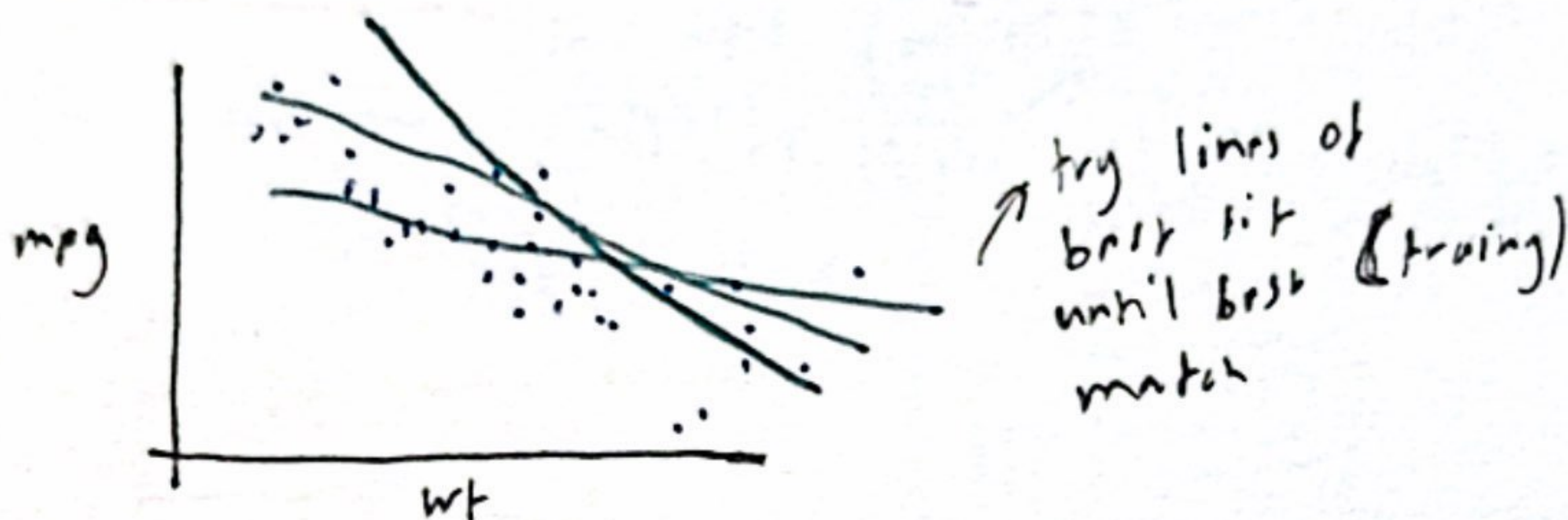


Model fitting (AKA training or learning)

- is the process of adjusting a models parameters to find the best match between the models predictions and the actual data

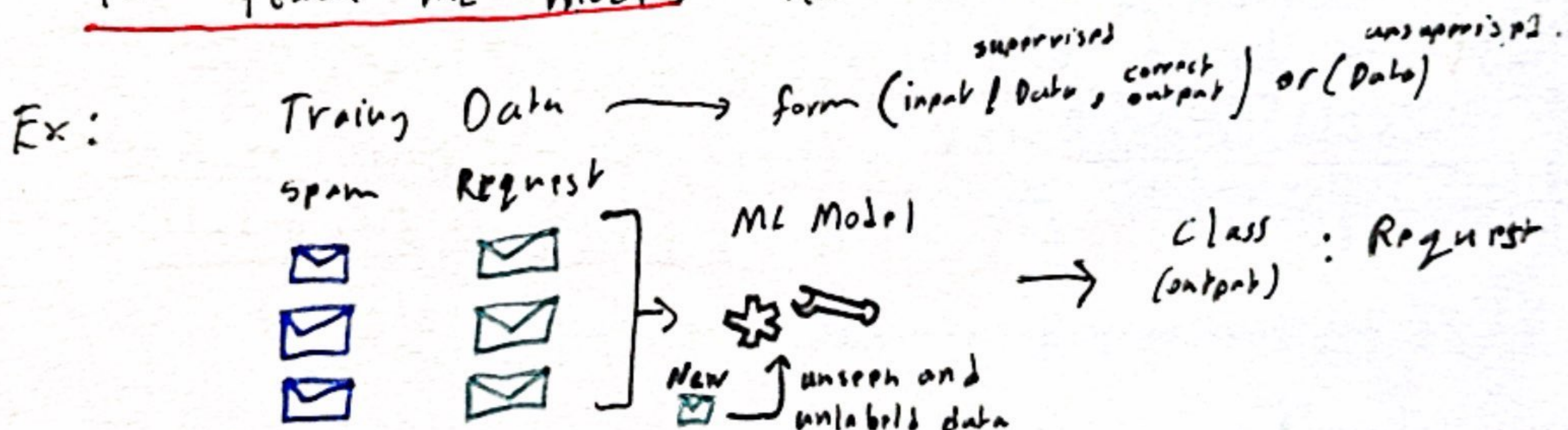
Ex: in linear Regression model fitting is trying out different lines until you find the line with the best fit



NOTE:
There are many types of fits
see later pages

Training Data (train set)

- is a carefully selected subset of data used to teach ML models how to make predictions



- the training data is 1000's of emails labeled as spam/non-spam used to teach the system to recognize characteristics of unwanted messages

Ex: in Images it could be 1000's of labeled

images of cats/dogs for ex and we would use that to teach the model what is cat vs dog it would learn to recognize it based on characteristics

- Ex cats have whiskers dogs have bungee ears etc ..

- Too much can be bad as at some point more data gives little improvement (diminishing returns), inc noise, inc computation, train time inc etc