**Decision Tree Classifier**

* **Under sampling significantly improves TP rate (Recall):** Notice how the TP rate jumps drastically for the under sampled datasets. This means the model is much better at correctly identifying individuals with heart disease after under sampling. For example, with "10 features", the TP rate goes from 0.0704 to 0.6563. This is a crucial improvement, especially in a medical context where identifying true positives is vital.
* **Under sampling comes at the cost of accuracy and precision:** While under sampling helps with recall, it reduces accuracy and precision. This is because the model becomes more sensitive to the minority class (heart disease), leading to more false positives.
* **More features don't always mean better performance:** In the original datasets, using all features resulted in the highest TP rate (0.2461) but the lowest accuracy (0.8540). This suggests that some features might be noisy or irrelevant, hindering the model's performance.
* **F1-score provides a balanced view:** The F1-score, which balances precision and recall, is highest for the under sampled datasets. This highlights the importance of considering both metrics, especially when dealing with imbalanced datasets.
* **ROC AUC is generally low:** The ROC AUC values are relatively low across all datasets, indicating that the model's overall ability to discriminate between classes is not very strong.

**Analysis by Dataset:**

* **10 features:** This dataset shows the most dramatic improvement with under sampling, with a significant increase in TP rate and F1-score.
* **20 features:** Similar to "10 features," under sampling improves TP rate and F1-score, but the improvement is slightly less pronounced.
* **All features:** Under sampling still helps with recall, but the accuracy and precision drop more significantly compared to the other datasets.

**Conclusions:**

* **Under sampling is crucial for improving the model's ability to detect heart disease.** However, it's important to be aware of the trade-off with accuracy and precision.
* **Feature selection plays a role in model performance.** Using all features doesn't necessarily lead to the best results.