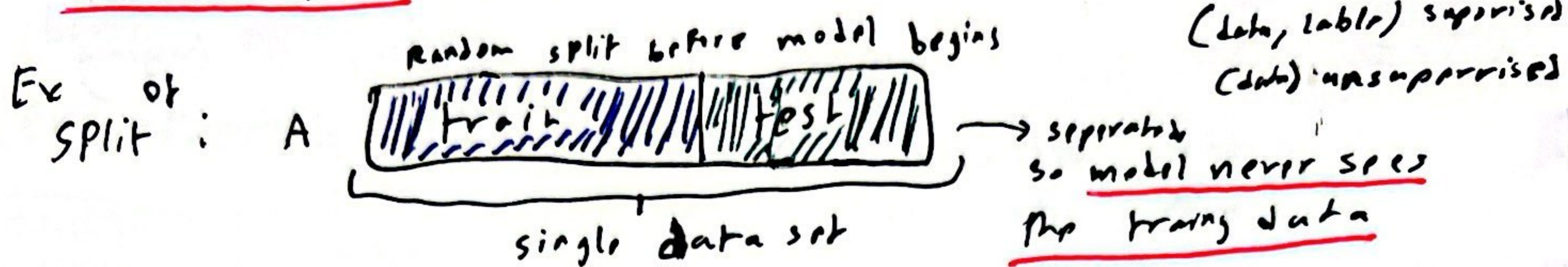


# Test Data (test set)

- is a separate collection of data used to evaluate how well a ML model performs on Ex of Unseen Data if it has not seen in training. Form:



- \* The test process helps verify if the model has truly learned to make good predictions, rather than memorizing or being wrong

- \* Data leakage is any inclusion of even parts of the test data in the training data



## Supervised Learning (70% of ML applications)

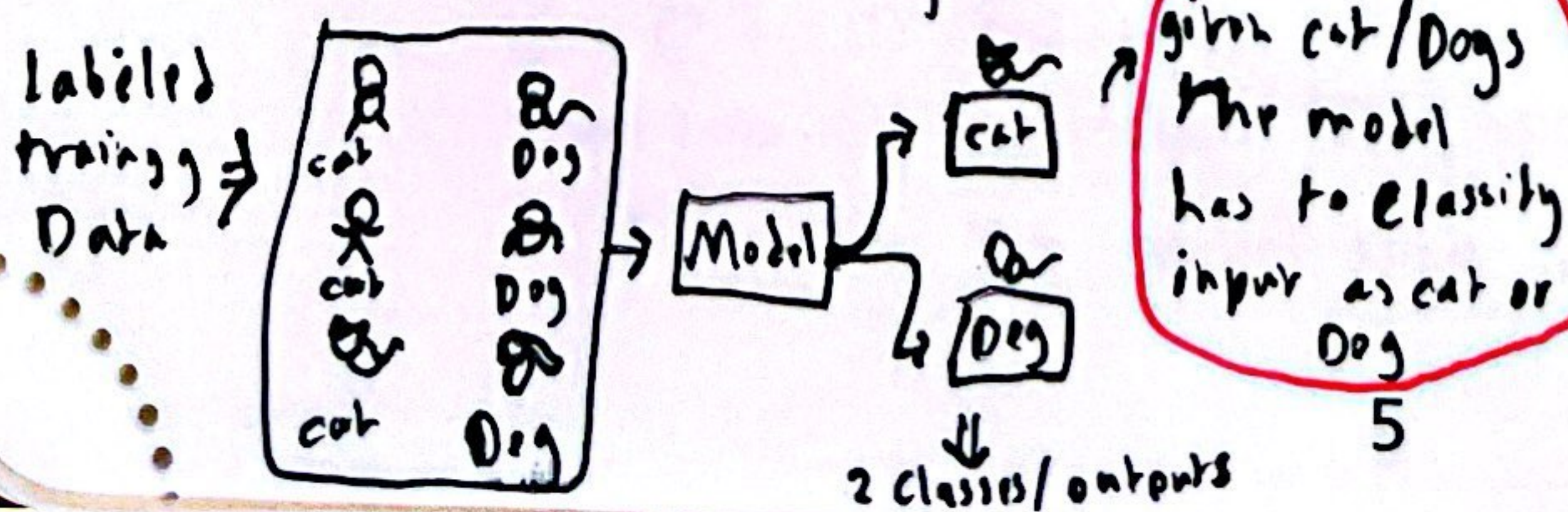
- is a branch of ML where models learn from labeled examples i.e. the true outcome/ target is known or provided. each Ex in training data has data + label (correct answer)

Ex: classifying dogs or cats see training data Ex's

## Unsupervised Learning (form: data) (Ex choose your own email labels)

- is a branch of ML where models learn to find patterns and structure in data without any labeled examples (correct answers, instead of being taught what to look for these models find and group relationships between data)

Ex (Supervised Learning)



Ex (Unsupervised)

Raw data no label

