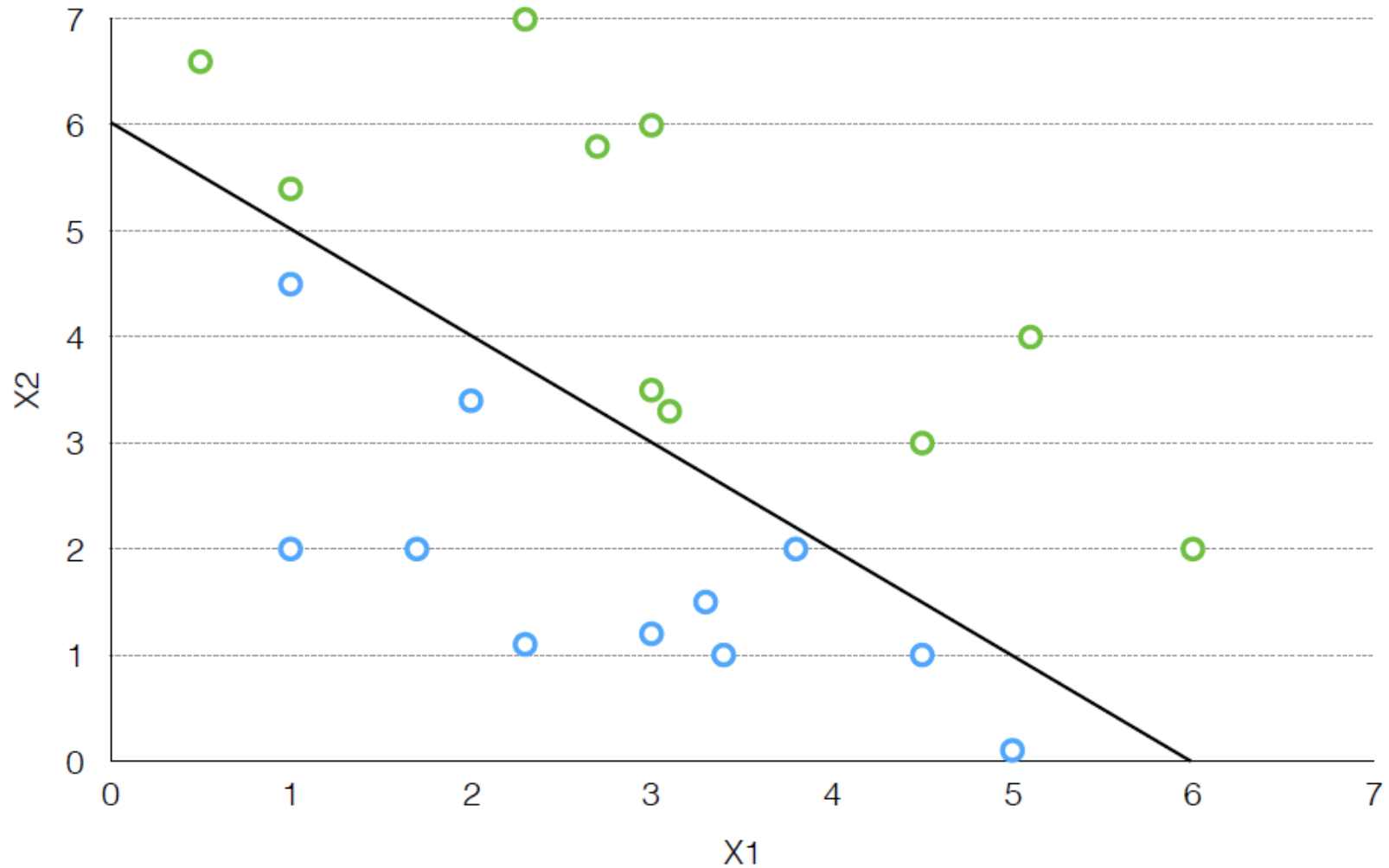
Gender	Years Exp.	Source 1	Source 2	Source 3	Phone Screen	On-site Interview
1	3	0	1	0	9	1
0	2	0	0	1	7.5	0
0	2	1	0	0	7	0
0	4	0	1	0	8.5	1
1	4	0	1	0	9.5	1
1	2	0	1	0	6.5	0
0	3	1	0	0	8	0
1	2	0	0	1	8	0
1	4	0	1	0	9	1
0	4	0	1	0	7	1


No Onsite Interview
Onsite Interview
X2
X1
0
1
2
3
4
5
6
7

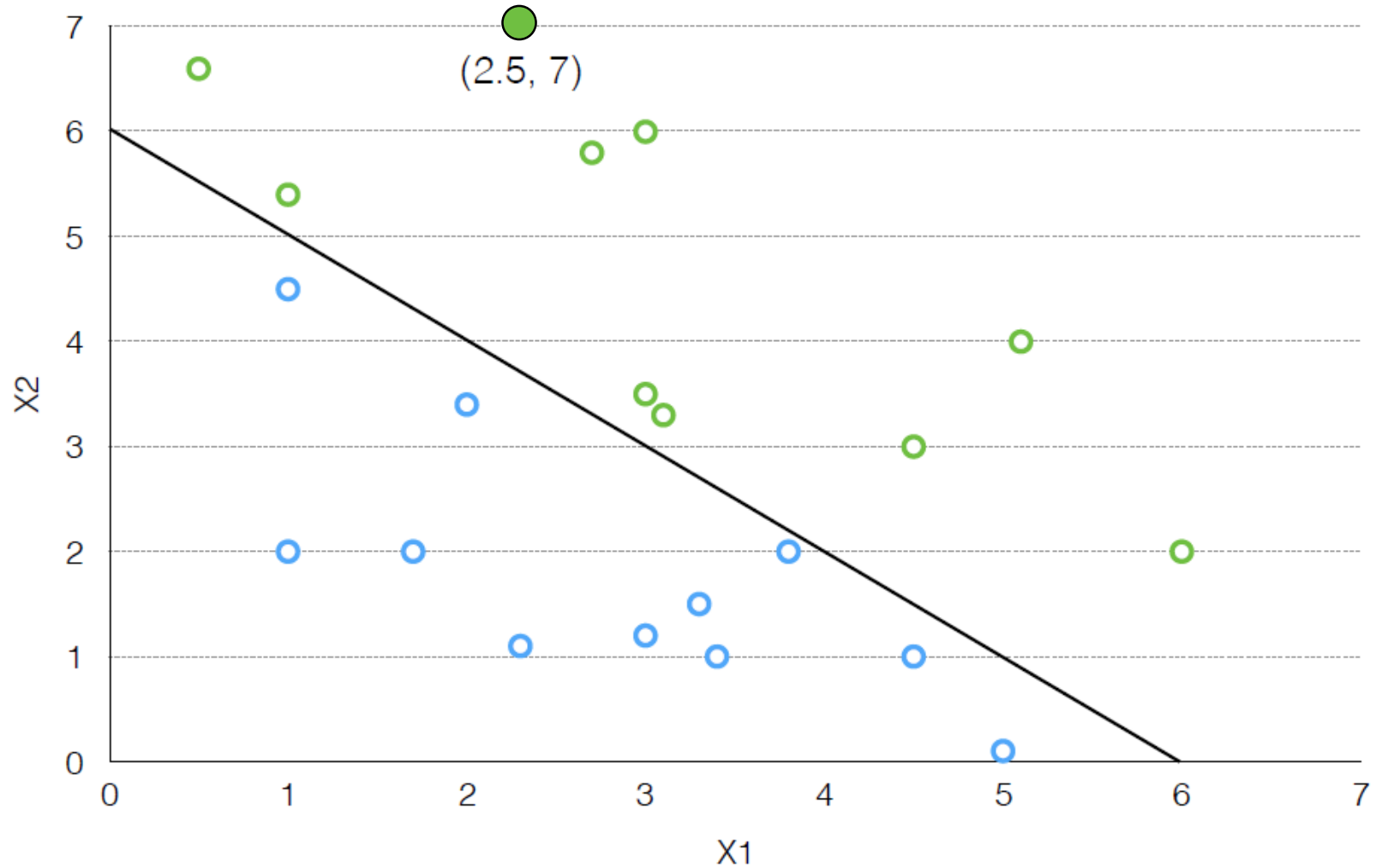
Decision Boundary = $-6 + 1x_1 + 1x_2$



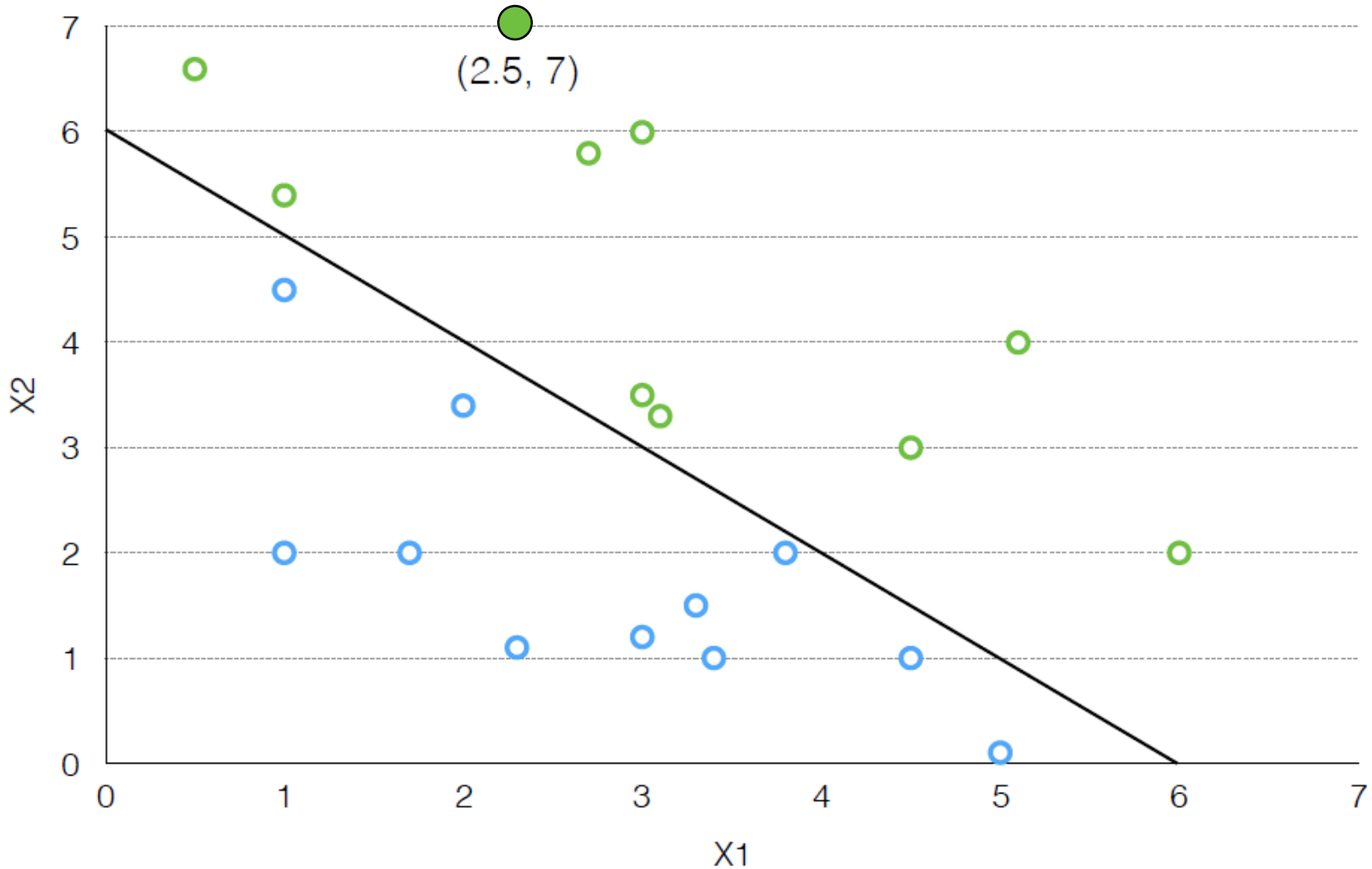
$$\hat{f}(X) = \frac{1}{1+e^{-(-6 + 1x_1 + 1x_2)}}$$



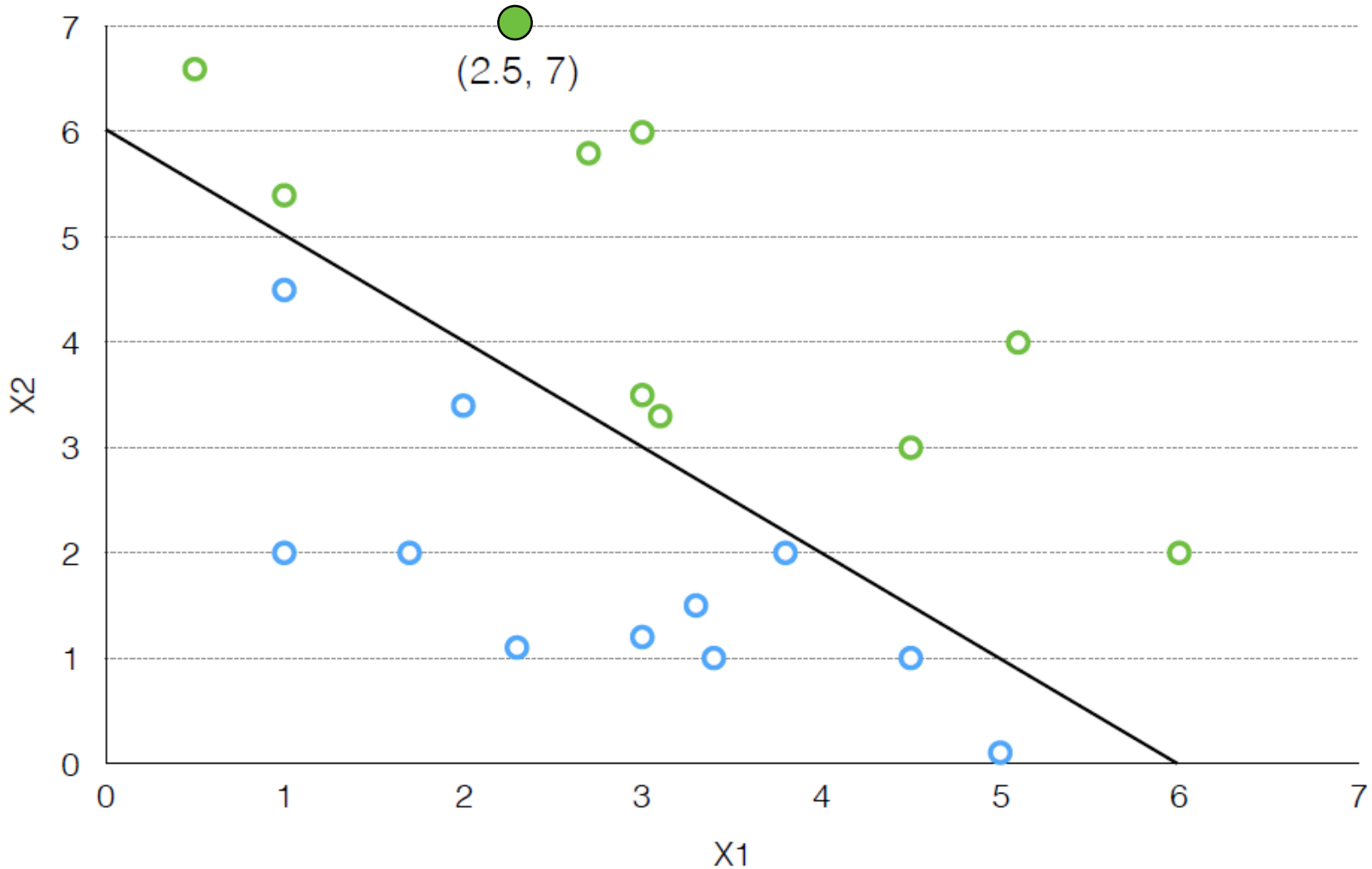
$$\frac{1}{1+e^{-(-6 + 1x_1 + 1x_2)}}$$



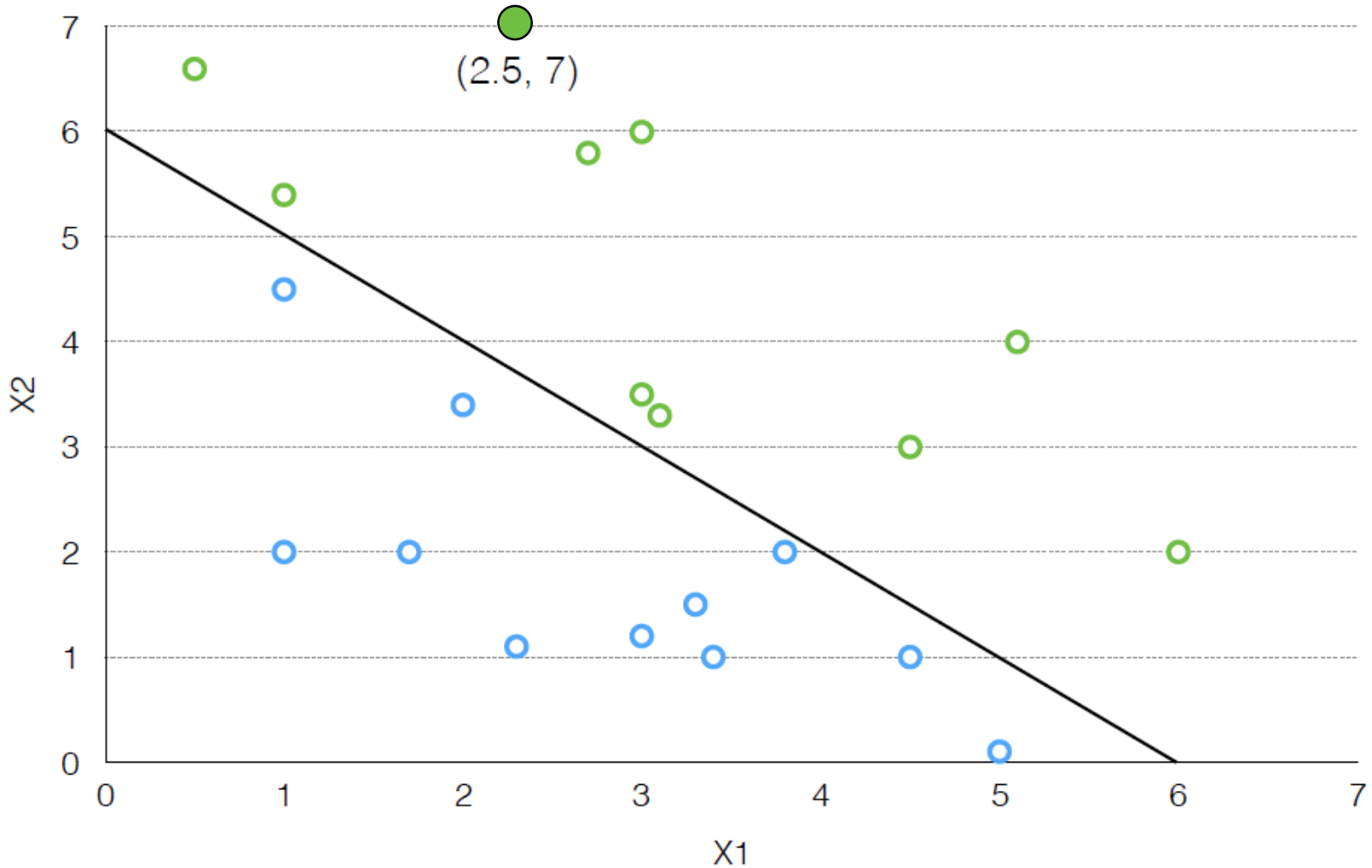
$$\frac{1}{1+e^{-(-6+2.5+7)}}$$



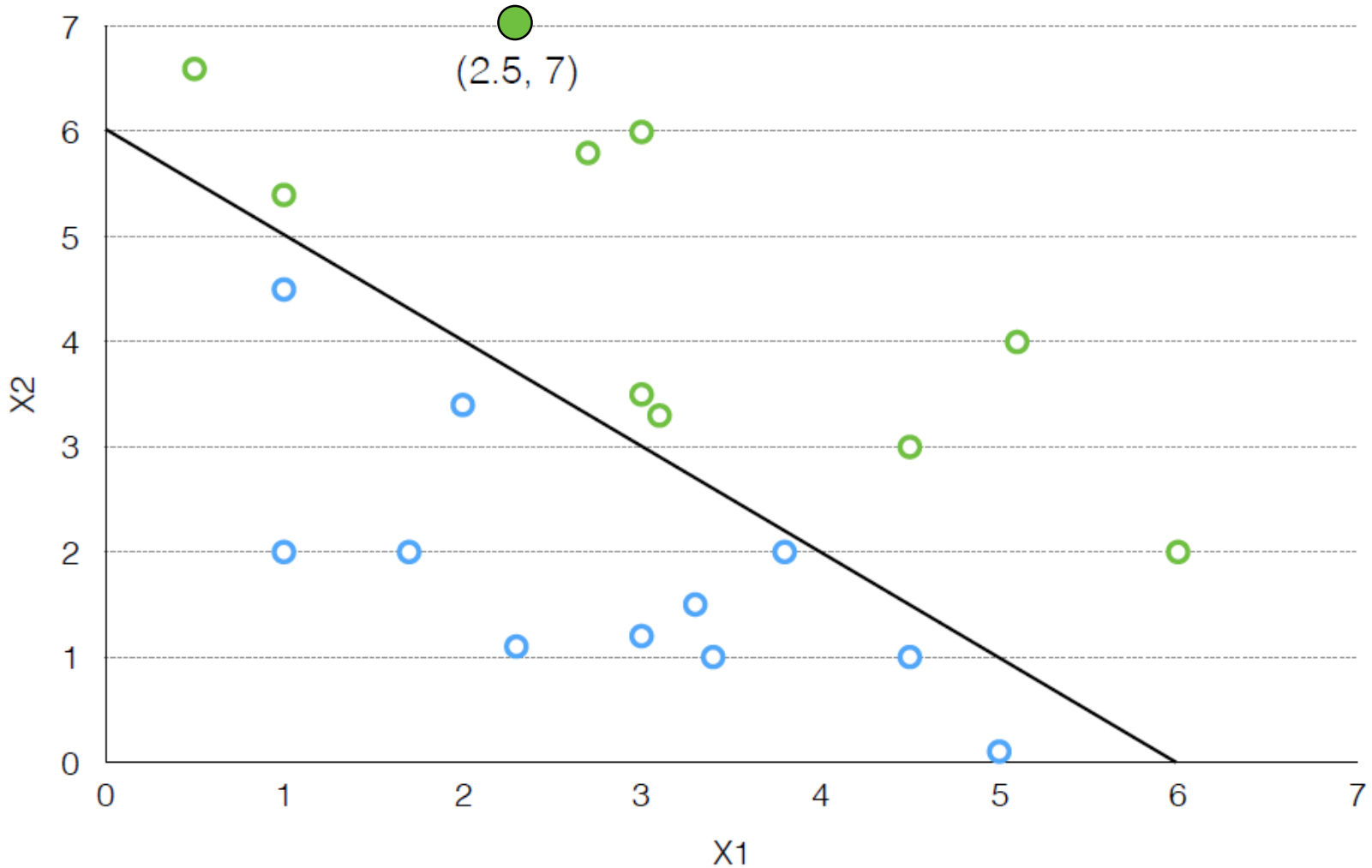
$$\frac{1}{1+e^{-(-6+2.5+7)}}$$



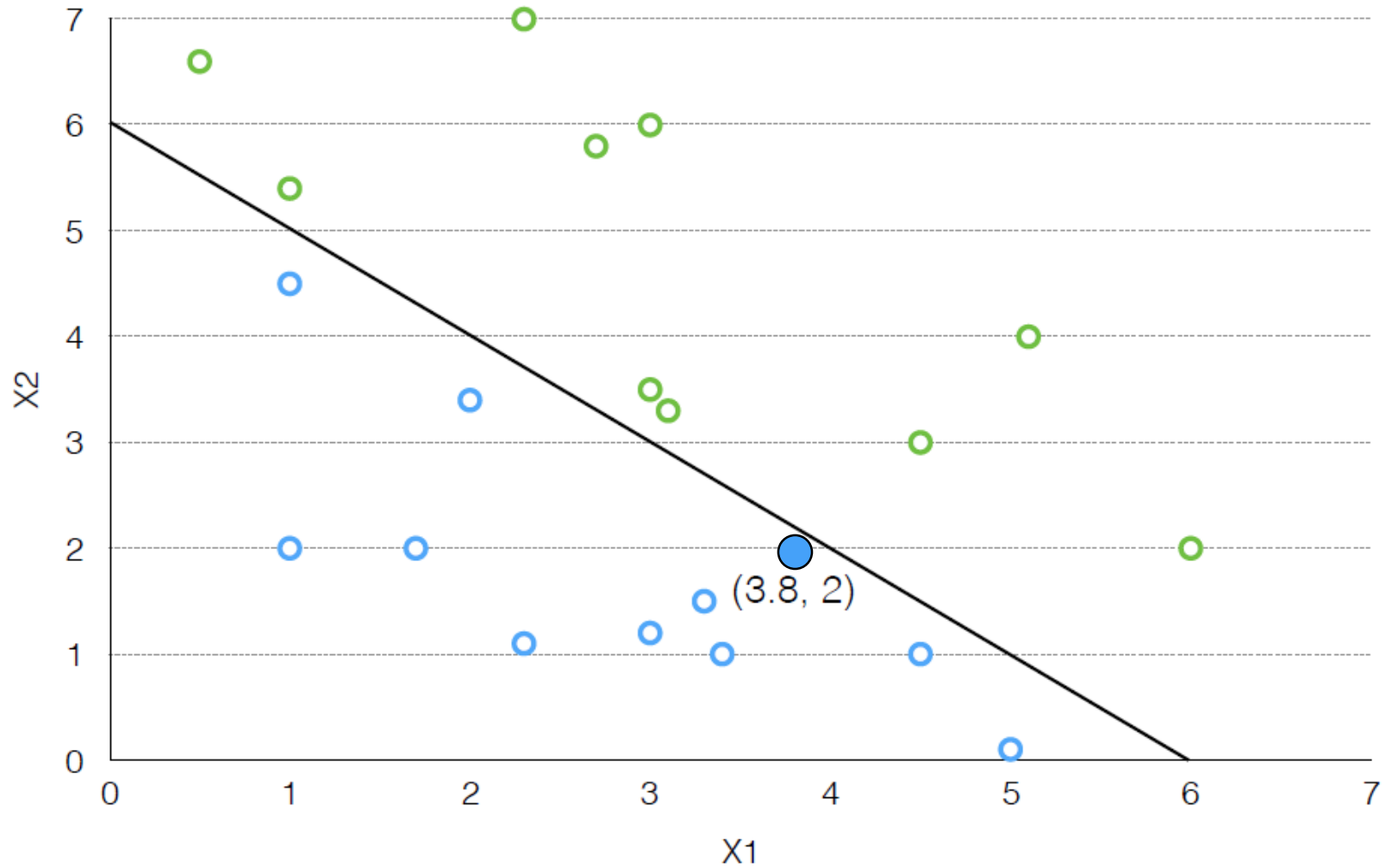
$$\frac{1}{1+e^{-(-6+9.5)}}$$



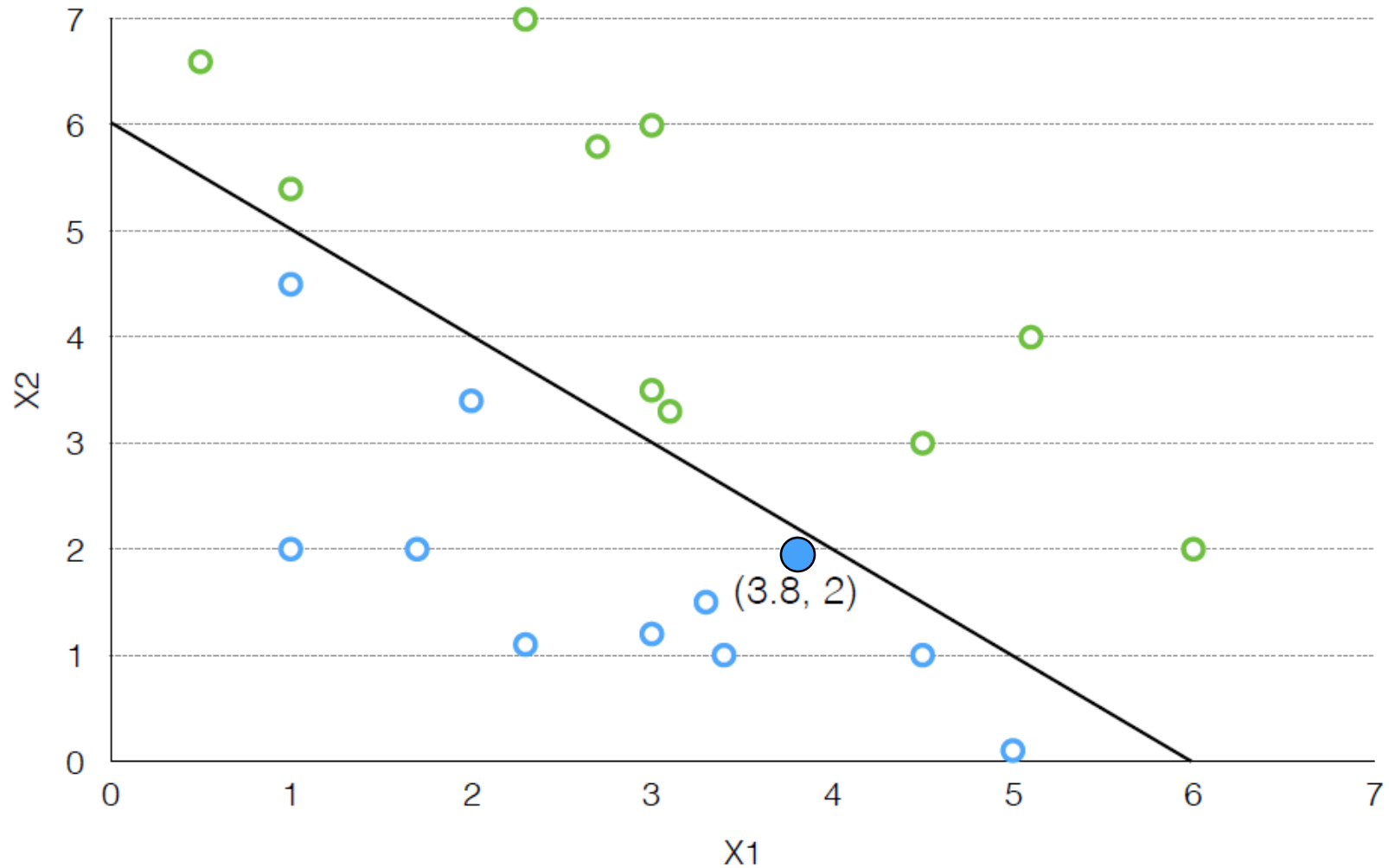
$$\frac{1}{1+e^{-(3.5)}} = .97 \text{ (probability)}$$



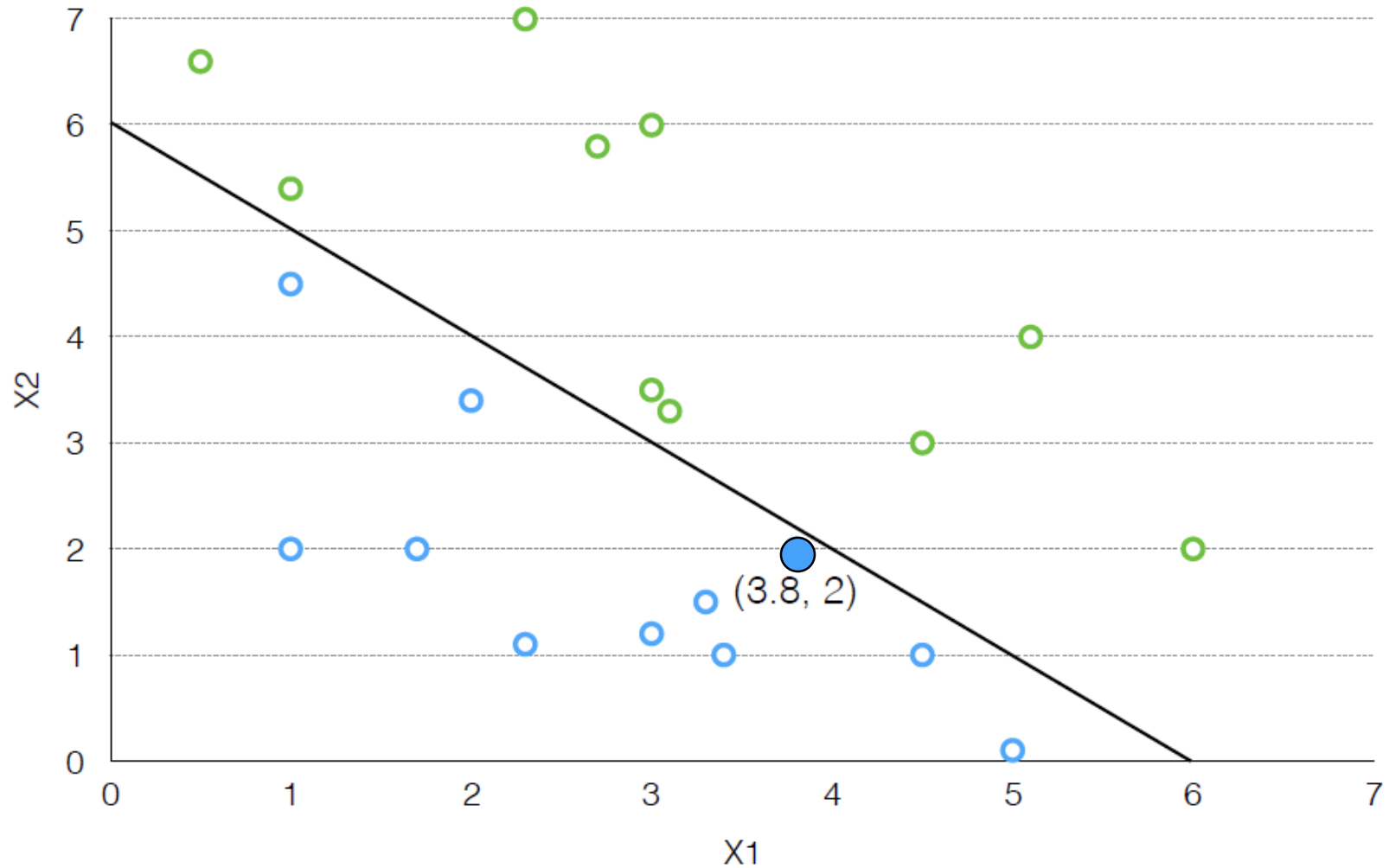
$$\frac{1}{1+e^{-(-6 + 1x_1 + 1x_2)}}$$



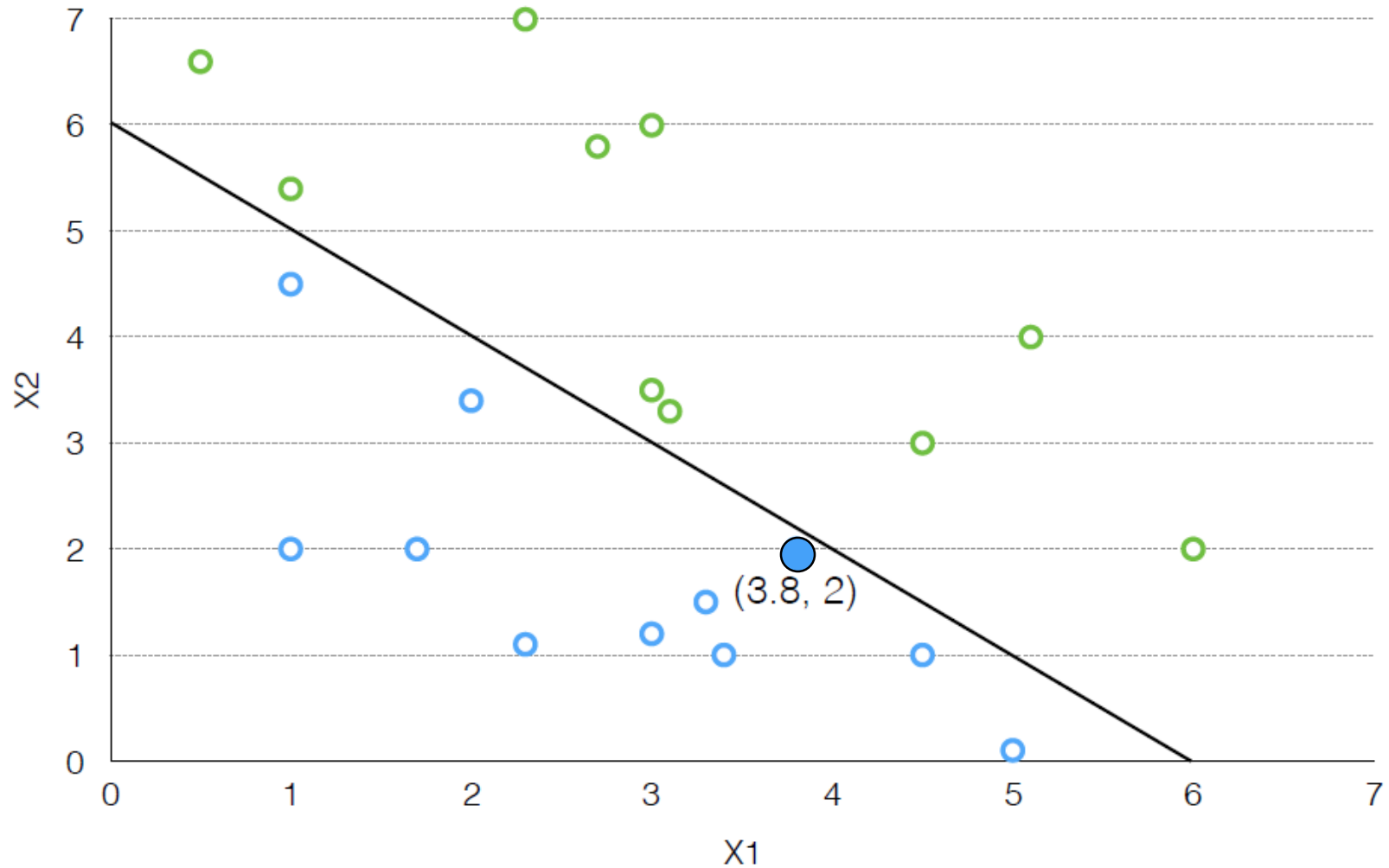
$$\frac{1}{1+e^{-(-6+3.8+2)}}$$



$$\frac{1}{1+e^{-(-6+5.8)}}$$



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



Support Vector Machine

Support Vector Machine

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

predict 1 when $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 >= 0$

predict 0 when $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 < 0$

Support Vector Machine

predict 1 when $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 >= 1$

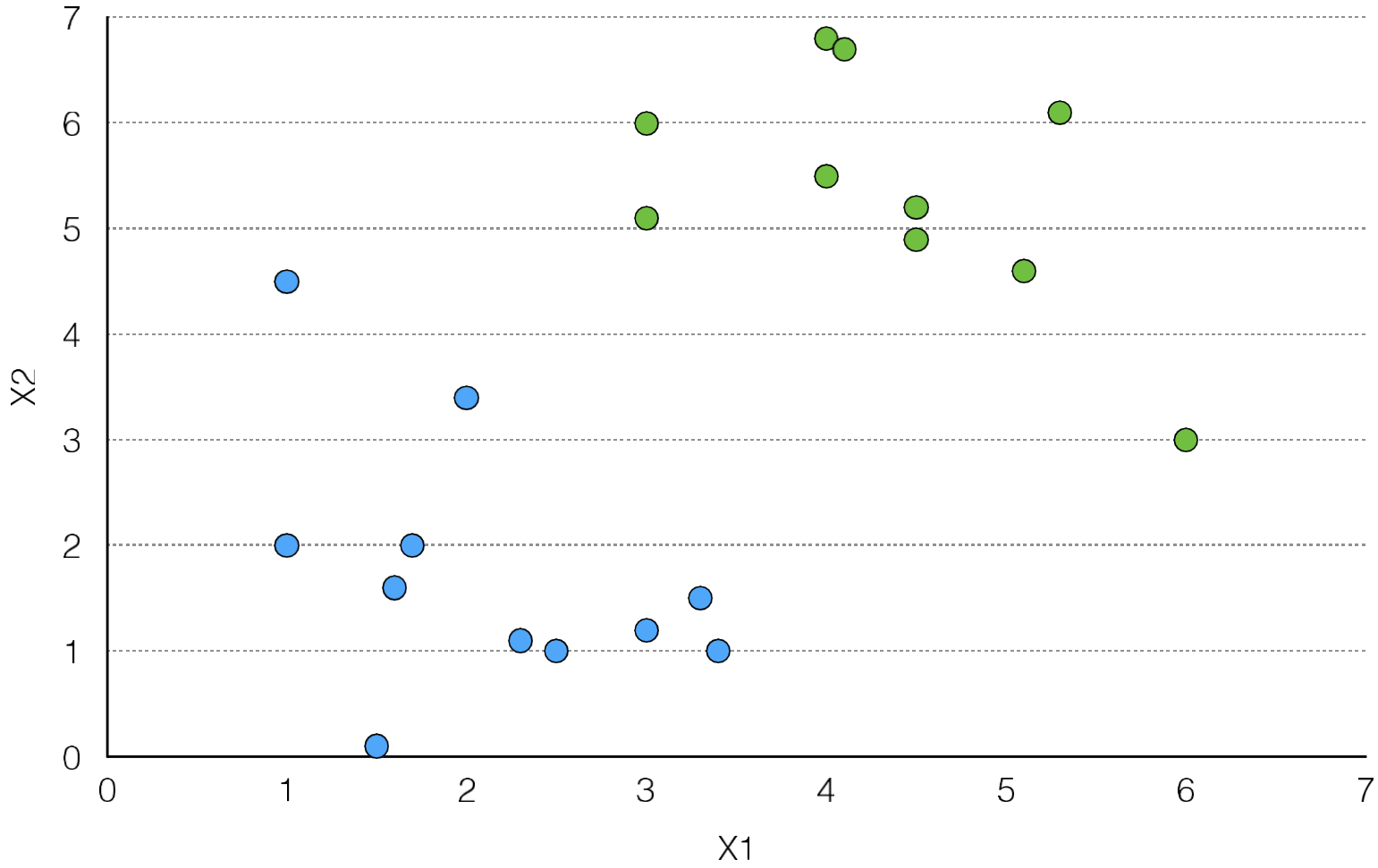
predict 0 when $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 < -1$

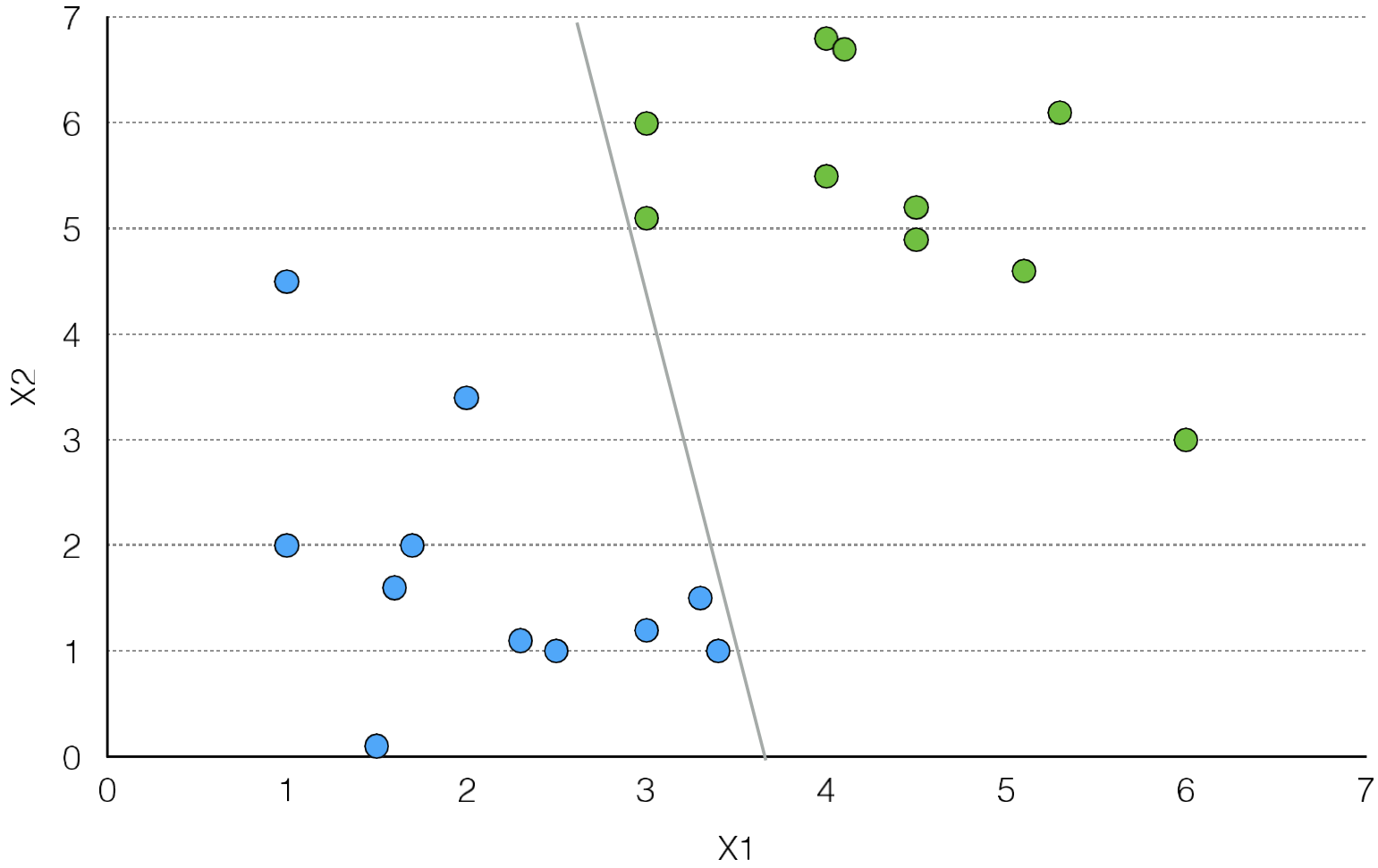
Support Vector Machine

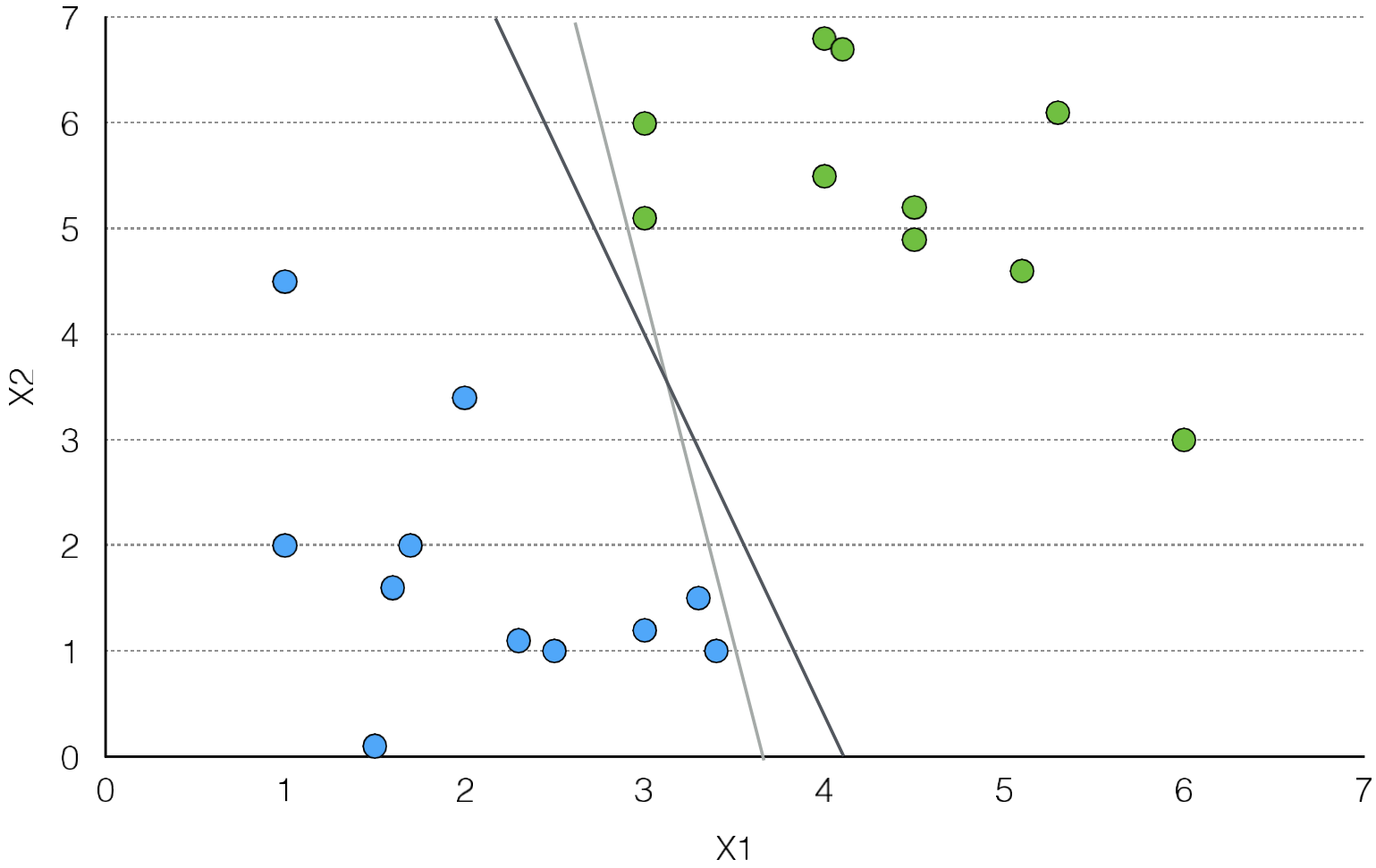
Large Margin Classifier

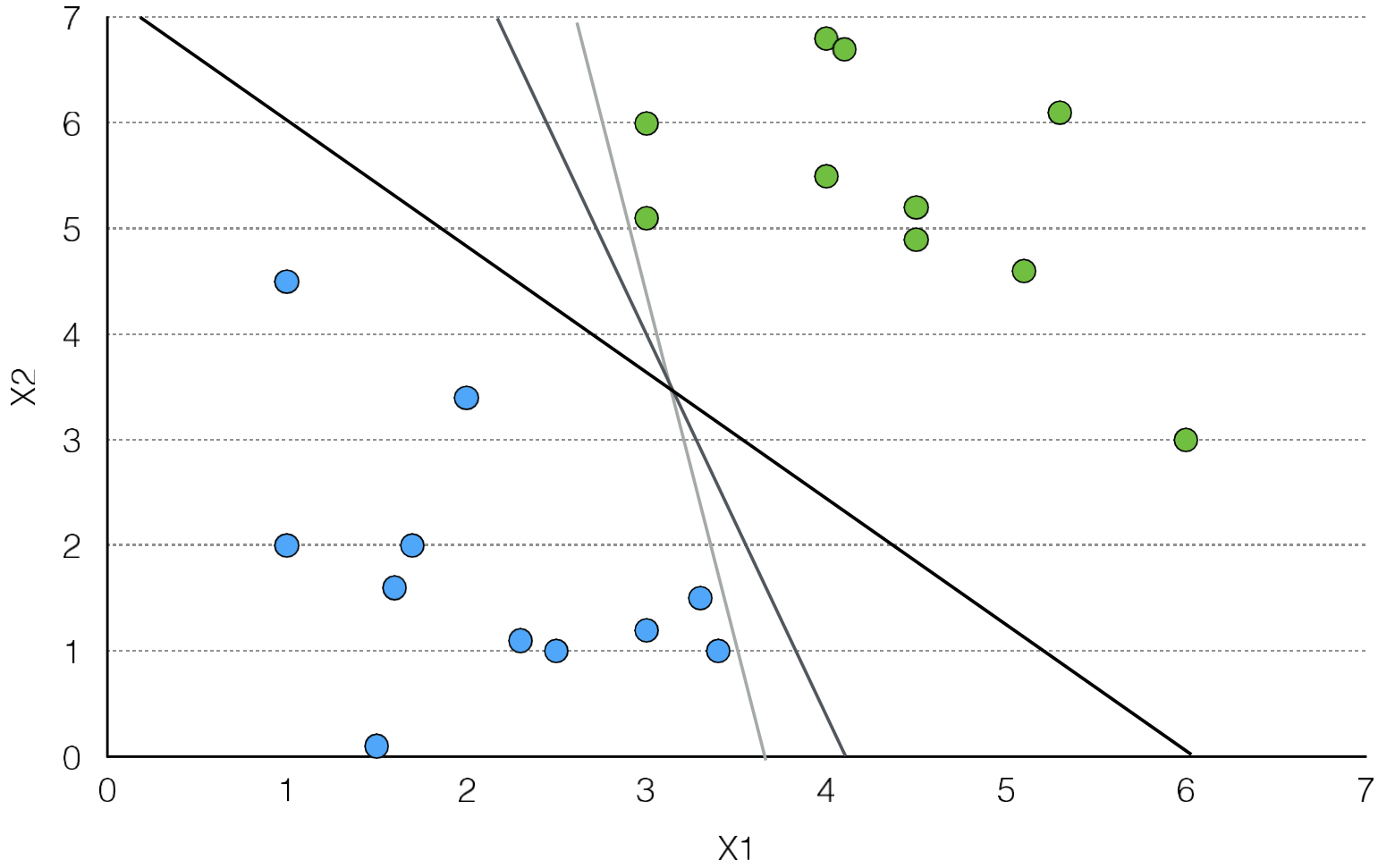
predict 1 when $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 >= 1$

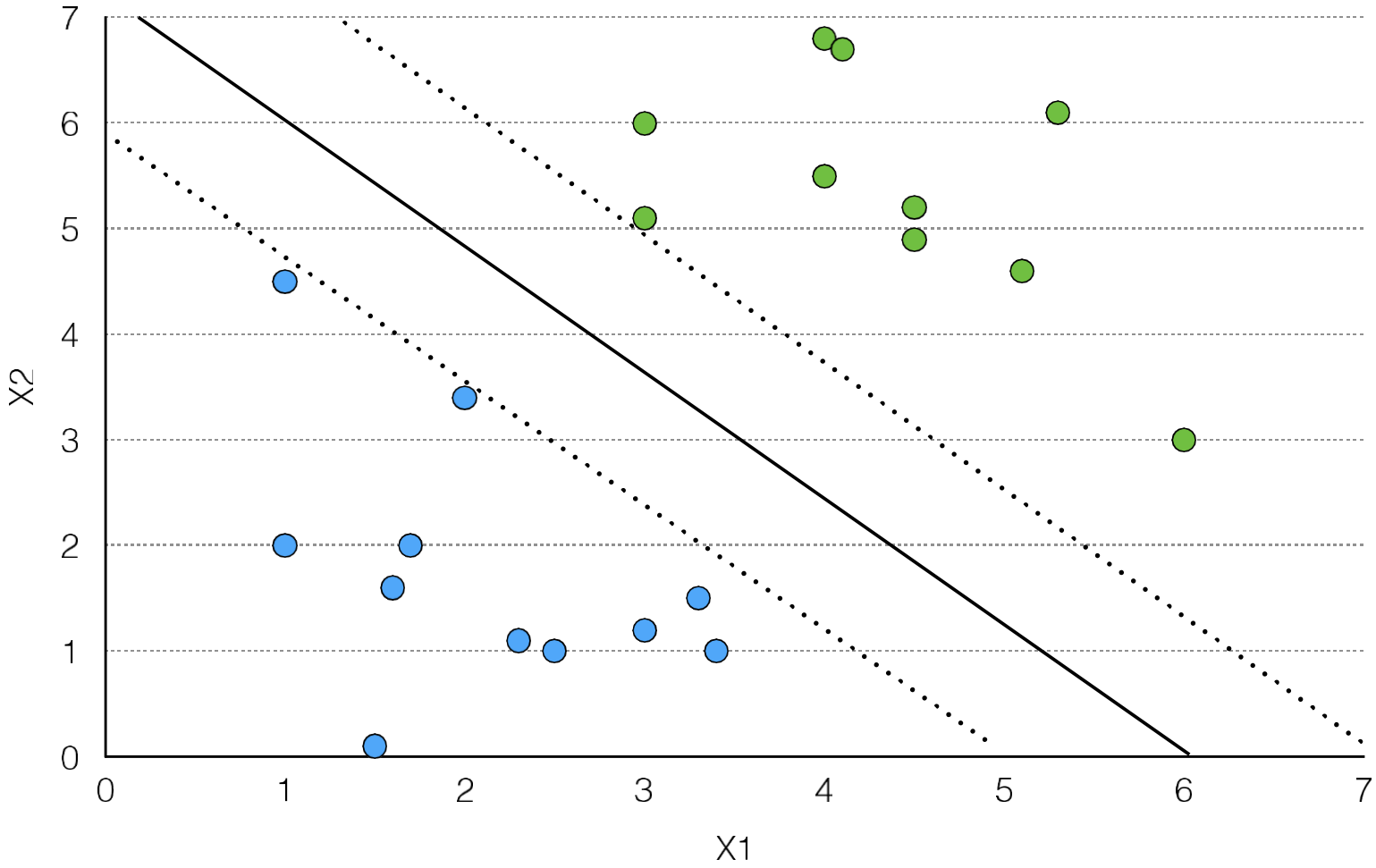
predict 0 when $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 < -1$

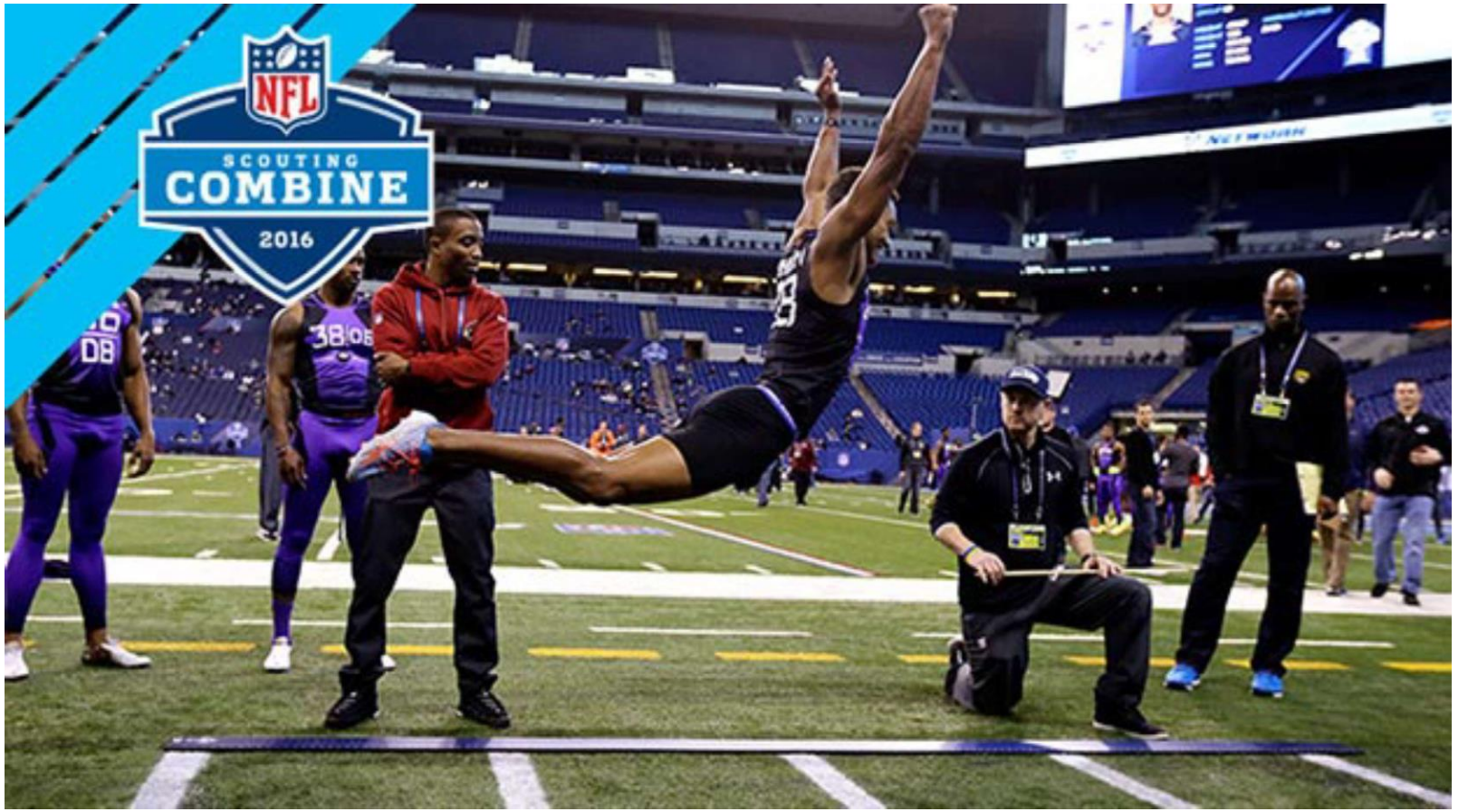

7
6
5
4
3
2
1
0
X2
0
1
2
3
4
5
6
7
X1


7
6
5
4
3
2
1
0
X2
0
1
2
3
4
5
6
7
X1

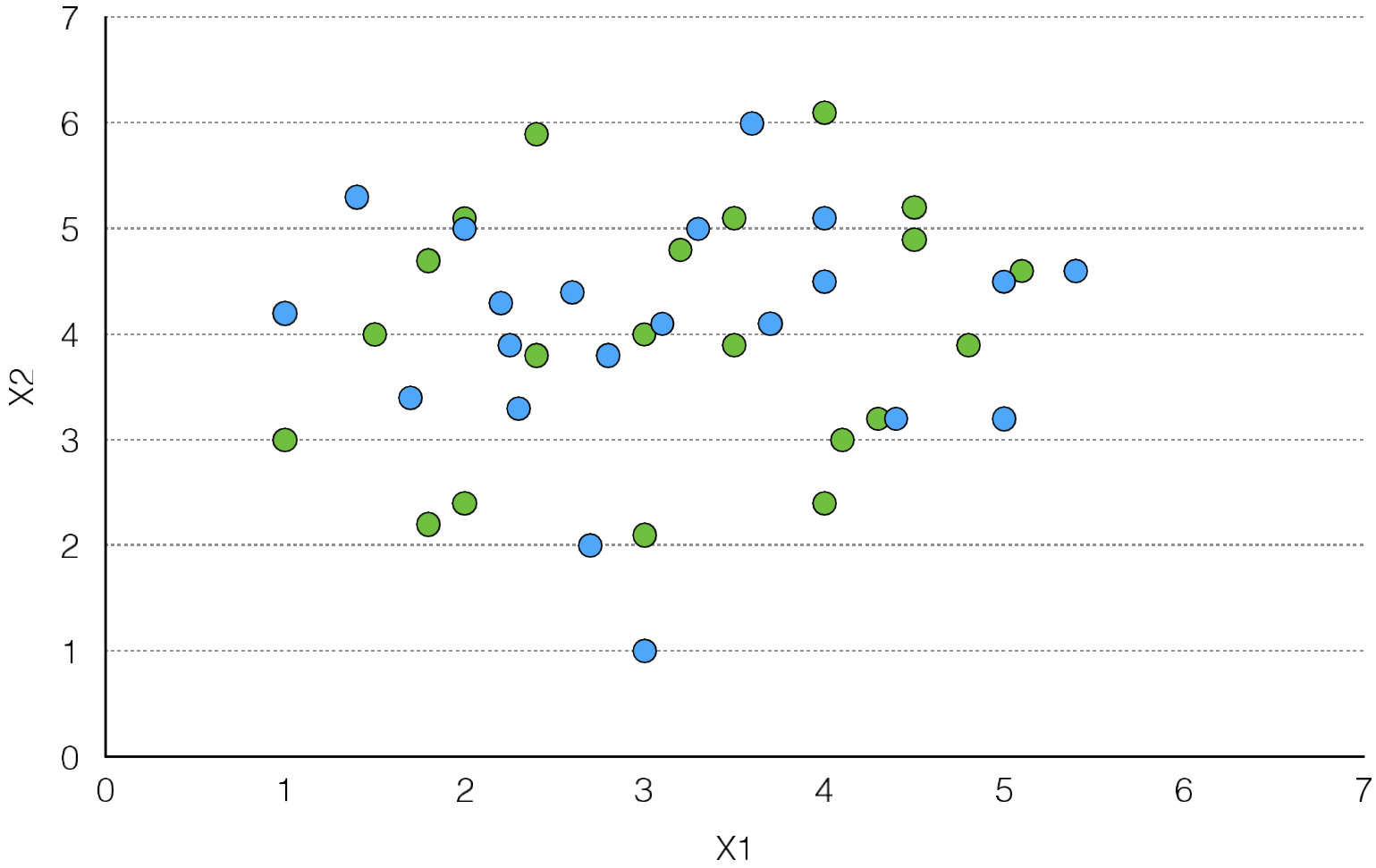

7
6
5
4
3
2
1
0
X2
0
1
2
3
4
5
6
7
X1


7
6
5
4
3
2
1
0
X2
0
1
2
3
4
5
6
7
X1


7
6
5
4
3
2
1
0
X2
0
1
2
3
4
5
6
7
X1


NFL
SCOUTING
COMBINE
2016

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


7
6
5
4
3
2
1
0
X2
0
1
2
3
4
5
6
7
X1

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

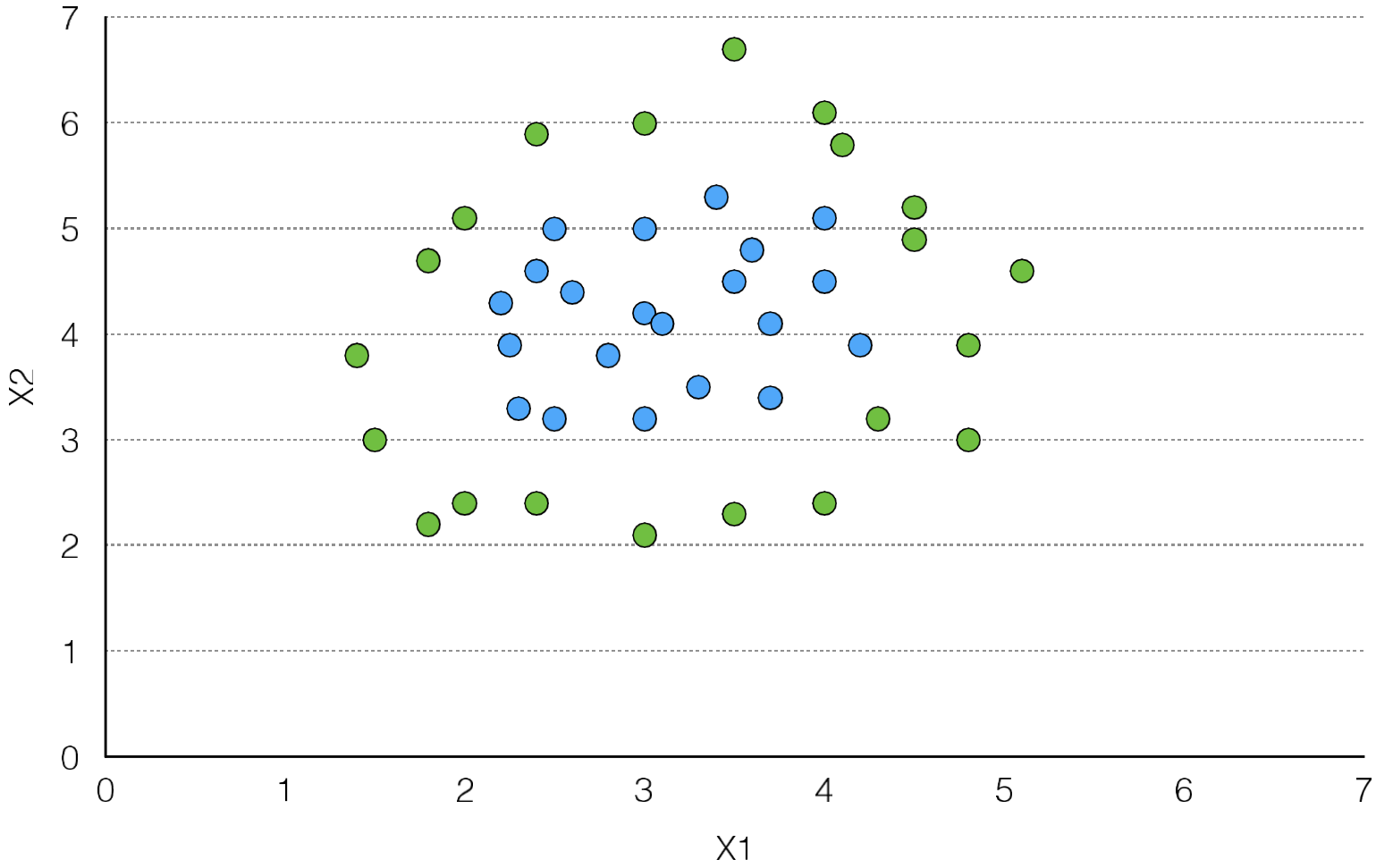

44
OL

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

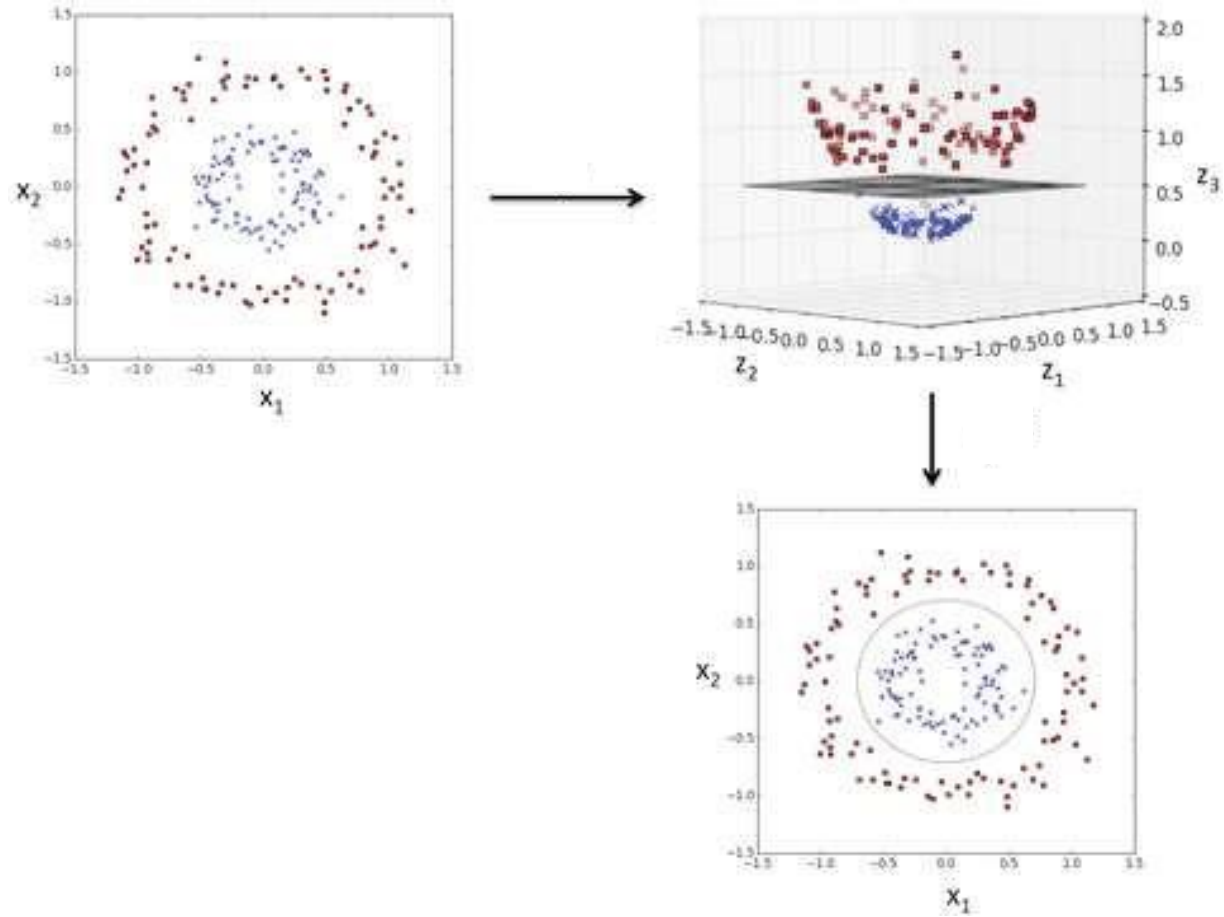
Feature Engineering

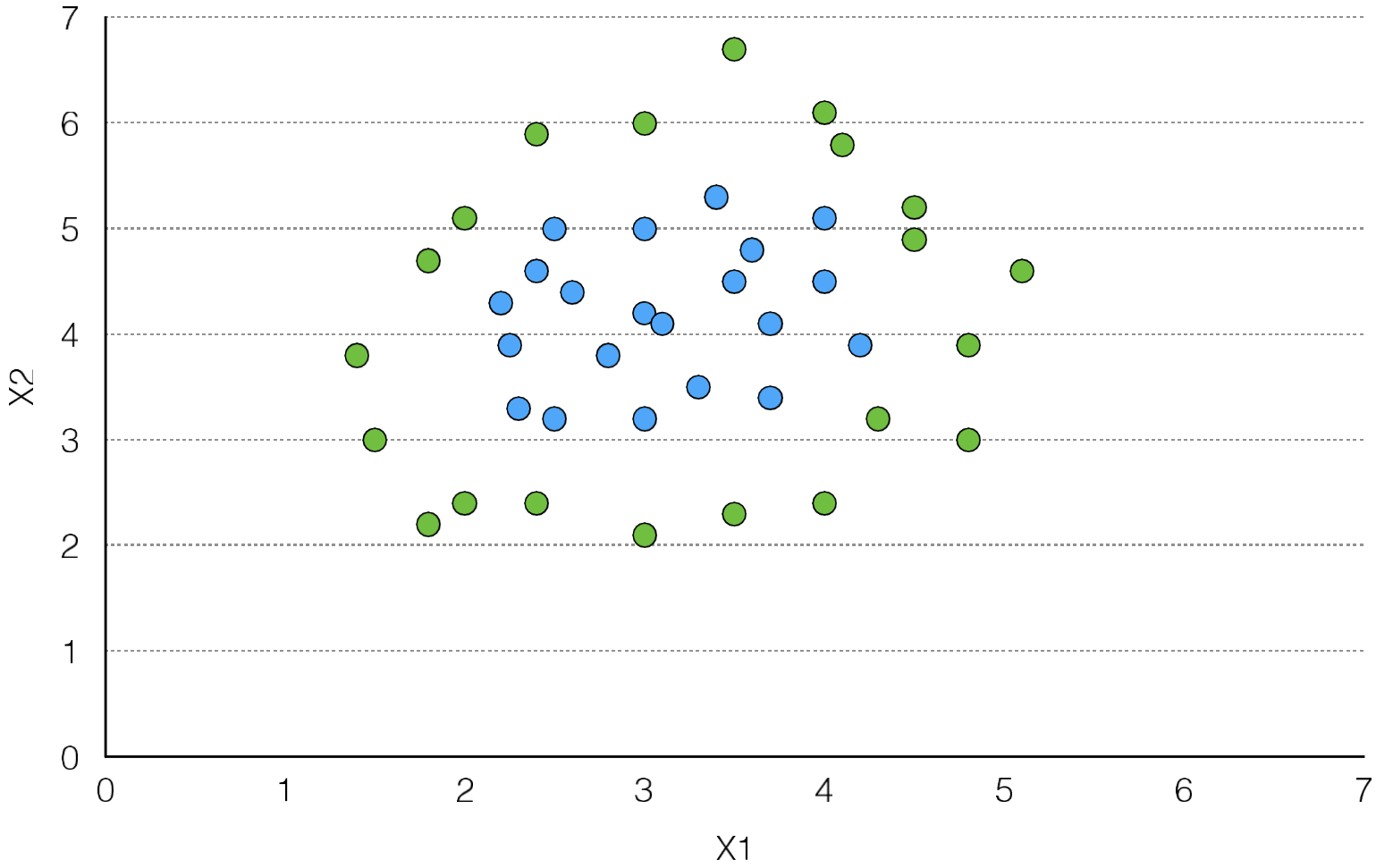
Speed-to-Size (40-yd/bsa)	BMI (wt/ht^2)	Drafted
2.16	37.2	1
2.06	33.9	1
2.02	26.3	0
1.97	28	1
2.23	31	0
2.00	35.9	0
2.03	26.4	1
1.99	28.6	1
1.85	35.2	0
2.03	33.5	1

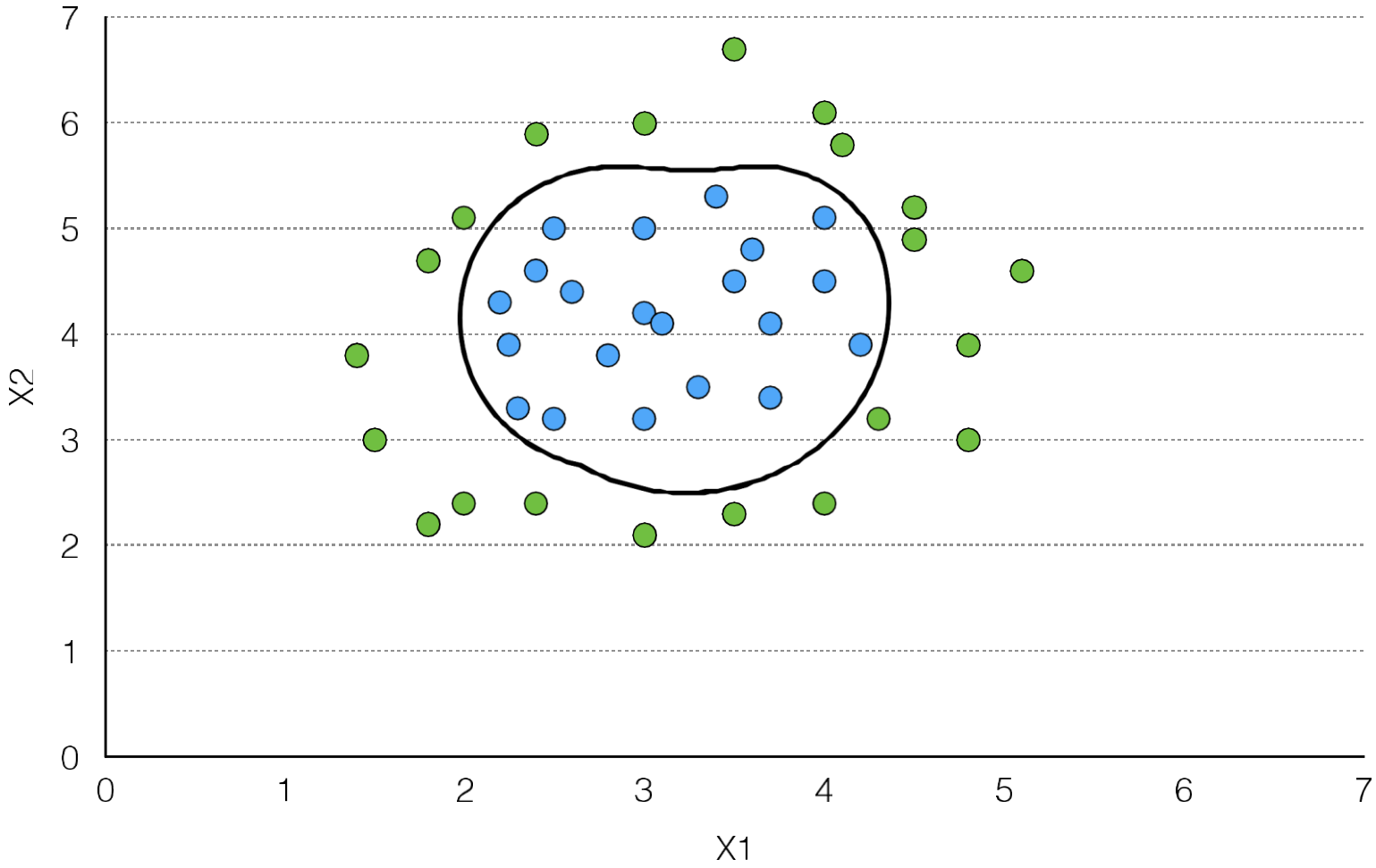

7
6
5
4
3
2
1
0
X2
0
1
2
3
4
5
6
7
X1

Kernel

Non-linear Classification

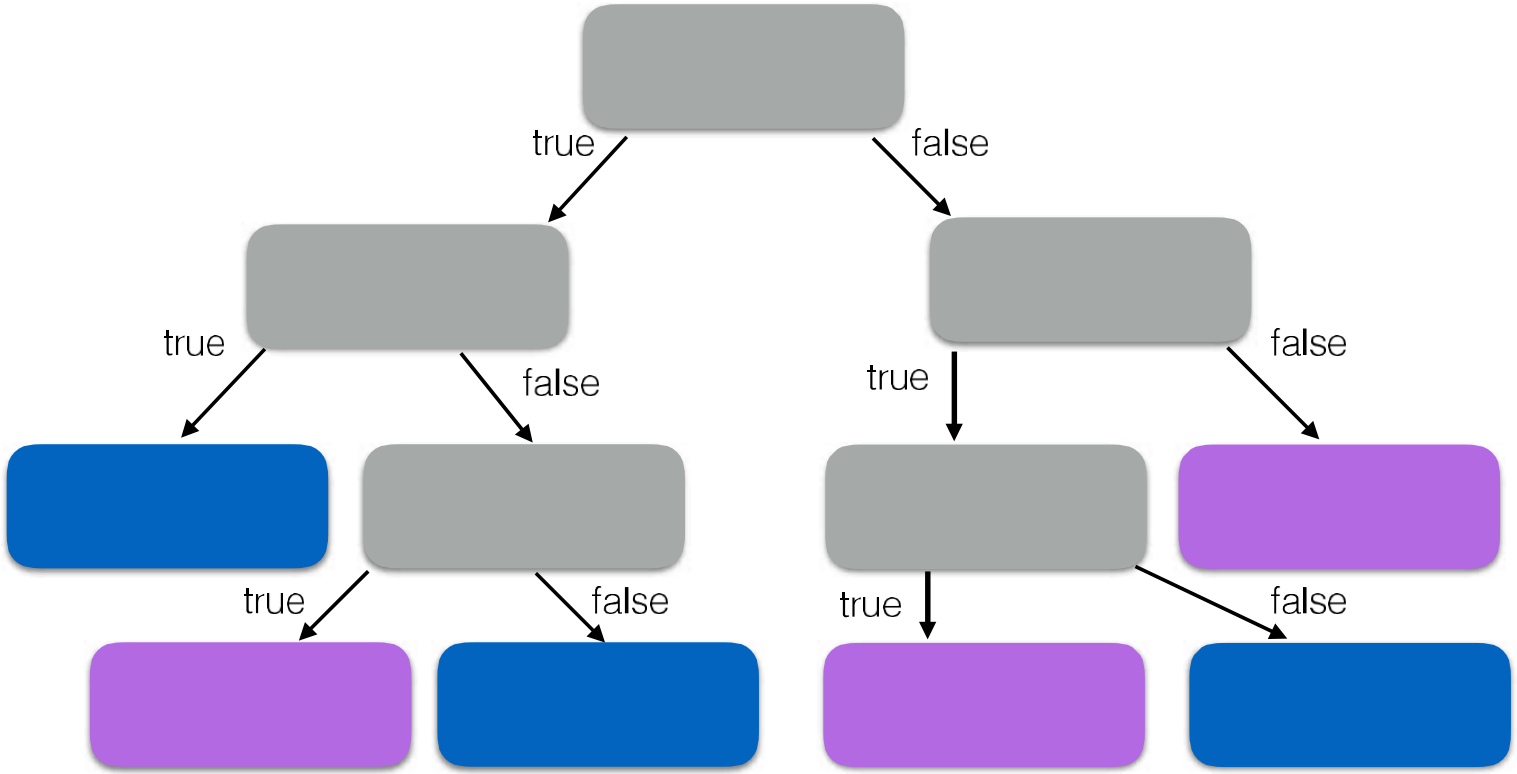

x2
x1
z3
z2
z1
x2
x1


7
6
5
4
3
2
1
0
X2
0
1
2
3
4
5
6
7
X1


7
6
5
4
3
2
1
0
X2
0
1
2
3
4
5
6
7
X1

Decision Tree

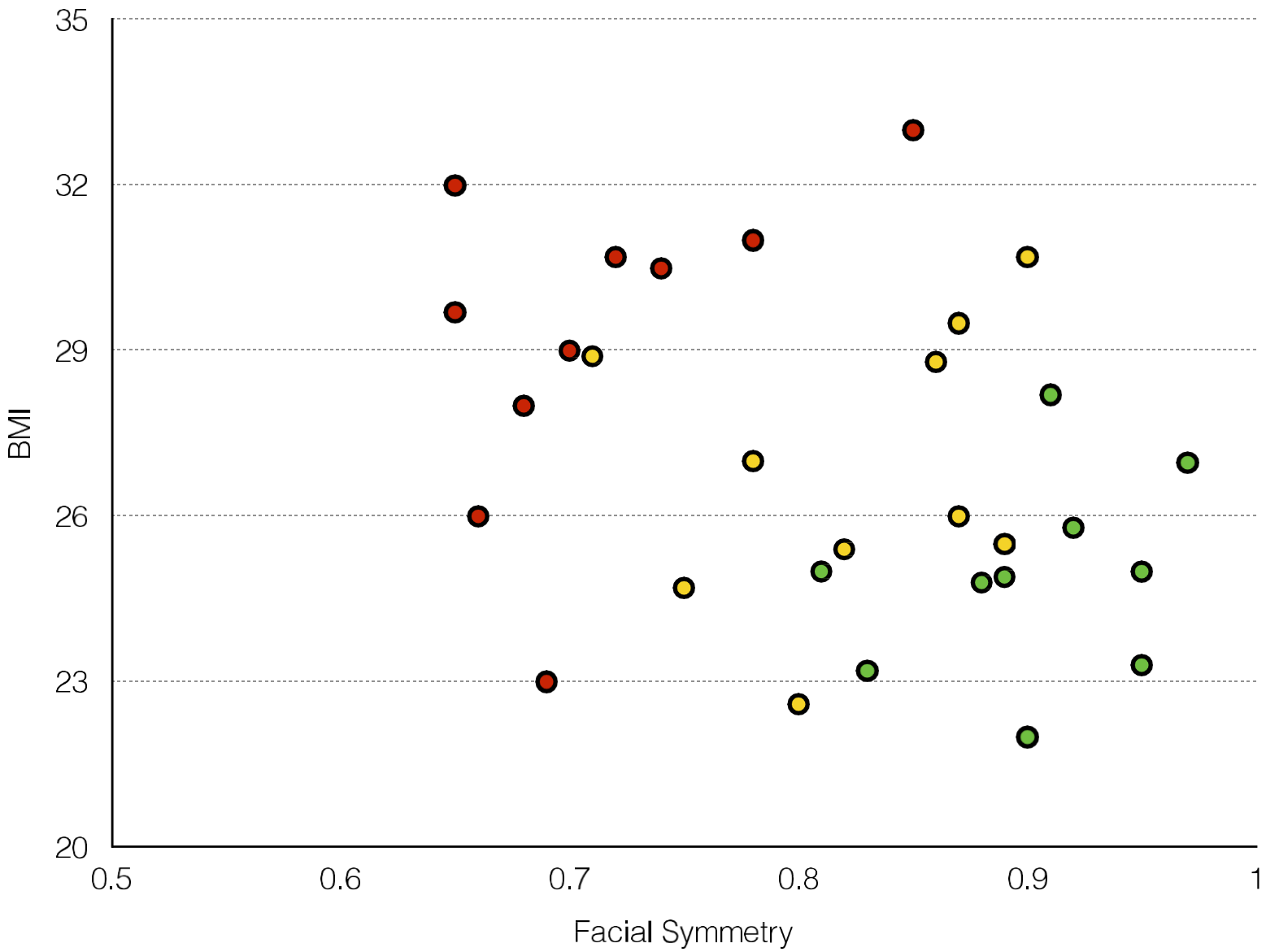
Decision Tree

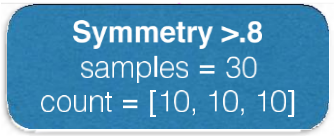


Short-term Attractiveness

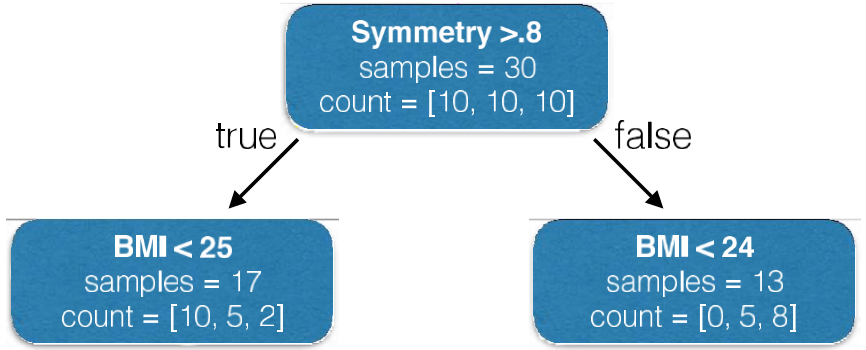
Short-term Attractiveness

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

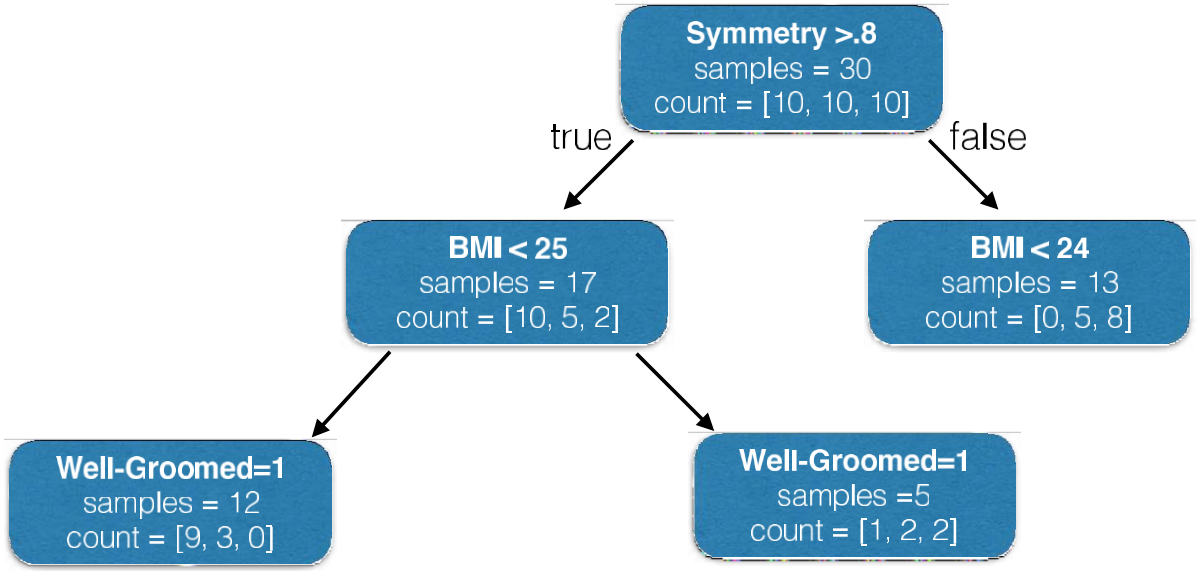

35
32
29
26
23
20
BMI
0.5
0.6
0.7
0.8
0.9
1
Facial Symmetry



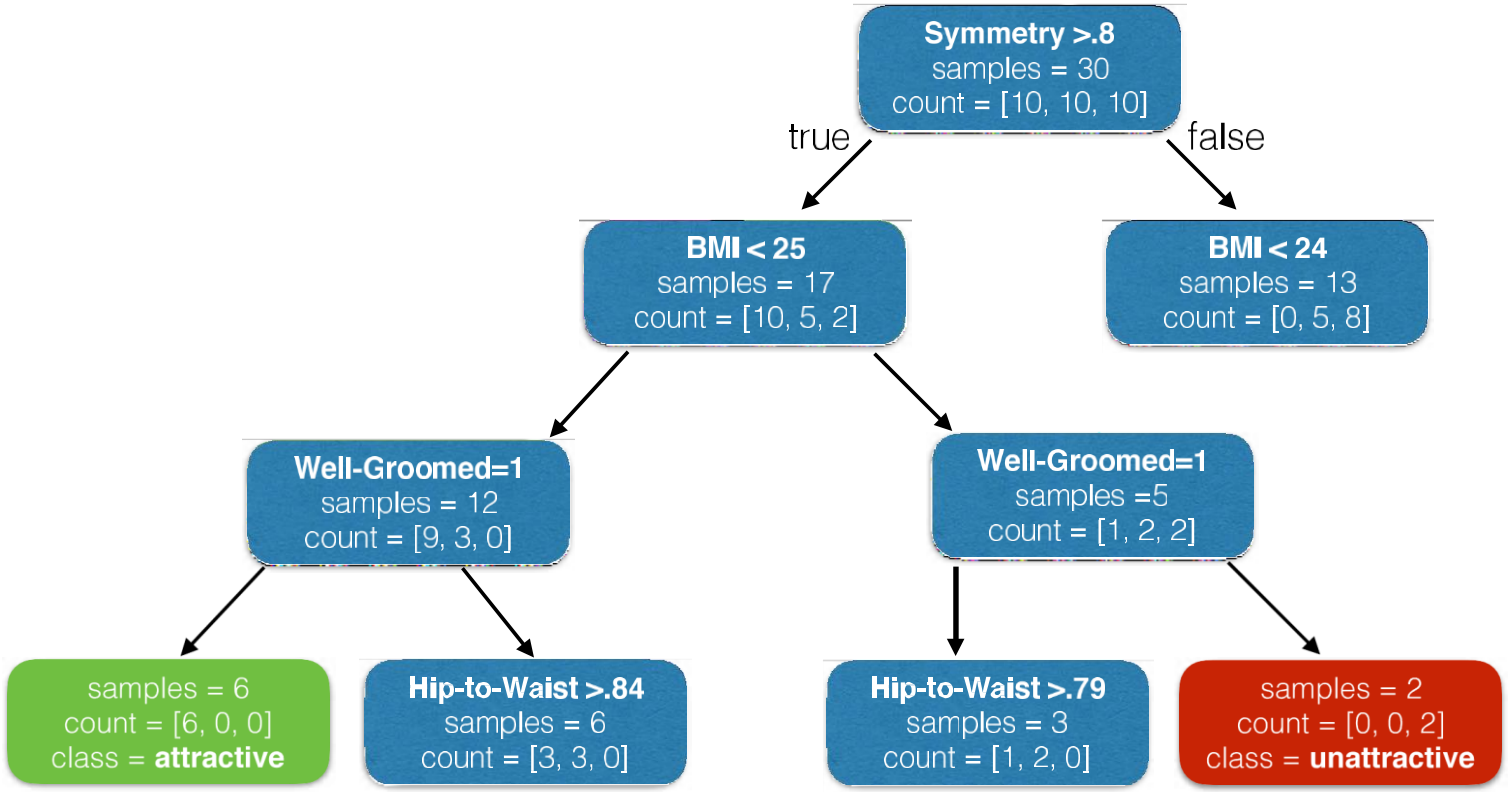
[att, ave, un]


Symmetry >.8
samples = 30
count = [10, 10, 10]
true
false
BMI < 25
samples = 17
count = [10, 5, 2]
BMI < 24
samples = 13
count = [0, 5, 8]

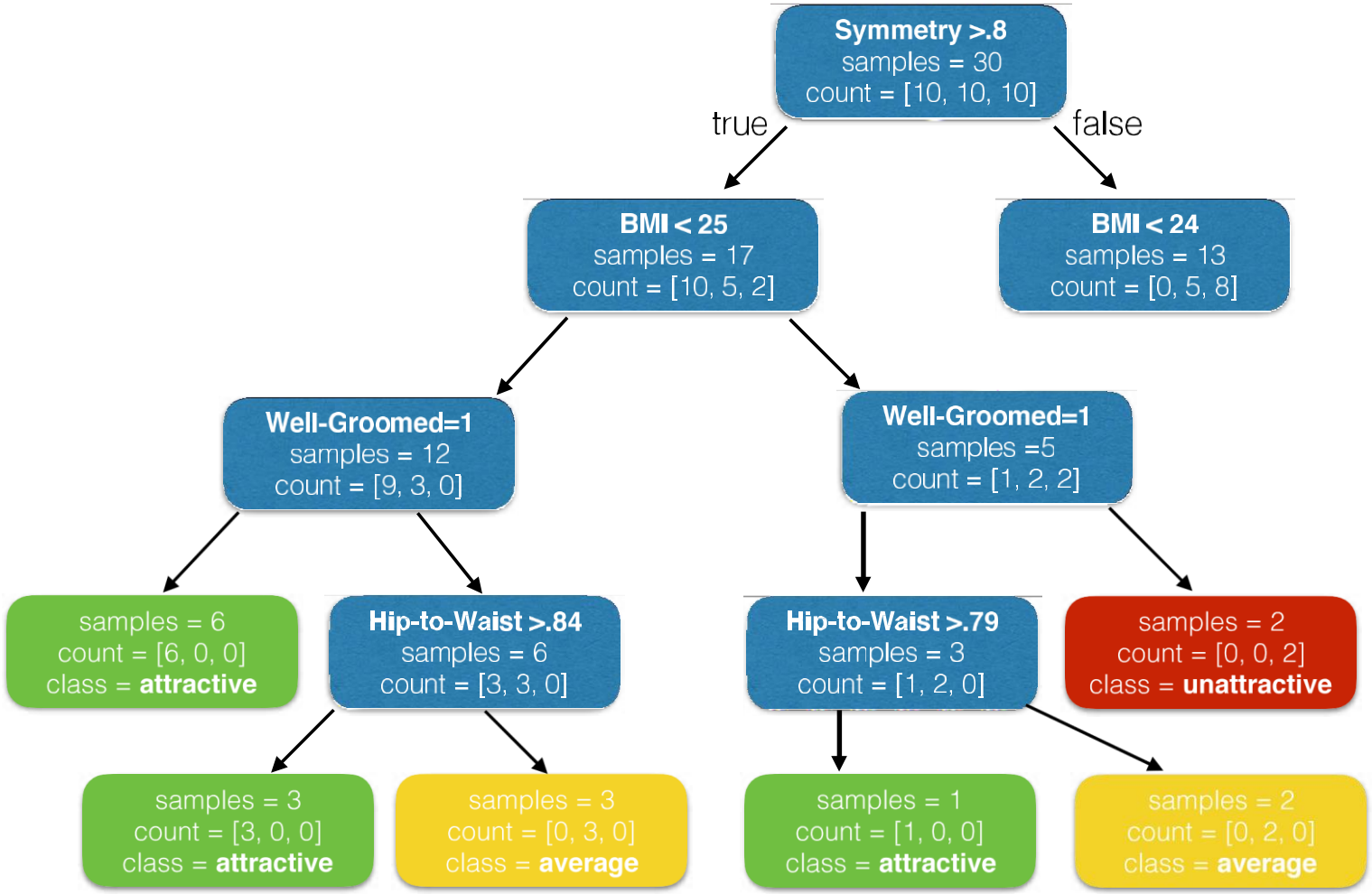
[att, ave, un]


Symmetry >.8
samples = 30
count = [10, 10, 10]
true
false
BMI < 25
samples = 17
count = [10, 5, 2]
BMI < 24
samples = 13
count = [0, 5, 8]
Well-Groomed=1
samples = 12
count = [9, 3, 0]
Well-Groomed=1
samples =5
count = [1, 2, 2]

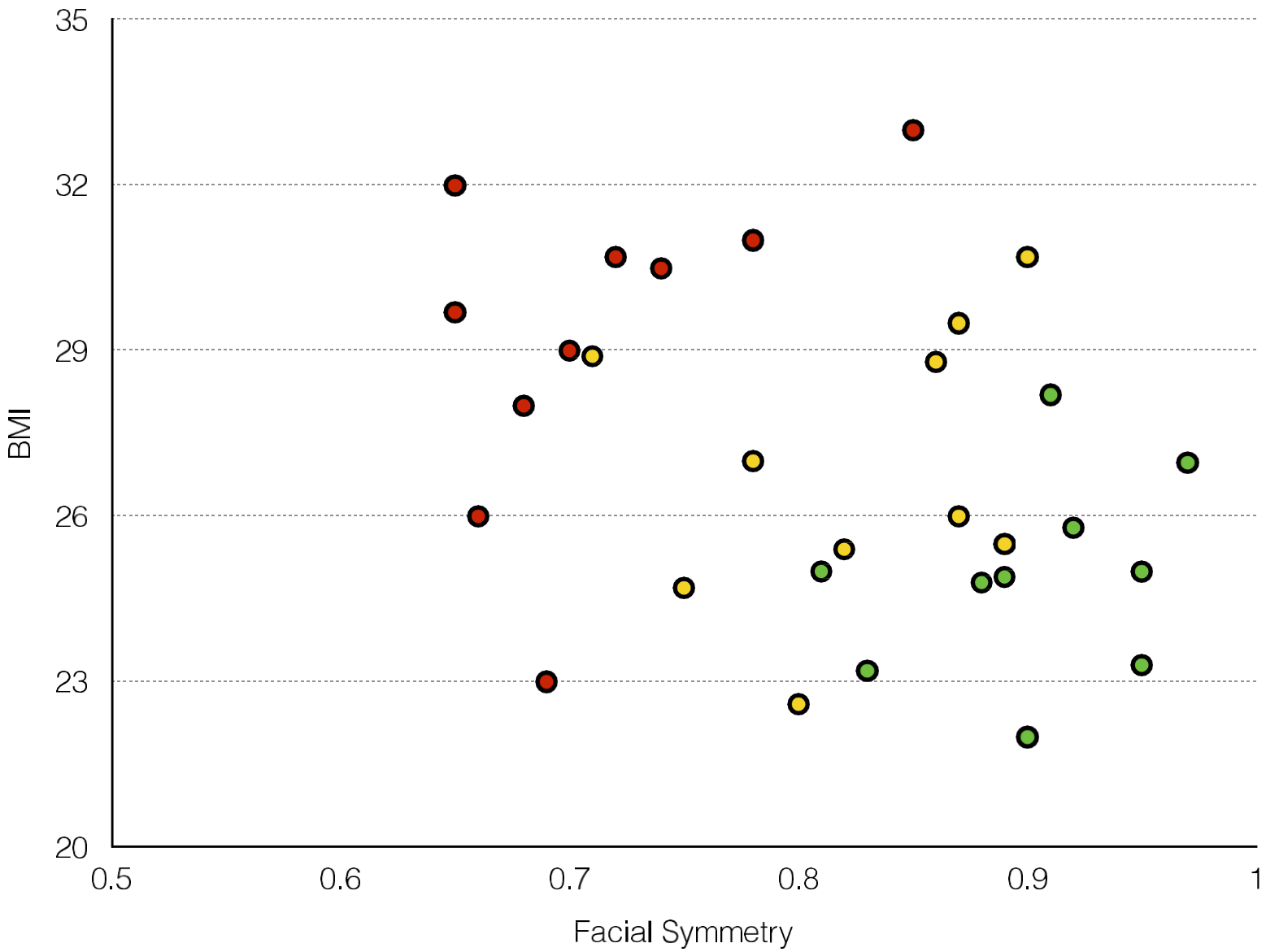
[att, ave, un]

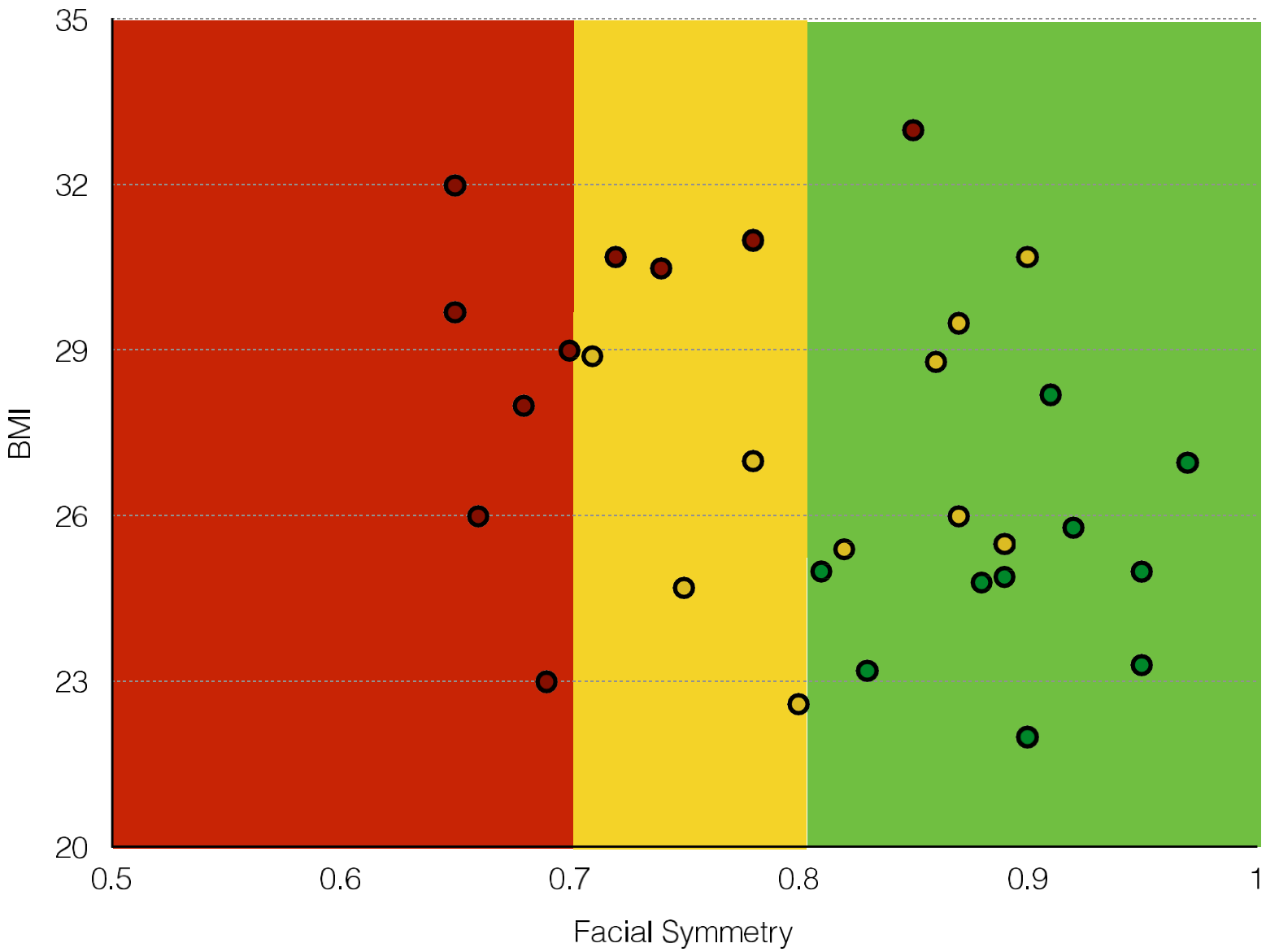

Symmetry >.8
samples = 30
count = [10, 10, 10]
true
false
BMI < 25
samples = 17
count = [10, 5, 2]
BMI < 24
samples = 13
count = [0, 5, 8]
Well-Groomed=1
samples = 12
count = [9, 3, 0]
Well-Groomed=1
samples =5
count = [1, 2, 2]
samples = 6
count = [6, 0, 0]
class = attractive
Hip-to-Waist >.84
samples = 6
count = [3, 3, 0]
Hip-to-Waist >.79
samples = 3
count = [1, 2, 0]
samples = 2
count = [0, 0, 2]
class = unattractive

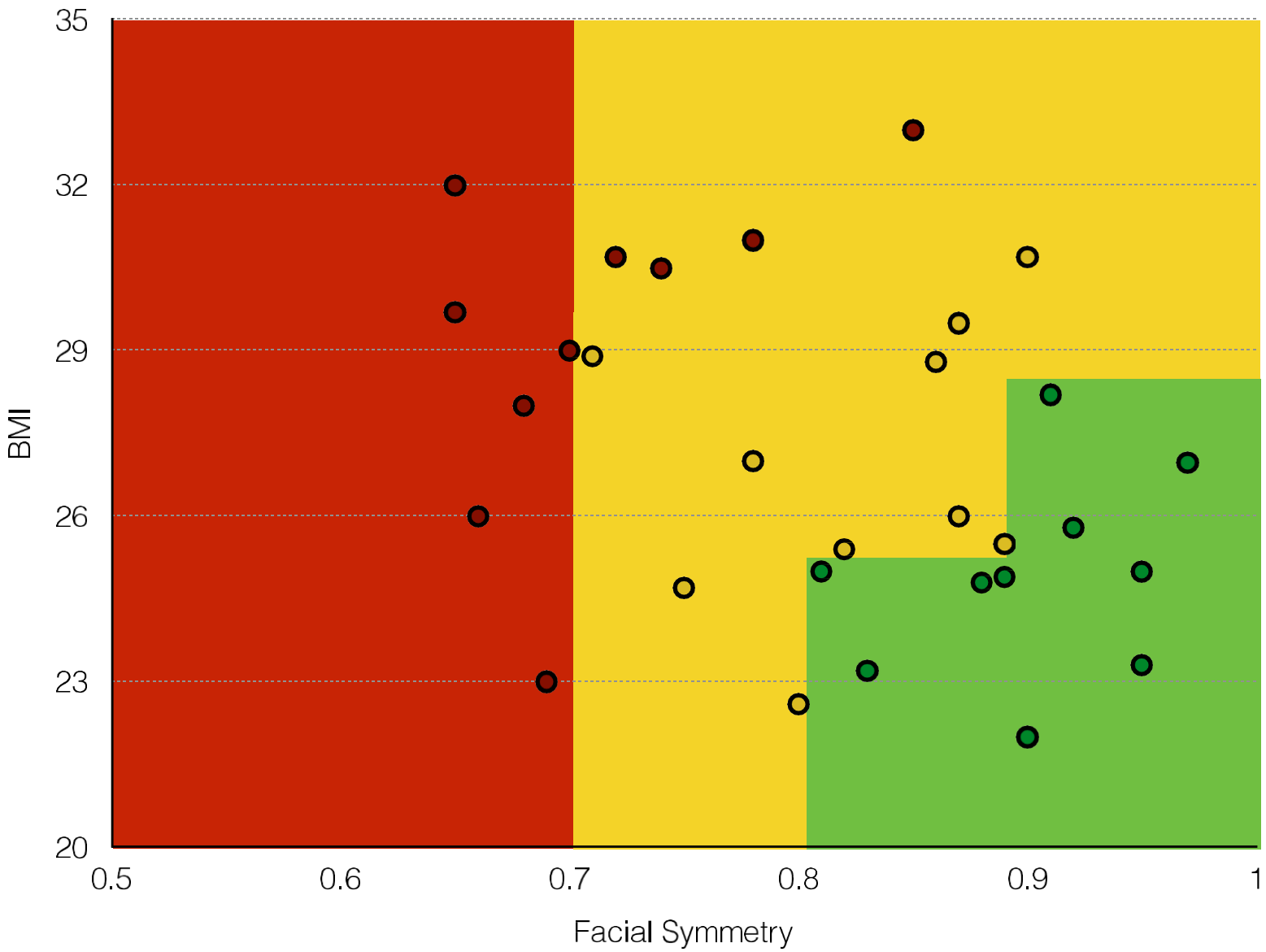
[att, ave, un]

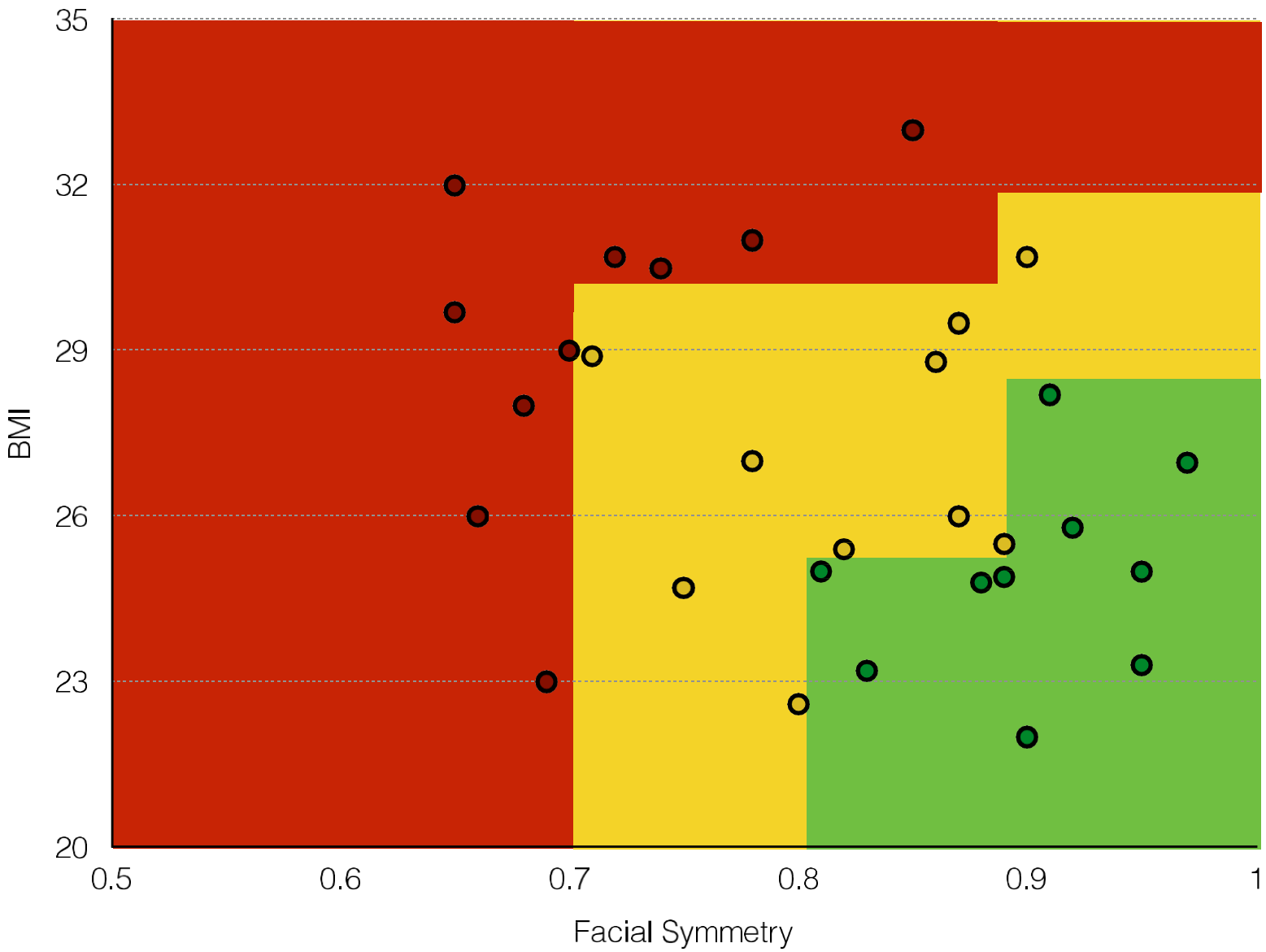

Symmetry >.8
samples = 30
count = [10, 10, 10]
true
false
BMI < 25
samples = 17
count = [10, 5, 2]
BMI < 24
samples = 13
count = [0, 5, 8]
Well-Groomed=1
samples = 12
count = [9, 3, 0]
Well-Groomed=1
samples =5
count = [1, 2, 2]
samples = 6
count = [6, 0, 0]
class = attractive
Hip-to-Waist >.84
samples = 6
count = [3, 3, 0]
Hip-to-Waist >.79
samples = 3
count = [1, 2, 0]
samples = 2
count = [0, 0, 2]
class = unattractive
samples = 3
count = [3, 0, 0]
class = attractive
samples = 3
count = [0, 3, 0]
class = average
samples = 1
count = [1, 0, 0]
class = attractive
samples = 2
count = [0, 2, 0]
class = average

[att, ave, un]


35
32
29
26
23
20
BMI
0.5
0.6
0.7
0.8
0.9
1
Facial Symmetry


35
32
29
26
23
20
BMI
0.5
0.6
0.7
0.8
0.9
1
Facial Symmetry


35
32
29
26
23
20
BMI
0.5
0.6
0.7
0.8
0.9
1
Facial Symmetry


35
32
29
26
23
20
BMI
0.5
0.6
0.7
0.8
0.9
1
Facial Symmetry

K-nearest Neighbor


WORK HARD
NAP HARD

	Number of Relationships	Grade of First Relationship
sample a	3	7
sample b	6	11
sample c	3	9
sample d	5	10

Euclidean Distance

point a = [a1, a2]

point b = [b1, b2]

Euclidean Distance

point a = [a1, a2]

point 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]

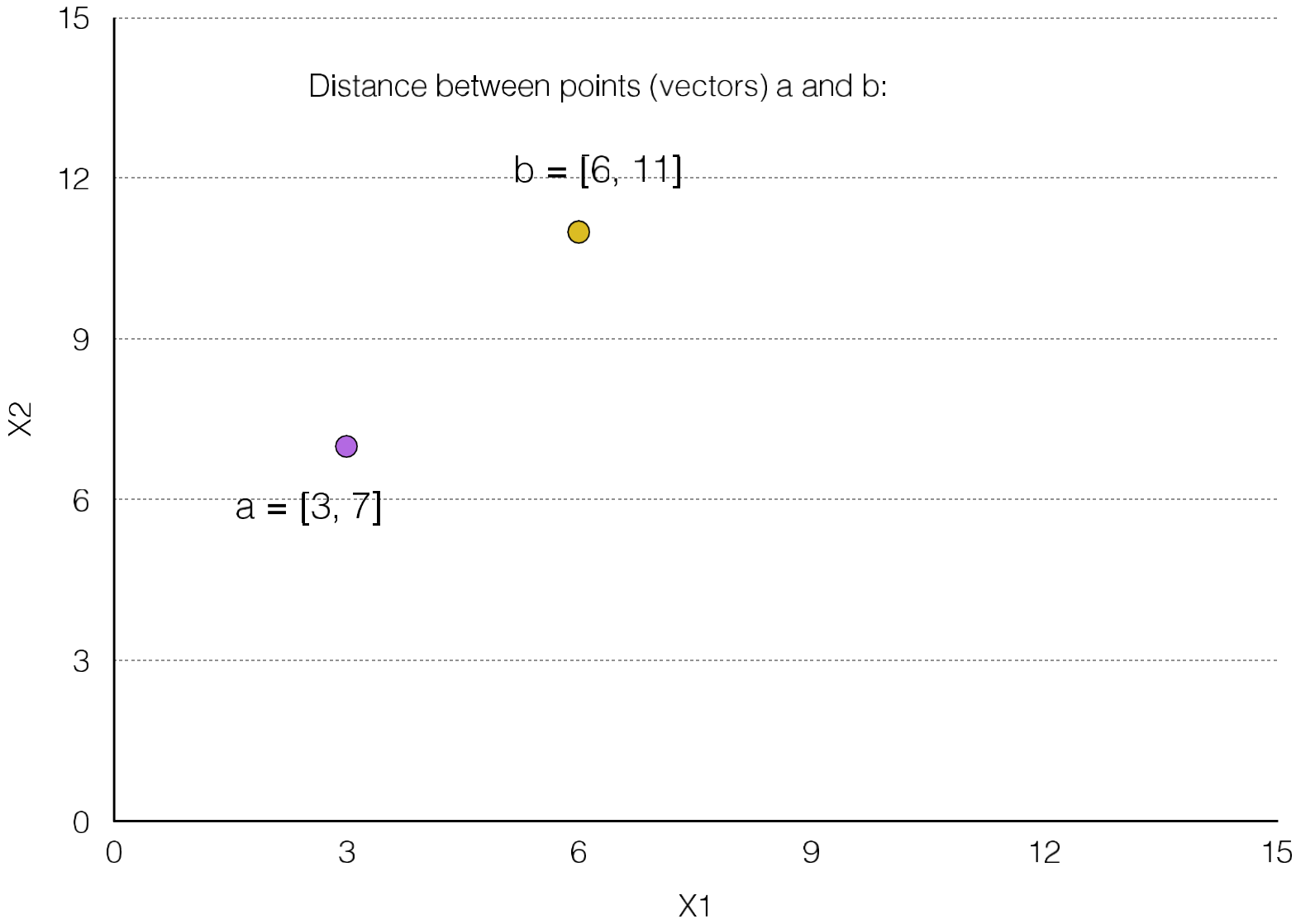
$$\sqrt{(-3)^2+(-4)^2}$$

$$\sqrt{9+16}$$

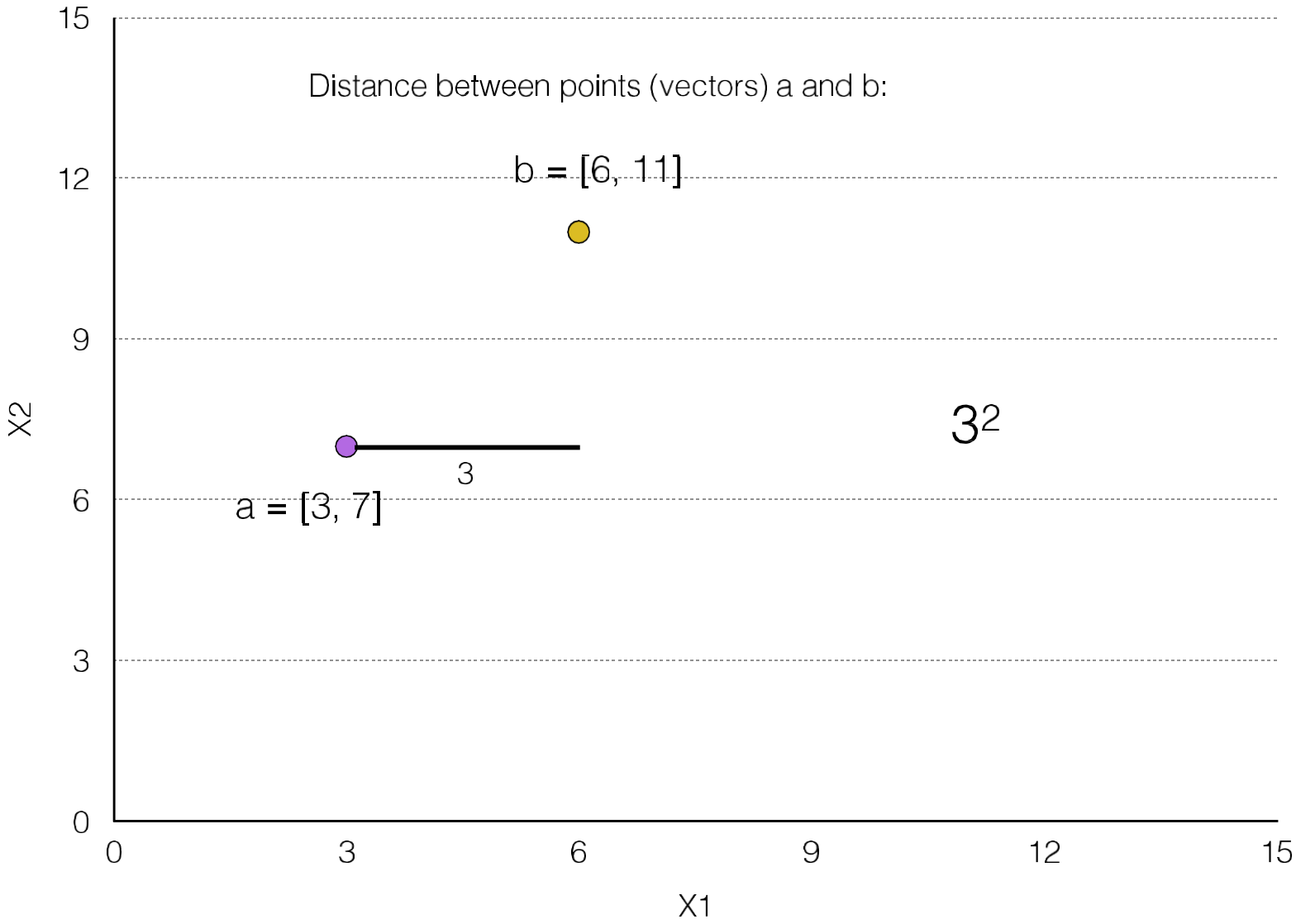
Distance between points (vectors) a and b:

$$\sqrt{25}=5$$

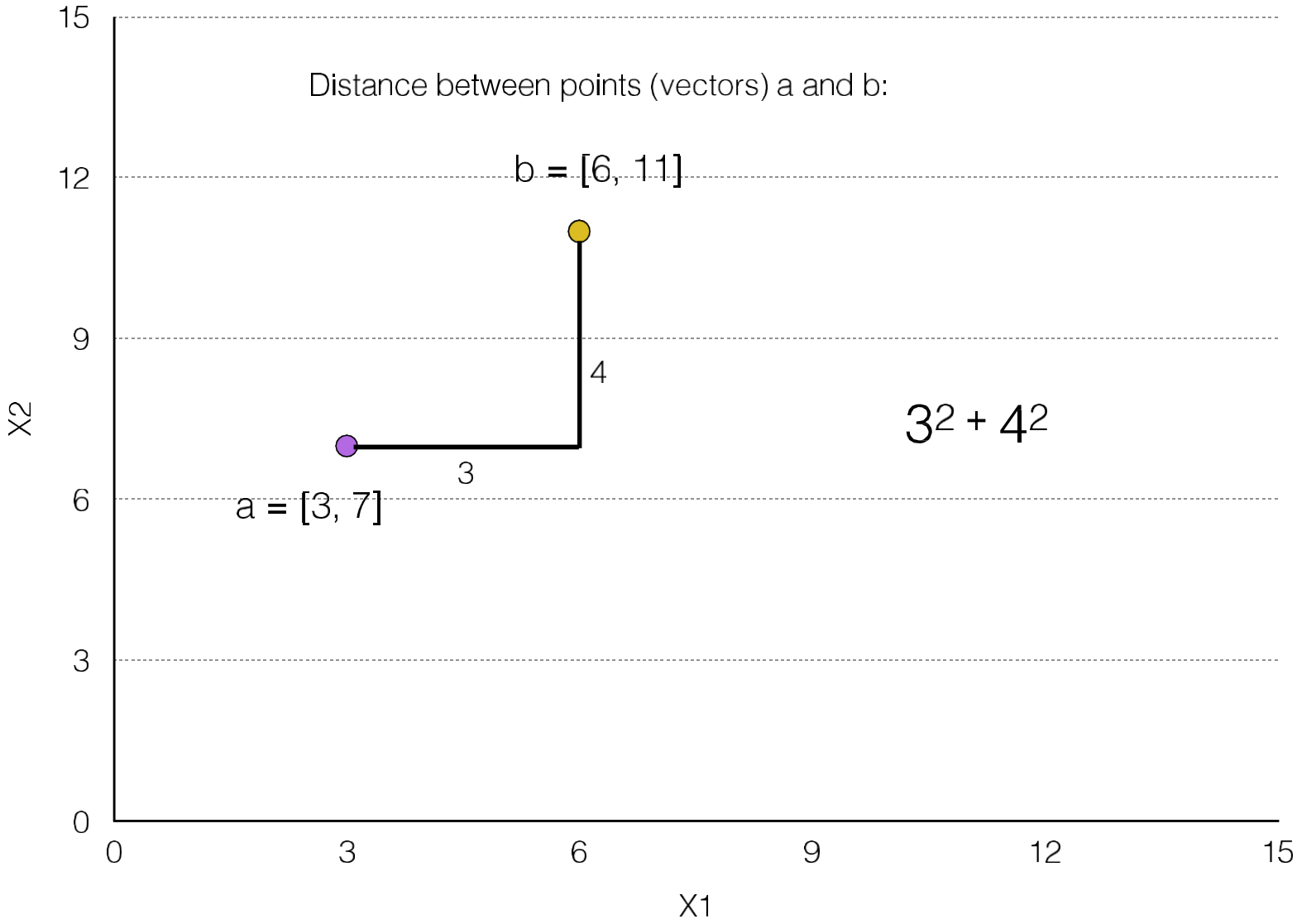
Euclidean Distance



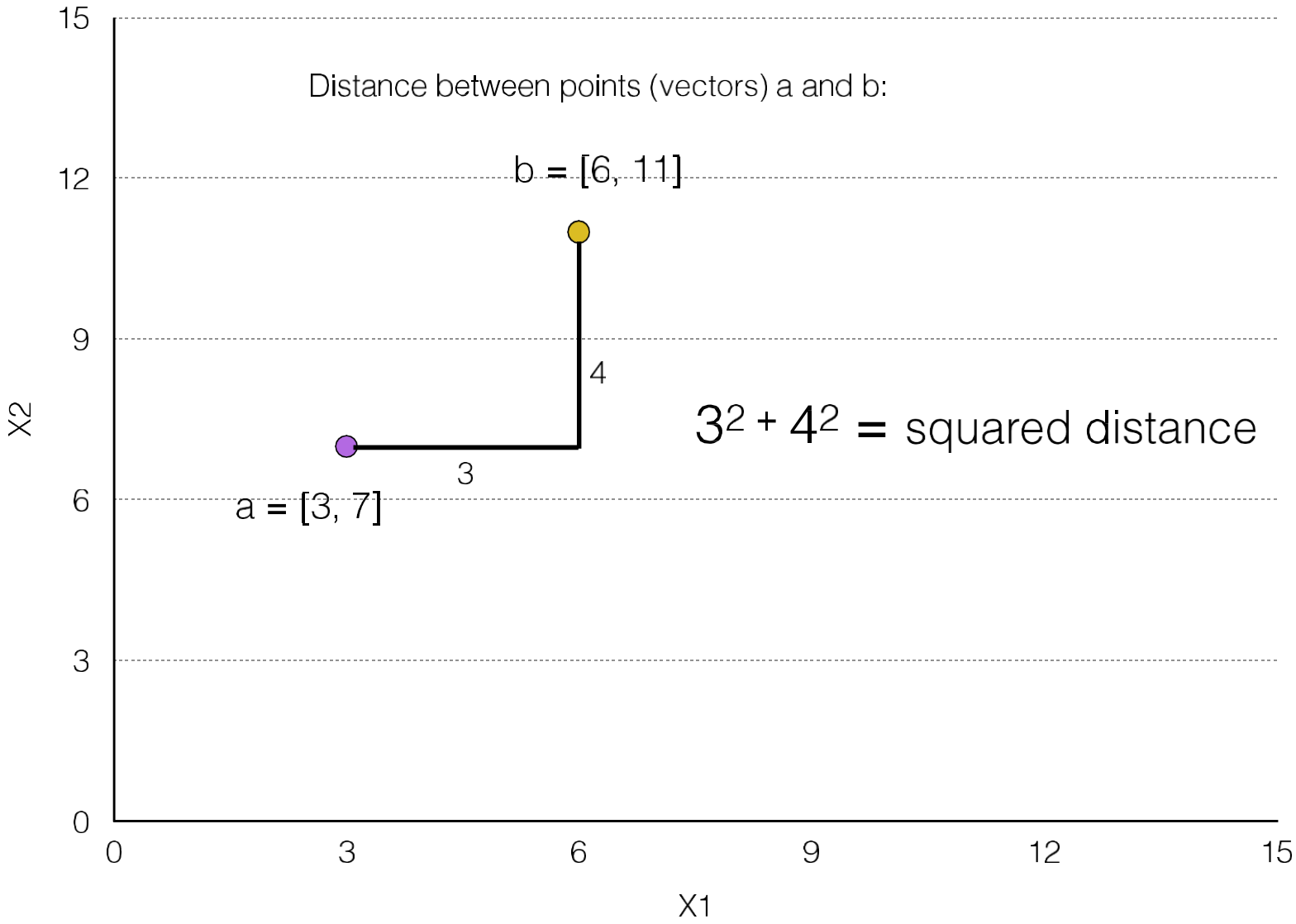
Euclidean Distance



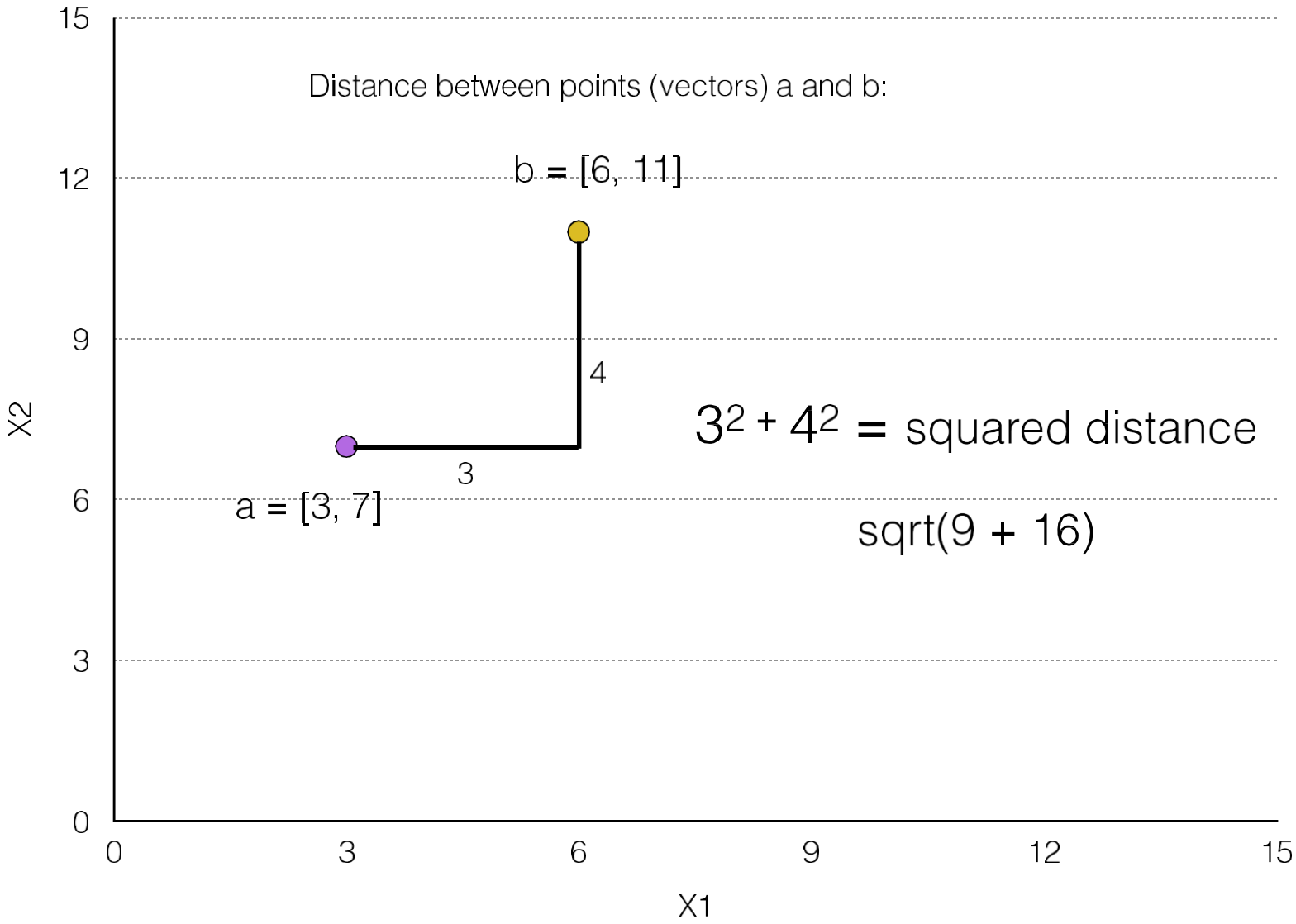
Euclidean Distance



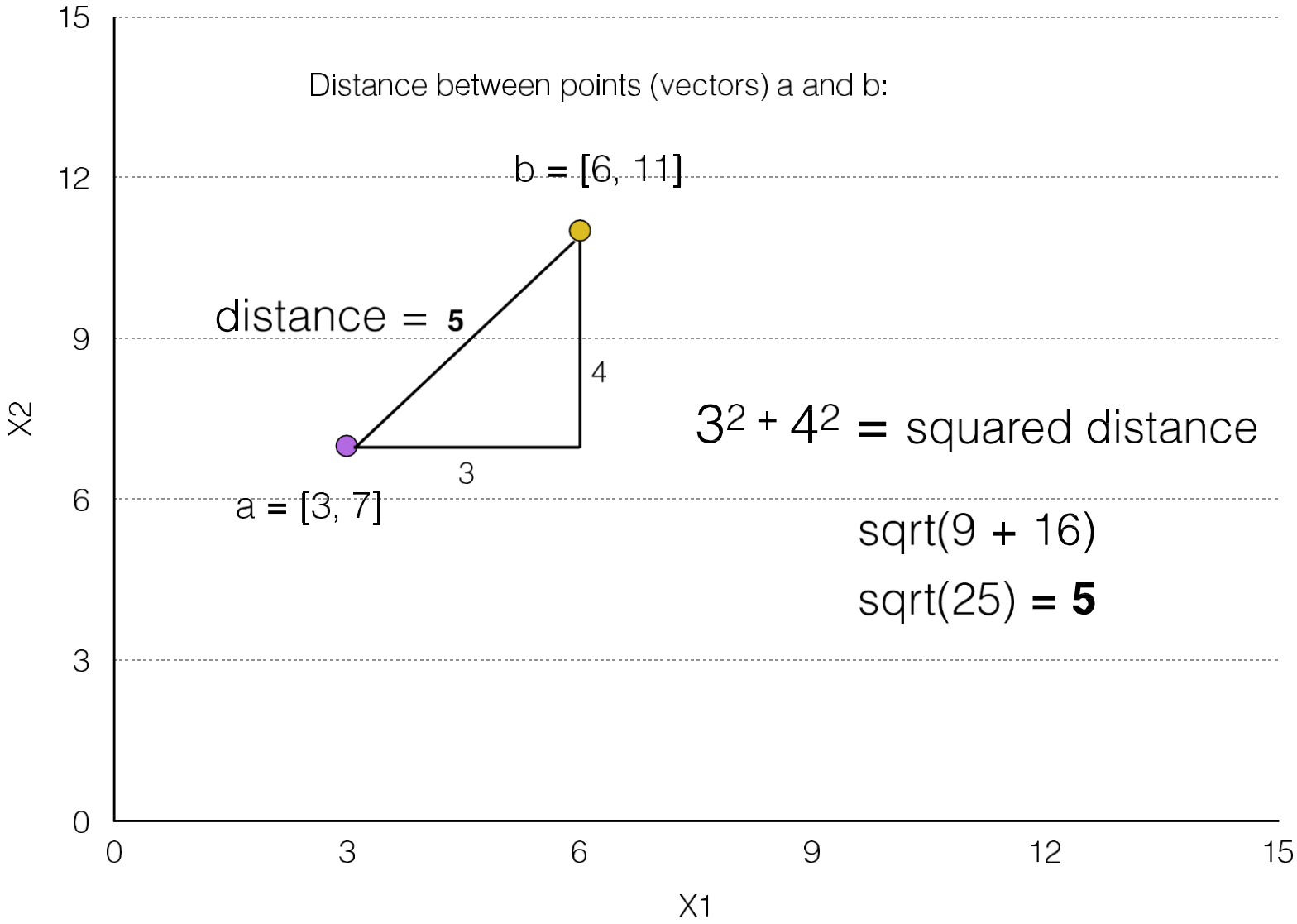
Euclidean Distance



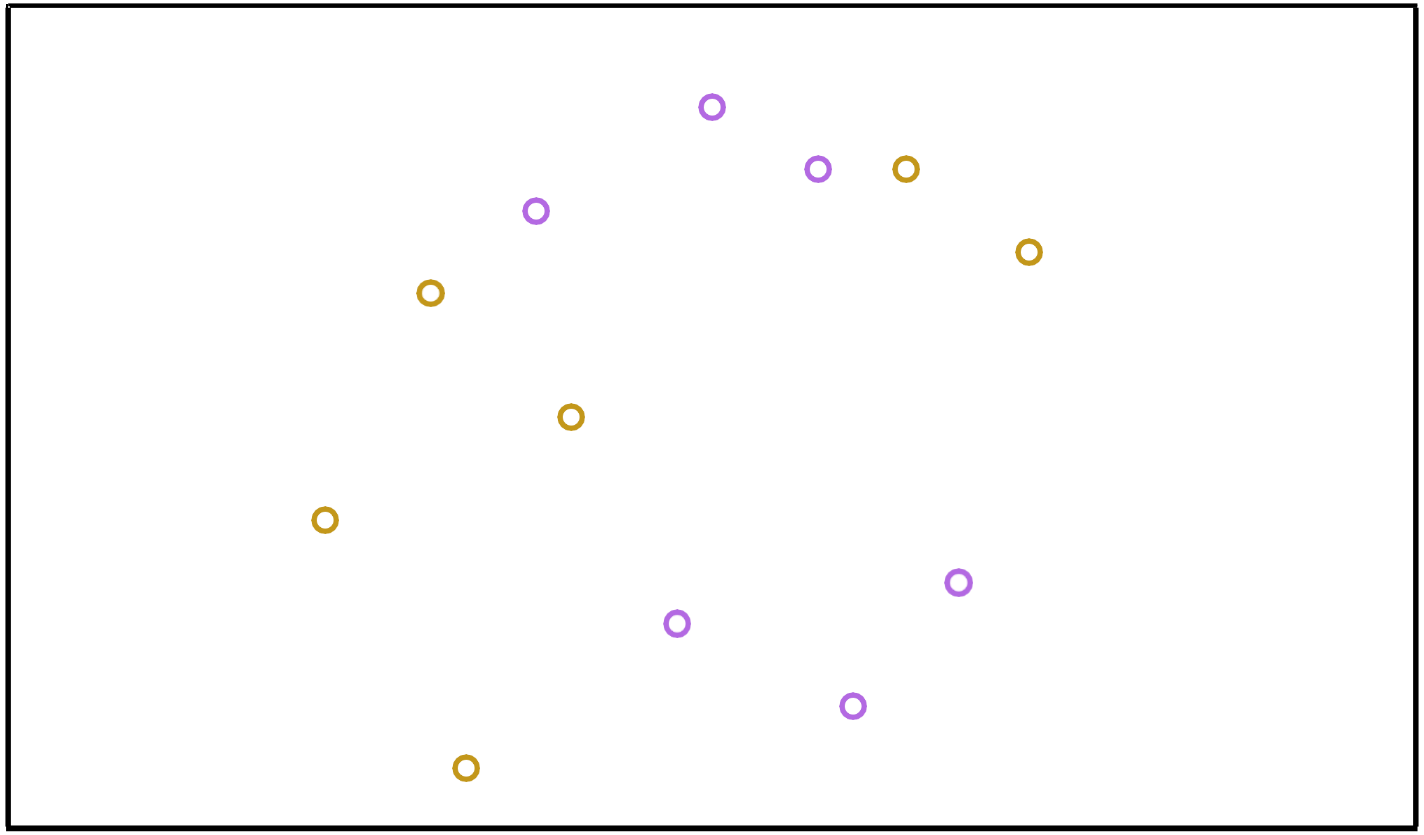
Euclidean Distance



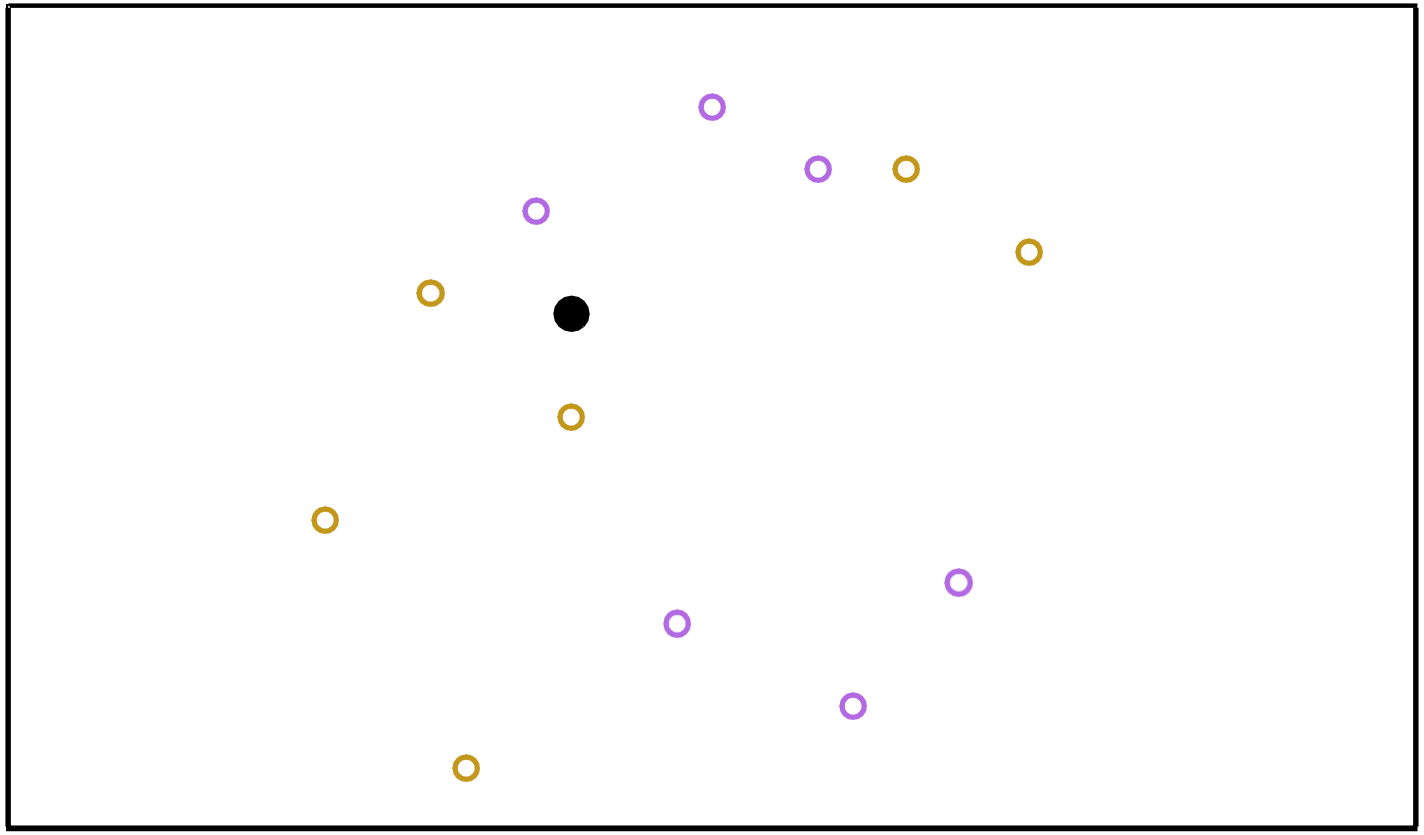
Euclidean Distance



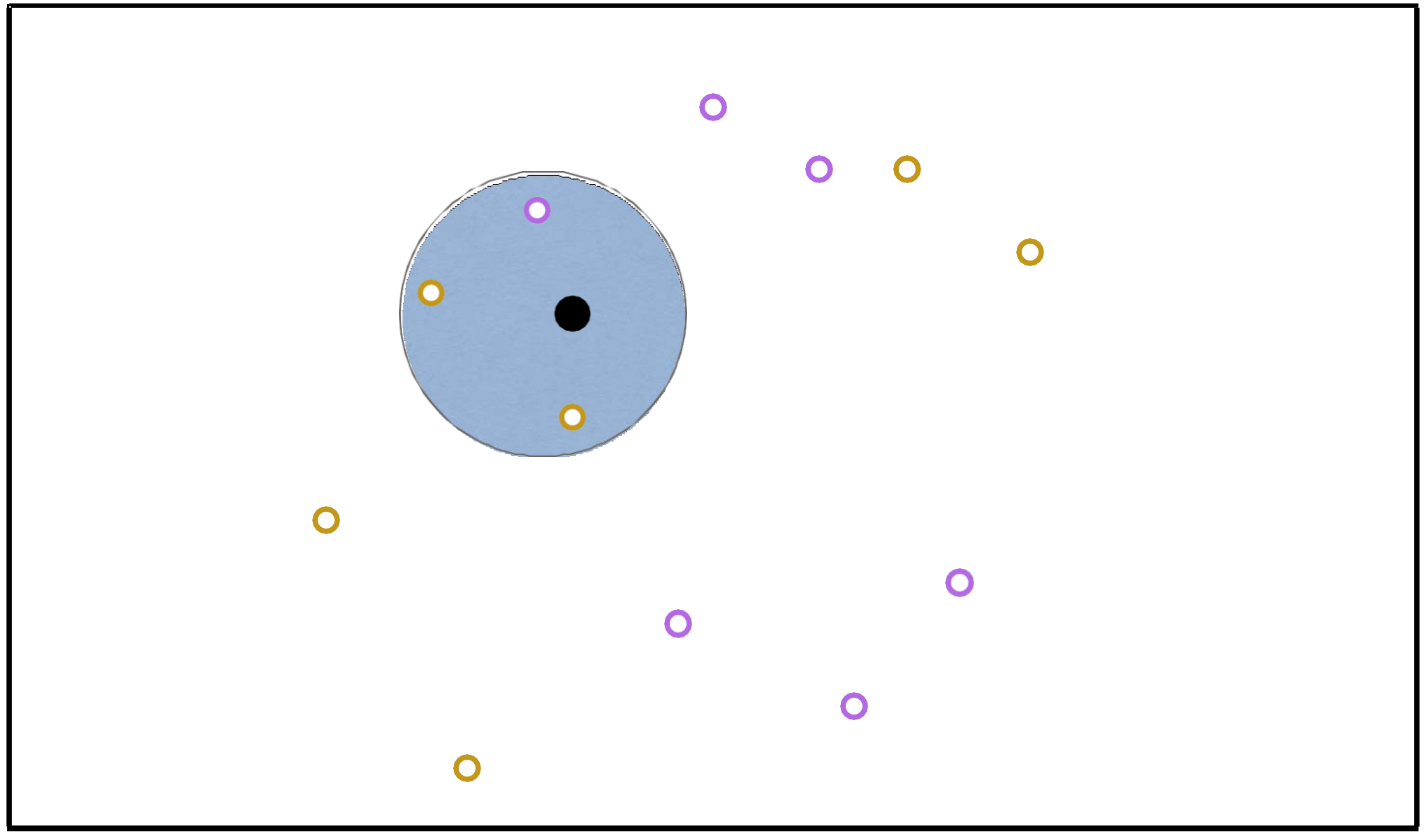
K-nearest Neighbor



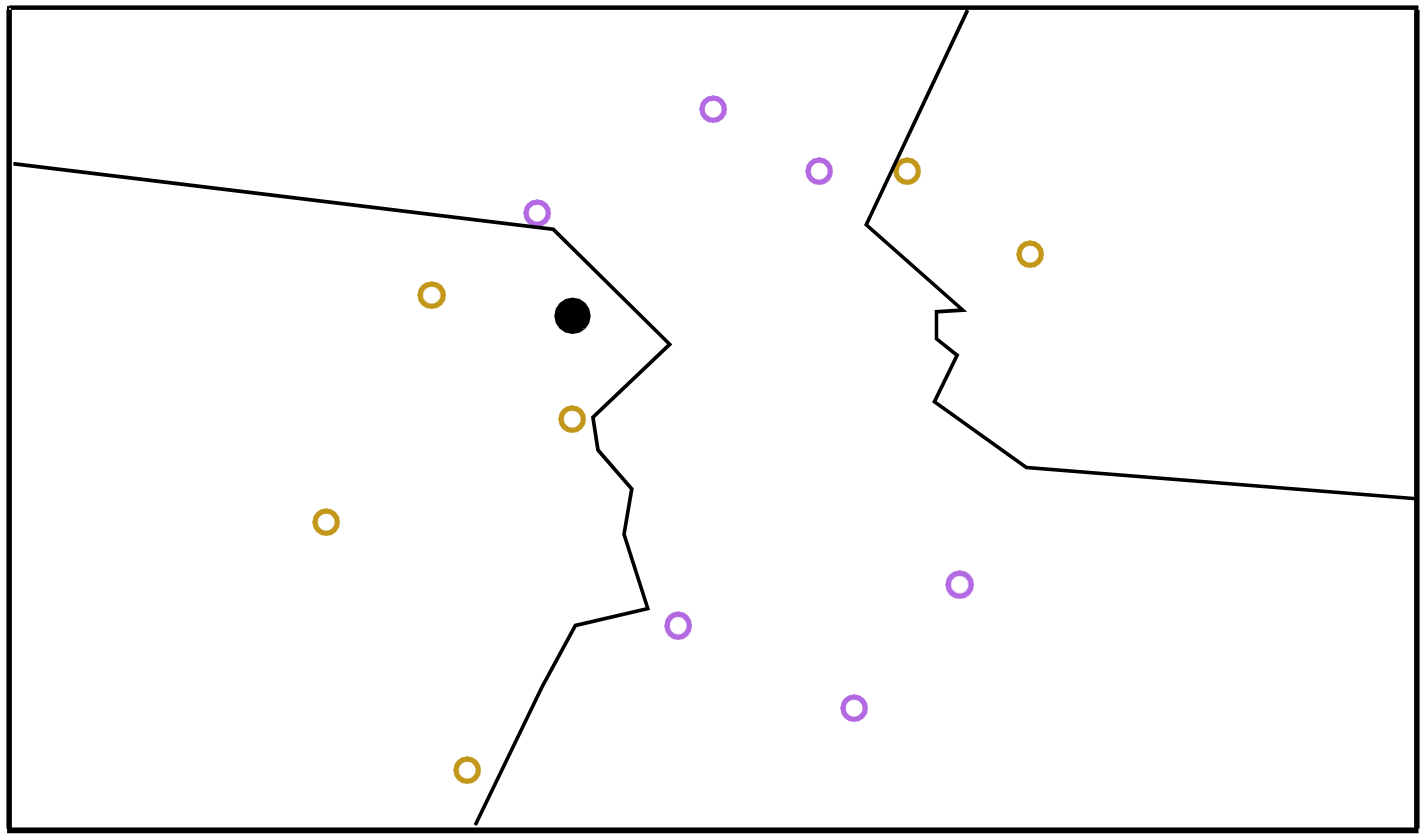
K-nearest Neighbor



K = 3



K = 3



Stays together

Breaks up

K = 3