

A pink awareness ribbon is centered in the background, forming a loop at the top and crossing itself in the middle.

# **BREAST CANCER**

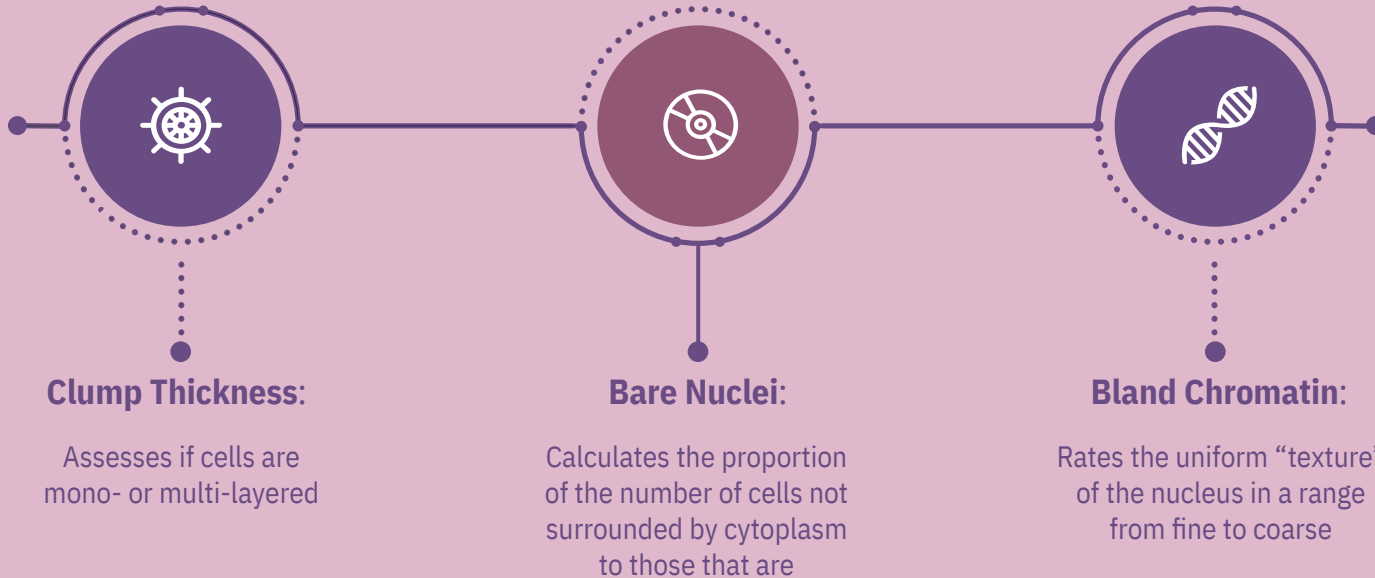
**PREDICTING BENIGN/MALIGNANT TUMORS**

# \$4,000,000,000

**ESTIMATED COST OF FALSE-POSITIVE  
MAMMOGRAMS AND OVERDIAGNOSIS OF  
BREAST CANCER\***

# TUMOR CHARACTERISTIC EXAMPLES

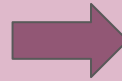
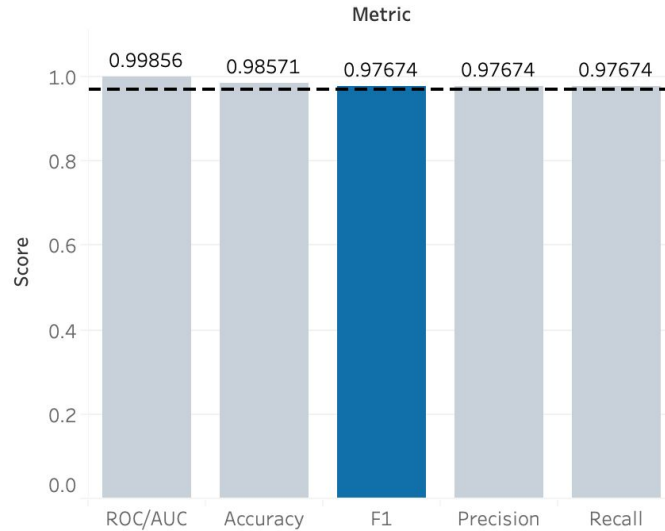
---



# LOGISTIC REGRESSION RESULTS

## F1 Score

All metrics scored above **0.97**



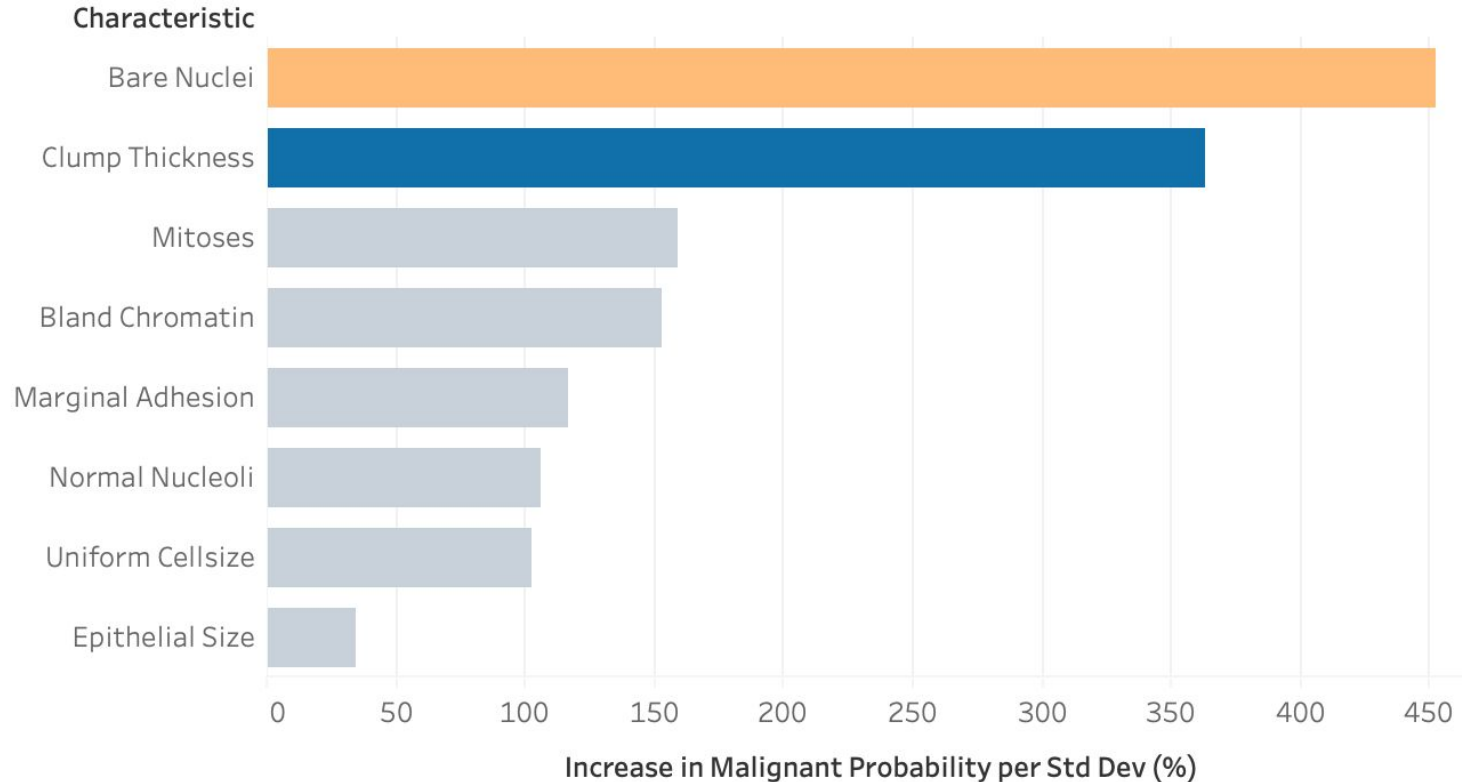
## Confusion Matrix

True	Benign	96	1
	Malignant	1	42
		Benign	Malignant
		Predicted	

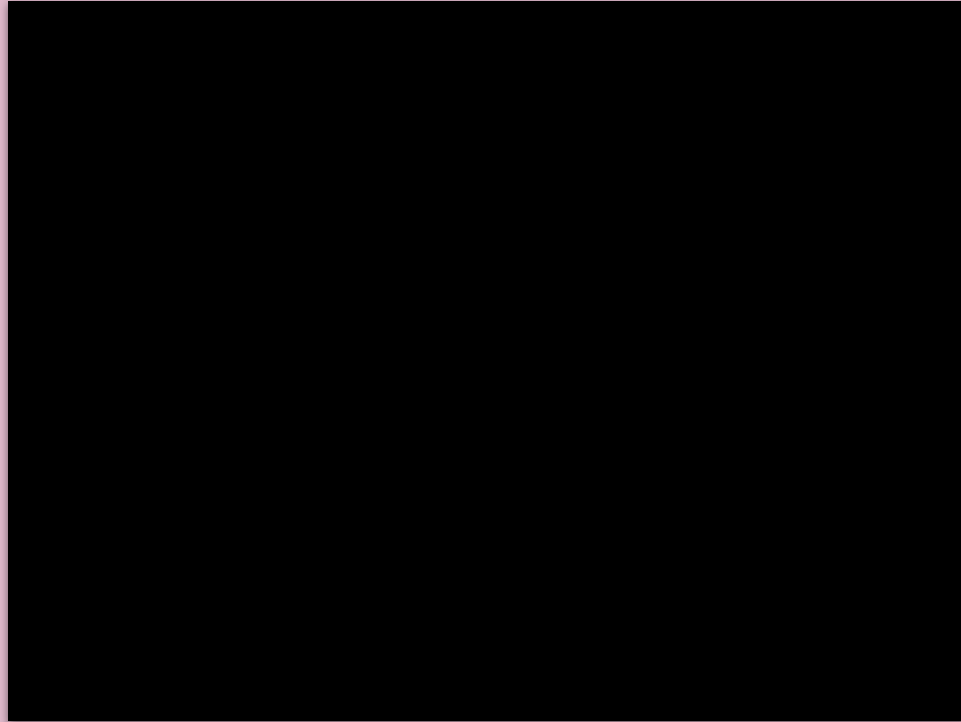
The Confusion Matrix shows the relationship between True and Predicted classes. The top row represents True Benign cases, with 96 correctly classified (True Positives) and 1 incorrectly classified as Malignant (False Positive, FP). The bottom row represents True Malignant cases, with 1 incorrectly classified as Benign (False Negative, FN) and 42 correctly classified (True Negatives). The FP and FN cells are highlighted with red borders.

## Model Coefficients

Cell **degradation** and tumor **growth** have the greatest effect on malignant probability



## Tumor Characteristic Effects Demo



# KEY TAKEAWAYS

---

01

Machine learning algorithms can significantly improve diagnoses

02

Saved costs can be distributed to more meaningful applications of healthcare



# FUTURE CONSIDERATIONS

---



## **Image Recognition**

Test machine learning and image recognition on mammograms

## **Cost of Tumor Sample**

Understand the overall costs of getting tumor sampling rather than mammogram

## **Application to Recent Research**

Apply classification model to more recent research and data



# THANKS

**Does anyone have any questions?**

[davideweon@gmail.com](mailto:davideweon@gmail.com)

# APPENDIX

---

Feature	Description
Clump thickness	Assesses if cells are mono- or multi-layered
Uniformity of cell size	Evaluates the consistency in size of cells in sample
Marginal adhesion	Quantifies how much cells on the outside of the epithelial tend to stick together
Single epithelial cell size	Relates to cell uniformity, determines if epithelial cells are significantly enlarged
Bare nuclei	Calculates the proportion of the number of cells not surrounded by cytoplasm to those that are
Bland chromatin	Rates the uniform “texture” of the nucleus in a range from fine to coarse
Normal nucleoli	Determines whether the nucleoli are small and barely visible or larger, more visible, and more plentiful
Mitoses	Describes the level of mitotic (cell reproduction) activity

# REFERENCES

---

DR. WILLIAM H. WOLBERG . *Breast Cancer Wisconsin (Original) Data Set*.  
UCI Machine Learning Repository

[https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+\(original\)](https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(original))

# CREDITS

---

- ◀ Presentation template by Slidesgo
- ◀ Icons by Flaticon
- ◀ Infographics by Freepik
- ◀ Images created by Freepik
- ◀ Text & Image slide photo by Freepik.com