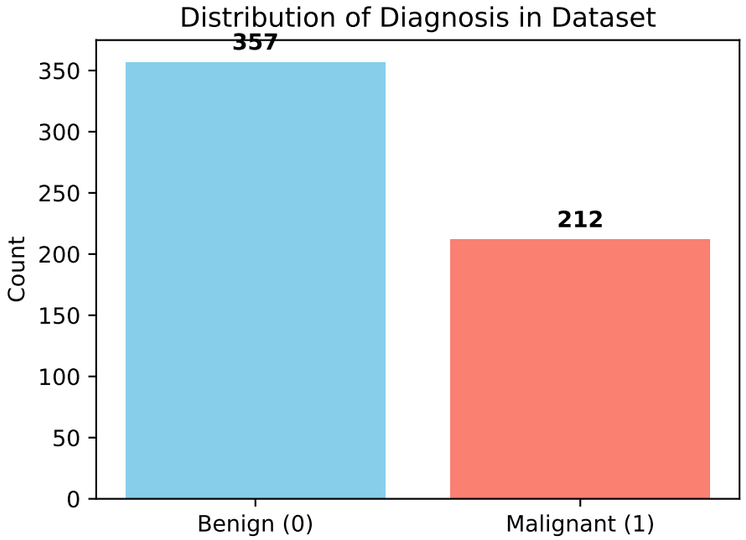
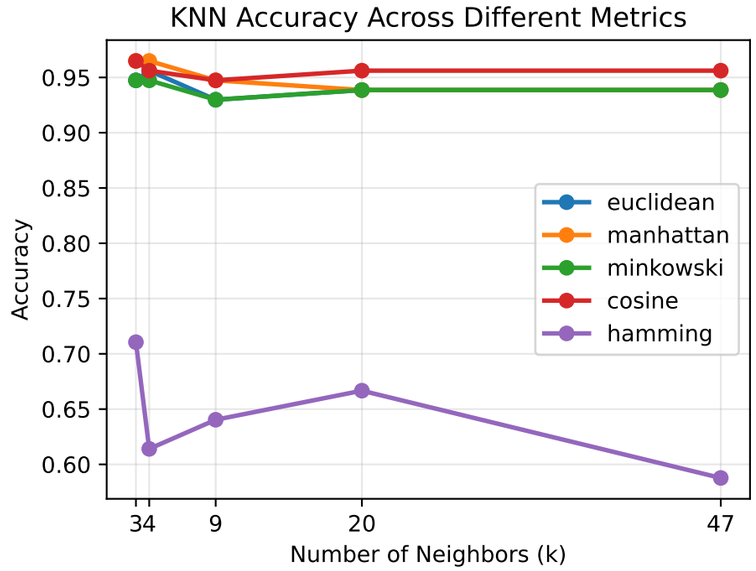
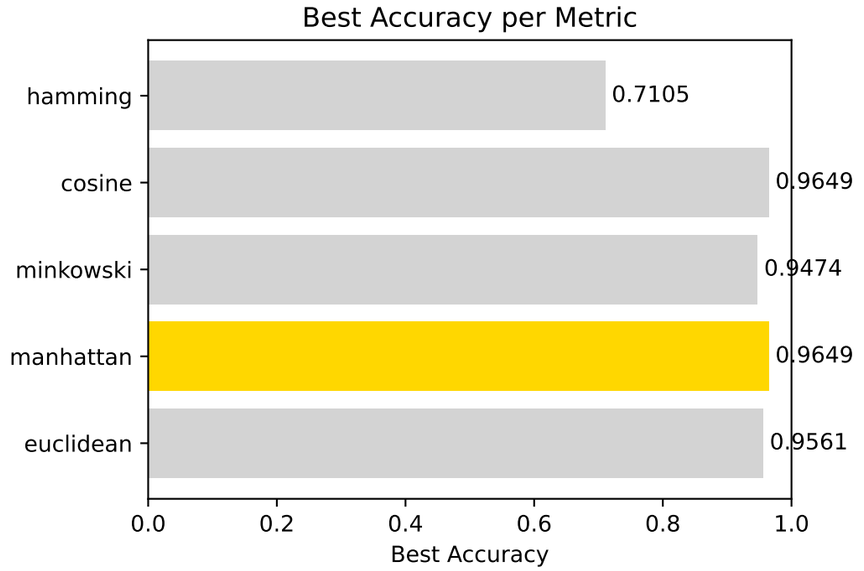
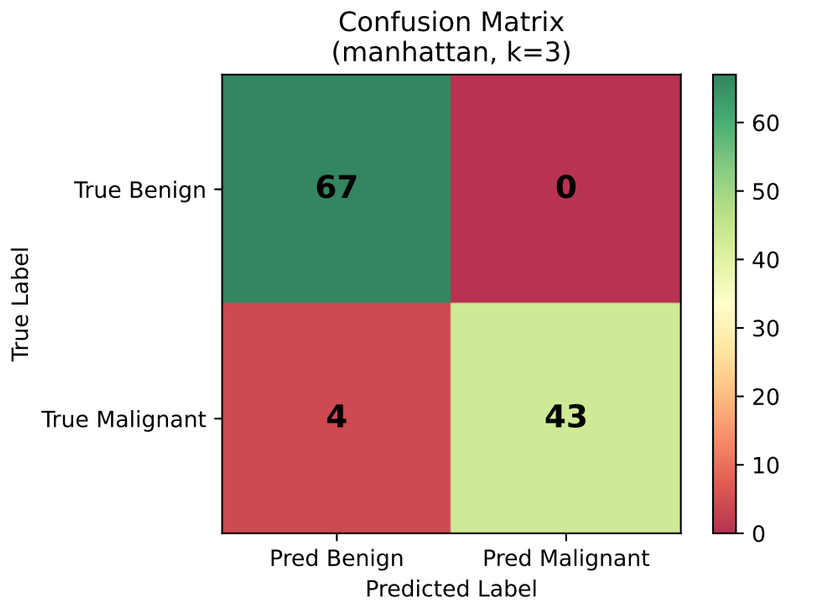
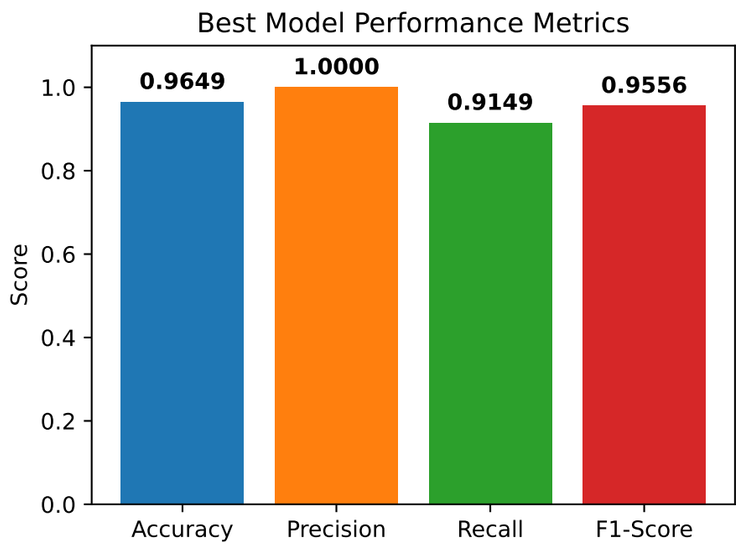
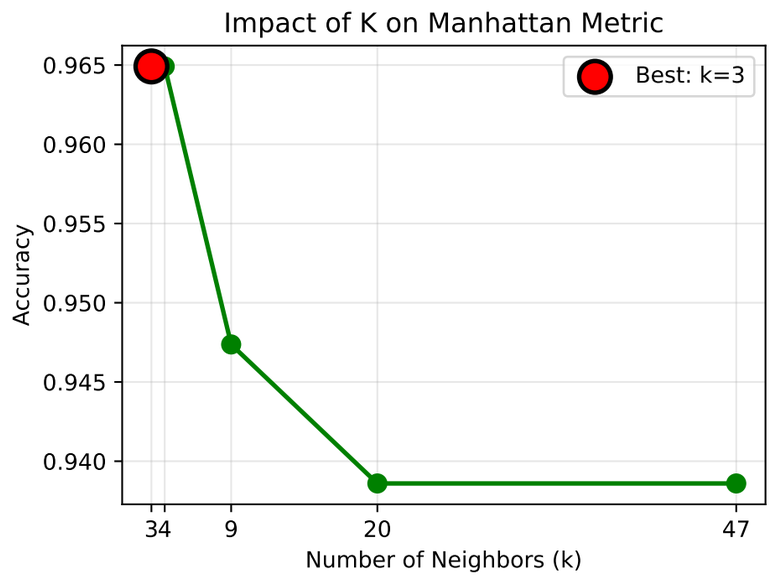

Distribution of Diagnosis in Dataset
Count
357
212
Benign (0)
Malignant (1)

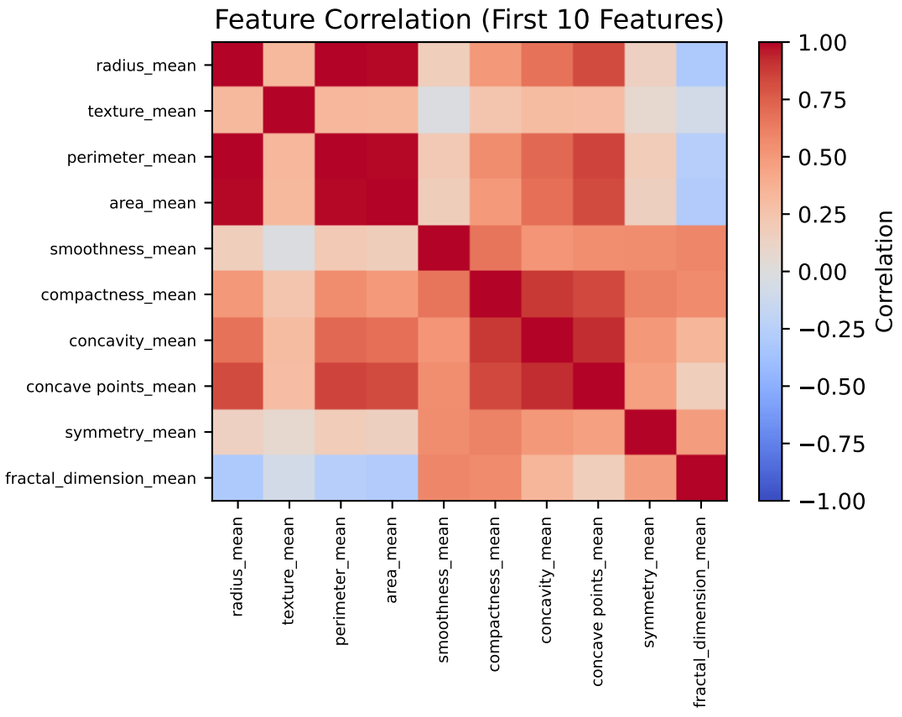

KNN Accuracy Across Different Metrics
Accuracy
Number of Neighbors (k)
euclidean
manhattan
minkowski
cosine
hamming

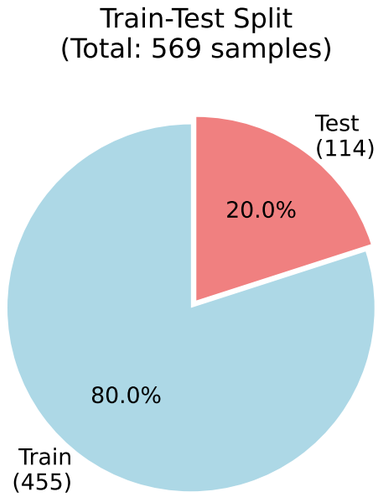

Best Accuracy per Metric
hamming
cosine
minkowski
manhattan
euclidean
0.7105
0.9649
0.9474
0.9649
0.9561
Best Accuracy

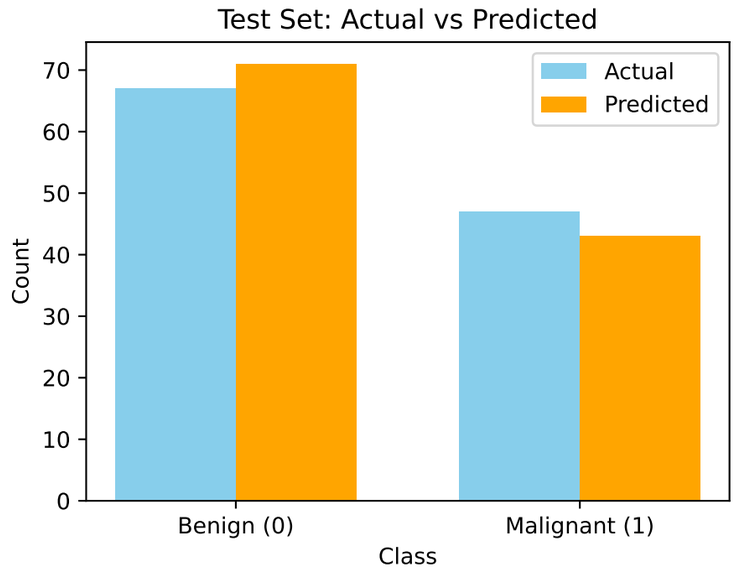

Confusion Matrix
(manhattan, k=3)
True Label
True Benign
True Malignant
67
0
4
43
Pred Benign
Pred Malignant
Predicted Label


Best Model Performance Metrics
Score
0.9649
1.0000
0.9149
0.9556
Accuracy
Precision
Recall
F1-Score


Impact of K on Manhattan Metric
Accuracy
Best: k=3
Number of Neighbors (k)


Feature Correlation (First 10 Features)
radius_mean
texture_mean
perimeter_mean
area_mean
smoothness_mean
compactness_mean
concavity_mean
concave points_mean
symmetry_mean
fractal_dimension_mean
Correlation


Train-Test Split
(Total: 569 samples)
Test
(114)
20.0%
80.0%
Train
(455)


Test Set: Actual vs Predicted
Actual
Predicted
Count
Benign (0)
Malignant (1)
Class

KNN CLASSIFIER - COMPREHENSIVE REPORT

DATASET INFORMATION:
• Total Samples: 569
• Training Samples: 455 (80.0%)
• Testing Samples: 114 (20.0%)
• Number of Features: 30
• Benign Cases: 357 (62.7%)
• Malignant Cases: 212 (37.3%)

MODEL EVALUATION SUMMARY:
• Metrics Tested: euclidean, manhattan, minkowski, cosine, hamming
• K Values Tested: [3, 4, 9, 20, 47]
• Total Configurations: 25

BEST MODEL CONFIGURATION:
• Distance Metric: MANHATTAN
• Optimal K: 3
• Test Accuracy: 0.9649 (96.49%)

DETAILED PERFORMANCE METRICS:
• Precision: 1.0000 (100.00%)
• Recall: 0.9149 (91.49%)
• F1-Score: 0.9556 (95.56%)

CONFUSION MATRIX BREAKDOWN:
• True Negatives (TN): 67 (Correctly predicted Benign)
• False Positives (FP): 0 (Benign predicted as Malignant)
• False Negatives (FN): 4 (Malignant predicted as Benign)
• True Positives (TP): 43 (Correctly predicted Malignant)

ALL METRICS COMPARISON:
• Euclidean   : Best Accuracy = 0.9561 at k=4
• Manhattan   : Best Accuracy = 0.9649 at k=3
• Minkowski   : Best Accuracy = 0.9474 at k=3
• Cosine   : Best Accuracy = 0.9649 at k=3
• Hamming   : Best Accuracy = 0.7105 at k=3

KEY INSIGHTS:
✓ Perfect Precision: No false positives (no benign cases misclassified as m
✓ High Recall (91.49%): Model catches most malignant cases
✓ Excellent Accuracy (96.49%): Model performs very well overall
⚠ Warning: 4 malignant case(s) misclassified as benign (critical error)

Report generated successfully!
Visualization included.