

# CANCER-PREDICTION



# TEAM-MEMBERS

boyina siddharda - e22cseu0168

boyina sri harsha - e22cseu0174

korada ganesh - e22cseu0170

kappala guna vardhan - e22cseu0175

# ABSTRACT

Cancer prediction is a critical area in healthcare, with the potential to save lives through early detection and risk assessment.

This project focuses on the development of an intuitive and user-friendly cancer prediction tool using Streamlit, a Python library for creating web applications

# INTRODUCTION

In this presentation, we will explore Enhancing Cancer Prediction using Python for probability and statistical analysis. We will discuss the potential of leveraging Python for improving cancer prediction models and the impact on healthcare outcomes.

# UNDERSTANDING CANCER-PREDICTION

Cancer prediction involves analyzing complex data to identify patterns and risk factors. Python provides powerful tools for advanced statistical analysis and probability modeling, enabling more accurate predictions and personalized treatment plans.

In this project we are using logistical regression model for our prediction purpose

# METHODS

## 1. Data Collection and Preparation:

Gather a dataset of patient records with relevant medical information and cancer diagnosis labels (malignant or benign).

Clean and pre-process the data to handle missing values, outliers, and data type inconsistencies.

## 2. Feature Selection:

Identify the most relevant features from the dataset that are most predictive of cancer diagnosis.

Eliminate redundant or irrelevant features to reduce computational complexity and improve prediction accuracy.

### 3. Data Splitting:

Divide the dataset into training and testing sets.

The training set is used to train the logistic regression model, and the testing set is used to evaluate its performance.

### 4. Model Training:

Train the logistic regression model using the training set.

This involves estimating the coefficients for the logistic regression equation that best fit the training data.

### 5. Model Evaluation:

Evaluate the performance of the trained logistic regression model using the testing set.  
Common metrics used for evaluation include accuracy, precision, recall, and F1 score.

# Libraries that are imported

```
import streamlit as st import joblib  
import pandas as pd  
import plotly.graph_objects as go  
import numpy as np  
import pandas as pd  
from sklearn.preprocessing import StandardScaler  
from sklearn.model_selection import train_test_split  
from sklearn.linear_model import LogisticRegression  
from sklearn.metrics import  
accuracy_score,classification_report import joblib
```

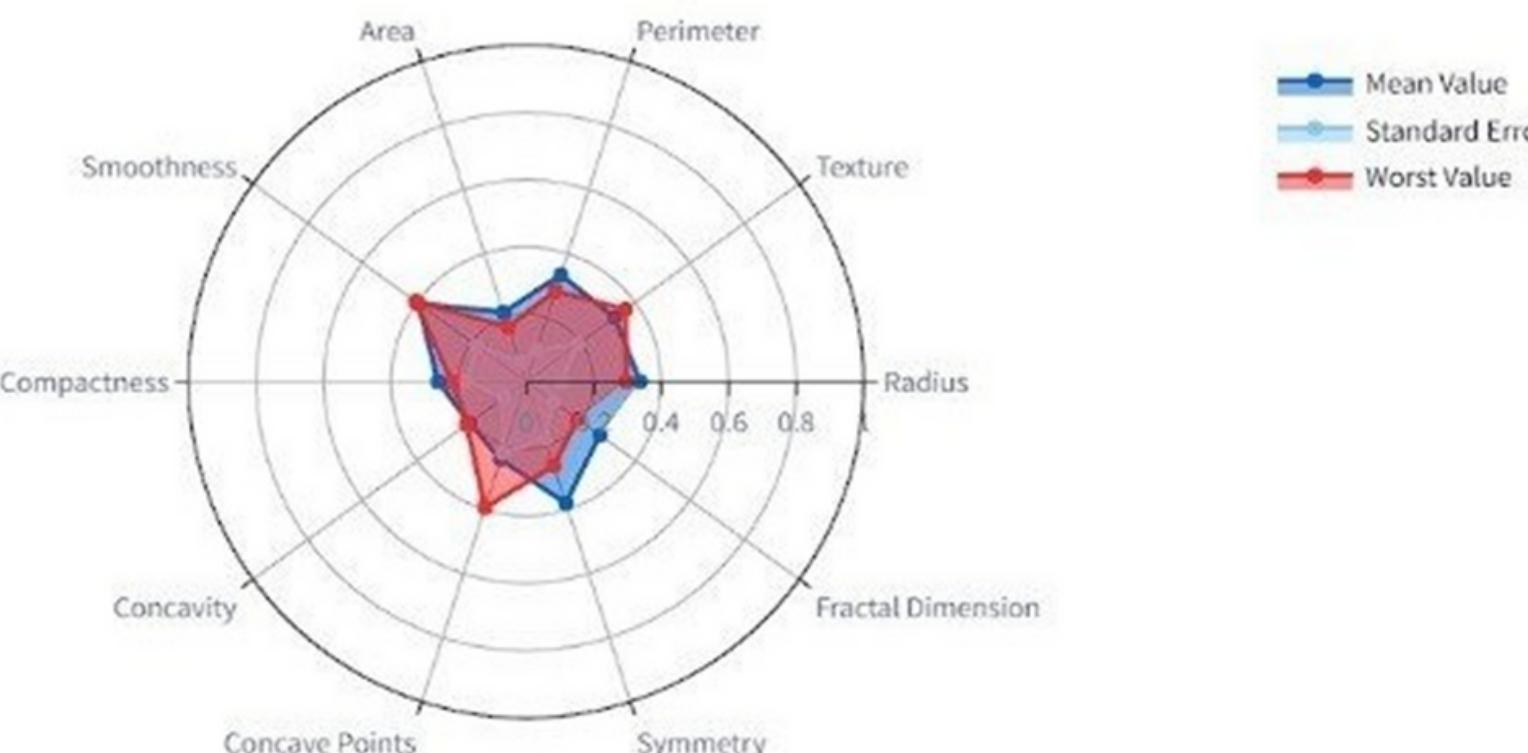
# RESULT

## Cell Nuclei Measurements



## Cancer Predictor

Please connect this app to your cytology lab to help diagnose breast cancer from your tissue sample. This app predicts using a machine learning model whether a breast mass is benign or malignant based on the measurements it receives from your cytology lab. You can also update the measurements by using the sliders in the sidebar.



## Cell cluster prediction

The cell cluster is:

Benign

Probability of being benign:

0.5809017618796077

Probability of being malignant:

0.4190982381203923

This app can assist medical professionals in making a diagnosis, but should not be used as a substitute for a professional diagnosis.

# CONCLUSION

In conclusion, leveraging Python for a probability and statistical analysis holds great promise in enhancing cancer prediction. By harnessing the power of Python, we can improve the accuracy of prediction models and ultimately contribute to more effective cancer diagnosis and treatment.

Thank you