What is Machine Learning?

algorithm and/or identifies a pattern
↳ not coded explicitly
↳ learns from the input data
↳ generalises to unseen data

What does it mean to be explicitly programmed?

- Defining rules for every possible scenario.
- Provide step-by-step instructions

Example: Spam detection
Gmail does it for you. How?

- Learning isn't involved.
  ↳ No training, no dependence on the training data

What makes a problem an ML problem? For which problems should ML come to mind?

→ Requires pattern recognition that is either unknown or too complex to define manually
Making predictions, i.e., generalizing to unseen data.
→ Unstructured data → text, image, graph
→ When decisions need context → based on history
→ Automation of manual, time consuming tasks
→ Above all, when you have data to train on.

Examples: Text recognition in images
Recommendation
Weather forecast
Identifying tumours in scans

Types of machine learning

- Supervised → Regression, classification
- Unsupervised → Clustering, association rules, dimensionality reduction

- Semi-supervised
  ↳ Abundance of unlabelled data
  ↳ Labelling data is difficult or time consuming

Steps
1. Train on available labelled data.
2. Use predictions on unlabelled data as the new labels.
3. Model improves over time.

Examples: Sentiment analysis

- Self-supervised
  ↳ No labelled data } Learn from the data itself
  ↳ Use a proxy task }
    ↳ Use next word prediction to teach grammar. → This is followed by fine tuning to specific tasks.
      ↳ sentiment
      ↳ Q/A
- Reinforcement
  ↳ Rewards & punishment
  ↳ closest to human learning
  ↳ Infinite possibilities
    ↳ Walking robots
    ↳ Self driving cars.