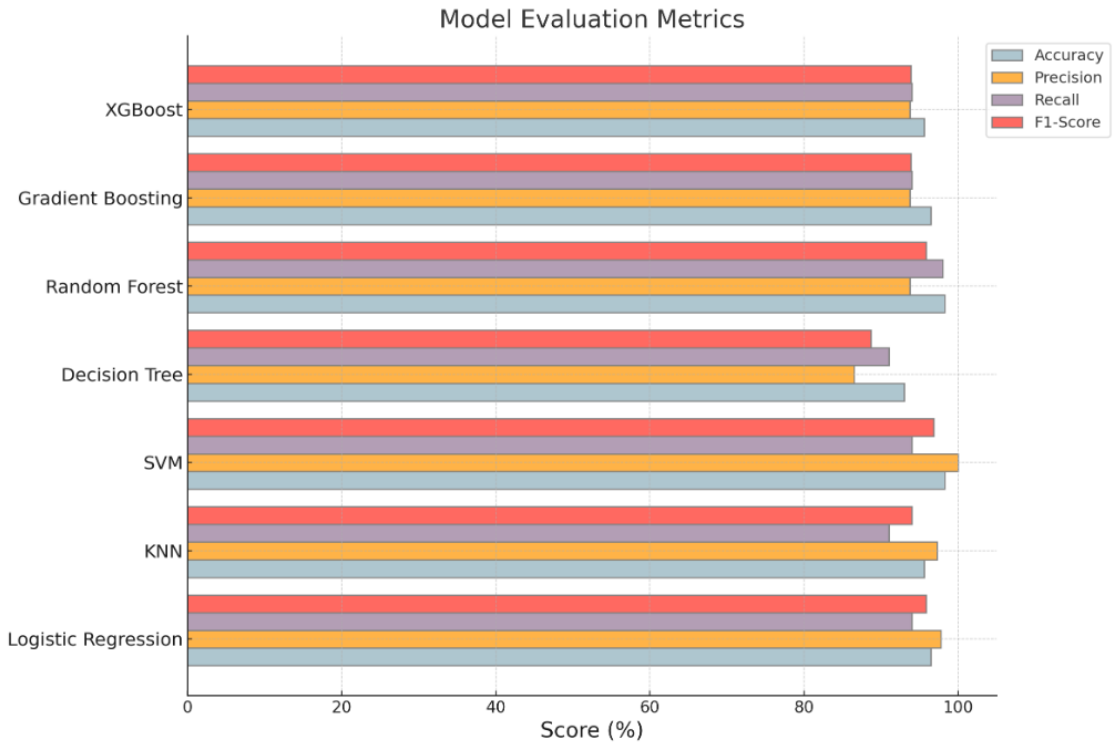
Próximos passos e melhoras no projeto “Breast Cancer Diagnostic Model”:

1) Atualizar o gráfico do README com valores corretos:

Model Evaluation

Each model was evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. Below are the results for each model:



2) Atualizar os valores da métricas do README e do arquivo notebook ipynb com valores corretos:

- **Logistic Regression**:
  - Accuracy: 96.49%
  - Precision: 97.78%
  - Recall: 94.00%
  - F1-Score: 95.87%
  - ROC-AUC: 0.96
- **K-Nearest Neighbors (KNN)**:
  - Accuracy: 95.61%
  - Precision: 97.27%
  - Recall: 91.00%
  - F1-Score: 94.00%
  - ROC-AUC: 0.95

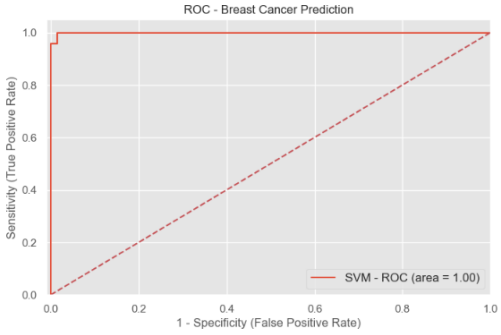
3) Refazer a ROC curve do modelo SVM porque não bate com o valor ROC do plot anterior:

```
Accuracy: 0.98
Precision: 0.98
ROC-AUC Score: 1.00
Classification Report:
              precision    recall  f1-score   support

           0       0.99      0.99      0.99        67
           1       0.98      0.98      0.98        47

    accuracy                           0.98       114
   macro avg       0.98      0.98      0.98       114
weighted avg       0.98      0.98      0.98       114

The profit using the model is USD 396000.00.
```



4) Fazer um Cross Validation descente

Cross Validation

```
[68…
# Validação Cruzada (Cross Validation)
print("\n--- Cross Validation ---")

from sklearn.model_selection import cross_val_score

# Função para realizar a validação cruzada e exibir os resultados
def perform_cross_validation(model, X, y, cv=10):
    cv_scores = cross_val_score(model, X, y, cv=cv)
    print(f"Cross Validation Scores: {cv_scores}")
    print(f"Mean CV Score: {cv_scores.mean()}")

# Regressão Logística
print("\nLogistic Regression:")
perform_cross_validation(log_reg, X, y)

# K-Nearest Neighbors (KNN)
print("\nK-Nearest Neighbors (KNN):")
perform_cross_validation(knn, X, y)

# Suporte Vector Machine (SVM)
print("\nSupport Vector Machine (SVM):")
perform_cross_validation(svc, X, y)

# Decision Tree Classifier
print("\nDecision Tree Classifier:")
```

5) No notebook ipynb, corrigir a descrição das métricas utilizadas nos modelos:

4.9.2 Evaluating the Gradient Boosting Classifier Model

The performance of the Gradient Boosting Classifier model is evaluated using various metrics:

- **Accuracy Score**: Measures the proportion of correctly predicted instances.
- **Confusion Matrix**: Summarizes the true positives, true negatives, false positives, and false negatives.
- **Classification Report**: Provides detailed metrics such as precision, recall, and F1-score for each class.

6) Revisar o notebook inteiro e depois refazer o README do projeto inteiro.