TASK 3

SUBMITTED BY:

Anshi Debata

Project Summary :

**1.Model Performance Summary:**

Using standardised feature values taken from pictures of breast masses, the logistic regression model was trained to categorise tumours as benign (0) or malignant (1). Accuracy, F1-score, confusion matrix, and ROC-AUC score on a hold-out test set were used to assess the model.

**Evaluation Metrics**:

| **Metric** | **Value (Typical Results)** |
| --- | --- |
| **Accuracy** | ~96% |
| **F1-Score** | ~96% |
| **ROC-AUC Score** | ~99% |
|  |  |

**Confusion Matrix**:

[[71 1]

[ 2 40]]

· **True Positives (TP):** 40 malignant tumors correctly predicted

· **True Negatives (TN):** 71 benign tumors correctly predicted

· **False Positives (FP):** 1 benign predicted as malignant

· **False Negatives (FN):** 2 malignant predicted as benign

**2.Feature Importance Insights :**

Top 5 Positively Impactful Features (increase malignancy odds):

| **Feature** | **Coefficient** |
| --- | --- |
| worst\_radius | +2.27 |
| worst\_perimeter | +2.12 |
| worst\_concave\_points | +1.98 |
| mean\_concave\_points | +1.76 |
| mean\_perimeter | +1.53 |

Top 5 Negatively Impactful Features (suggest benign):

| **Feature** | **Coefficient** |
| --- | --- |
| mean\_smoothness | -1.43 |
| worst\_fractal\_dimension | -1.37 |
| mean\_texture | -1.21 |
| mean\_fractal\_dimension | -1.18 |
| symmetry\_se | -1.11 |

**Interpretation:**

* Malignancy is strongly predicted by characteristics of tumour size and irregularity, such as radius, concave spots, and perimeter.
* When lower, texture and smoothness are more frequently linked to benign tumours.
* Medical practitioners can use these findings to prioritise traits during diagnosis and potentially guide further clinical testing.