c.1 Which feature extraction method produced the best results for each classifier? Why do you think this is the case?

GLCM features gave the best results in Naive Bayes

PCA features gave the best results in Logistic Regression

For Naive Bayes, GLCM worked slightly better — probably because it captures texture information that matches the model's assumption that features are somewhat independent.

For Logistic Regression, PCA features gave the highest accuracy, since PCA removes redundant noise and keeps the most important components for a linear decision boundary.

In contrast, Histogram features didn't perform as well because they only describe color distributions, not textures or spatial patterns, which are crucial for medical images like skin lesions.

c.2 Compare the performance of Naive Bayes, Logistic Regression, and LDA. Which model performed best and under what conditions?

Logistic Regression came out on top overall — especially with PCA features (0.8144 accuracy). It handles correlated data well and doesn't rely on the strong independence assumptions that Naive Bayes does.

Naive Bayes still did decently, especially with GLCM features (0.7500), but its simple assumptions limit it a bit.

LDA performed somewhere in between. It worked okay on raw pixels (0.7708), but it wasn't as strong as Logistic Regression.

c.3 How does LDA perform on the raw pixel data compared to the other models using extracted features? Explain why this might be a poor strategy for images.

LDA did okay on raw pixel data (0.7708), but not as well as the models that used extracted features.

That's because raw pixels are high-dimensional and highly correlated. LDA assumes all classes share the same covariance matrix — which clearly doesn't hold for images.

In short, it struggles with the complexity and correlation in pixel-level data, so it can overfit or produce unstable boundaries.

c.4 The ViT model is a complex deep learning model. What are the advantages and disadvantages of using it solely as a feature extractor for a simple model like Naive Bayes?

Pros: ViT (Vision Transformer) features are really rich — they capture global structure and texture without manual feature design. You can reuse pretrained weights, which saves a ton of time.

Cons: Those features are very high-dimensional and heavily correlated, which breaks Naive Bayes' independence assumption. Plus, generating ViT embeddings is slow and GPU-heavy.

c.5 Summarize your findings on the relationship between feature engineering and model performance for this image classification task.

Feature engineering is extremely important — good features can significantly improve model performance. These engineered features capture structure and texture information that raw pixels or simple color histograms fail to represent.