Breast Cancer Classification Model

Project Objective

Develop a machine learning–based system to classify breast cell samples as benign (non-cancerous) or malignant (cancerous).
To enhance diagnostic speed, consistency, and decision support for pathologists.

Model Performance

Key Results

Metric	Result	Interpretation
Overall Accuracy	96.4%	Correctly classifies 96 out of every 100 cases
AUC Score	0.997	Near-perfect ability to distinguish between classes
Benign Recall	97%	Correctly identifies 97% of benign samples
Malignant Recall	95%	Correctly identifies 95% of malignant samples

Summary:

Out of 137 test samples, 132 were classified correctly and 5 were misclassified.

Clinical Implications

This system functions as a decision support tool, not a standalone diagnostic method.
All predictions should be validated by pathologists and integrated with standard diagnostic workflows.
The 3 false negatives (2.2%) emphasize the importance of human oversight.

This system ensures:

- A near-instant preliminary analysis of cytology data.
- Reduces diagnostic variability between reviewers.
- Assists clinicians in identifying high-risk cases.
- Improves triage and resource allocation.

ROC Curve Insights

- **AUC = 0.997** demonstrates clear distinction between benign and malignant cases.
- The ROC curve closely follows the top-left corner, indicating high sensitivity and specificity.

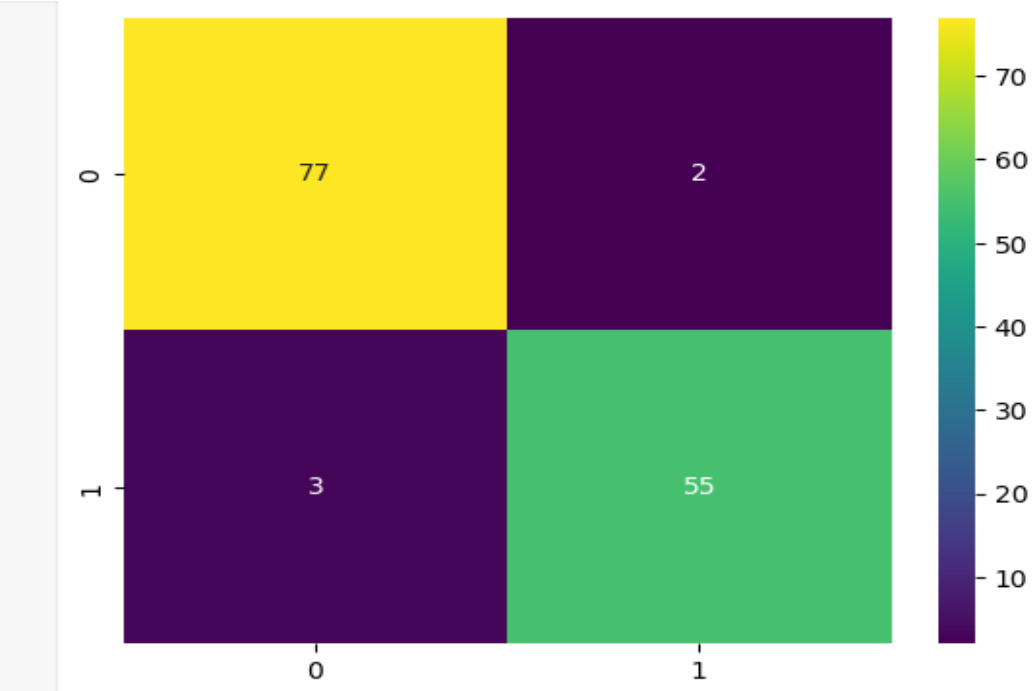
Technical Overview

1. Data: Analyzed 683 patient samples with 9 diagnostic features
2. Model: Used Support Vector Machine (SVM) with RBF kernel and Identified 78 support vectors as key decision boundaries

Recommendations

1. Pilot Deployment- integrate the model into existing diagnostic workflow in shadow mode and collect performance metrics and clinician feedback.
2. System Integration- embed within laboratory information systems and provide real-time classification suggestions and alerts.
3. Optimization- ensure a retraining of the model with new patient data and monitor performance via a validation dashboard.

Confusion Matrix Summary



- True Negatives (77) – Correctly identified benign samples.
- True Positives (55) – Correctly identified malignant samples.
- False Positives (2) – Benign misclassified as malignant.
- False Negatives (3) – Malignant missed.

Value Proposition

Operational Gains

- Faster diagnostic turnaround times.
- Prioritization of high-risk cases.
- Reduced workload variability among staff.
- Enhanced diagnostic consistency and traceability.

Patient Impact

- Quicker preliminary results.
- Shorter wait times for review.
- More efficient and confident care pathways.

Conclusion

This **Breast Cancer SVM Classification Model** achieves 96.4% accuracy and 0.997 AUC, showing exceptional reliability in differentiating benign and malignant samples.

When integrated as a clinical decision support tool, it can significantly:

- Improve diagnostic efficiency
- Enhance workflow consistency
- Support faster, more confident patient management