

A decorative graphic on the left side of the slide consists of white lines and circles on a blue gradient background, resembling a circuit board or neural network structure.

# CLINICAL DECISION SUPPORT WITH MACHINE LEARNING: PREDICTING BREAST CANCER OUTCOMES

BY: ASIANA HOLLOWAY

COURSE: SAT5141 - CLINICAL DECISION SUPPORT & AI MODELING



# ENHANCING CLINICAL DECISION- MAKING WITH AI

---

## Why Breast Cancer Prediction?

---

### Importance of early detection

---

### Role of ML in oncology

---

### How prediction assists CDSS

# DATASET OVERVIEW

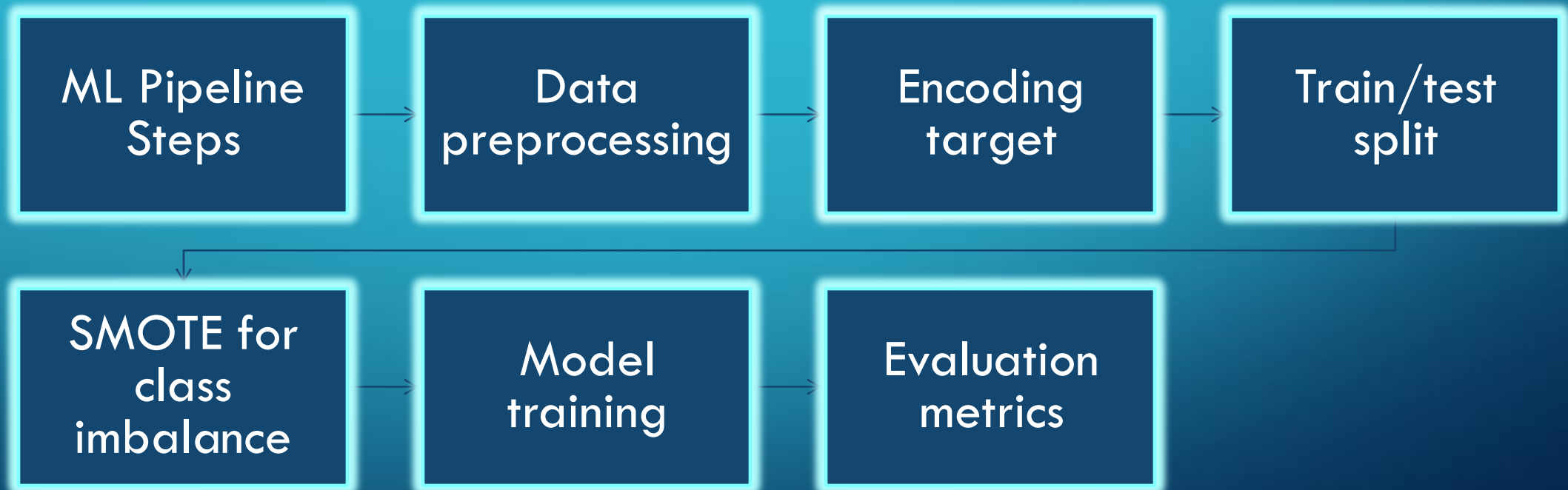
## Wisconsin Diagnostic Breast Cancer Dataset

- 569 samples
- 30 imaging features
- Target: Benign vs. Malignant
- No missing values
- Sourced from Kaggle

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactness_mean	concavity_mean	concave points_mean	...	texture_worst	perimeter_worst	area_worst	smoothnes
0	842302	M	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.3001	0.14710	...	17.33	184.60	2019.0	
1	842517	M	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.0869	0.07017	...	23.41	158.80	1956.0	
2	84300903	M	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.1974	0.12790	...	25.53	152.50	1709.0	
3	84348301	M	11.42	20.38	77.58	386.1	0.14250	0.28390	0.2414	0.10520	...	26.50	98.87	567.7	
4	84358402	M	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.1980	0.10430	...	16.67	152.20	1575.0	

5 rows x 33 columns

# METHODOLOGY OVERVIEW



# MODELS USED



Machine Learning  
Models Compared



Logistic Regression



Decision Tree



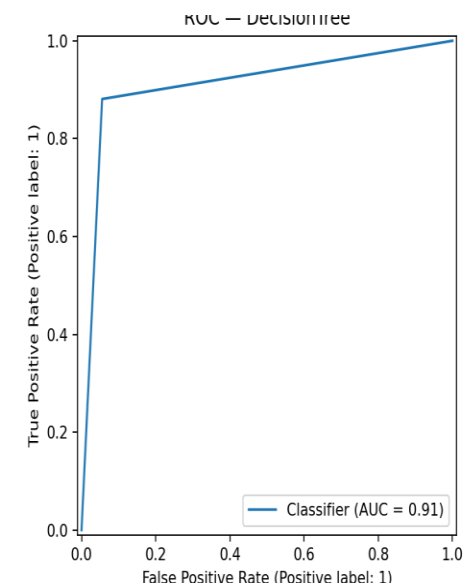
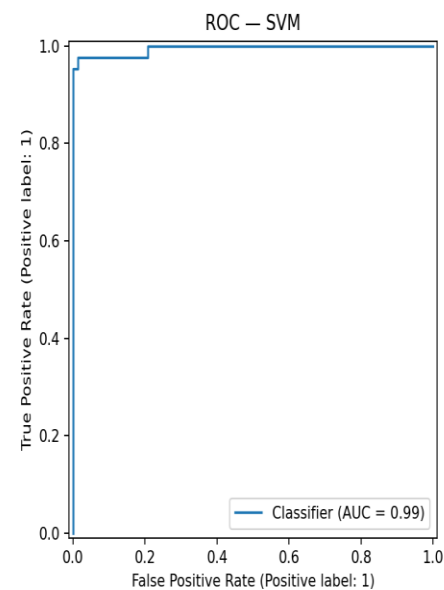
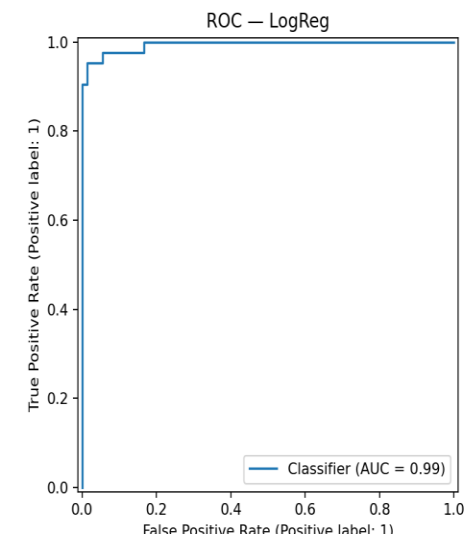
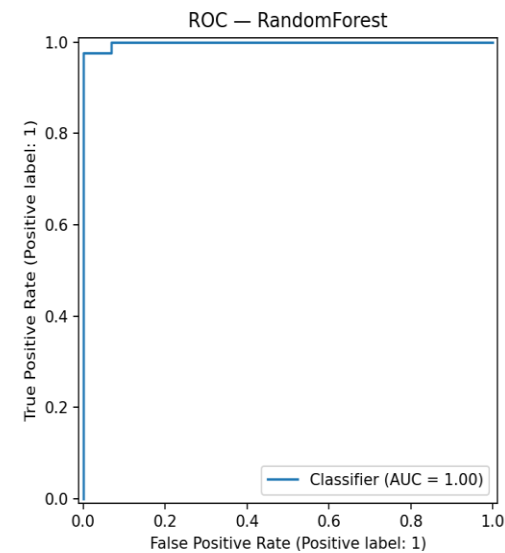
Random Forest



Support Vector  
Machine

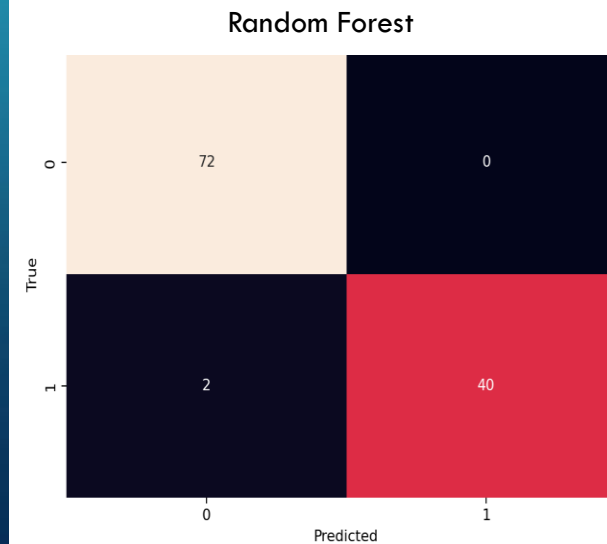
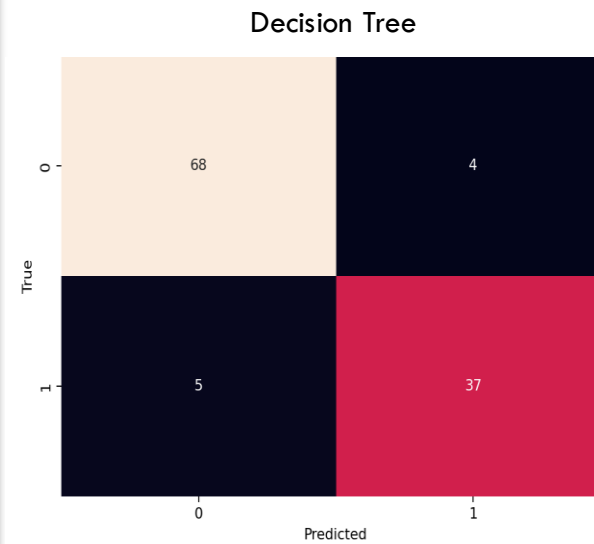
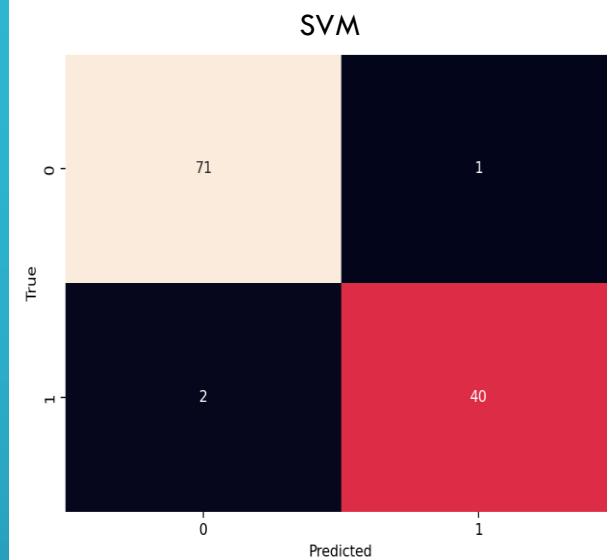
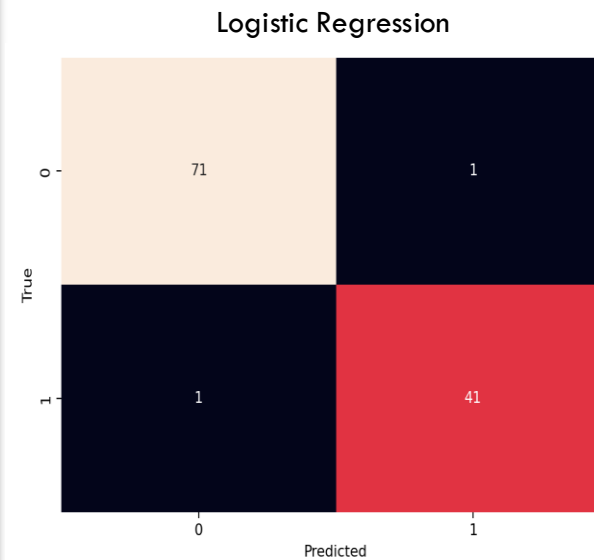
# ROC CURVE COMPARISON

Model AUC Performance



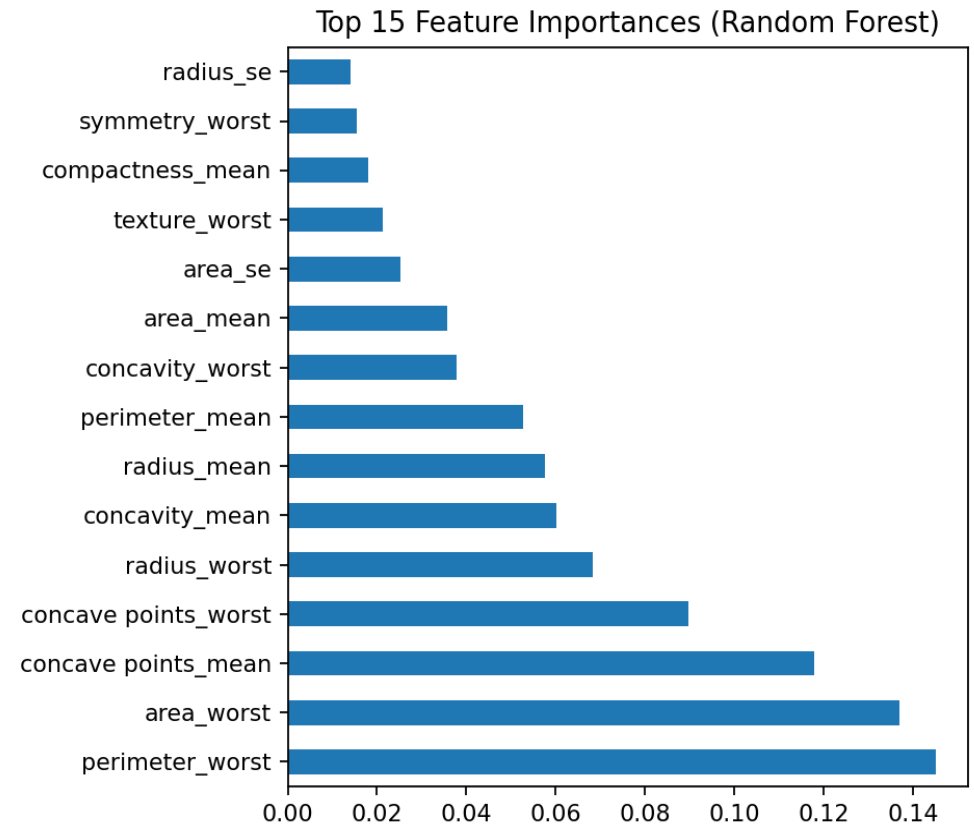
# CONFUSION MATRICES

Classification Accuracy & Errors



# FEATURE IMPORTANCE (RANDOM FOREST)

TOP PREDICTORS OF  
MALIGNANCY





# CLINICAL INTERPRETATION

How ML Supports Clinical Decision-Making

- ML identifies high-risk tumors
- Reduces diagnostic errors
- Improves triage for follow-up imaging
- Provides interpretable insights



# IMPLEMENTATION & EXECUTION GUIDE

How to Access the Notebook, Data, and Results

- Google Colab notebook
- Dataset included
- Results folder with all outputs
- README explains steps



## SUMMARY & CONCLUSIONS

---

### Key Findings

---

SVM & Random Forest performed best

---

High AUC & low false negatives

---

ML enhances CDSS for oncology

---

Important predictive features identified

# REFERENCES

- Dataset: <https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data>
- Abdel-Zaher, A. M., & Eldeib, A. M. (2016). Breast cancer classification using deep belief networks. Expert Systems with Applications.
- Chaurasia, V., & Pal, S. (2017). Data mining techniques to predict and resolve breast cancer survivability. IJCSMC.
- Jović, A., Brkić, K., & Bogunović, N. (2015). Feature selection methods with applications to bioinformatics. MIPRO.
- Rajendran, K., Jayabalan, M., & Thiruchelvam, V. (2020). Predicting breast cancer via supervised ML. IJACSA.