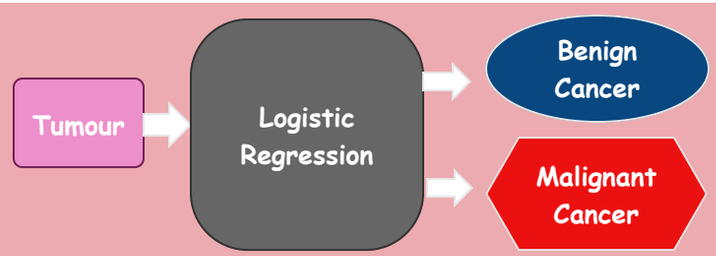
Proprietary & All Rights Reserved

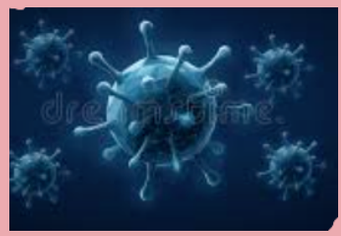
© 2025 Sweety Seelam. All rights reserved.

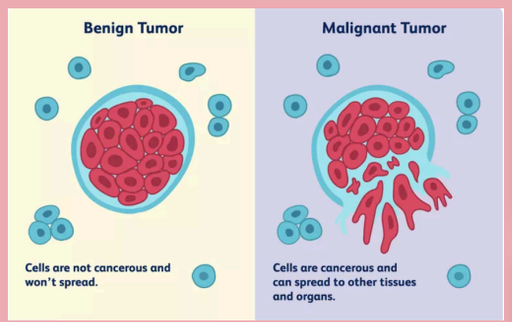
This project, including its source code, trained models, datasets (where applicable), visuals, and dashboard assets, is protected under copyright and made available for educational and demonstrative purposes only.

Unauthorized commercial use, redistribution, or duplication of any part of this project is strictly prohibited.

Breast Cancer Prediction Risk Overview







114
Total Patients

98.2%
Accuracy %

41
Malignant Predictions

38.9%
Average Risk Probability (%)

Breast Cancer Risk Prediction – Overview

This dashboard provides a clear, interpretable, and trustworthy view of breast cancer risk predictions made using a machine learning model trained on diagnostic clinical features.

Key Performance Indicators (KPIs) such as accuracy, average predicted risk, and malignant case counts reflect how well the model is performing.

SHAP values explain the impact of each clinical feature on the model's

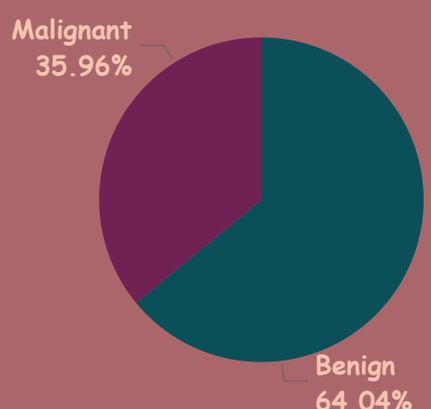
Reset Filters

Predicted Diagnosis Typ...
- Select all
- Benign
- Malignant

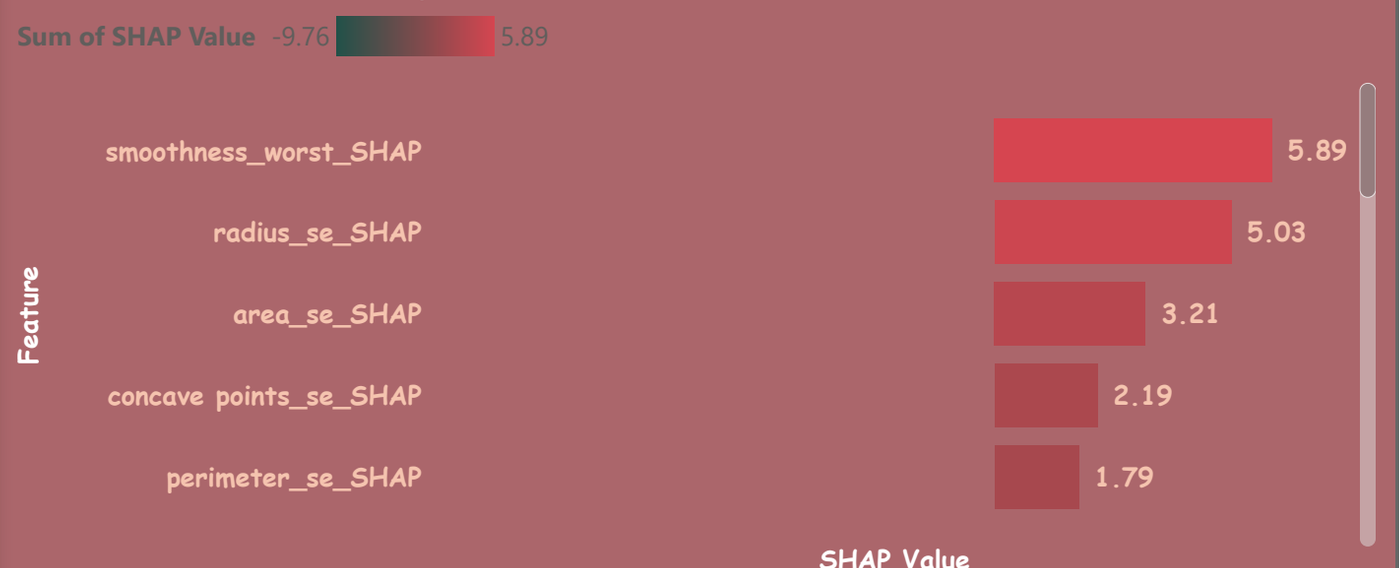
Risk Probability (%)
0.0% 100.0%

Actual Diagnosis Types
- Select all
- Benign
- Malignant

Predicted Diagnosis: Benign vs. Malignant



Top SHAP Drivers of Prediction (Global)



Patient Level Risk Insights

Select a Patient

- Select all
- 1
- 2
- 3
- 4

Reset Filters

Predicted Diagnosis Typ...

- Select all
- Benign
- Malignant

Risk Probability (%)

0.0% 100.0%

Actual Diagnosis Types

- Select all
- Benign
- Malignant

Breast Cancer Risk Prediction – Patient Level Risk Insights

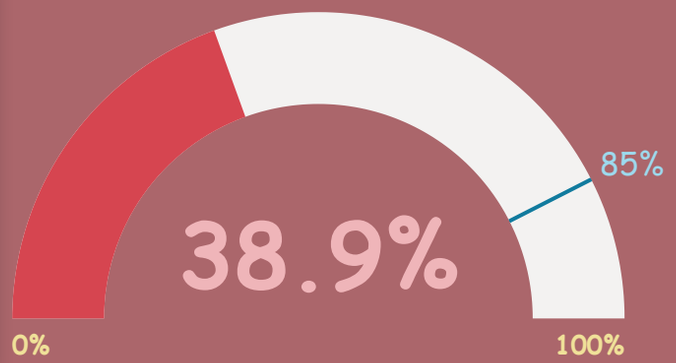
📌 **Interpretation**

This section shows the **AI-predicted breast cancer risk** for the selected patient based on their diagnostic features.

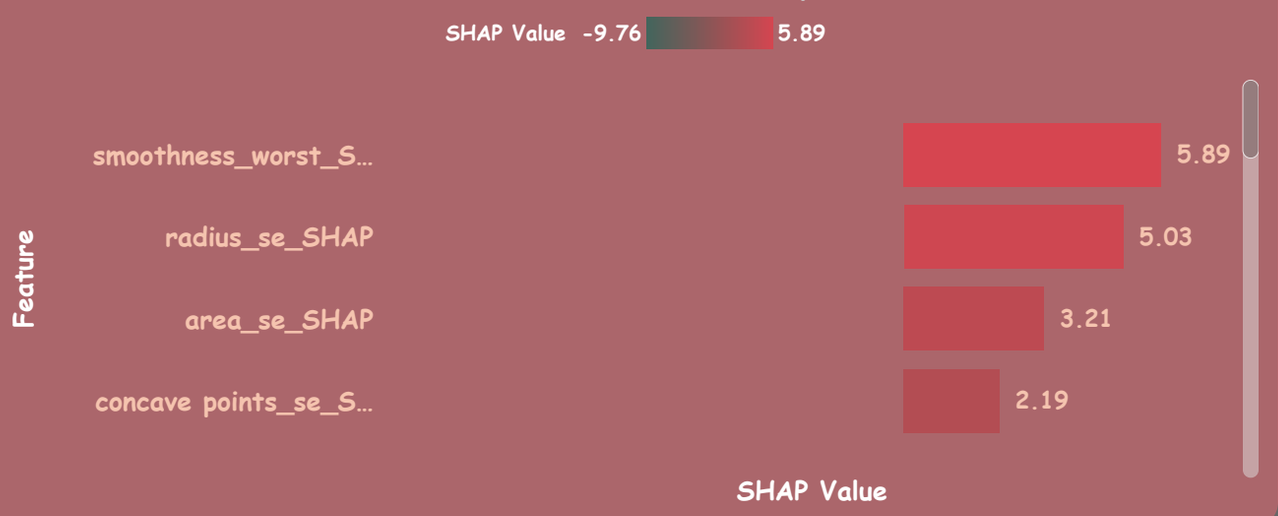
- The **Risk Probability (%)** displayed in the **gauge chart** represents the model's **confidence** that this patient may have **malignant** cancer.
- The **bar chart** (SHAP values) explains *why* the model made that prediction:

1. Features **pushing the risk higher** (positive SHAP values) are likely indicators of malignancy.

Predicted Risk Level (%)



SHAP Feature Contributions per Patient



Patient Summary Table

Patient ID	Actual Diagnosis Types	Predicted Diagnosis Types	Prediction_Probability
1	Benign	Benign	16.9%
2	Malignant	Malignant	99.8%
3	Malignant	Malignant	93.7%
4	Benign	Benign	2.6%
5	Benign	Benign	0.5%
6	Malignant	Malignant	100.0%
7	Malignant	Malignant	100.0%
8	Malignant	Malignant	88.5%
9	Benign	Benign	42.9%
10	Benign	Benign	1.3%

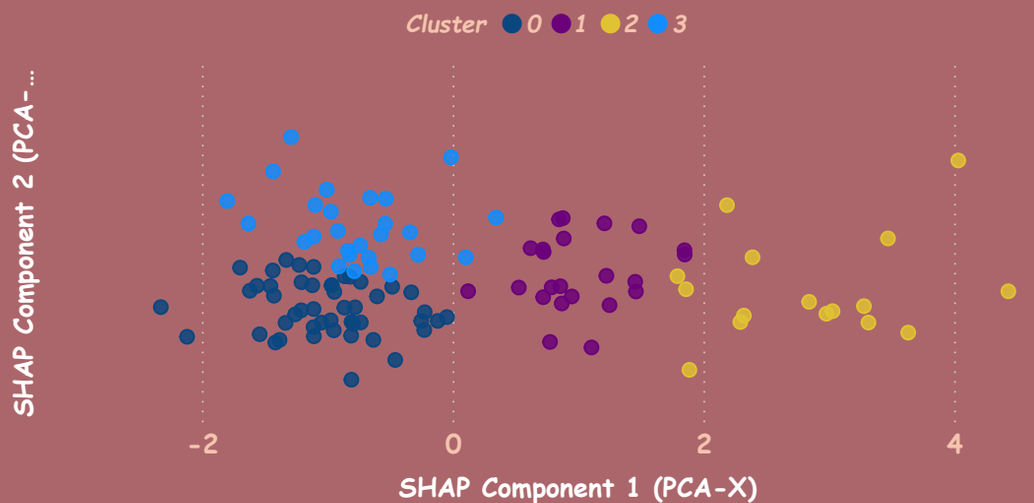
Clustered Patient Segments

SHAP Behavioral Profiles

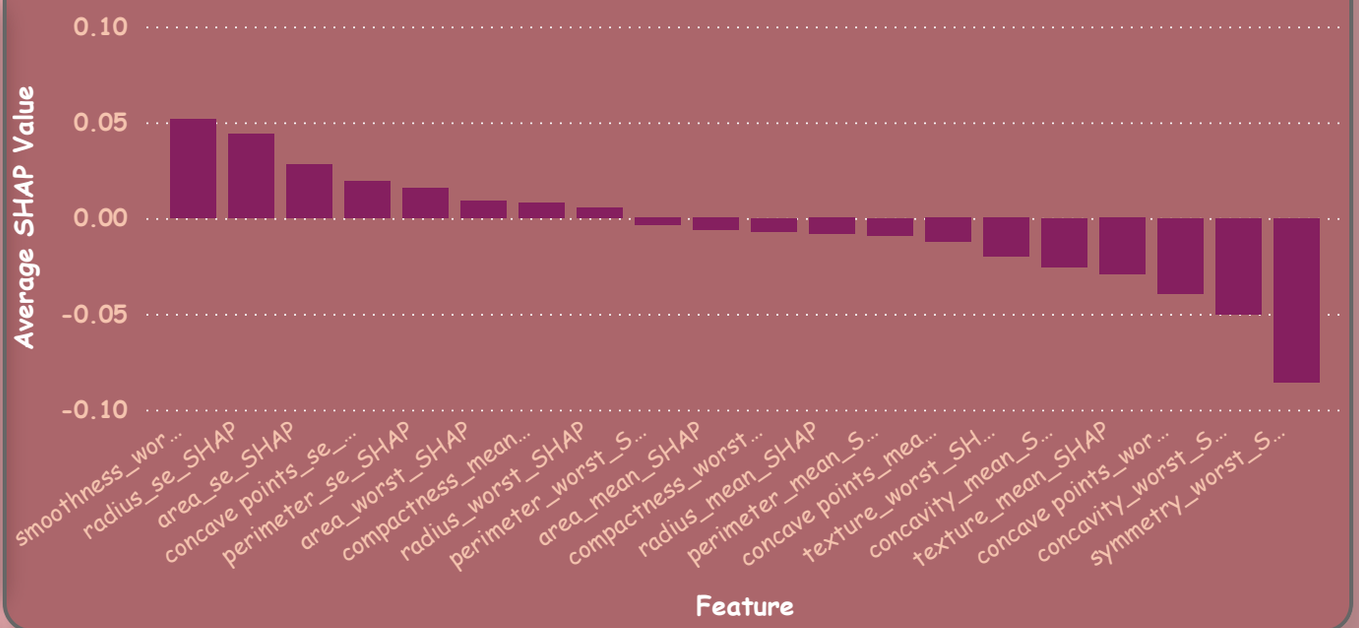
texture_worst_SHAP
Top SHAP Feature by Value

114
Total Patients In Segment

Patient Segmentation Based on SHAP Behavioral Profiles



Top SHAP Feature in Each Cluster



Breast Cancer Risk Prediction – Clustered Patient Segments

Interpretation

- Each cluster represents a group of patients with similar SHAP contribution patterns - meaning the model made predictions for similar reasons.
- SHAP shows how each feature contributed to the model's prediction – whether it pushed the risk up or down.
- The scatterplot shows spatial separation via PCA, and the bar chart highlights which features contributed most to predictions within the selected cluster.
- These insights help define behavioral segments and support tailored

Select Cluster

0 | 1 | 2 | 3

Actual Diagnosis Types

- Select all
- Benign
- Malignant

Reset Filters

Risk Probability (%)

0.0% — 100.0%

Predicted Diagnosis Typ...

- Select all
- Benign
- Malignant

Patient ID

- Select all
- 1
- 2
- 3
- 4
- 5

Breast Cancer Risk Prediction

✅ Conclusion

The breast cancer prediction model demonstrates **exceptional performance and explainability** through a combination of Logistic Regression and SHAP value interpretation. Based on clinical data from **114 patients**, the model:

- Achieves a **98.2% prediction accuracy**, correctly identifying **malignant and benign cases** with minimal error.
- Predicts **41 malignant cases**, suggesting it can flag high-risk patients early with high confidence.
- Outputs an **average risk probability of 38.9%**, indicating a realistic balance between sensitivity and specificity.
- Uses SHAP to surface the **top 5 most influential features** globally:

💼 Business Impact

Implementing this solution in a real-world clinical or diagnostic setting could deliver **tangible monetary and operational benefits**:

1.Reduce Misdiagnosis Costs

Hospitals may spend **$11,000–$17,000 per patient** on delayed or incorrect breast cancer diagnoses. This model minimizes that by improving early detection accuracy to **over 98%**, preventing such errors.

2.Early Detection = Lower Treatment Costs

Early-stage detection typically costs **40–60% less** in treatment compared to late-stage cancer. If even **10% more cancers are caught early**, a hospital could save **$500K+ annually** for

📌 Business Recommendations

If adopted by a health-tech firm, hospital network, or diagnostic lab, the following strategic actions are recommended:

1.Integrate into Diagnostic Workflows

Use the model as a **decision support tool** — triaging patients who need priority diagnostic testing or radiologist review.

2.Deploy Patient-Level Risk Dashboards

Implement the SHAP-powered dashboards (like your Page 2) for real-time visual insight into **why a patient is flagged** as high-risk. This promotes **shared decision-making** with patients

📖 Project Storytelling

🎗 "Explaining Breast Cancer Predictions: AI-Powered Risk Scoring with Global & Patient-Level SHAP Interpretability"

In the fight against breast cancer, early and accurate diagnosis saves lives. This project introduces a transparent machine learning solution that not only predicts whether a tumor is likely malignant, but also **explains why**.

Using a logistic regression model trained on structured diagnostic features (e.g., radius, smoothness, concavity), the system outputs predictions for **114 patients**, reaching **98.2% accuracy** and flagging **41 malignant cases**. Beyond raw predictions, the integration of **SHAP values** at both global and patient levels enables doctors to visualize the **specific features**