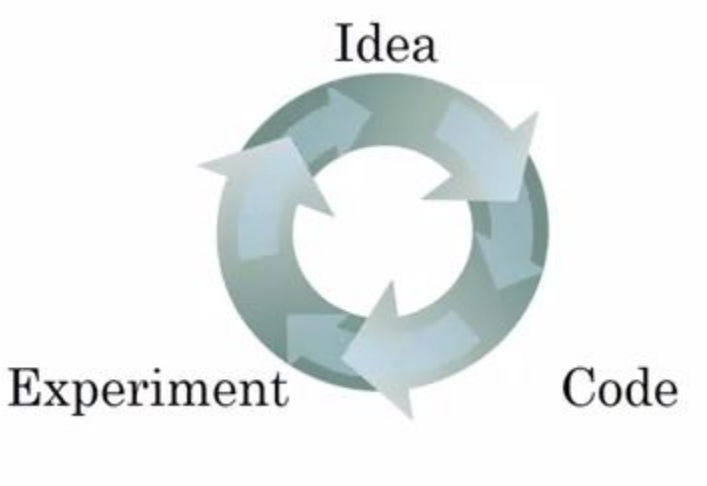
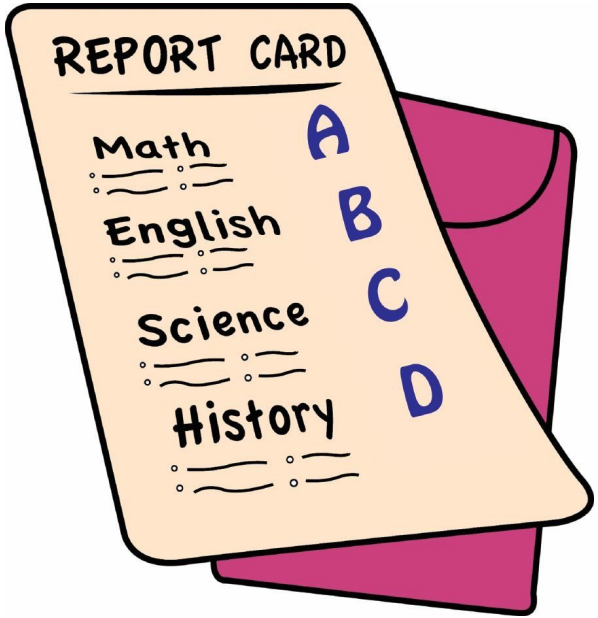
Diagnosing Machine Learning Models

Model life cycle



Evaluation metrics



Classification problems

Attributes

tongue_length_cm = 100

ear floppiness=10

nose_cat=very_wet

Prediction

Dog

✓

Binary classification

Predicted	Expected	Score
Cat	Cat	✓
Dog	Cat	✗
Dog	Dog	✓
Cat	Dog	✗

Confusion Matrix

Confusion matrix

	Expected	
Predicted	**True Positive** Eg. Predicts a pregnant woman as pregnant.	**False Positive** **(Type I Error)** Eg. Predicts a non-pregnant woman as pregnant.
	False Negative **(Type II Error)** Eg. Predicts a pregnant woman as non-pregnant.	**True Negative** Eg. Predicts a non-pregnant woman as not pregnant.

Classification metrics

Number of correct prediction in relation to the total number of predictions

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

	Expected	
	True Positives 40	**False Positives** 20
	False Negatives 10	**True Negatives** 30

$$\frac{40+30}{40+30+20+10} = 0.7$$

Classification metrics

Recall- How good our model is at *retrieving* a specific class

$$recall = \frac{TP}{TP+FN}$$

Precision- How *particular* our model can be with regards to a specific class

$$precision = \frac{TP}{TP+FP}$$

F1 score- *Weighted average* of recall and precision

$$F1 = 2 \bullet \frac{precision \bullet recall}{precision + recall}$$

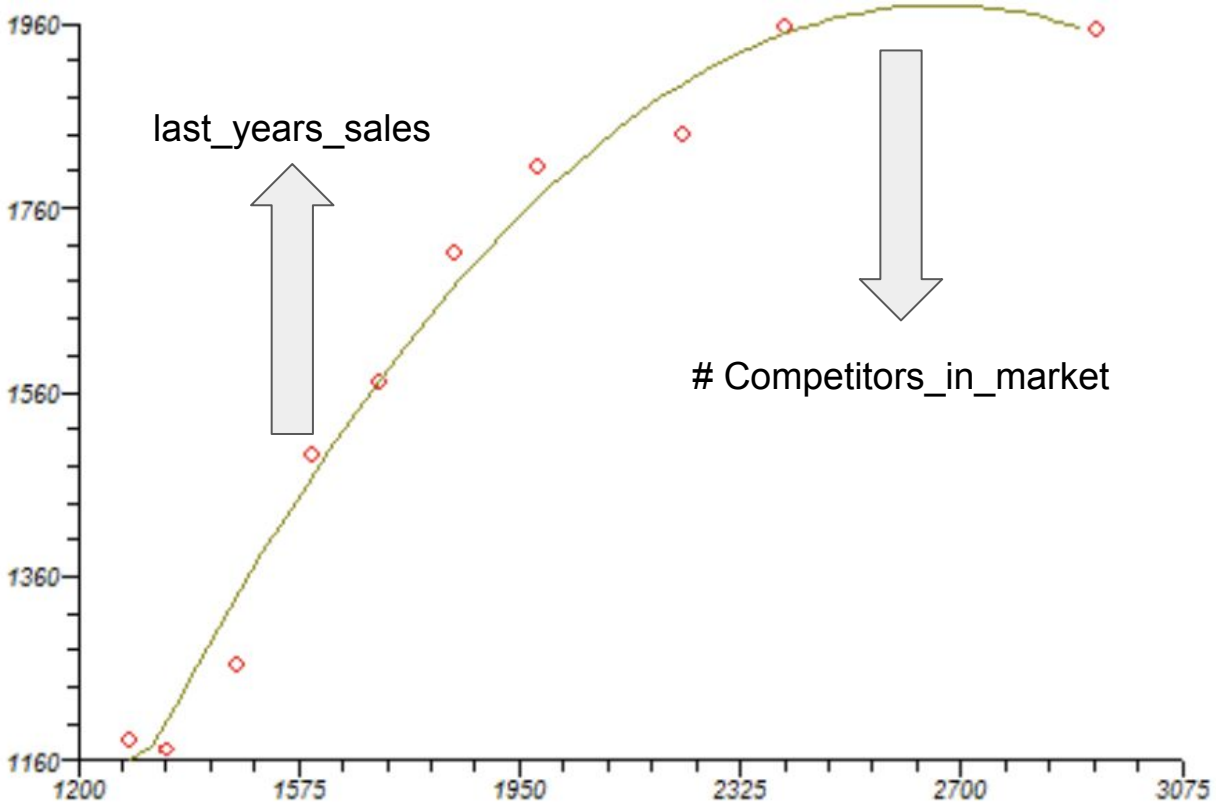
Precision-recall trade-off

Recall

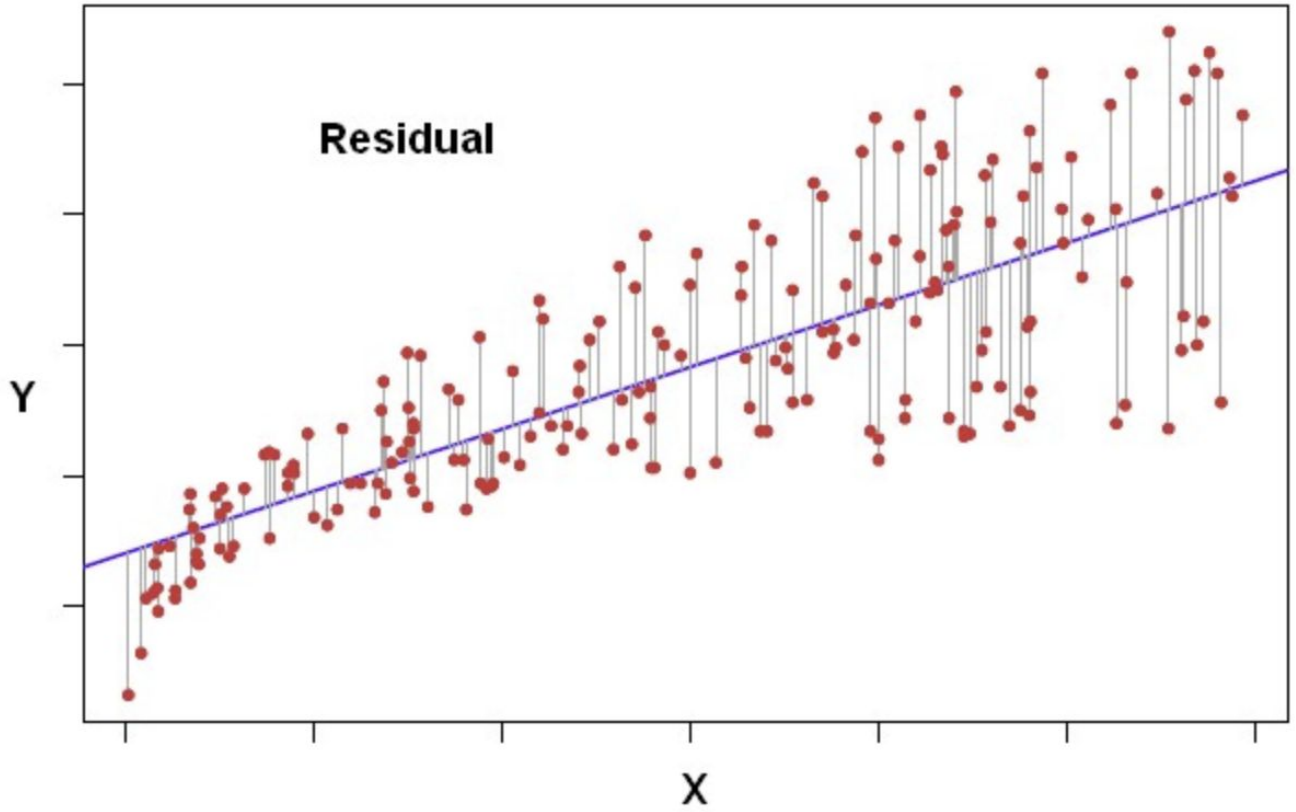
VS

Precision

Regression problems



Regression metrics



Regression metrics

Mean Absolute Error- Average of the *absolute* differences between the predicted and expected values

$$MAE = \frac{SUM(predicted\ values - expected\ values)}{Total\ number\ of\ samples}$$

Lower values indicate a better model

Regression metrics

Root Mean Square Error- Square root of the average of the square differences

$$\sqrt{\frac{SUM(predicted\ values - expected\ values)2}{Total\ number\ of\ samples}}$$

Lower values indicate a better model