Gurnsewak Singh

29/Jan/2026

Assessment 1

1) a) Large K with noisy data

2) c) They always overfit

3) d) Improving feature selection

4) d) model becomes linear

5) a) Target variable is distributed

6) c) Sigmoid

7) a) Precision

8) d) overfitting

9) c) Because distance calc. depends on scale

10) a) KNN

	Predicted Fraud	Predicted not fraud
Actual Fraud	120 (TP)	30 (FN)
Actually not fraud	50 (FP)	800 (TN)

=

$$\text{Accuracy} = \frac{TP}{TP+FN} = \frac{120}{}$$

$$\frac{TP+TN}{\text{Total}}$$

$$= \frac{120+800}{1000}$$

$$= \frac{92}{100}$$

$$= 0.92$$

a) Accuracy $= 92\%$

c) $$\text{Recall} = \frac{TP}{TP + FN}$$

$$= \frac{120}{120 + 30}$$

$$= \frac{120}{150}$$

$$= 0.8$$

b) $$\text{Precision} = \frac{TP}{TP + FP}$$

$$= \frac{120}{120 + 50} = \frac{120}{170}$$

$$= 0.706$$

d) $$\text{F1 score} = \frac{2(\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$$

$$= \frac{2(0.706 \times 0.8)}{0.706 + 0.8}$$

$$= \frac{1.129}{1.506}$$

$$= 0.74$$

c) No, Although accuracy is more than 90% but precision is 70 percent which is not a good score for model to get accepted.