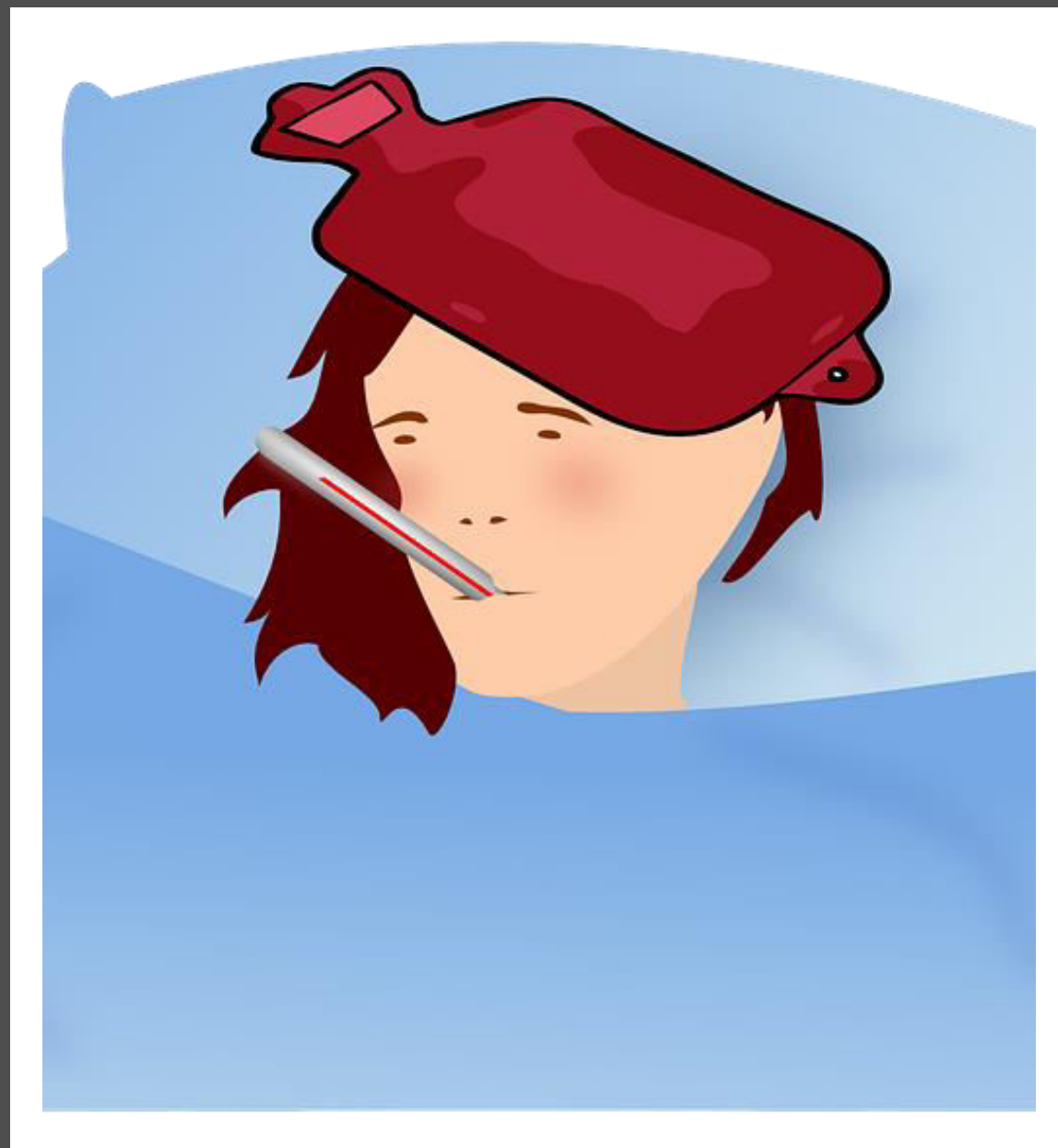


Wrong Evaluation Metric







1. Emotional and Psychological Burden

2. Financial Burden

3. Possible Health Deterioration due to Unrequired Treatment

Potential Death



Evaluation Metric

Classification

Regression

Evaluation Metric

Emergency Vs Non-Emergency Classification

Non – Emergency(0) = 1361

Emergency (1) = 991

		Actual	
		Y	N
Predicted	Y	TP	FP
	N	FN	TN

		Actual	
		Y	N
Predicted	Y	TP = 900	FP = 61
	N	FN =91	TN = 1300

= 93.5 %

Evaluation Metric

Emergency Vs Non-Emergency Classification

Non – Emergency(0) = 1361

Emergency (1) = 991

		Actual	
		Y	N
Predicted	Y	TP	FP
	N	FN	TN

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

Evaluation Metric - Classification

Accuracy

Balanced Dataset, Equal Importance to Positive and Negative Errors

		Actual	
		Y	N
Predicted	Y	TP	FP
	N	FN	TN

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

Evaluation Metric

Fraud Detection

No Fraud(0) = 1000

Fraud(1) = 5

		Actual	
		Y	N
Predicted	Y	TP = 0	FP = 0
	N	FN = 5	TN = 1000

Accuracy = 99.5 %

Evaluation Metric - Classification

Recall

Ability of a model to find all the data points of interest(Fraud in this case)

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

Evaluation Metric

Fraud Detection

No Fraud(0) = 1000

Fraud(1) = 5

		Actual	
		Y	N
Predicted	Y	TP = 0	FP = 0
	N	FN = 5	TN = 1000

Accuracy = 99.5 %

$$recall = \frac{TP}{TP + FN} = 0$$

Evaluation Metric

Fraud Detection

No Fraud(0) = 1000

Fraud(1) = 5

		Actual	
		Y	N
Predicted	Y	TP = 5	FP = 1000
	N	FN = 0	TN = 0

Accuracy = 0.5 %

$$recall = \frac{TP}{TP + FN} = 1$$

Evaluation Metric - Classification

Precision

Proportion of Data points that the model says relevant are actually relevant.

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

Evaluation Metric

Fraud Detection

No Fraud(0) = 1000

Fraud(1) = 5

		Actual	
		Y	N
Predicted	Y	TP = 5	FP = 1000
	N	FN = 0	TN = 0

Accuracy = 0.5 %

$$precision = \frac{TP}{TP + FP} = 5/1005$$

Evaluation Metric

Fraud Detection

No Fraud(0) = 1000

Fraud(1) = 5

		Actual	
		Y	N
Predicted	Y	TP = 0	FP = 0
	N	FN = 5	TN = 1000

Accuracy = 99.5 %

$$precision = \frac{TP}{TP + FP} = 0$$

Evaluation Metric

Fraud Detection

No Fraud(0) = 1000

Fraud(1) = 5

Actual

Y N

Predicted

	Y	N
Y	TP = 1	FP = 0
N	FN = 4	TN = 1000

Accuracy = 99.6 %

$$precision = \frac{TP}{TP + FP} = 1$$

Evaluation Metric - Classification

F1-Score

Equal Importance of False Positives and False Negatives.

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

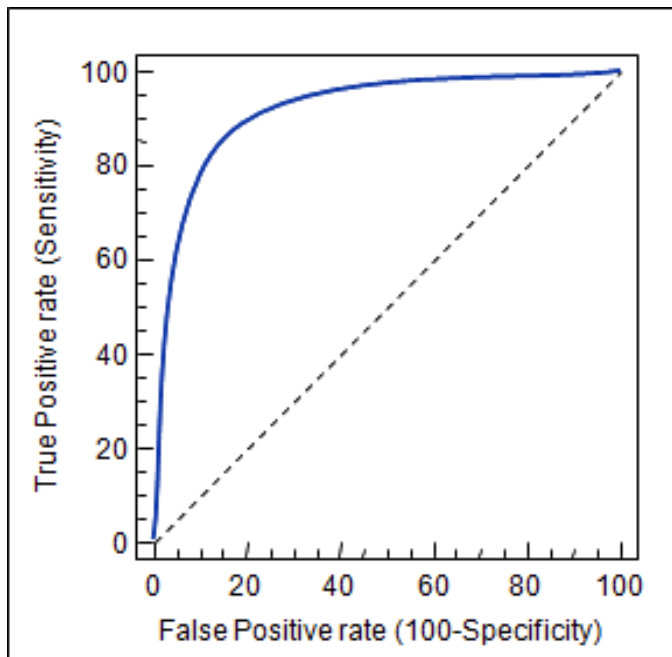
$$precision = \frac{TP}{TP + FP}$$

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

Evaluation Metric - Classification

AUC-ROC

When you want to increase your True Positive Rate as well as decrease your False positive Rate.



$$\text{TPR (sensitivity)} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{FPR (1-specificity)} = \frac{\text{FP}}{\text{TN} + \text{FP}}$$

Evaluation Metric - Classification

Black and White

Actual	Predicted
0	1
0	0
1	0
1	1

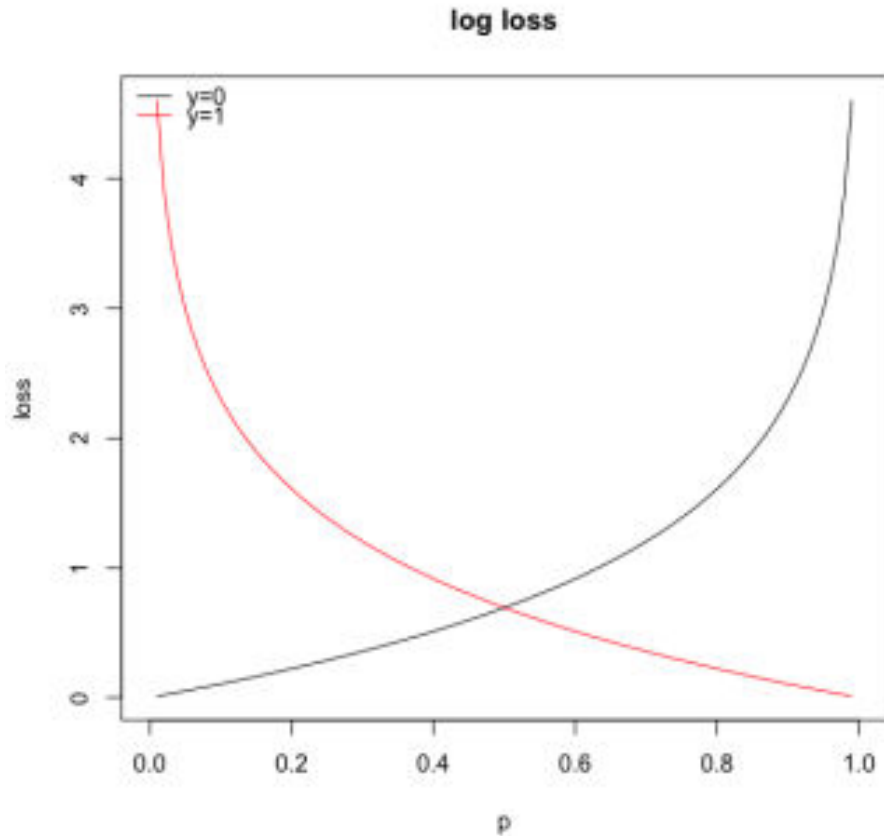
Gray

Actual	Predicted
0	0.2
0	0.7
1	0.8
1	0.5

Evaluation Metric - Classification

Log-Loss

Penalize model in exponential form on the basis of the confidence of the model.



$$\text{LogLoss} = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

Thank You

Evaluation Metric - Regression

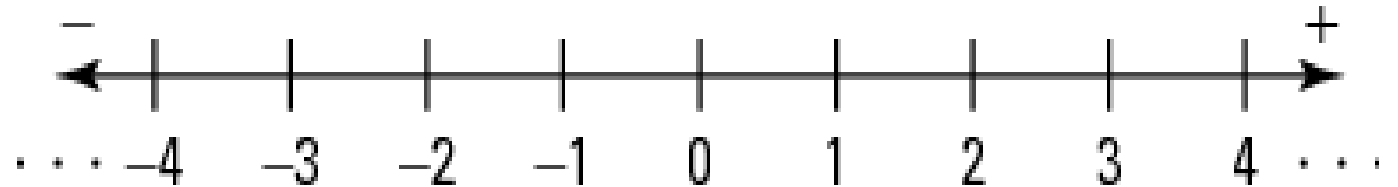
Actual	Predicted	Difference	Absolute Difference
50	49	1	1
40	42	-2	2
20	22	-2	2
40	43	-3	3
42	38	4	4
36	33	3	3
23	26	-3	3
233	231	2	2
		0	20

Evaluation Metric - Regression

MAE

1. Equal weight to Over Prediction and Under Prediction (Direction of Prediction)
2. Penalty in Proportion to Prediction Values.

$$MAE = \frac{1}{n} \sum_{j=1}^n \|y_j - \hat{y}_j\|$$



Evaluation Metric - Regression

Actual	Predicted
50	49
40	42
20	22
40	43
42	38
36	33
23	26
233	533

Error
1
-2
-2
-3
4
3
-3
-300

-302

Absolute Error
1
2
2
3
4
3
3
300

318

Evaluation Metric - Regression

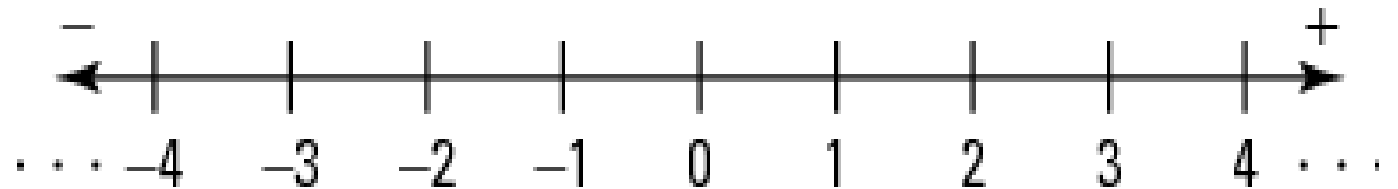
Actual	Predicted	Difference	Squared Error
50	49	1	1
40	42	-2	4
20	22	-2	4
40	43	-3	9
42	38	4	16
36	33	3	9
23	26	-3	9
233	533	-300	90000
		-302	90052

Evaluation Metric - Regression

RMSE

1. Not affected by Over Prediction or Under Prediction (Direction of Prediction)
2. Larger Errors to be penalized much heavier than Smaller Errors

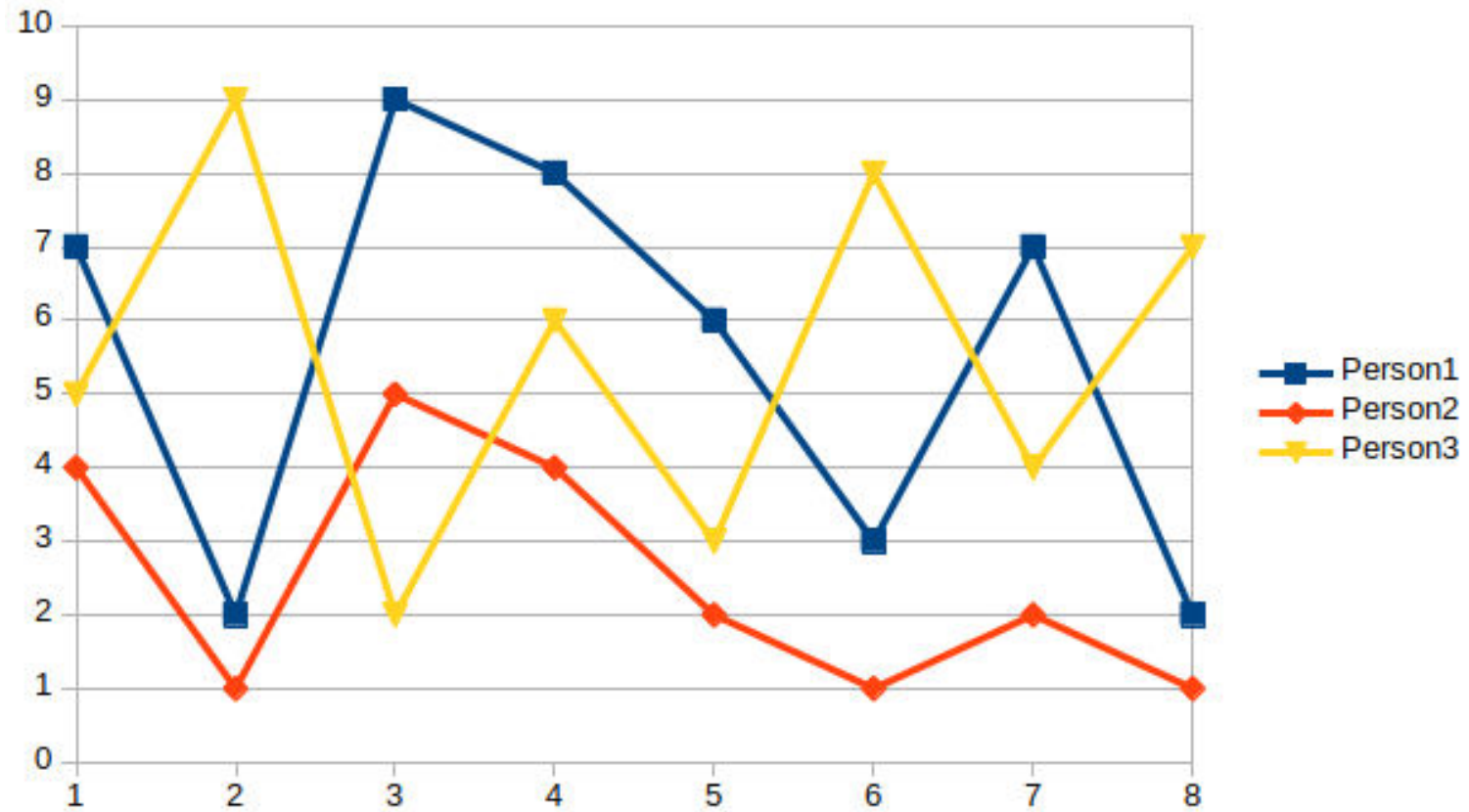
$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2}$$



Evaluation Metric - Regression

Movie Genres	Person 1	Person 2	Person 3
Horror	7	4	5
Comedy	2	1	9
Sci-Fi	9	5	2
Adventure	8	4	6
Crime	6	2	3
Drama	3	1	8
War	7	2	4
Western	2	1	7

Evaluation Metric - Regression



Evaluation Metric - Regression

Cosine Similarity

Similarity between shape of Prediction and Actual is Important rather than the difference between them.

$$\frac{\sum_{j=1}^n y_j \hat{y}_j}{\sqrt{\frac{1}{n} \sum_{j=1}^n y_j^2} \sqrt{\frac{1}{n} \sum_{j=1}^n \hat{y}_j^2}}$$

Optimising Evaluation Metric – Loss Functions

Evaluation Metrics

Accuracy

Precision

Recall

F1-score

Log Loss

Loss Functions

Binary
Cross - Entropy

Categorical
Cross - Entropy

Optimising Evaluation Metric – Loss Functions

Evaluation Metrics

MAE

RMSE

Cosine Similarity

Loss Functions

mean_absolute_
error

mean_squared_
error

cosine_proximity

Thank You