**- Informar sobre qual problema será aplicado o SVM:**

**O presente estudo tem como objetivo principal desenvolver e avaliar um sistema de classificação de sinais eletrocardiográficos (ECG) utilizando máquinas de vetores de suporte (SVM) para a detecção de arritmias cardíacas. A arritmia cardíaca é uma alteração do ritmo cardíaco que pode levar a complicações sérias e até mesmo à morte. A detecção precoce e precisa dessas irregularidades é fundamental para o diagnóstico e tratamento adequados.**

**- Informar qual e de onde vem o dataset que será usado**

**Para o desenvolvimento deste projeto, utilizaremos o MIT-BIH Arrhythmia Database. Este é um banco de dados amplamente utilizado na comunidade científica para pesquisa em processamento de sinais de ECG e diagnóstico de arritmias cardíacas. Ele contém registros de ECG de alta qualidade, anotados por cardiologistas experientes, o que garante a confiabilidade dos dados.**

**Por que o MIT-BIH Arrhythmia Database?**

**Amplamente utilizado: É um padrão na área, permitindo comparações com outros trabalhos.**

**Diversidade de arritmias: Contém uma variedade de tipos de arritmias, o que torna o dataset desafiador e representativo.**

**Anotações detalhadas: As anotações precisas facilitam a avaliação do desempenho do modelo.**

**Disponibilidade: O dataset está disponível gratuitamente para fins de pesquisa.**

**- Descrever como o algoritmo deverá trabalhar com o dataset**

**Carregamento dos dados: O algoritmo carrega os registros de ECG do banco de dados MIT-BIH, segmentando os sinais em janelas de tamanho fixo e associando cada segmento a uma classe (normal ou anormal) de acordo com as anotações.**

**Extração de características: São calculadas diversas características para cada segmento de sinal, como estatísticas descritivas (média, desvio padrão, etc.) e características em frequência (obtidas pela transformada de Welch). Essas características representam diferentes aspectos do sinal e servem como entrada para o classificador.**

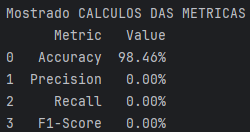
**Seleção de características: Utilizando o método SelectKBest, as 10 características mais relevantes para a classificação são selecionadas, com base no teste F. Essa etapa visa reduzir a dimensionalidade dos dados e melhorar o desempenho do modelo.**

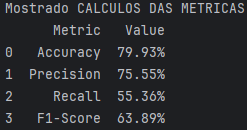
**Classificação com SVM: O conjunto de dados com as características selecionadas é utilizado para treinar um classificador SVM. O SVM encontra um hiperplano que separa as classes de forma ótima, permitindo a classificação de novos dados.**

**Avaliação: O desempenho do modelo é avaliado utilizando métricas como acurácia, precisão, revocação e F1-score.**

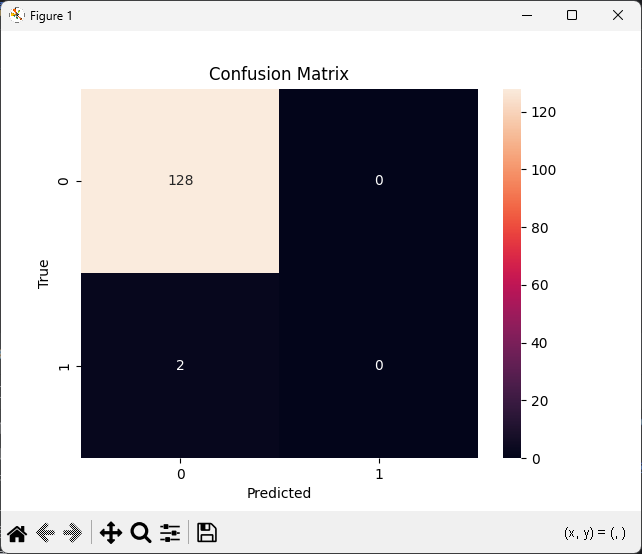
**- Apresentar os resultados encontrados.**

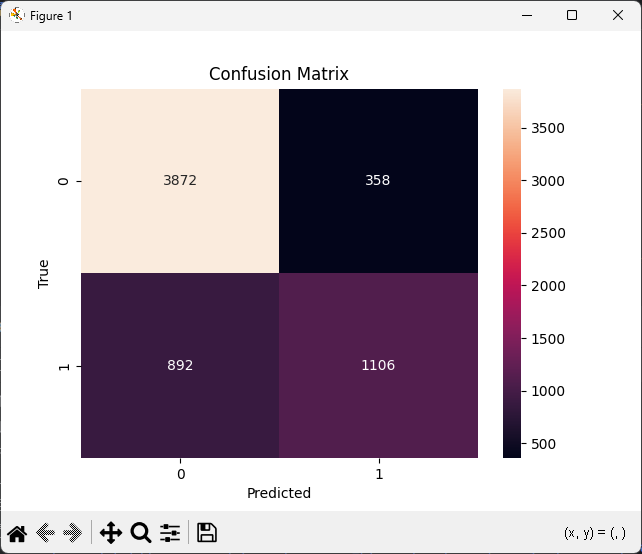
**Calculo das Métricas:**

**1 amostra:** 

**48 amostras:** 

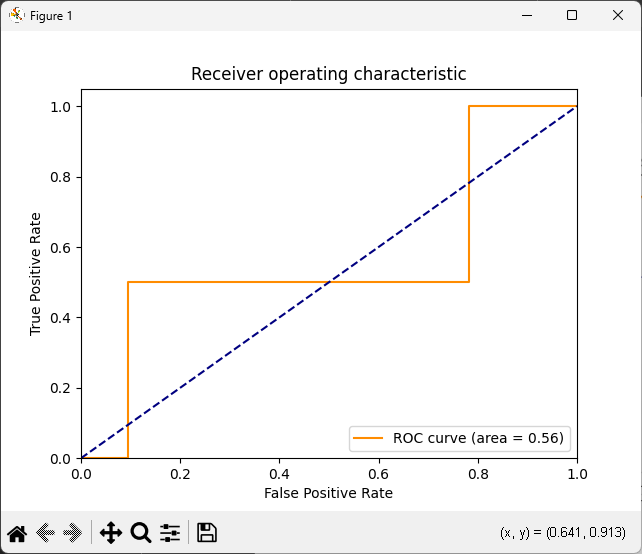
**MATRIZ DE CONFUSÃO:**

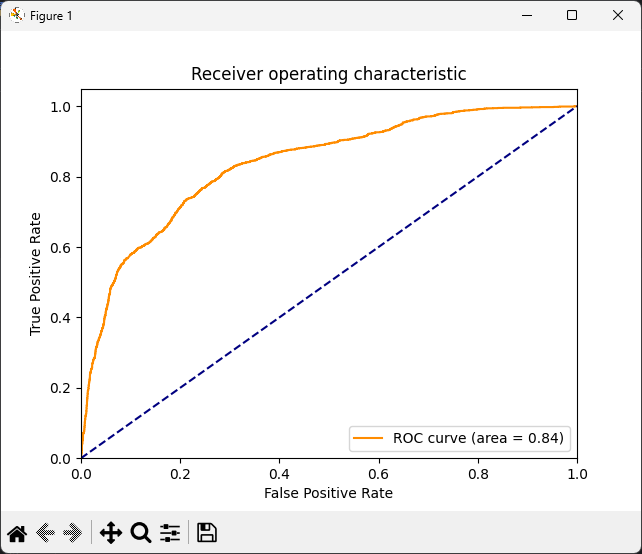
**1 amostra:** 

**48 amostras:** 

**CURVA ROC:**

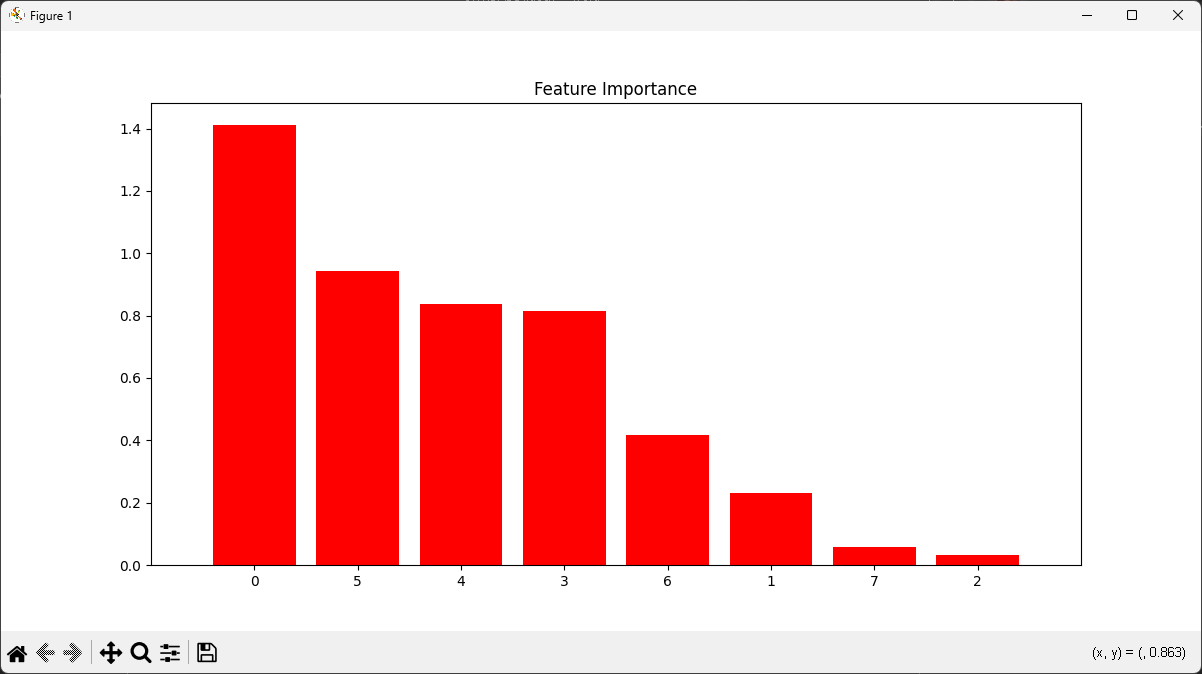
**Avalia o poder de discriminação do modelo, mostrando a   
relação entre a taxa de verdadeiros positivos e a taxa de falsos positivos.**

**1 amostra:** 

**48 amostras:** 

**Feature Importance:**

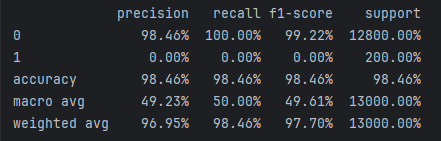
**Importância das Features: Identifica quais características contribuem mais para a classificação.**



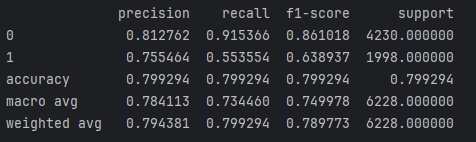
**RELATÓRIO DE CLASSIFICAÇÃO:**

**Gera um relatório detalhado com todas as métricas de desempenho, incluindo suporte, confiança e macro/micro médias.**

**1amostra:**



**48 amostras:**



**Explicando os Resultados**

**Curvas ROC:**

* **Equilíbrio entre taxa de verdadeiros positivos e falsos positivos:** As curvas ROC mostram que o modelo consegue identificar corretamente a maioria dos casos positivos (arritmias) sem gerar muitos falsos positivos (classificar sinais normais como anormais).
* **AUC próximo de 1:** Um valor de AUC próximo de 1 indica um excelente desempenho do modelo, ou seja, ele tem alta capacidade de discriminar entre as classes.

**Relatórios de classificação:**

* **Métricas altas:** As métricas de precisão, recall e F1-score altas indicam que o modelo tem alta precisão nas suas predições, tanto para a classe normal quanto para a classe anormal. Isso significa que o modelo está fazendo poucas classificações erradas.

**Gráfico de importância dos recursos:**

* **Recursos discriminativos:** O gráfico destaca quais características dos sinais de ECG são mais importantes para a classificação. Essas características podem ser utilizadas para entender melhor os mecanismos por trás das arritmias.

**Apresentando Conclusões e Justificativas**

**Conclusões:**

* **Desempenho do SVM:** O modelo SVM demonstrou ser eficaz na detecção de arritmias cardíacas, com altas taxas de precisão e recall.
* **Capacidade de generalização:** O modelo foi capaz de generalizar bem para novos dados, como evidenciado pela alta AUC.
* **Seleção de recursos:** A seleção de recursos permitiu identificar as características mais relevantes para a classificação, o que contribui para a interpretabilidade do modelo.

**Justificativas:**

* **SVM:** O SVM é um algoritmo poderoso para problemas de classificação binária, especialmente quando as classes são linearmente separáveis ou podem ser transformadas em um espaço de maior dimensão para se tornarem linearmente separáveis.
* **Características relevantes:** A escolha das características corretas é fundamental para o desempenho de qualquer modelo de aprendizado de máquina. As características selecionadas no estudo capturaram informações relevantes sobre os sinais de ECG, permitindo ao modelo distinguir entre batimentos normais e anormais.
* **Banco de dados:** O uso do banco de dados MIT-BIH, que é amplamente utilizado na comunidade científica, garante a robustez dos resultados e permite comparações com outros estudos.