

# Introduction to Neural Networks for Natural Language Processing

# Preliminaries

# Preliminaries

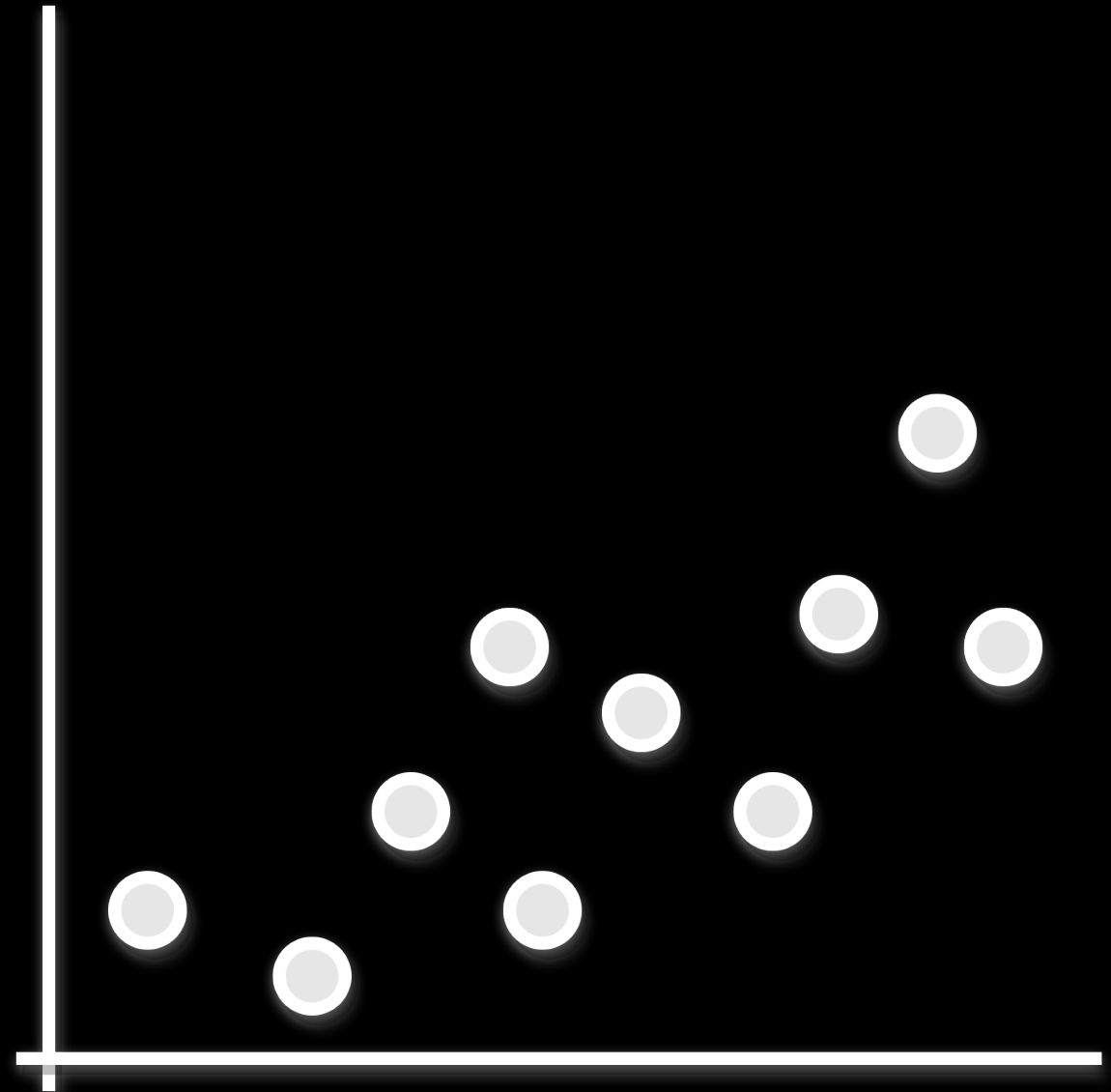
- What is it we want to do?
- Visualizing the problems.
- Regression & Classification.
- Notebook introduction: Algebra for Neural Networks.

# What is it we want to do?

- Before jumping into neural networks, we have to understand what they are for, **what we want to do.**
- Since we assume a **blank slate**, we will not focus on Natural Language Processing yet!
- Instead, we will start working on something simple like **just a few points drawn as you would in paper.**
- What we want to do is **discover trends** or **put each of them in groups.**

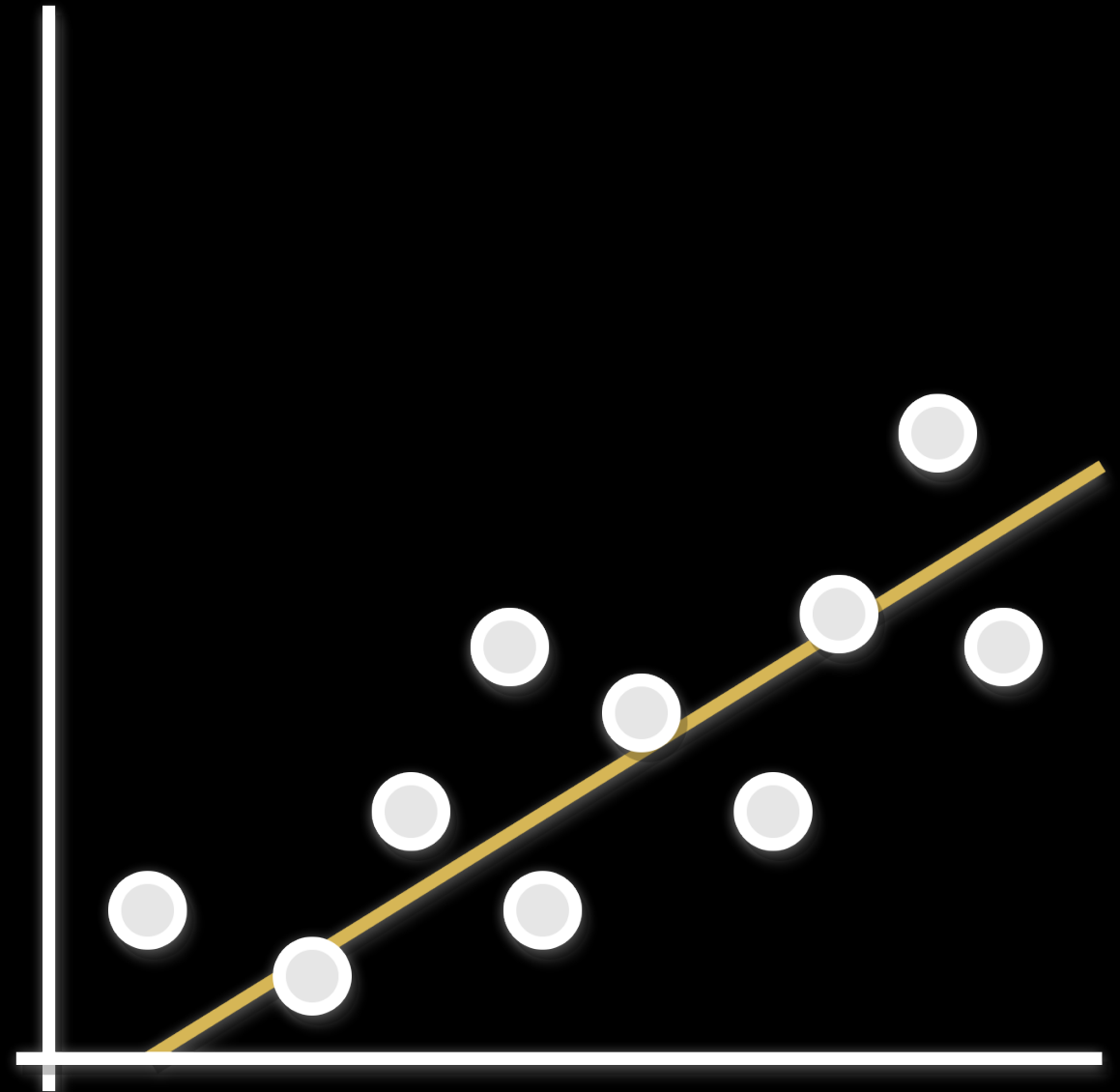
# Visualizing the problems

- We must take a quick look at what we mean by that!
- Consider the few points on the right, just plotted against two given axes.



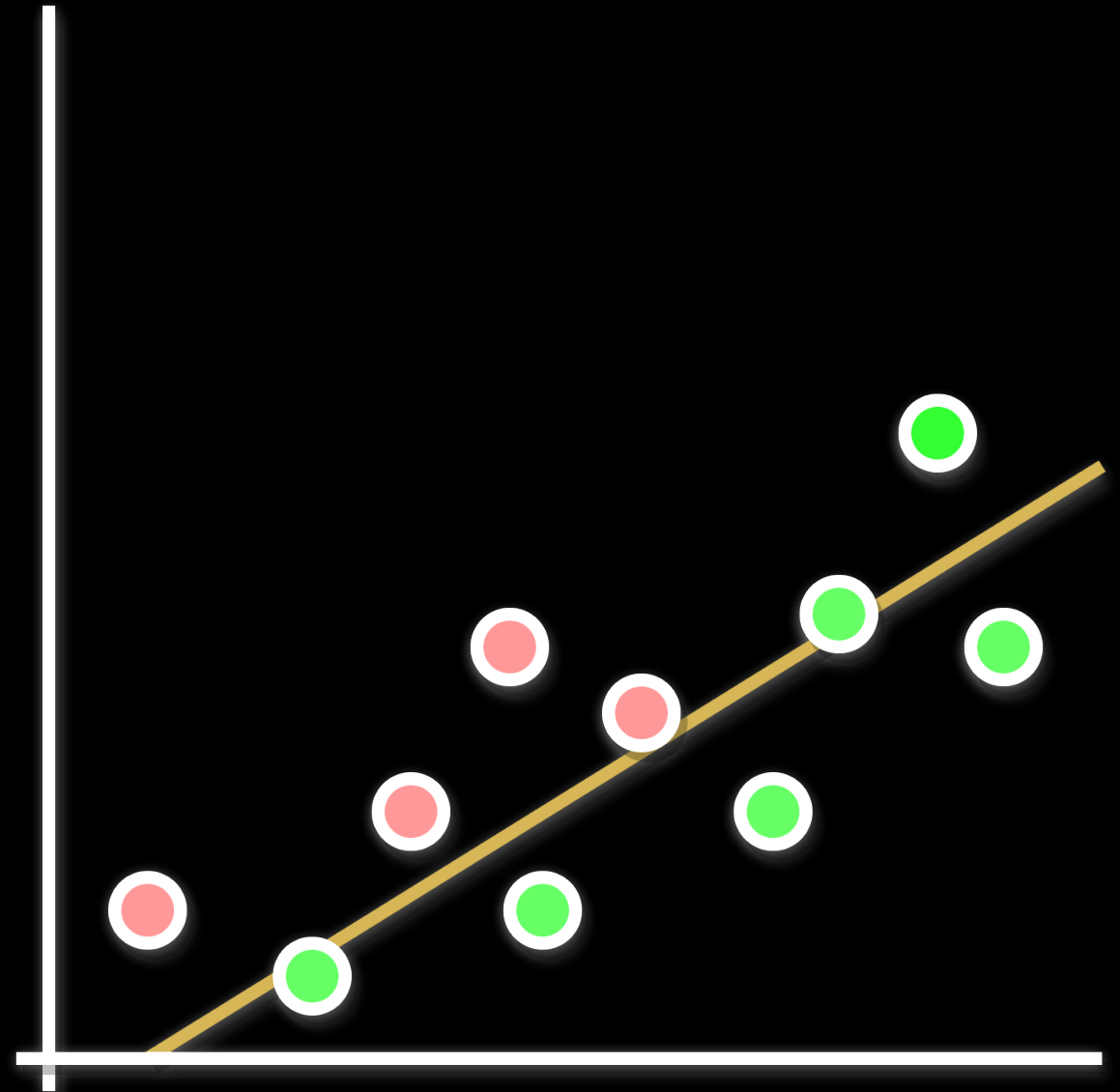
# Visualizing the problems

- We must take a quick look at what we mean by that!
- Consider the few points on the right, just plotted against two given axes.
- We may want to **find the trend that underlies them, as a line.**



# Visualizing the problems

- We must take a quick look at what we mean by that!
- Consider the few points on the right, just plotted against two given axes.
- We may want to **identify points belonging to different groups**, like red and green.



# Regression & Classification

- These two problems are respectively **regression** and **classification**.
- In regression, we want to **predict some quantity or property for an input instance**, given its attributes.
- In classification, we want to **predict the group or category for an input instance**, given its attributes.



# Regression & Classification

- An example of a **regression problem** is to predict the height of a tree given its species, age and the acidity of the soil.
- An example of a **classification problem** is to identify text as being 'negative', 'neutral' or 'positive' given the words contained in it.
- In both cases, **we want to automatically find a function** that takes attributes from trees, texts or any other kind of data and **produces an output that solves our problem.**

# Regression & Classification

- Typical text problems are **classification problems**, so during the course we won't cover regression explicitly.
- When we were visualising the problems, we hinted a way of doing classification by just **drawing lines to separate different groups**.
- However, you ask yourself: **how do you even draw a line between different texts?** A text is really not a point... and what is a point?

# Notebook introduction: Algebra for Neural Networks

- At this point, we have gone full circle and we really must ask:

## What is a 'point'?

- We want to understand what we mean by 'a point', and **the best way to understand something is to build it.**
- That is what we will do in the first programming notebook of the course... **Let's go!**

# Introduction to Neural Networks for Natural Language Processing