

Week 1: Introduction

Matthew Caldwell

COMP0088 Introduction to Machine Learning • UCL Computer Science • Autumn 2023

Course Outline

1. Introduction

- 2. Linear Models
- 3. Non-Parametric Models
- 4. Linear Models Revisited
- 5. Neural Networks
- 6. More Neural Networks

Supervised
Learning

- 7. Clustering & Dimensionality Reduction
- 8. Mixture Models & Expectation Maximisation

Unsupervised
Learning

- 9. Deep Learning Applications & Pitfalls
- 10. Fun & Games

Prerequisites

- Linear Algebra
 - Calculus
 - Probability & Statistics
 - Python Programming
- NB: This is not a course in any of these subjects.
Our focus is **pragmatic**: these are tools we need to get stuff done.

Format

- Pre-recorded lectures
 - 1–2 hours in ~15 minute chunks
- In person recap / Q&A
 - Friday, 17:00–18:00
- In person practical labs
 - Friday 10:00–12:00 / 12:00–14:00
- Optional drop in – “office hour”
 - Friday 15:00–16:00, online

► Video and slides available the previous week
► Watch at your own pace
► ...but preferably **before** lab sessions!

► Ask questions live or in advance on forum
► Latter might get priority

► These are the heart of the module!
► Exercises available the previous week

► One-to-one or small group
► For “private” questions
► Q&A session better for most topics

Assessment

- Exam (100% overall)
 - Online
 - 3 hours
 - Open book
 - Handwrite or type – you may also need to sketch some answers
 - We will work through some practice questions towards the end of term
 - Fair warning: the exam is not easy, but we'll do our best to make it fair

Ethos

- People learn in different ways
- You are responsible for your own engagement
 - If you can't stand to watch the videos, at least read the slides
 - Whatever you do, do it before Friday!
- ML is a practical discipline mediated by programming
 - Doing the practical exercises will really help your understanding
 - If you shun programming you won't get the most from ML
 - But: there is no coding in the exam

Week 1 Recap

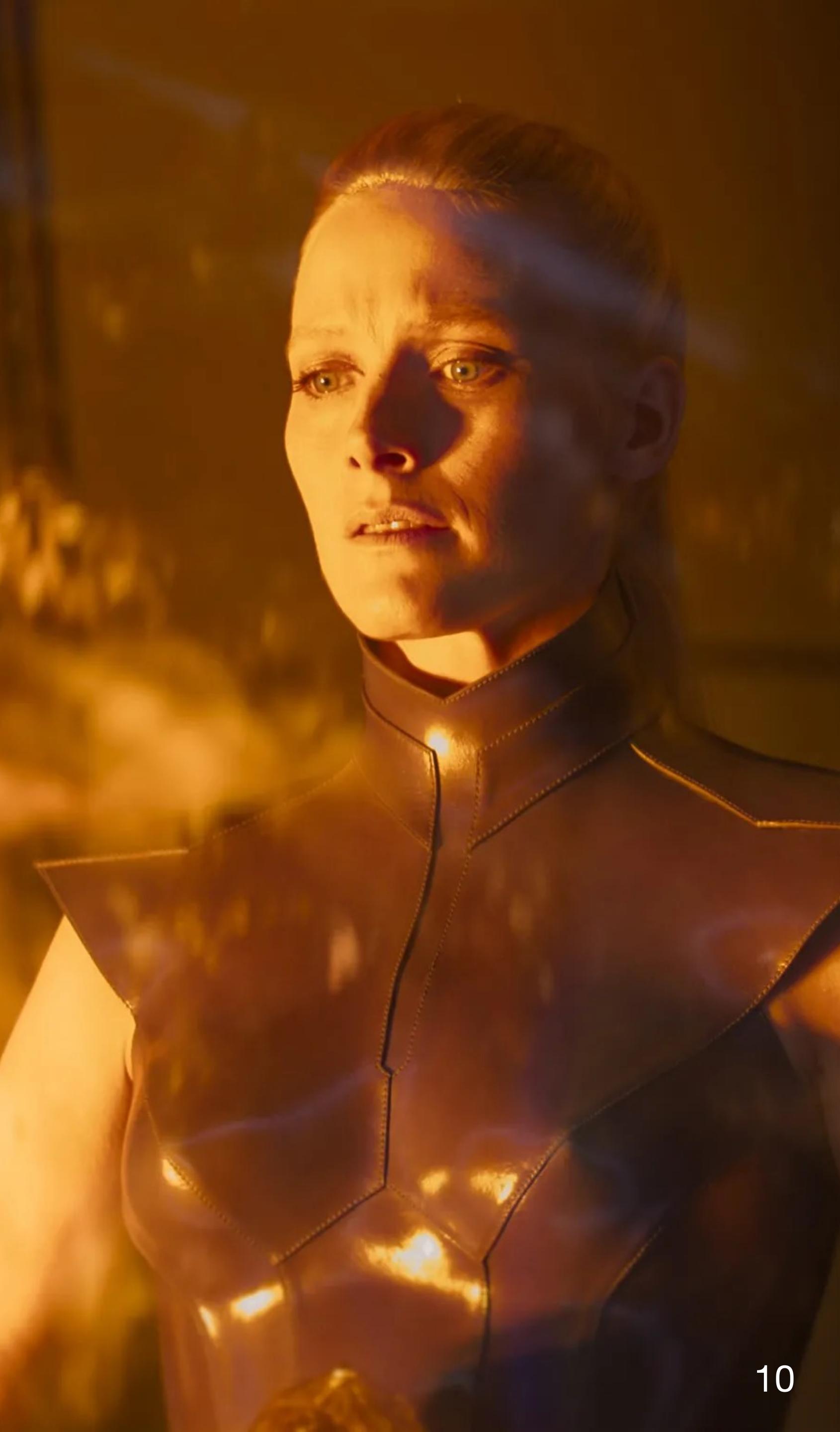
The Three Laws of Machine Learning

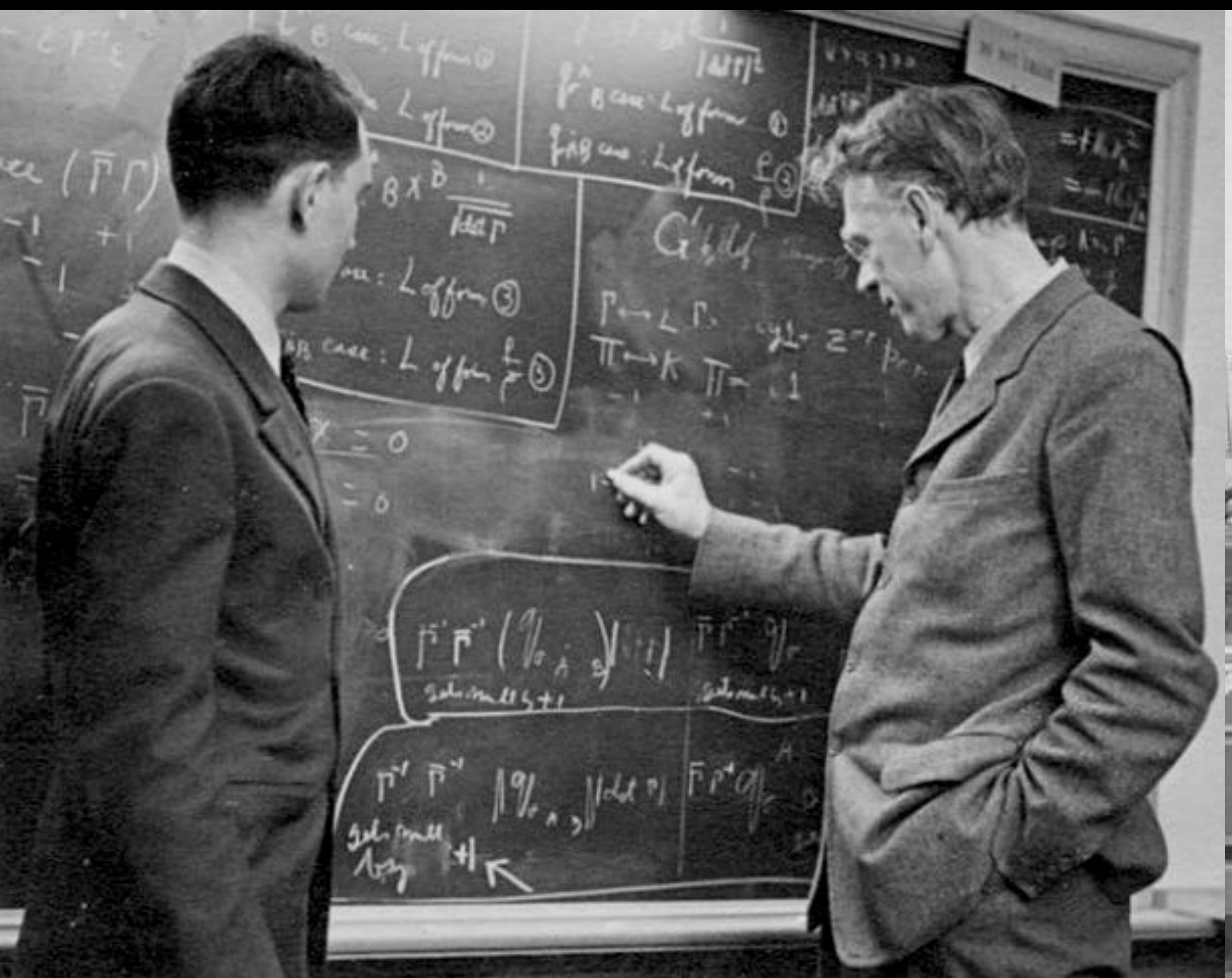
The Zeroth Law of Machine Learning

**Laws are just models
and all models are wrong**

The First Law of Machine Learning

Machine Learning is parameter fitting





learn | lə:n |

verb (past and past participle **learned** | lə:nt, lə:nd | or mainly British **learnt** | lə:nt |) [with object]

1 gain or acquire knowledge of or skill in (something) by study, experience, or being taught: *they'd started learning French* | [with infinitive] : *she is learning to play the piano* | [no object] : *we learn from experience.*

- commit to memory: *I'd learned too many grim poems in school.*
- become aware of (something) by information or from observation: [with clause] : *I learned that they had eaten already* | [no object] : *the trading standards office learned of the illegal network.*

2 archaic or informal teach (someone): 'That'll learn you,' he chuckled | [with object and infinitive] : *we'll have to learn you to milk cows.*

ChatGPT

Write a limerick about Jeremy Bentham

**Jeremy Bentham, a thinker so bright,
In ethics, he shed a clear light.**

**Utilitarian he'd preach,
The greatest good within reach,
For happiness, he'd gladly fight!**

MidJourney

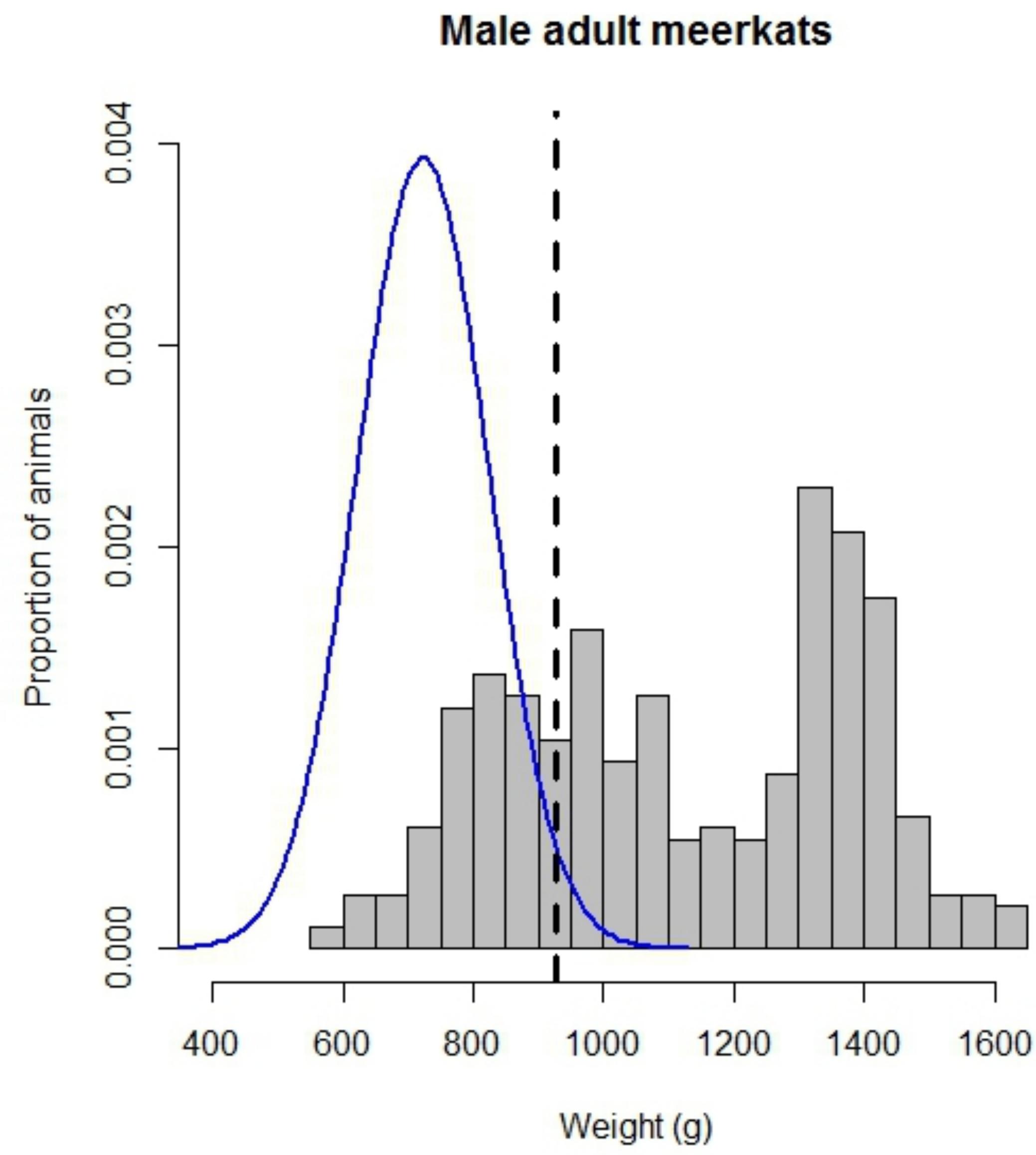
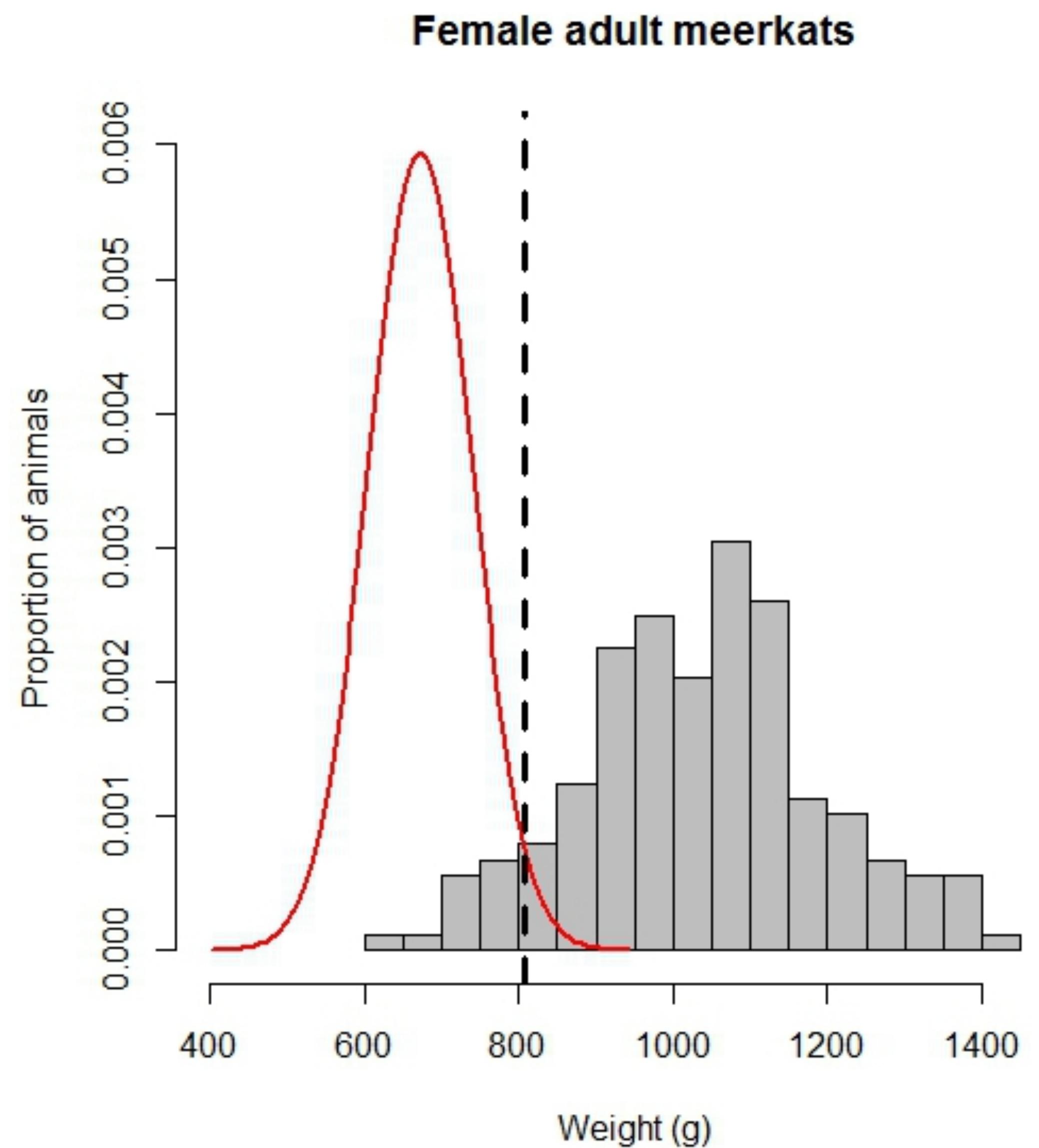
**/imagine Boris Johnson appearing
on Strictly Come Dancing after
leaving politics in disgrace**



The First Law of Machine Learning

Machine Learning is parameter fitting

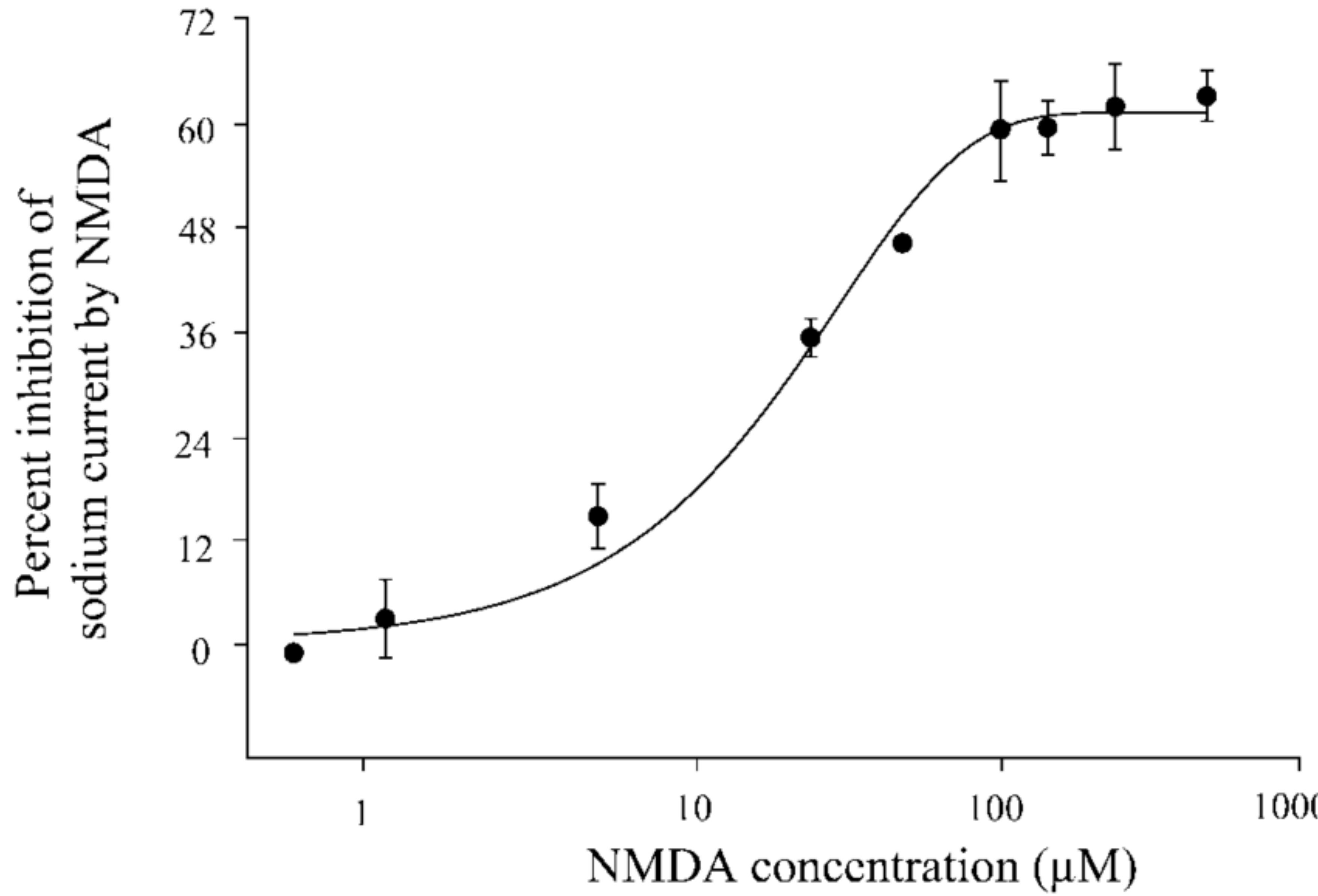


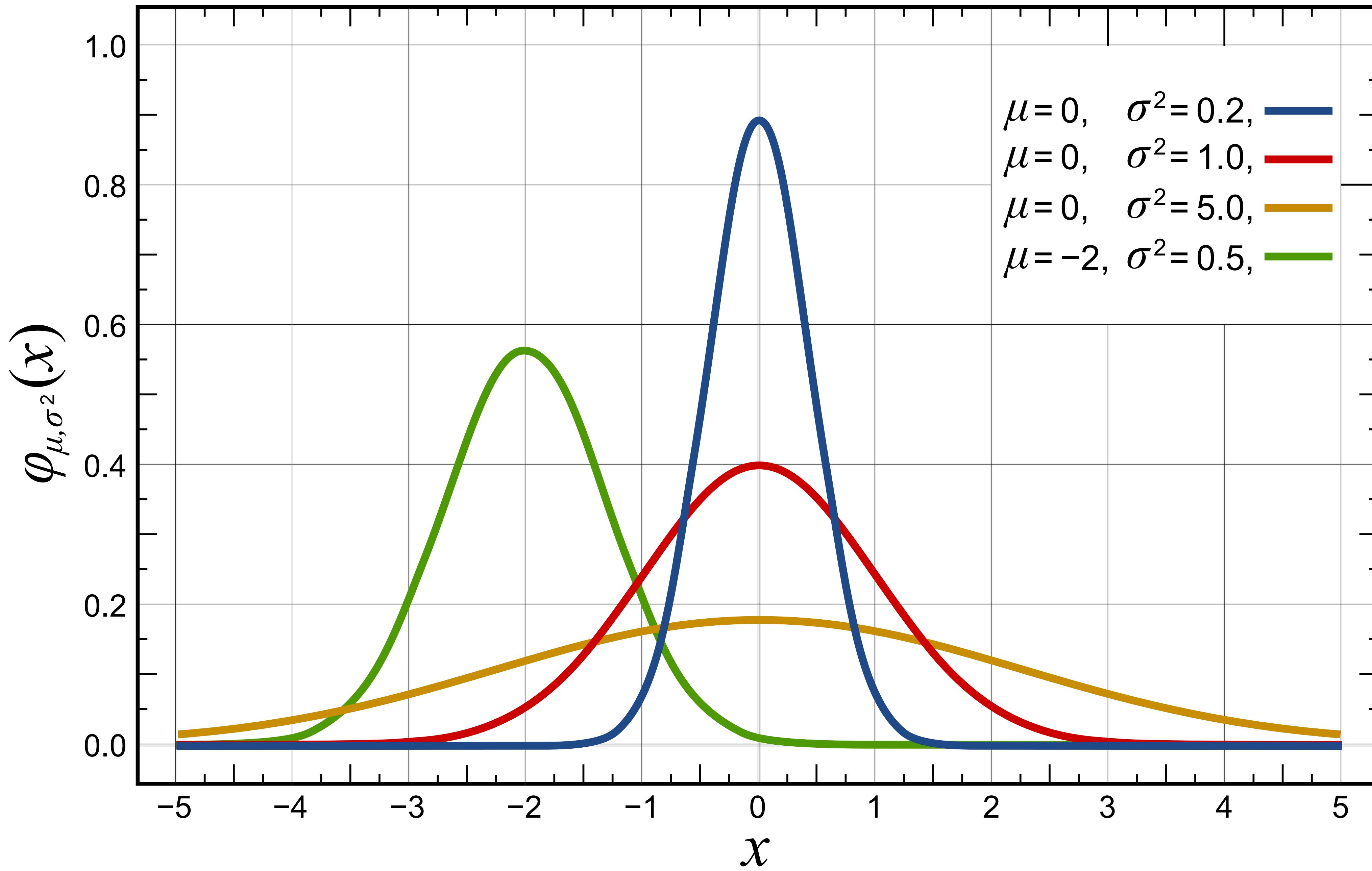


Higher average incomes are associated with higher life expectancy

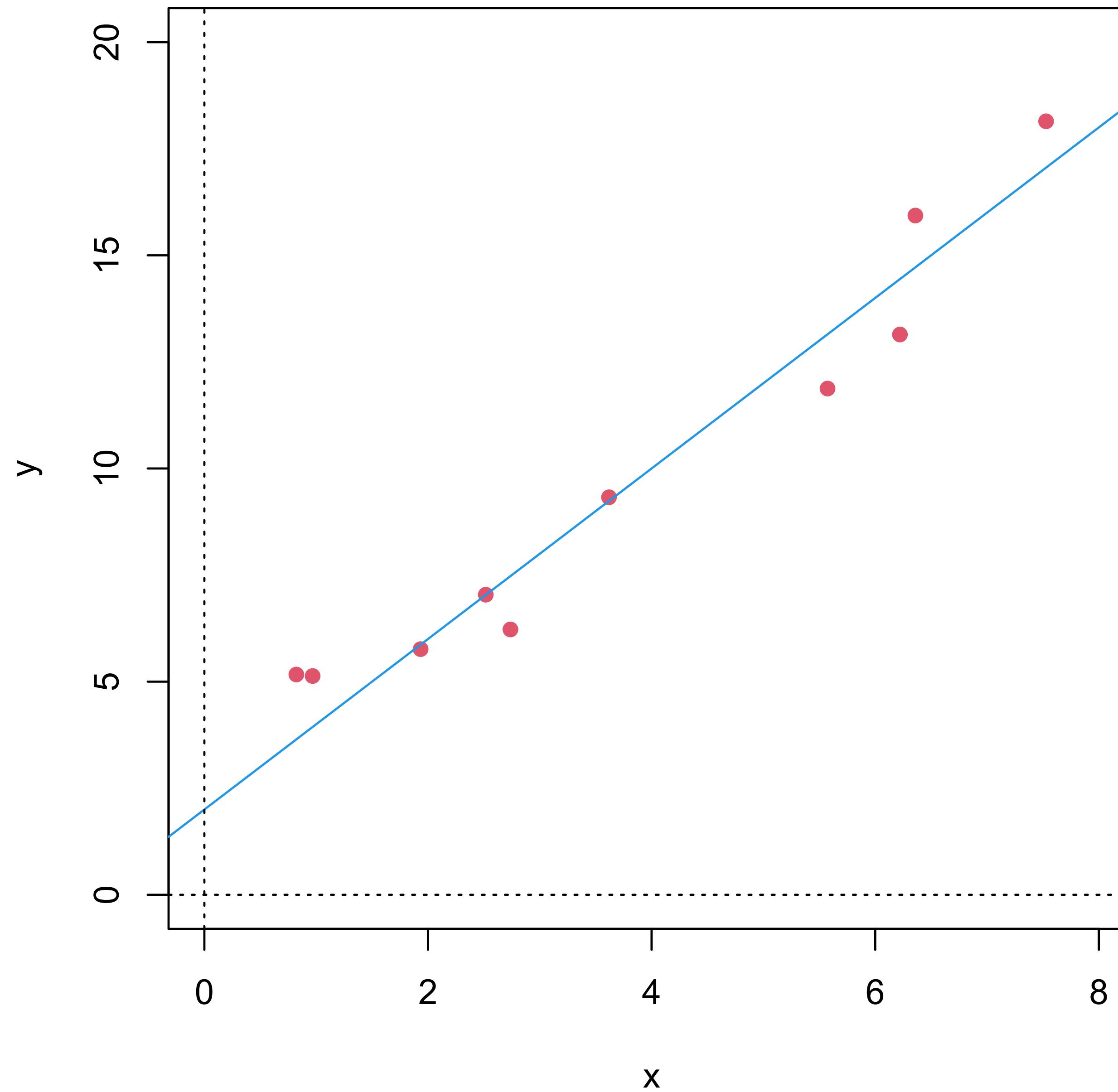
Relationship between life expectancy and net household income for men and women, by neighbourhood: England, 2016–20







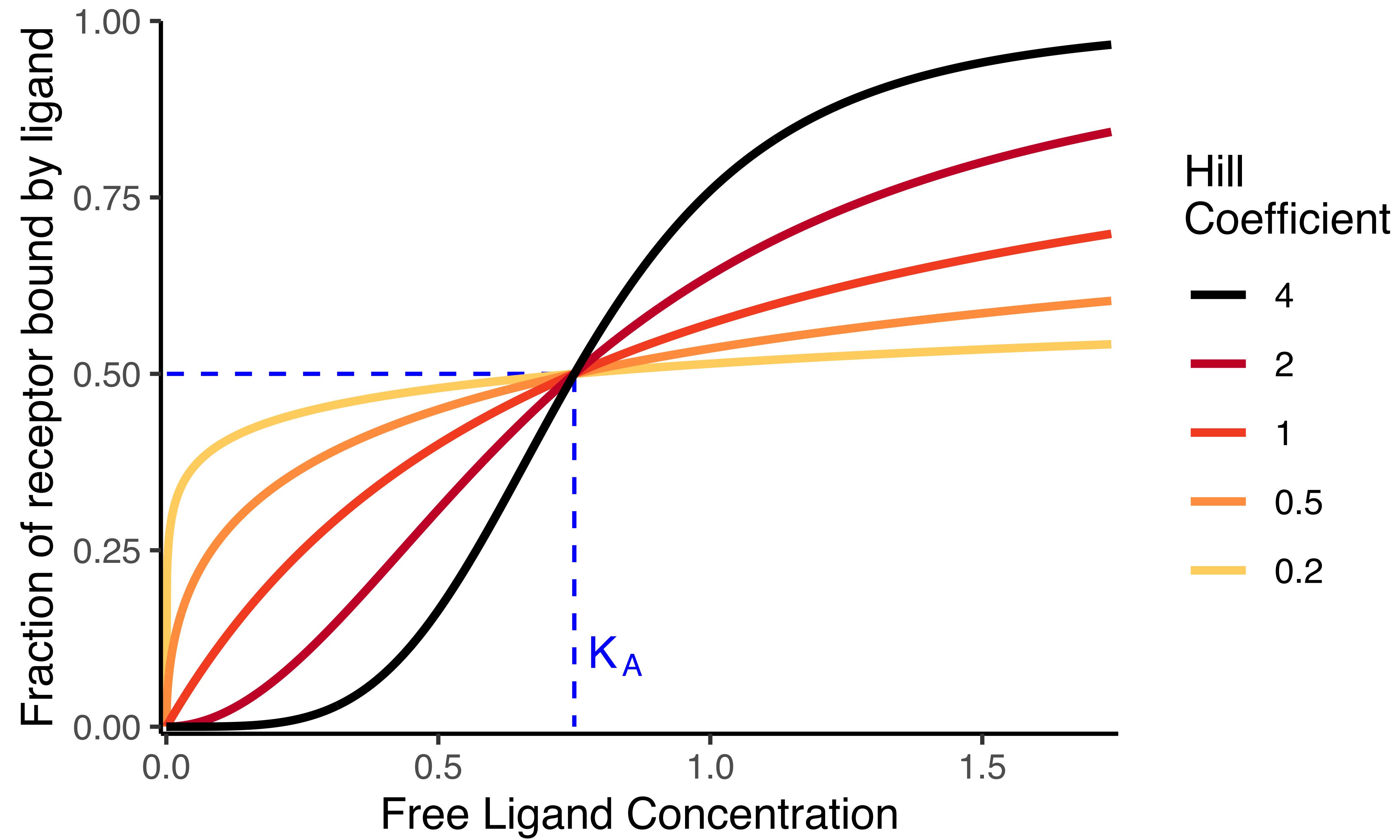
Simple Linear Model



$$y = ax + b$$

$a \rightarrow \text{slope}$

$b \rightarrow \text{intercept}$

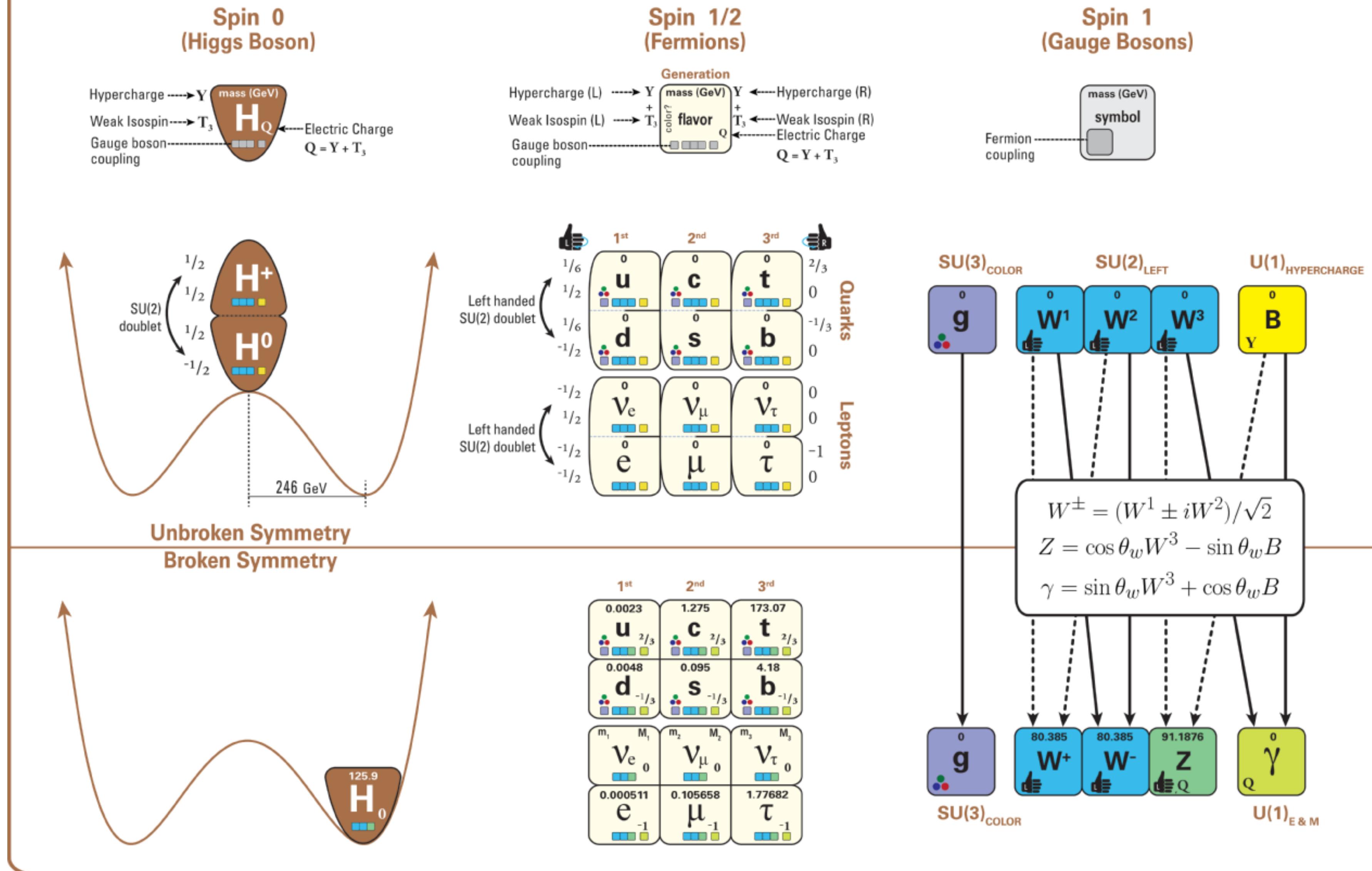


$$C(S_t, t) = N(d_+)S_t - N(d_-)Ke^{-r(T-t)}$$

$$d_+ = \frac{1}{\sigma\sqrt{T-t}} \left[\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t) \right]$$

$$d_- = d_+ - \sigma\sqrt{T-t}$$

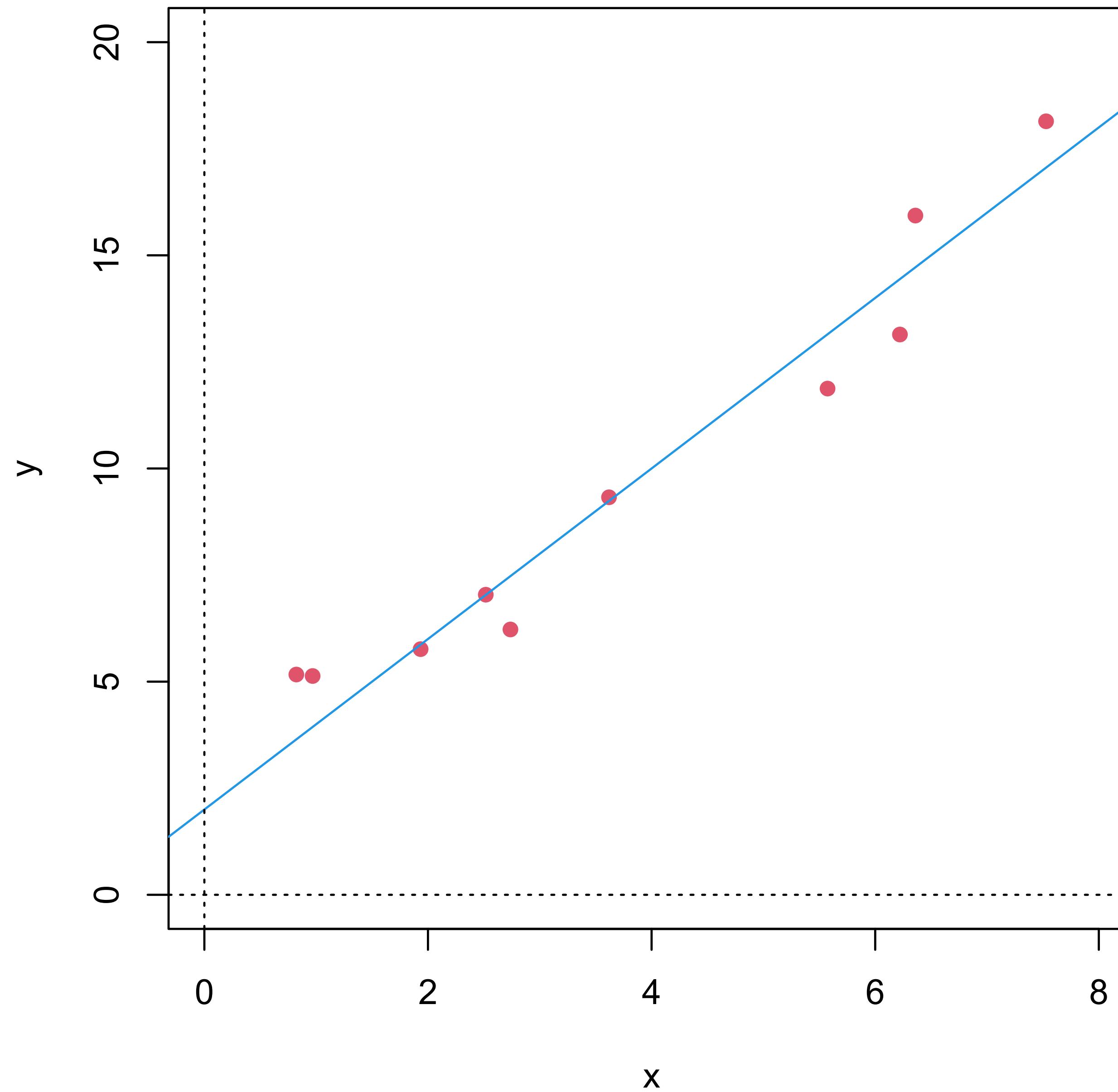
The Standard Model of Particle Physics



ChatGPT: billions of parameters

MidJourney: $10^8 - 10^{11}$

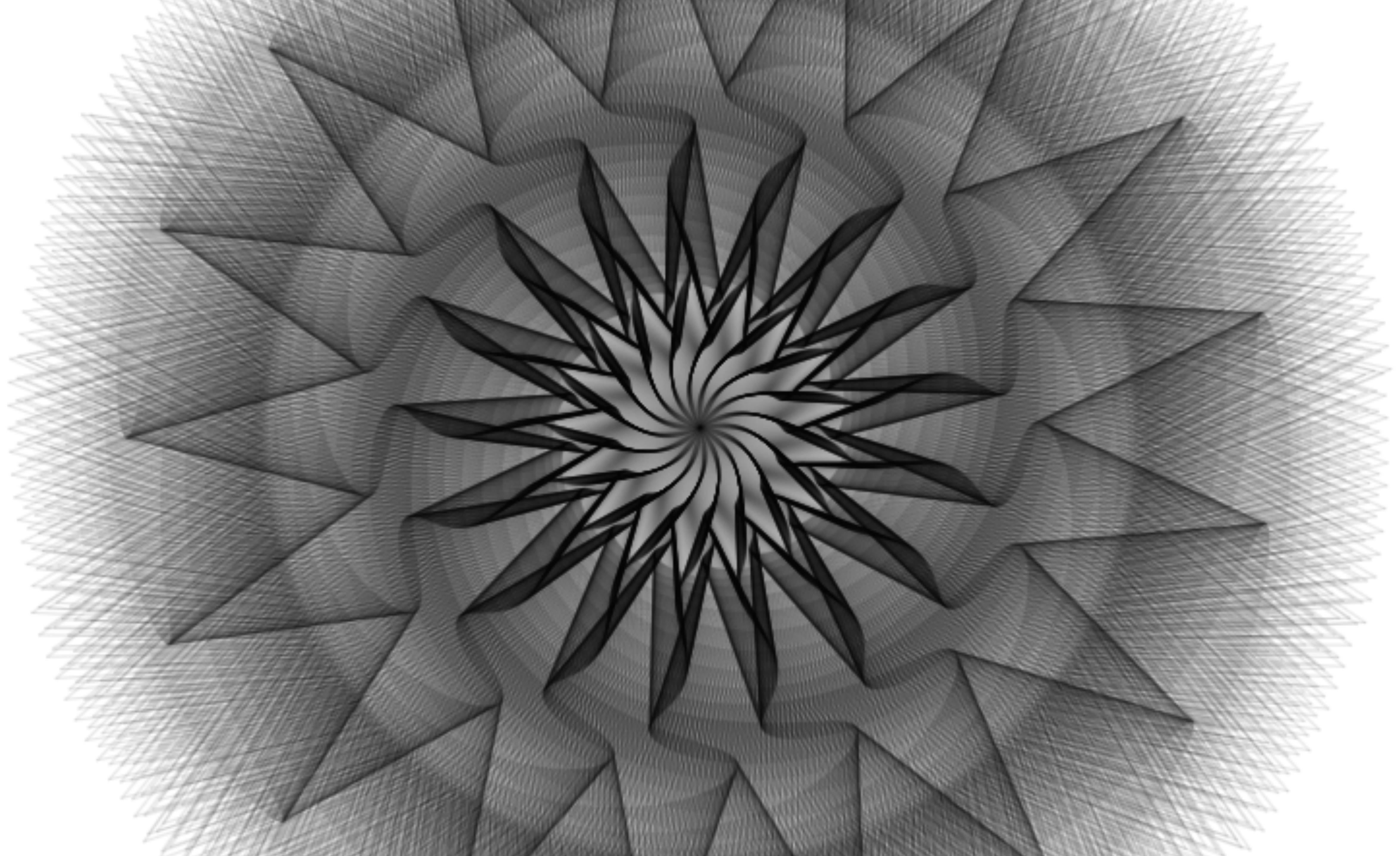
Simple Linear Model

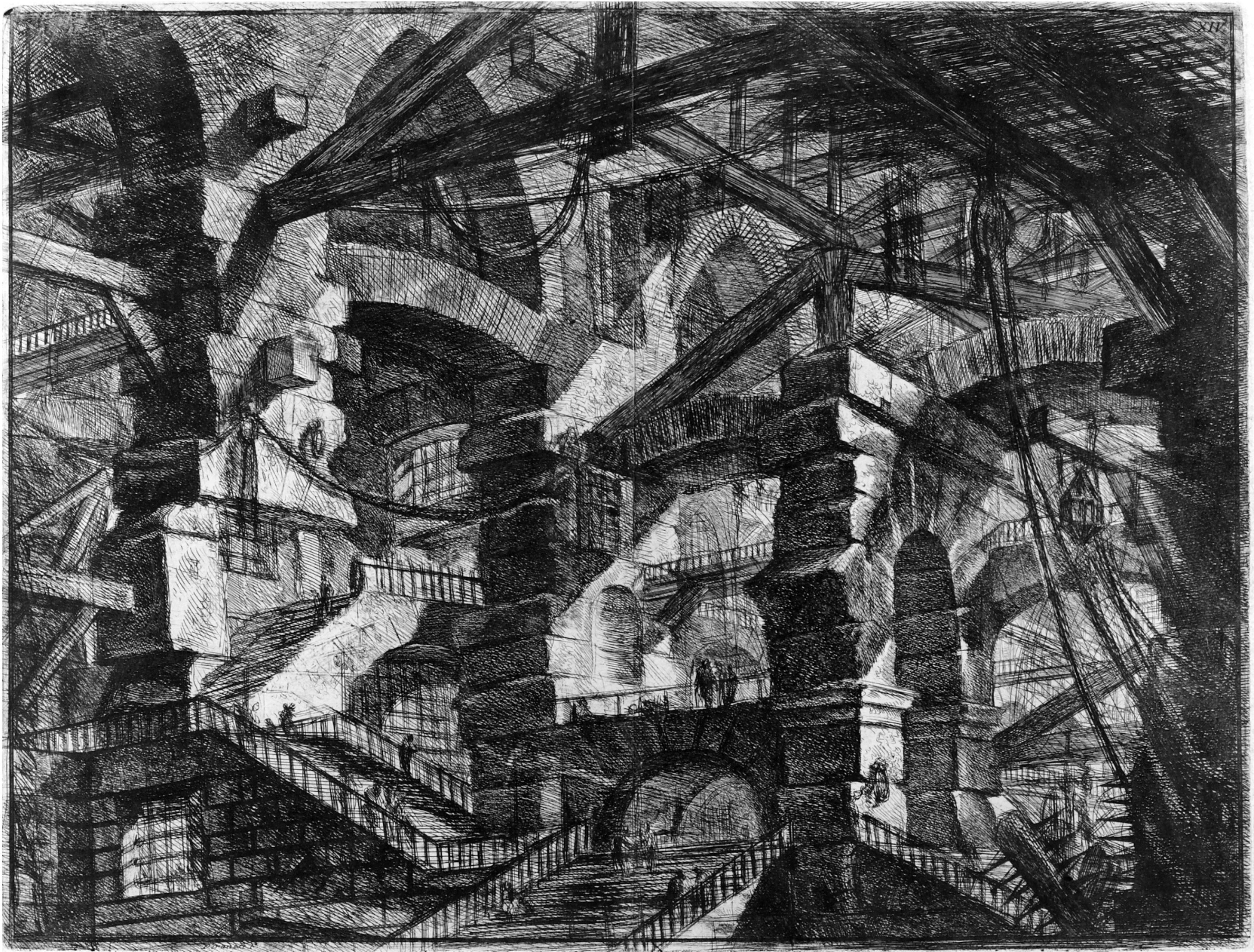


$$y = ax + b$$

$a \rightarrow \text{slope}$

$b \rightarrow \text{intercept}$





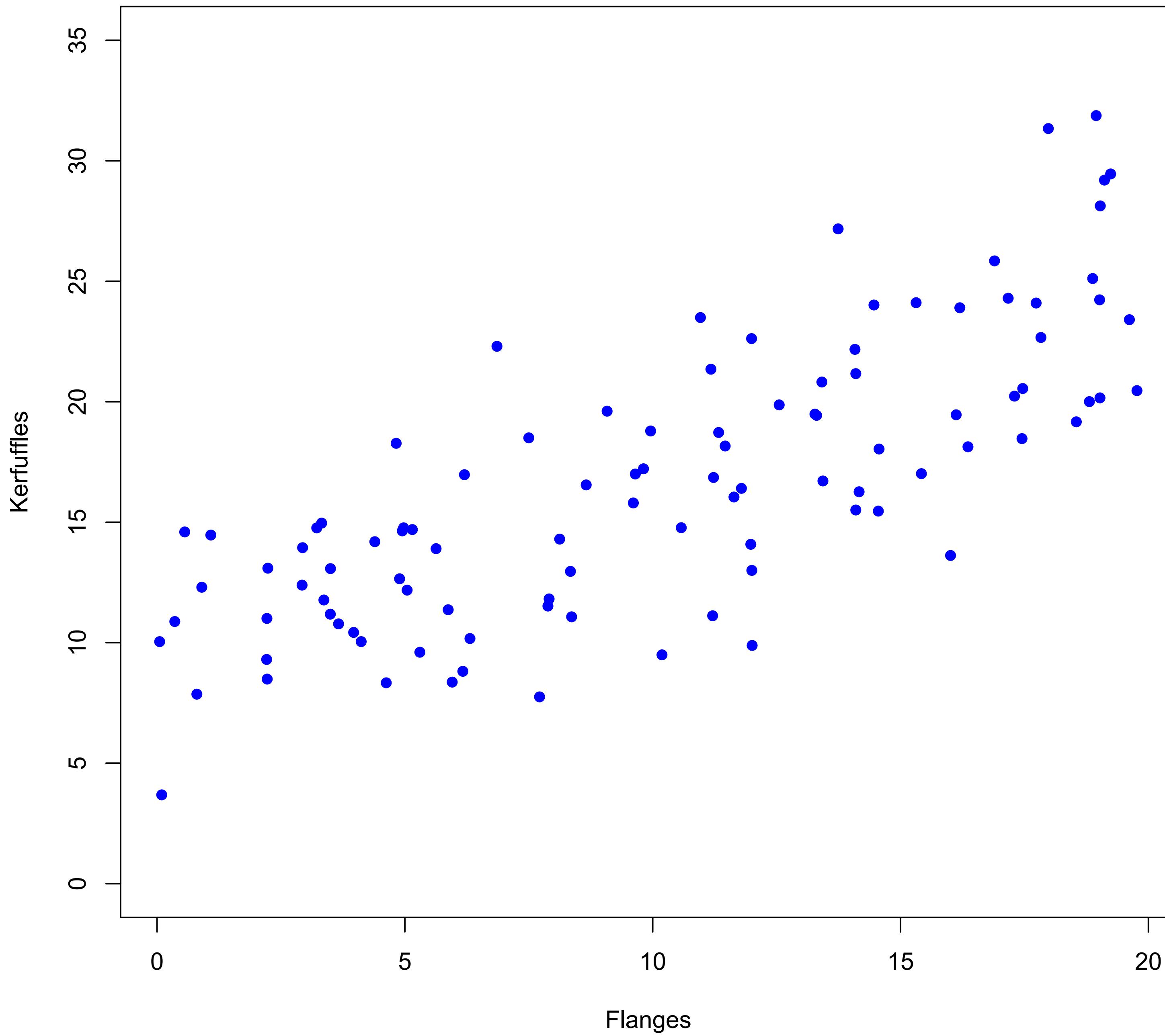
With four parameters I can fit an elephant, and with five I can make him wiggle his trunk

— *John Von Neumann*

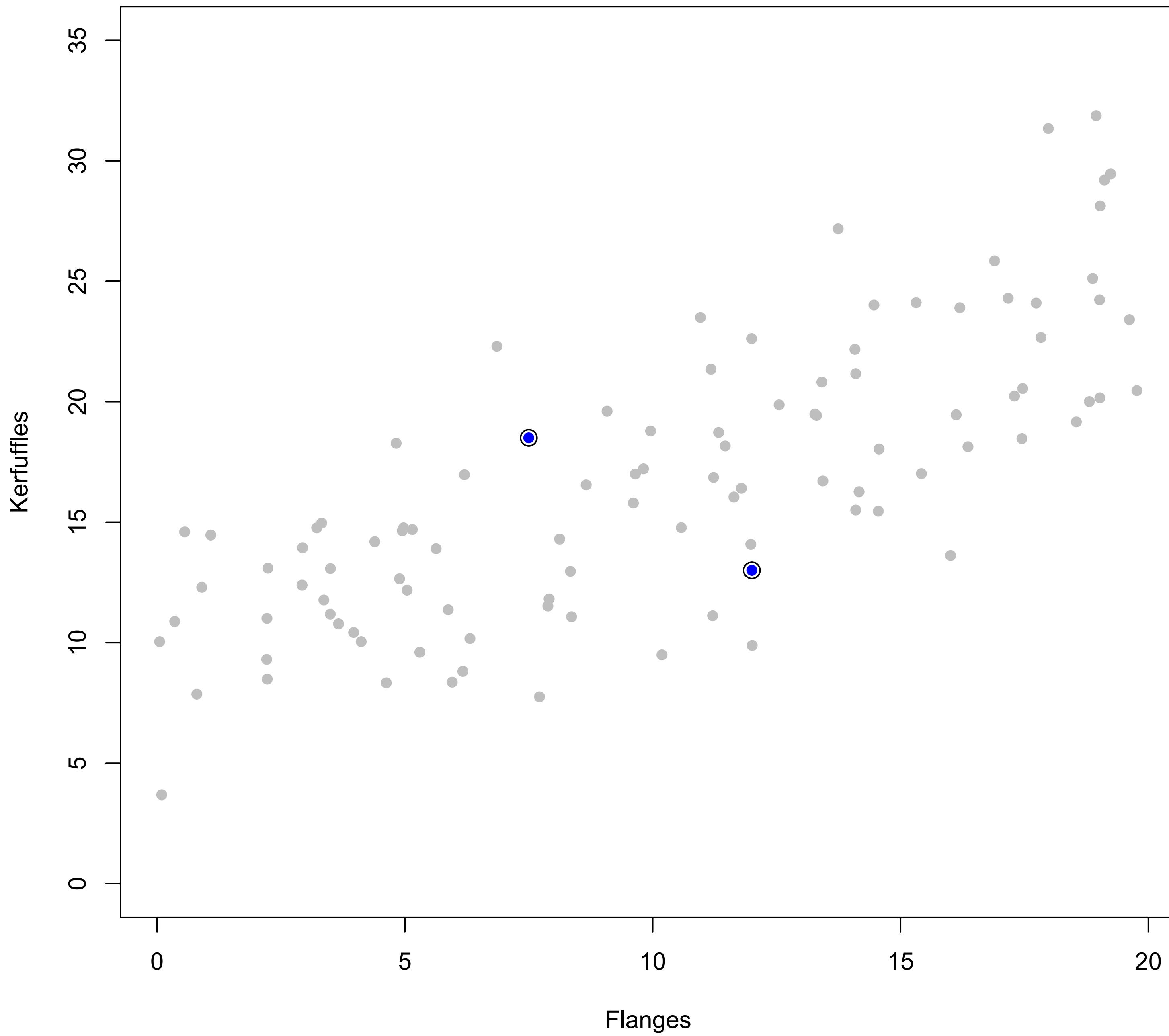
WHAT IS GOING ON??

**Is it learning what you *think* it's learning?
Is it learning what you *want*?**

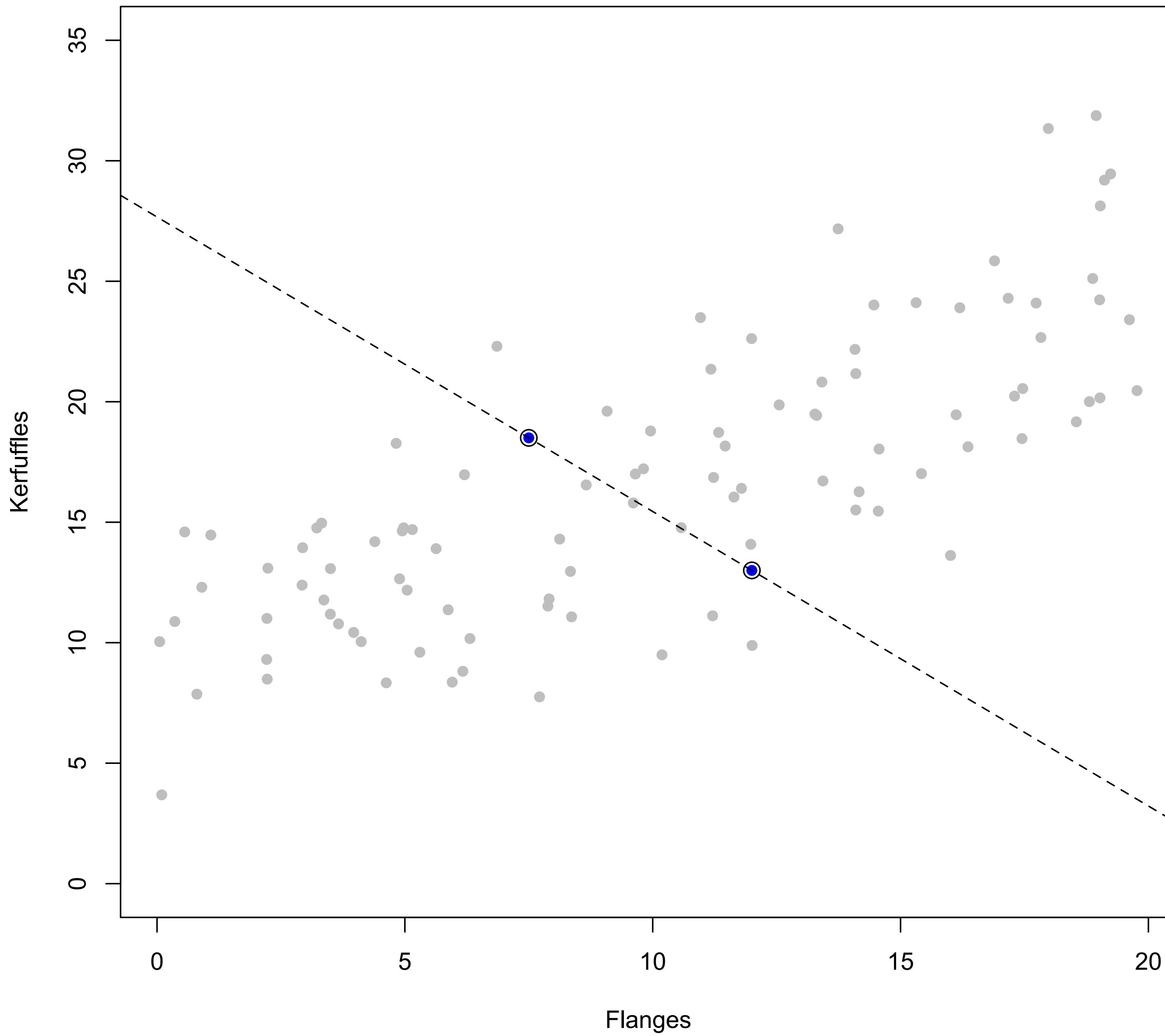
Kerfuffle-Flange Dependency



Kerfuffle-Flange Dependency



Kerfuffle-Flange Dependency



The Second Law of Machine Learning

Generalisation is everything

y ← — — **x**

$$y \equiv f(x)$$

$$y \equiv f(x, \theta)$$

$$L(y, f(x, \theta))$$

$$L(\mathbf{f}, \theta, \mathbf{x}, \mathbf{y})$$

$$L(f, \theta, \{X[, Y]\})$$

$$\operatorname{argmin}_{\theta} L(f, \theta, \{X[, Y]\})$$

$$\theta^* = \operatorname*{argmin}_{\theta} L(f, \theta, \{X[, Y]\})$$

$$\theta^* = \operatorname*{argmin}_{\theta} L(f, \theta, \{X[, Y]\})$$

$$\theta^* = \operatorname*{argmin}_{\theta} L(\mathbf{f}, \theta, \{\mathbf{X}, \mathbf{Y}\})$$

$$\theta > \{x[, y]\}$$

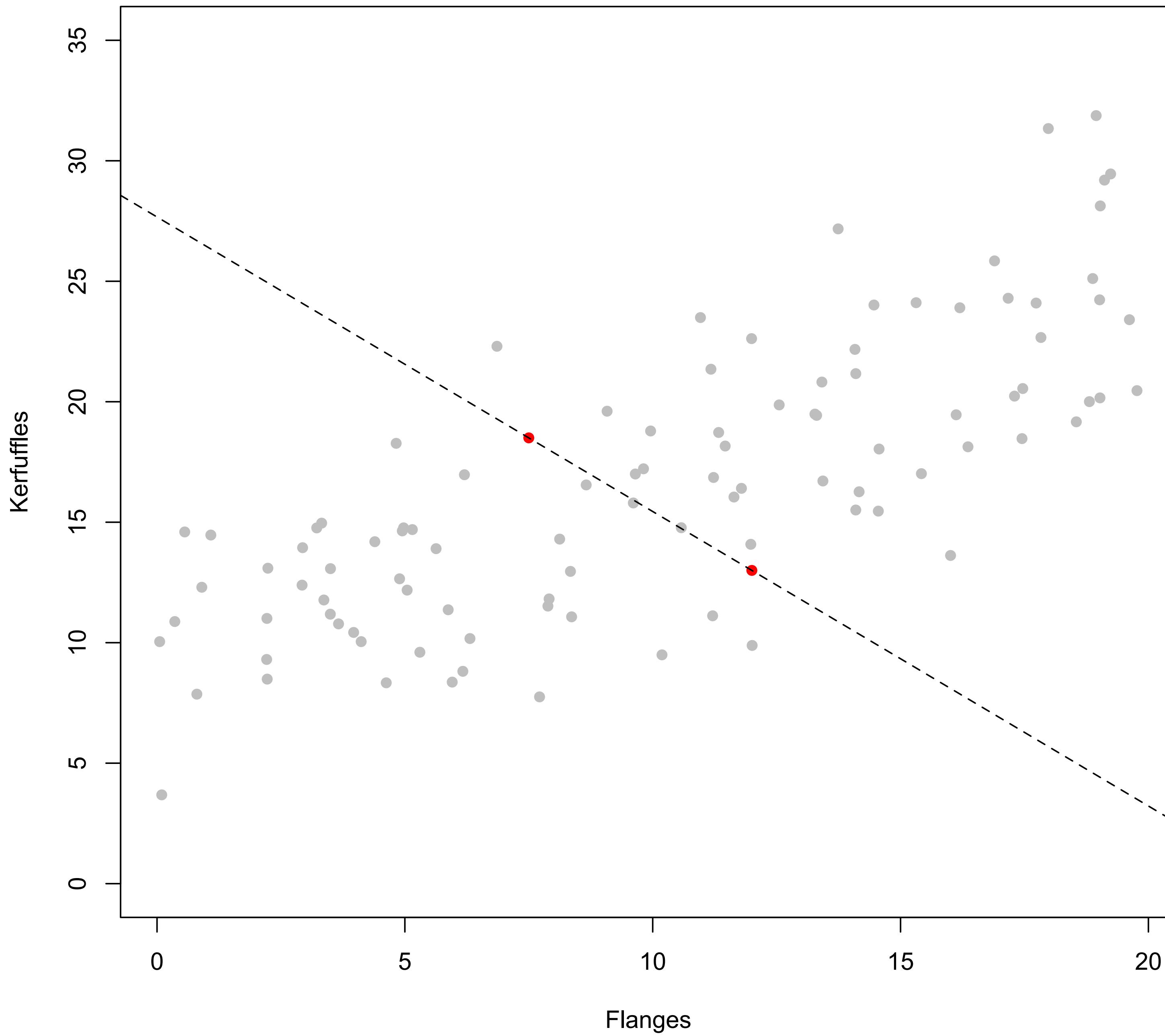
$$\theta^* = \operatorname*{argmin}_{\theta} L(f, \theta, \{X[, Y]\})$$

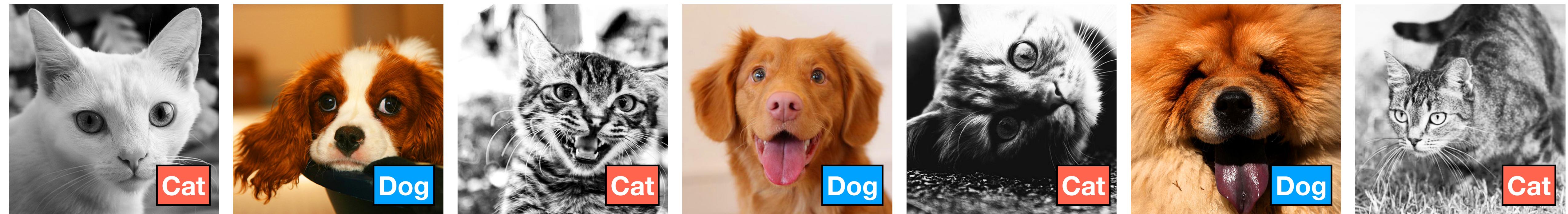


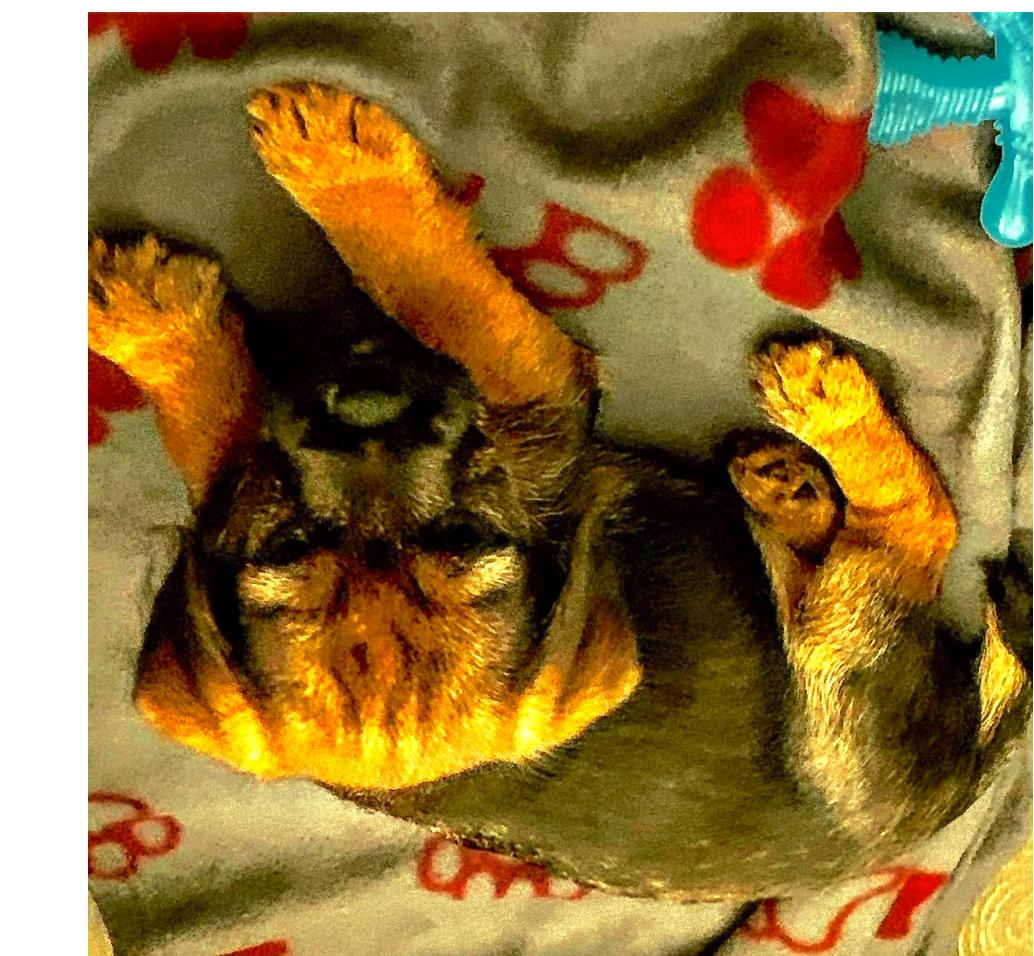
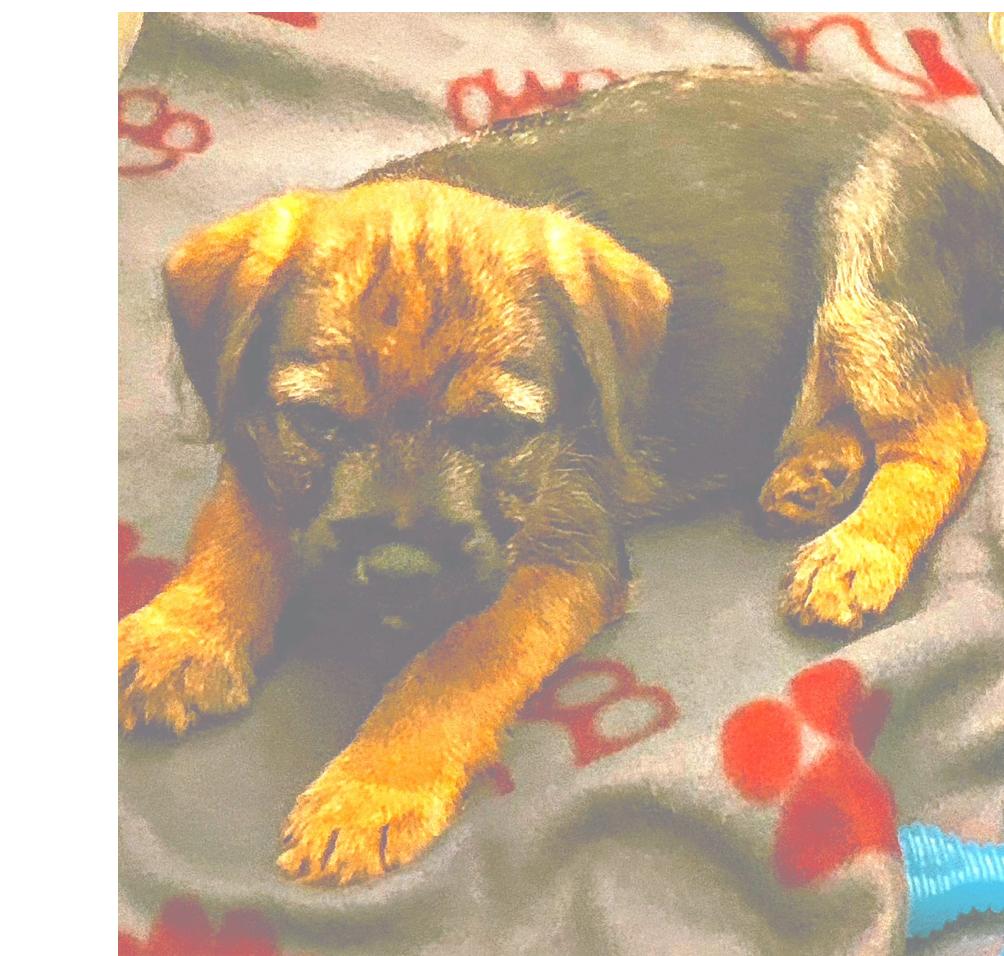
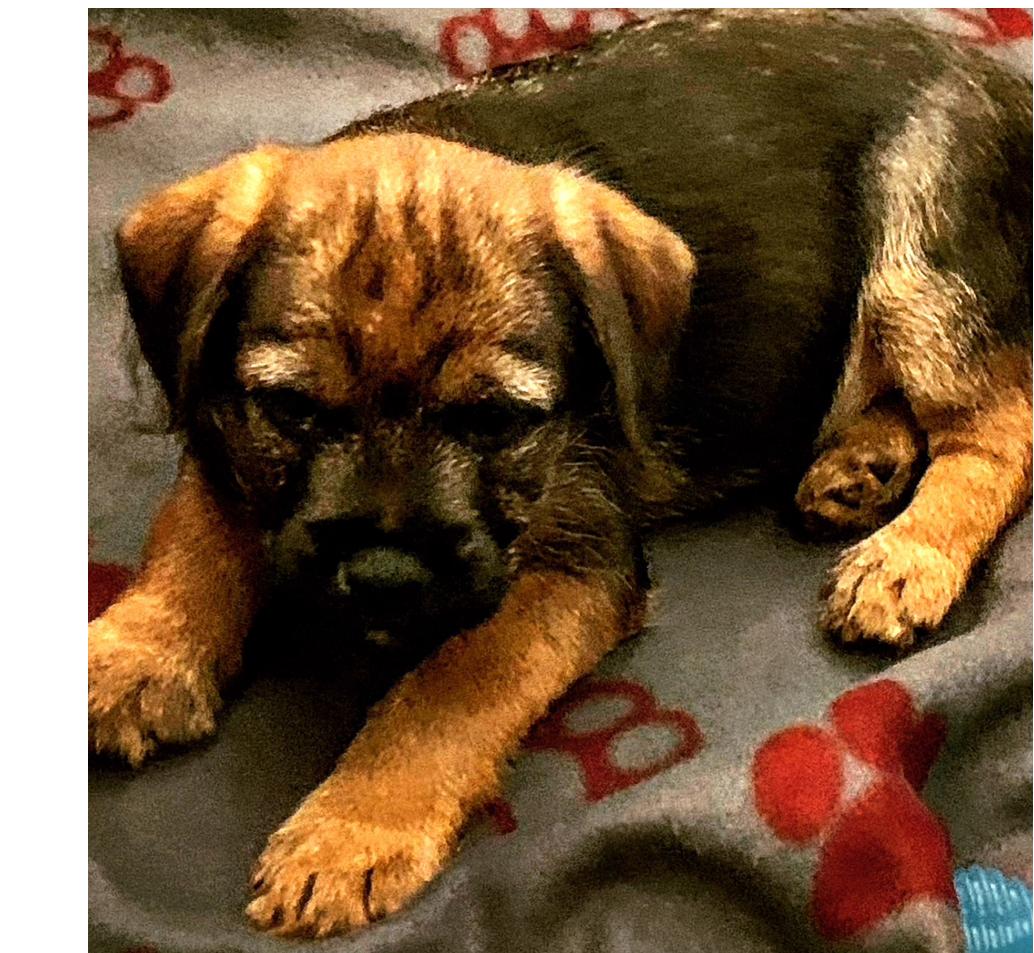
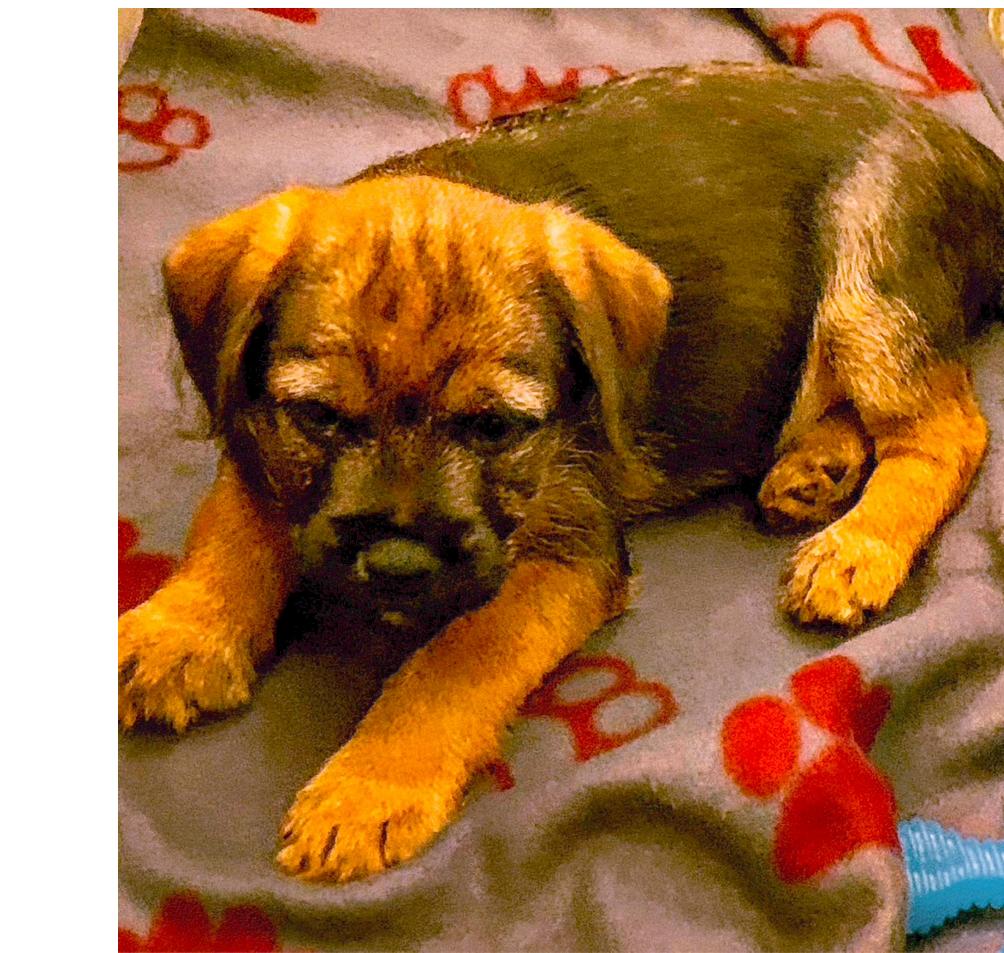
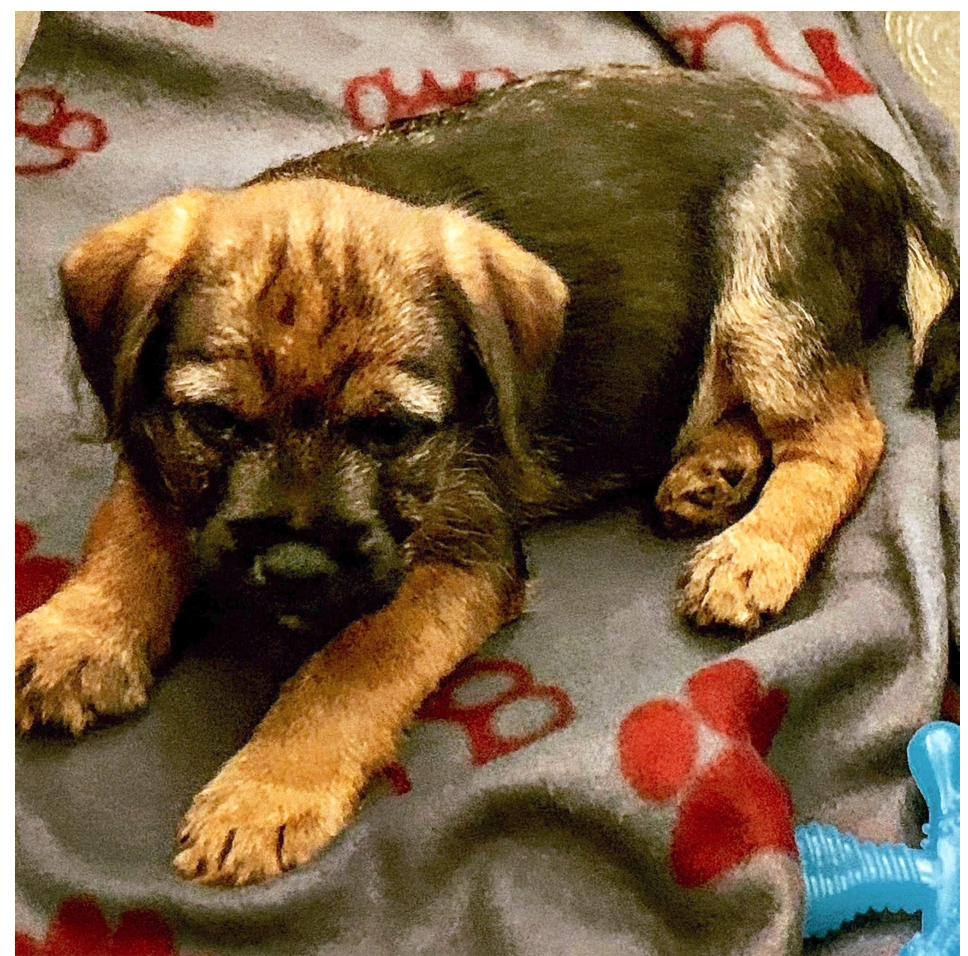
$$\theta^* = \operatorname*{argmin}_{\theta} L(f, \theta, \{X[, Y]\})$$

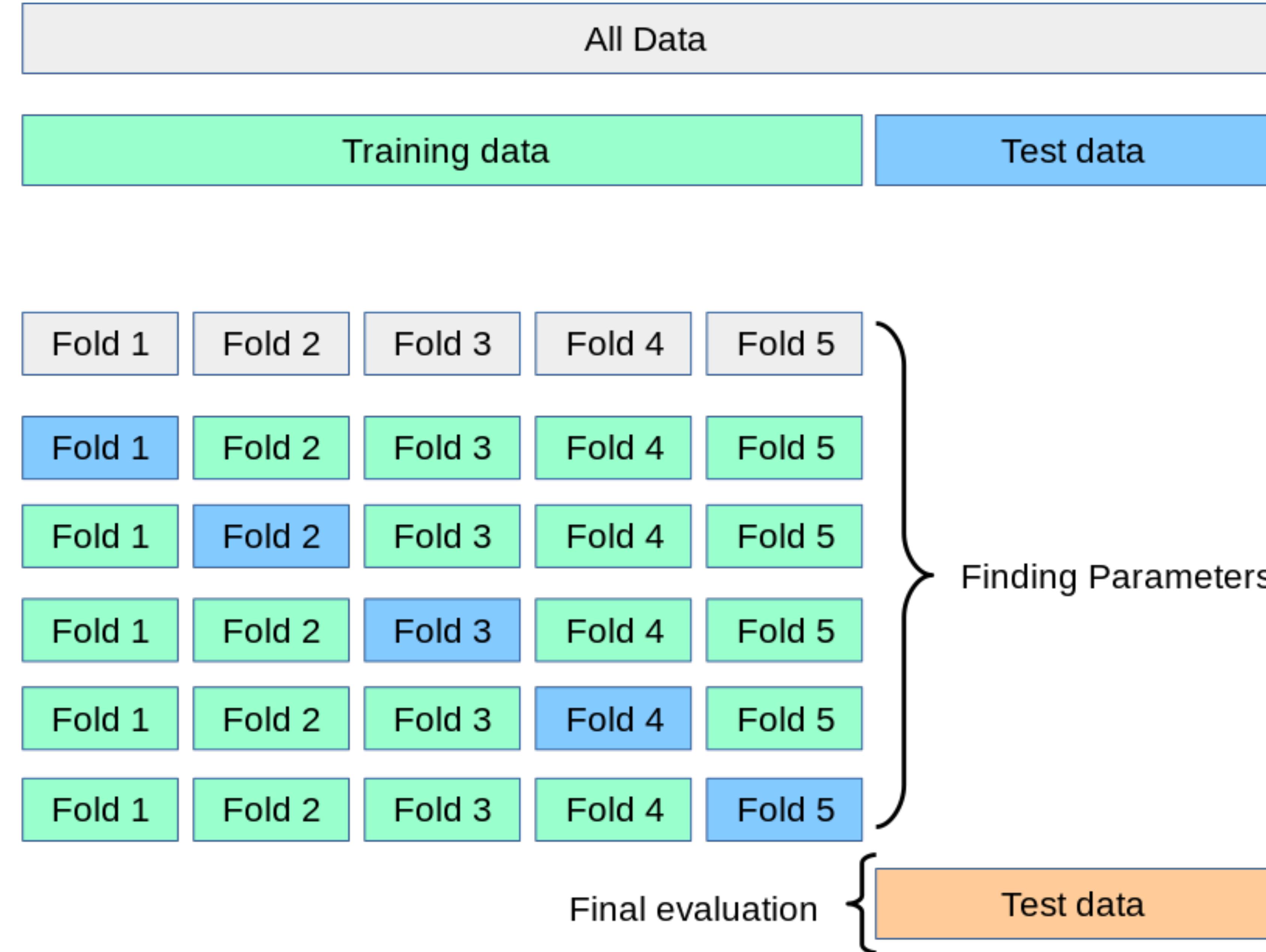
$$\theta^* = \operatorname*{argmin}_{\theta} L(\mathbf{f}, \theta, \{\mathbf{X}, \mathbf{Y}\})$$

Kerfuffle-Flange Dependency

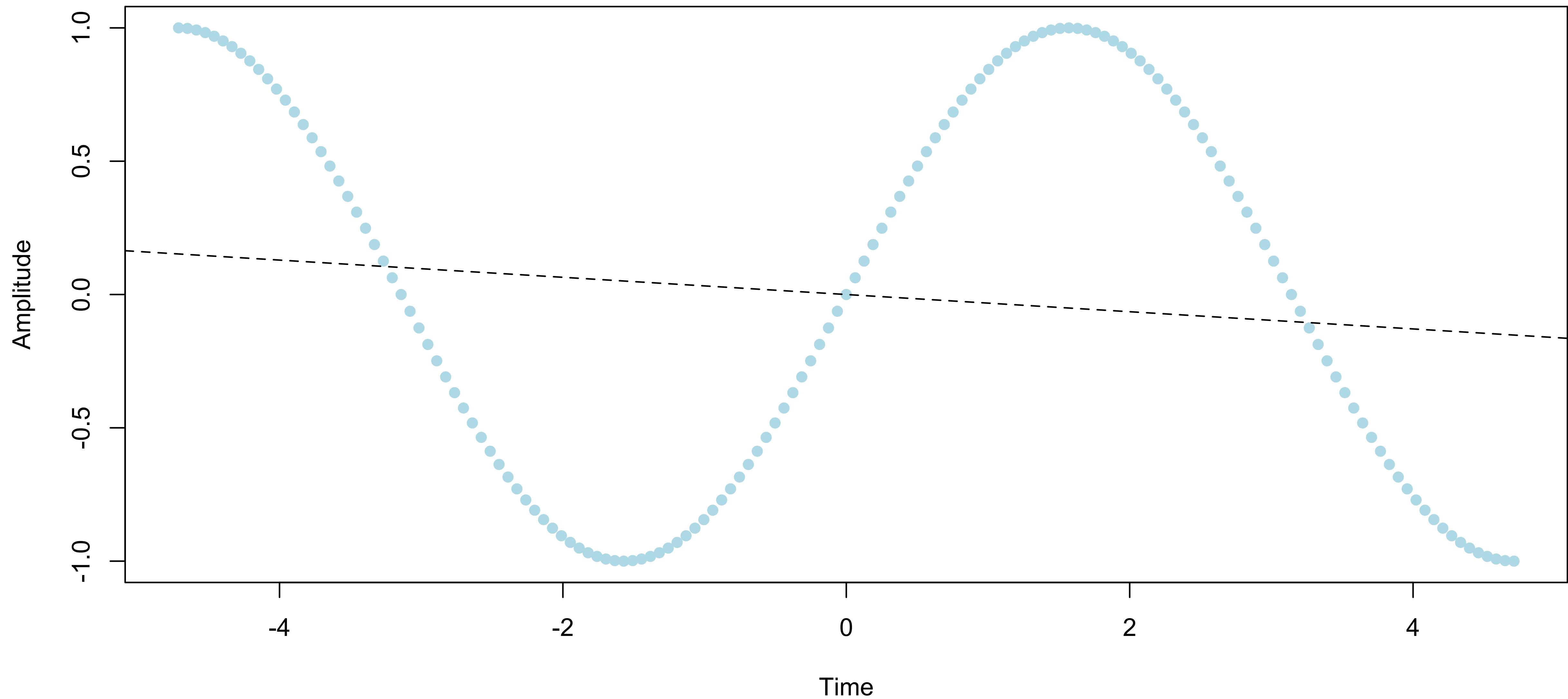




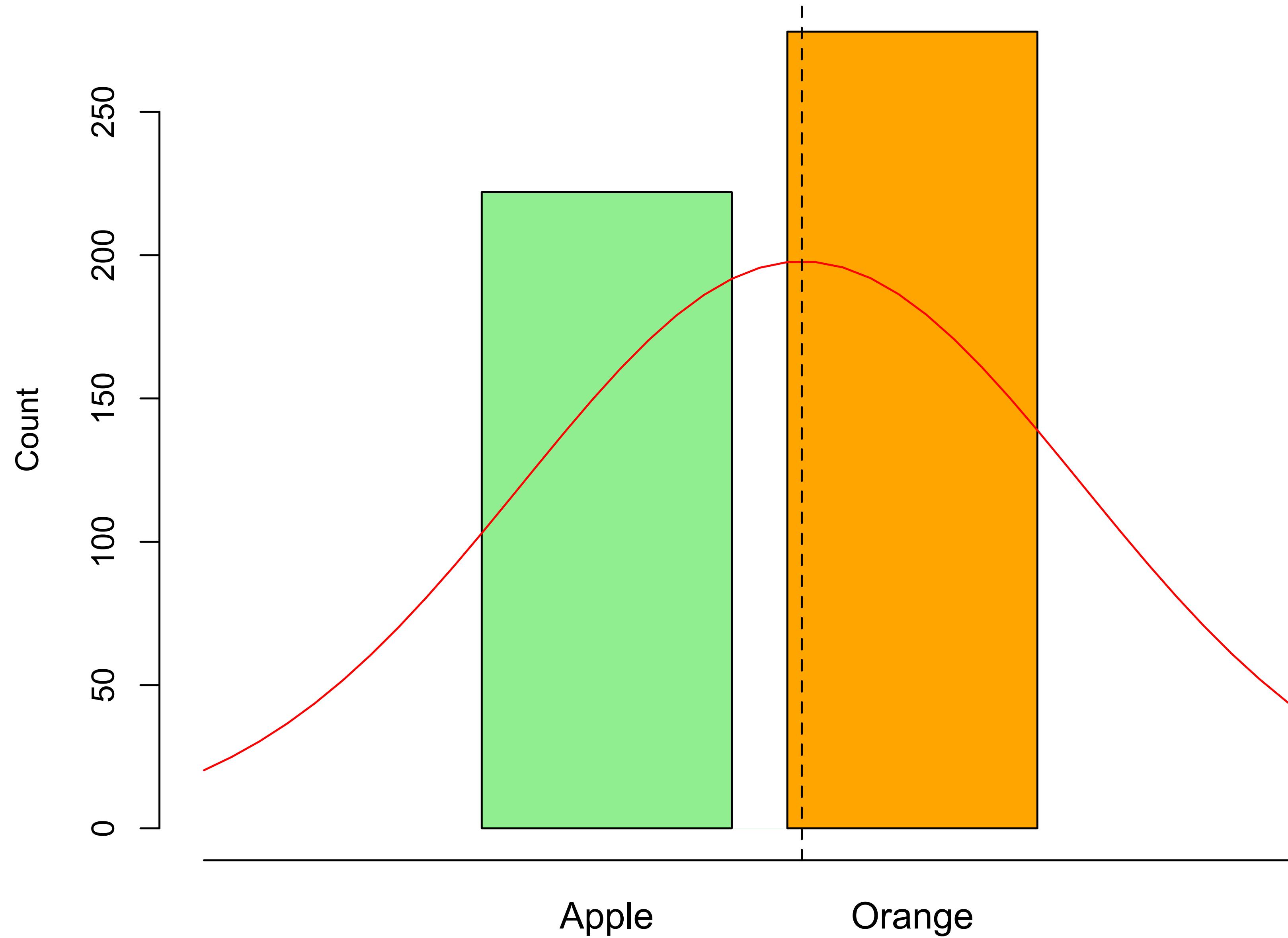




$$\theta^* = \operatorname*{argmin}_{\theta} L(\mathbf{f}, \theta, \{\mathbf{X}, \mathbf{Y}\})$$



Favourite Fruit



What's Your Personality Type?

Use the questions on the outside of the chart to determine the four letters of your Myers-Briggs type.
For each pair of letters, choose the side that seems most natural to you, even if you don't agree with every description.

1. Are you outwardly or inwardly focused? If you:

- Could be described as talkative, outgoing
- Like to be in a fast-paced environment
- Tend to work out ideas with others, think out loud
- Enjoy being the center of attention

then you prefer

E
Extraversion

- Could be described as reserved, private
- Prefer a slower pace with time for contemplation
- Tend to think things through inside your head
- Would rather observe than be the center of attention

then you prefer

I
Introversion

ISTJ

Responsible, sincere, analytical, reserved, realistic, systematic. Hardworking and trustworthy with sound practical judgment.

ISFJ

Warm, considerate, gentle, responsible, pragmatic, thorough. Devoted caretakers who enjoy being helpful to others.

INFJ

Idealistic, organized, insightful, dependable, compassionate, gentle. Seek harmony and cooperation, enjoy intellectual stimulation.

INTJ

Innovative, independent, strategic, logical, reserved, insightful. Driven by their own original ideas to achieve improvements.

2. How do you prefer to take in information? If you:

- Focus on the reality of how things are
- Pay attention to concrete facts and details
- Prefer ideas that have practical applications
- Like to describe things in a specific, literal way

then you prefer

S
Sensing

- Imagine the possibilities of how things could be
- Notice the big picture, see how everything connects
- Enjoy ideas and concepts for their own sake
- Like to describe things in a figurative, poetic way

then you prefer

N
Intuition

ESTP

Outgoing, realistic, action-oriented, curious, versatile, spontaneous. Pragmatic problem solvers and skillful negotiators.

ESFP

Playful, enthusiastic, friendly, spontaneous, tactful, flexible. Have strong common sense, enjoy helping people in tangible ways.

ENFP

Enthusiastic, creative, spontaneous, optimistic, supportive, playful. Value inspiration, enjoy starting new projects, see potential in others.

ENTP

Inventive, enthusiastic, strategic, enterprising, inquisitive, versatile. Enjoy new ideas and challenges, value inspiration.

ESTJ

Efficient, outgoing, analytical, systematic, dependable, realistic. Like to run the show and get things done in an orderly fashion.

ESFJ

Friendly, outgoing, reliable, conscientious, organized, practical. Seek to be helpful and please others, enjoy being active and productive.

ENFJ

Caring, enthusiastic, idealistic, organized, diplomatic, responsible. Skilled communicators who value connection with people.

ENTJ

Strategic, logical, efficient, outgoing, ambitious, independent. Effective organizers of people and long-range planners.

3. How do you prefer to make decisions? If you:

- Make decisions in an impersonal way, using logical reasoning
- Value justice, fairness
- Enjoy finding the flaws in an argument
- Could be described as reasonable, level-headed

then you prefer

T
Thinking

- Base your decisions on personal values and how your actions affect others
- Value harmony, forgiveness
- Like to please others and point out the best in people
- Could be described as warm, empathetic

then you prefer

F
Feeling

4. How do you prefer to live your outer life? If you:

- Prefer to have matters settled
- Think rules and deadlines should be respected
- Prefer to have detailed, step-by-step instructions
- Make plans, want to know what you're getting into

then you prefer

J
Judging

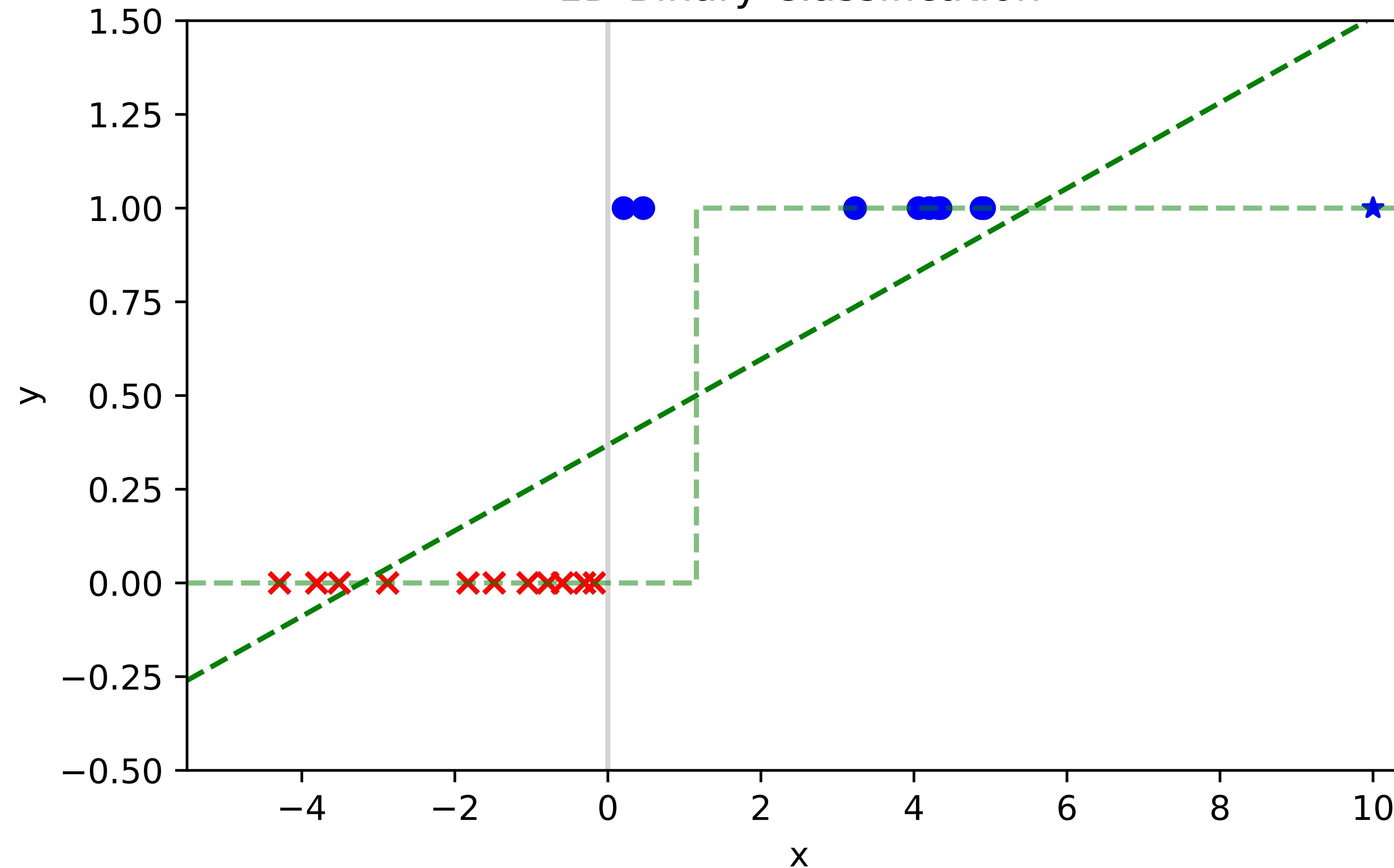
- Prefer to leave your options open
- See rules and deadlines as flexible
- Like to improvise and make things up as you go
- Are spontaneous, enjoy surprises and new situations

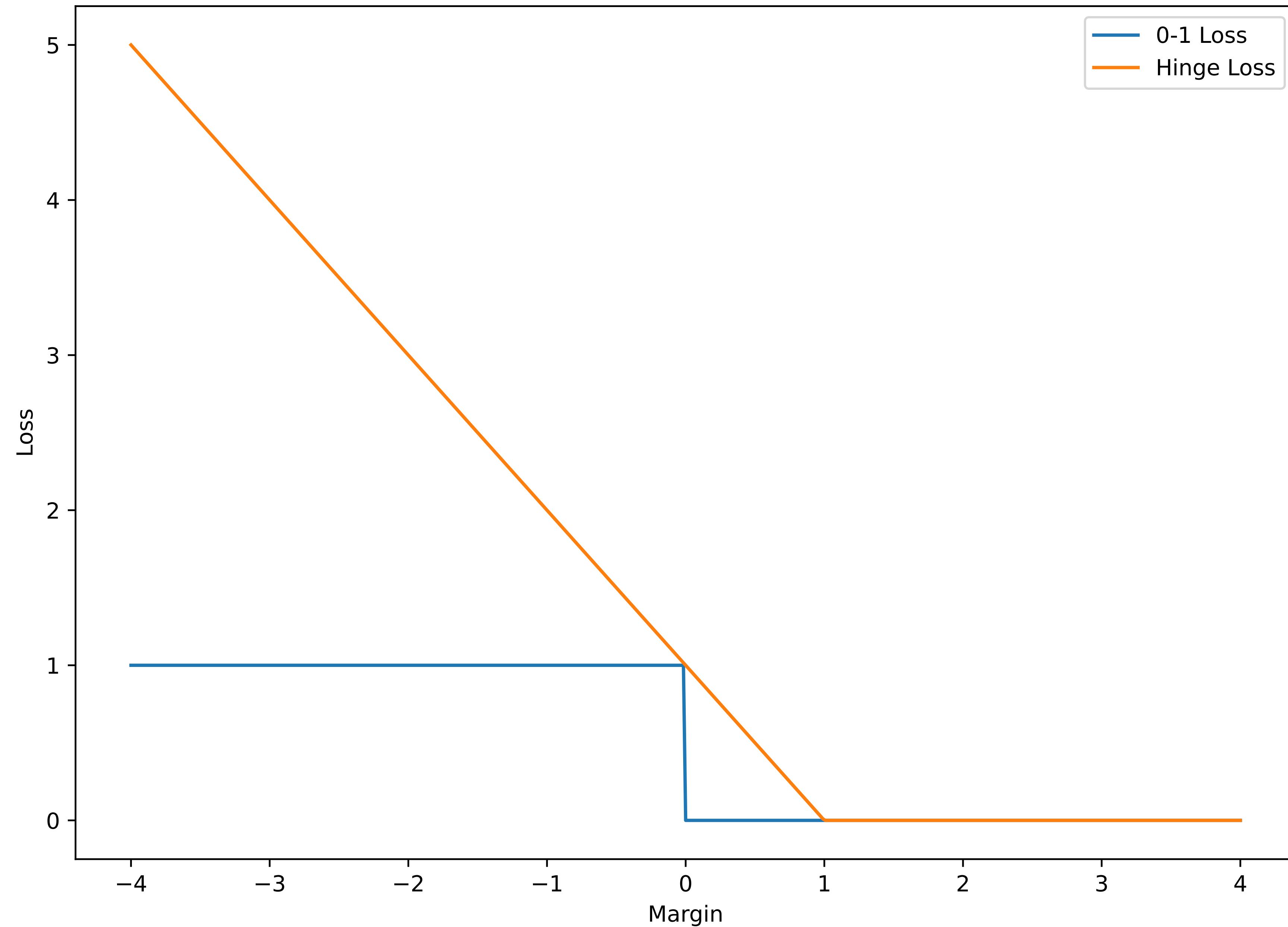
then you prefer

P
Perceiving

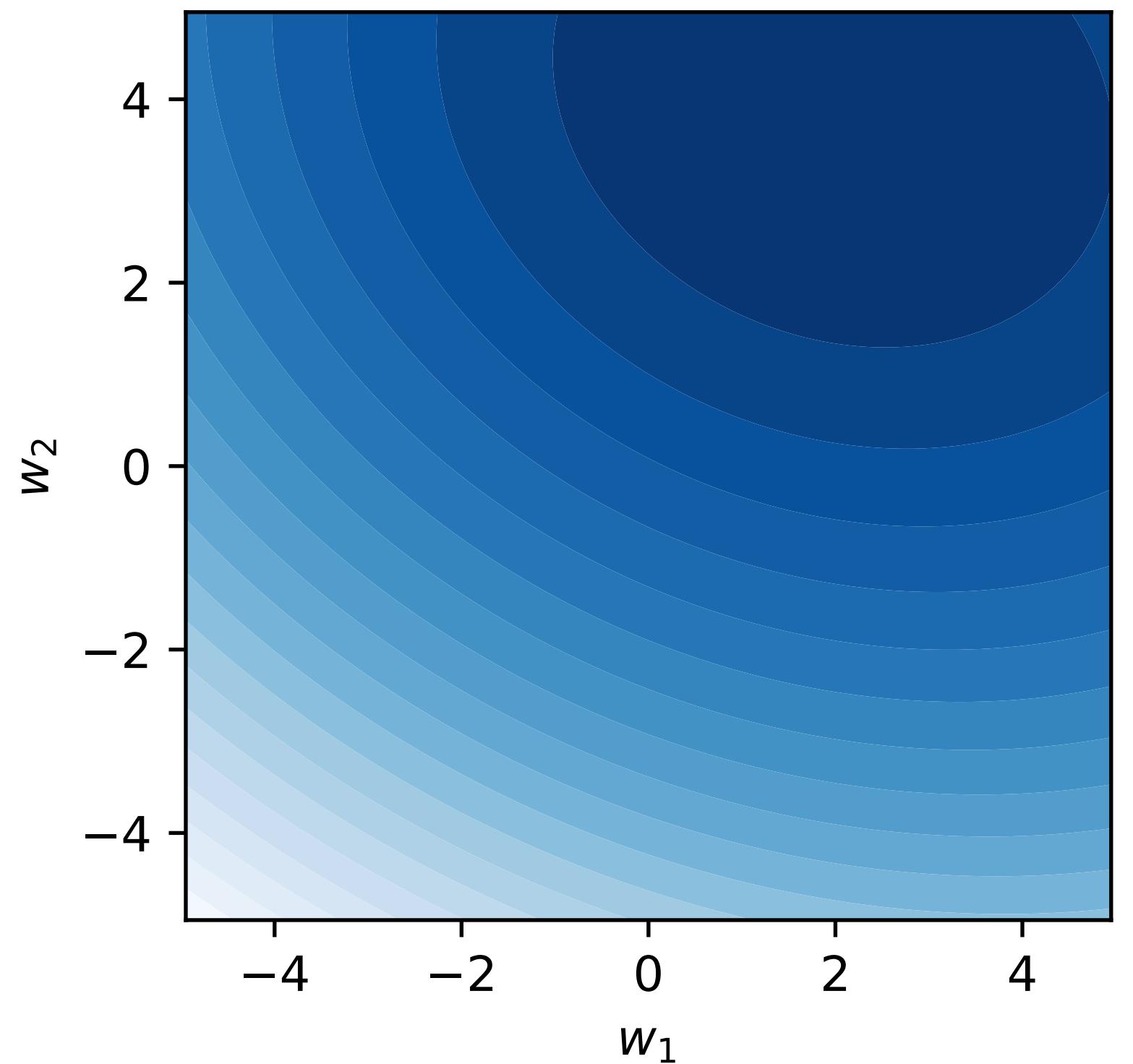
$$\theta^* = \operatorname*{argmin}_{\theta} L(f, \theta, \{X[, Y]\})$$

1D Binary Classification

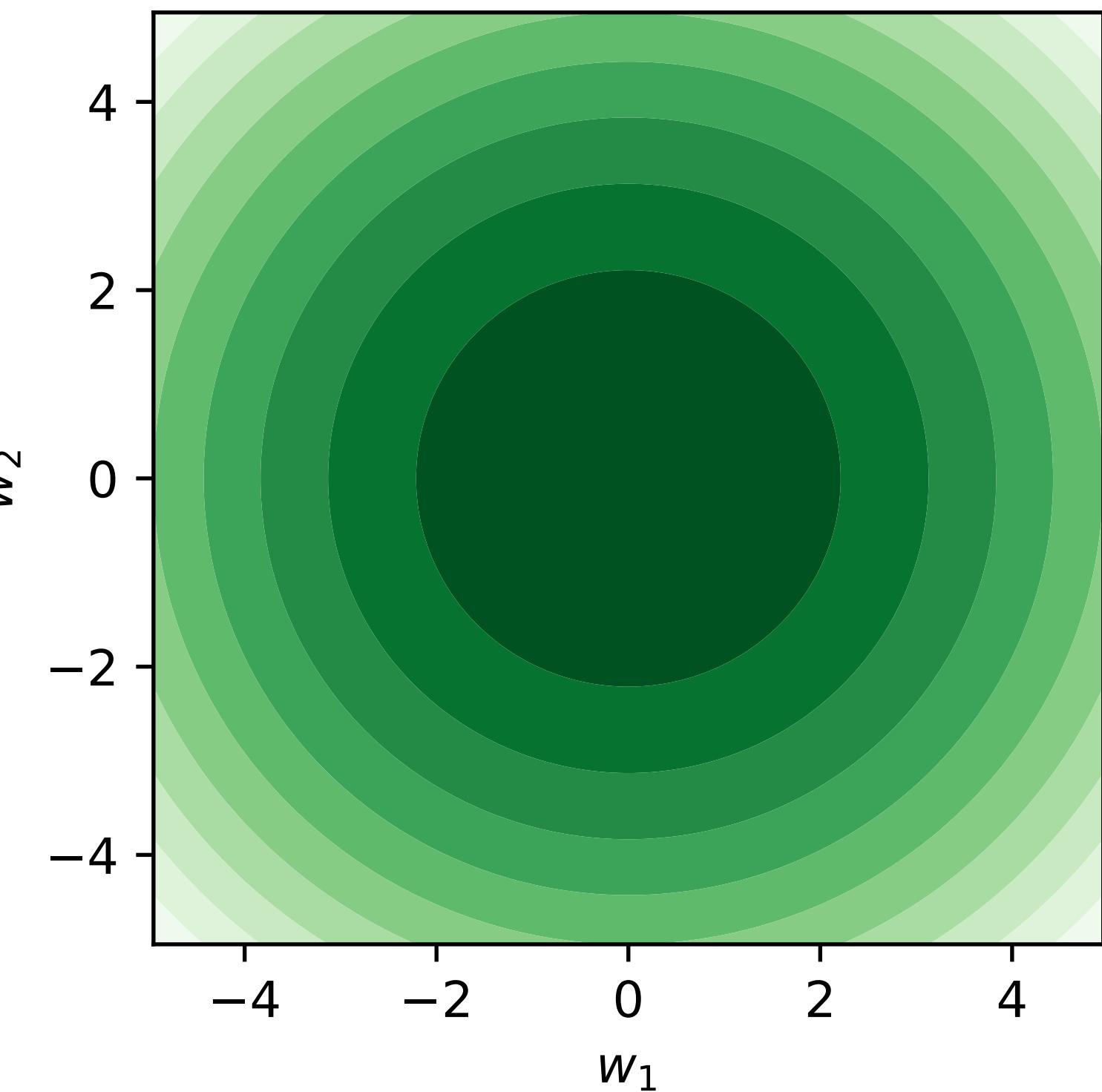




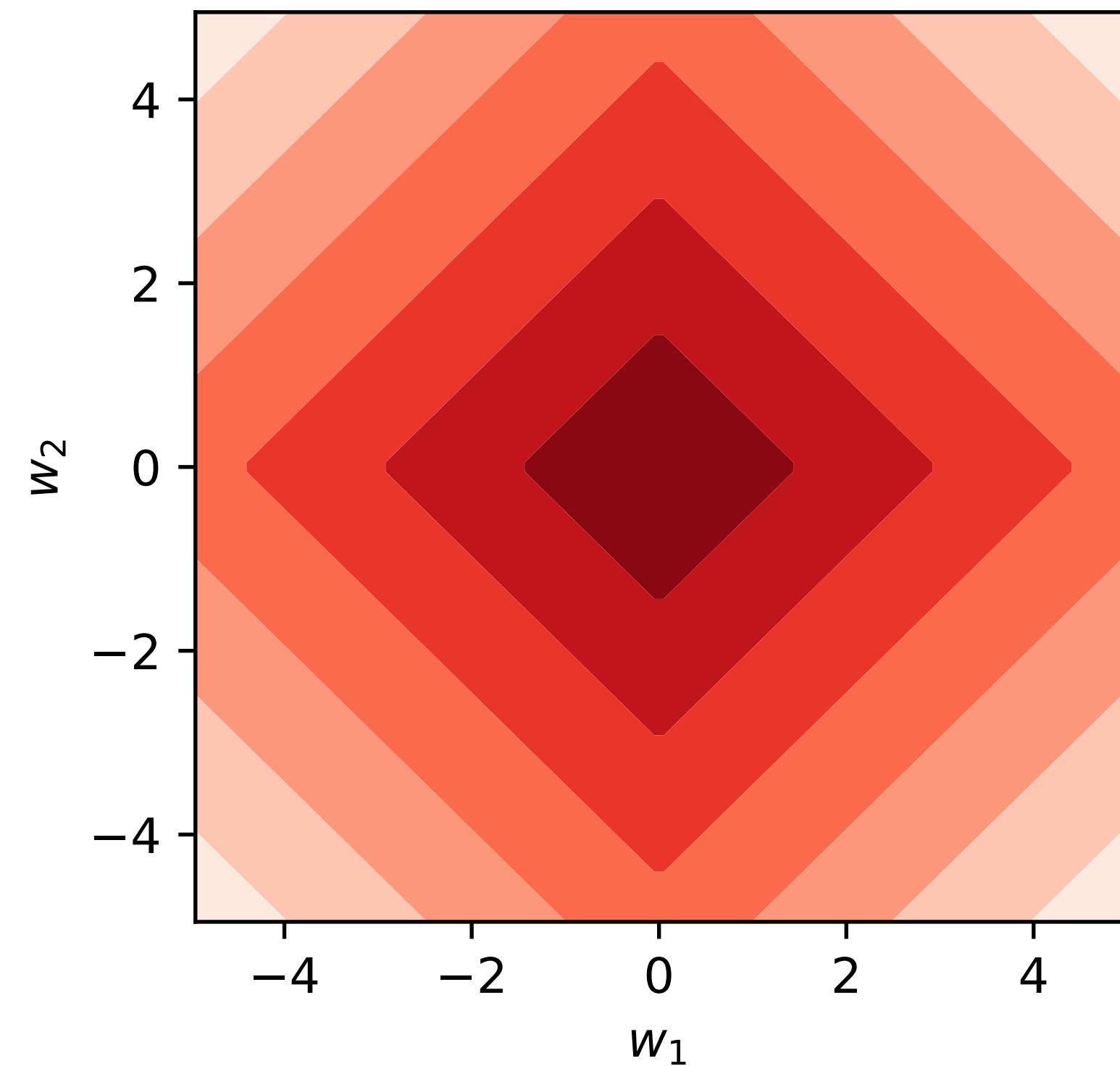
OLS Loss: $\|\hat{\mathbf{y}} - \mathbf{y}\|^2$



Ridge Penalty: $\|\mathbf{w}\|^2$



Lasso Penalty: $\|\mathbf{w}\|_1$





The Third Law of Machine Learning

Machine Learning is not magic



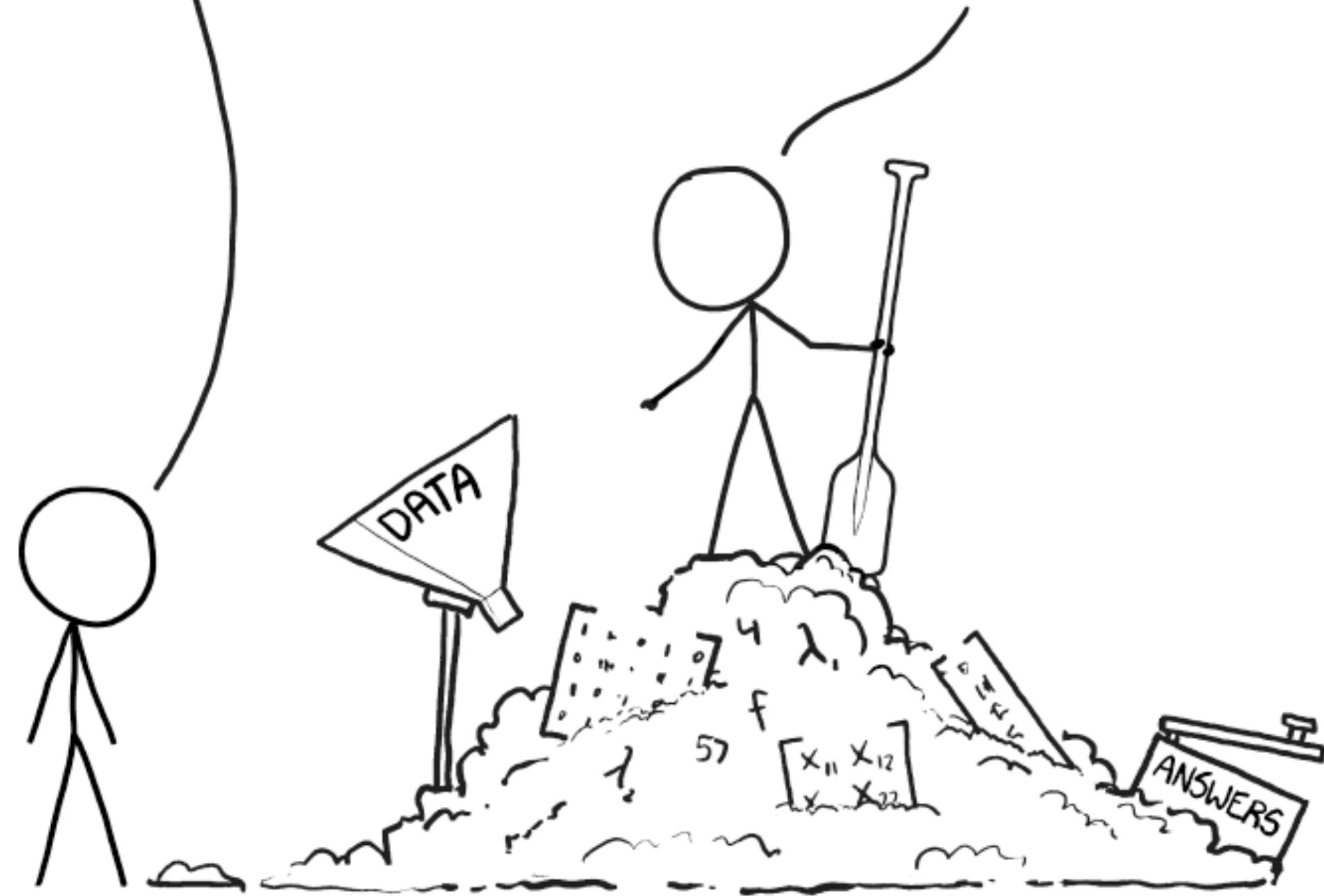


THIS IS YOUR MACHINE LEARNING SYSTEM?

YUP! YOU POUR THE DATA INTO THIS BIG
PILE OF LINEAR ALGEBRA, THEN COLLECT
THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL
THEY START LOOKING RIGHT.





Questions?

Next week: Linear Models

