Problem1：

1. Model with the highest recall: Depth 3.0 (recall=1.0)  
   Reason: As the depth of the model increases, the complexity of the decision tree increases, and the training data can be better fitted, thereby improving the recall rate.  
   2. Model with the lowest precision: Depth 1.0 (precision=0.5)  
   Reason: The Depth 1.0 model is too simple to accurately distinguish between positive and negative classes, resulting in a large number of false positives and reduced accuracy.  
   3. Model with the best F1 score: Depth 3.0 (F1 score = 1.0)  
   Reason: The Depth 3.0 model achieves the best balance between complexity and performance, which can fully identify positive samples and avoid false positives.  
   4.Micro：  
   Globally calculates metrics and directly counts TP, FP, and FN for all categories. Applies to datasets with unbalanced categories.  
   Macro：  
   Metrics are calculated independently for each category and then taken as an unweighted average. Suitable for scenarios where all categories are equally important.  
   Weighted：  
   Metrics are calculated independently for each category, and then weighted average based on the sample size of the category. Suitable for scenarios where the categories are unbalanced but require trade-offs.

Problem2：

Entropy: 0.929

Gini Index: 0.452

Misclassification Error: 0.345

Information Gain: 0.579

First Split Feature: Uniformity of Cell Size, Threshold: 2.5

Problem3：

Original data model:

F1 Score: 0.9302, Precision: 0.9302, Recall: 0.9302

FP: 3, TP: 40, FPR: 0.0423, TPR: 0.9302

PCA1 model (only the first principal component):

F1 Score: 0.8889, Precision: 0.9474, Recall: 0.8372

FP: 2, TP: 36, FPR: 0.0282, TPR: 0.8372

PCA2 model (the first two principal components):

F1 Score: 0.9425, Precision: 0.9318, Recall: 0.9535

FP: 3, TP: 41, FPR: 0.0423, TPR: 0.9535