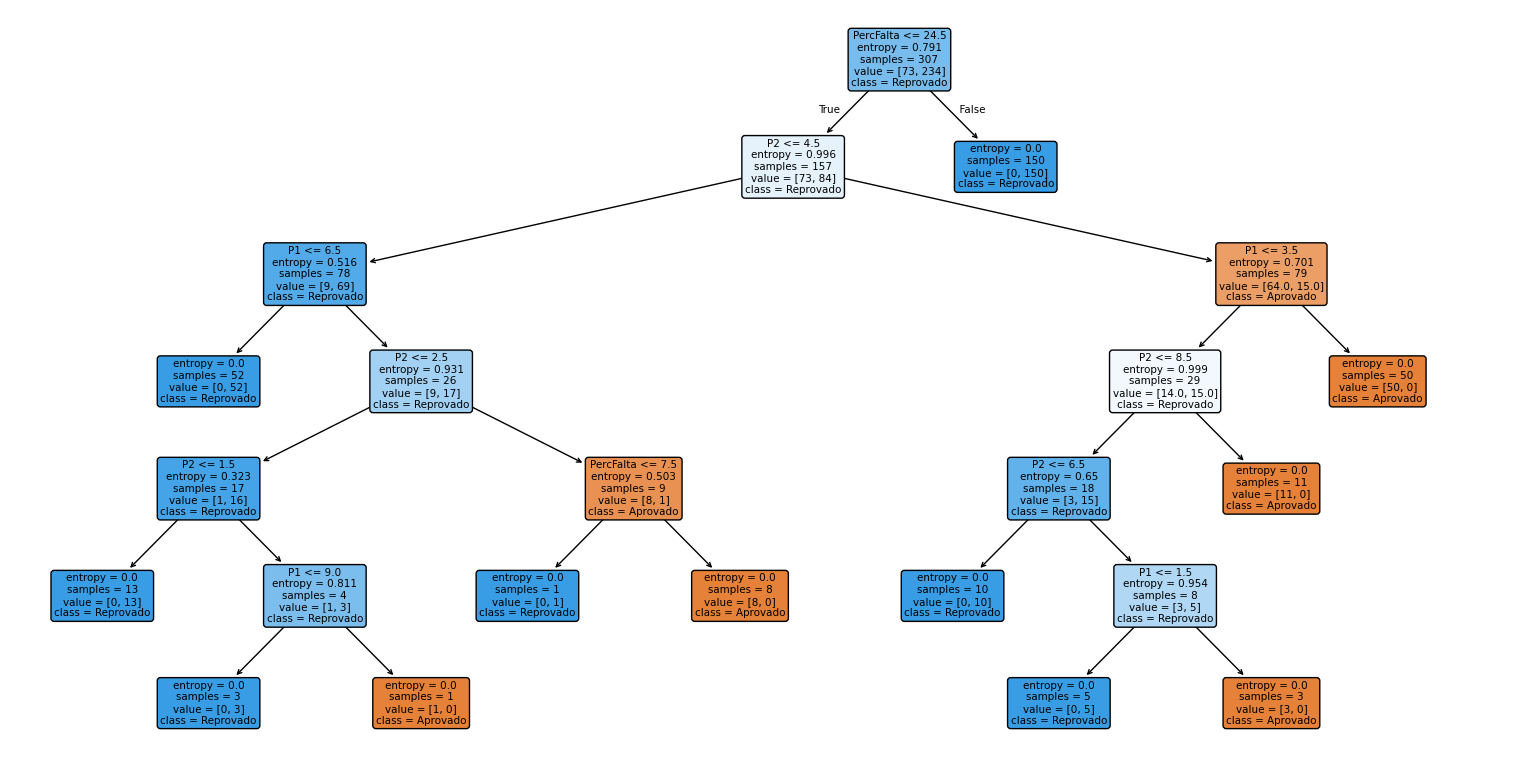
# Atividade #03 - Aprendizado Indutivo - Árvore de Decisão

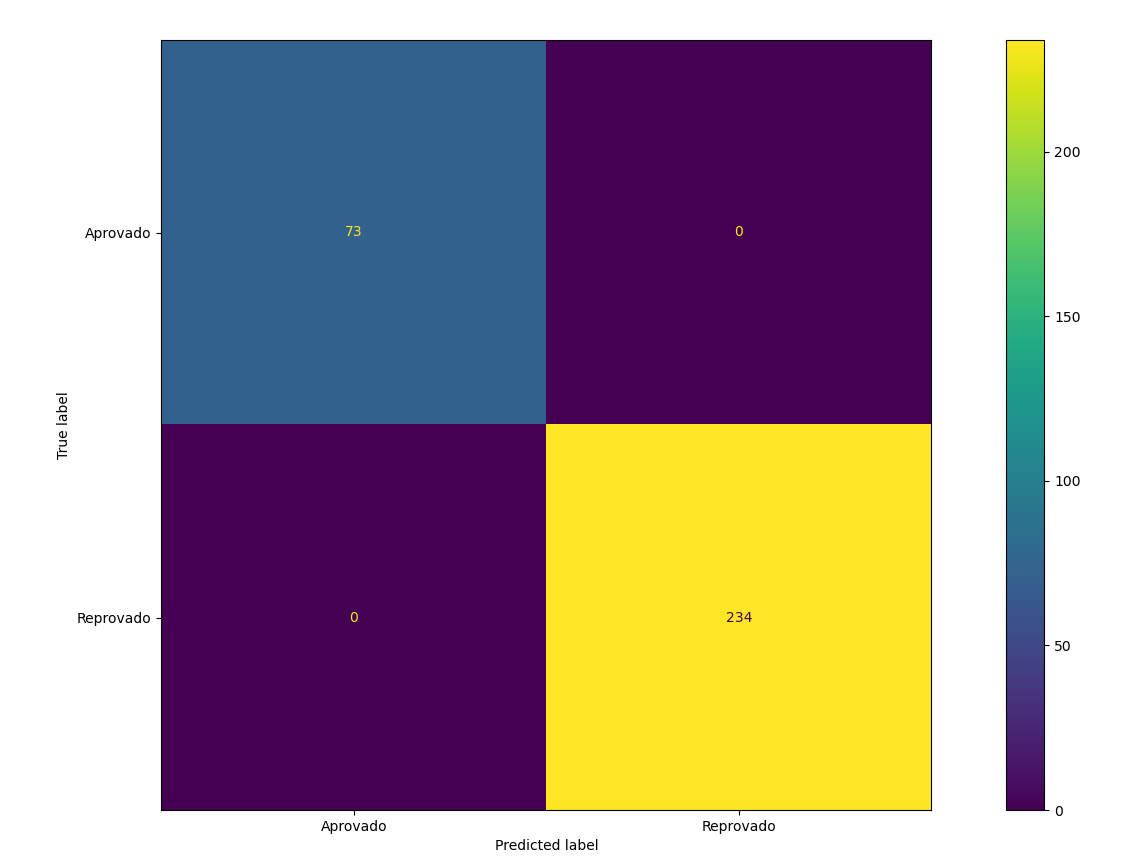
A base de dados utilizada para a realização do laboratório foi a base *CriterioProvas.arff* que comtempla os itens P1 (prova 1), P2 (prova 2), PercFalta (percentual de falta) e o resultado (Aprovado ou Reprovado). Para o aluno passar de semestre ele deve obter um percentual de faltas superior a 25% e a média de ambas as provas resultando em 5, seguindo esses critérios o aluno é aprovado.

Na base de dos foi encontrado algumas inconsistências se formos comparar algumas linhas, porém foi ignorado esse fato e assumido que ela estava totalmente certa. As linhas 24, 26 e 40 revelam que a média de p1 e p2 deve ser 5 e a linha 38 nega isso.

*Imagem 1- Árvores encontradas (CriterioProvas.arff)*

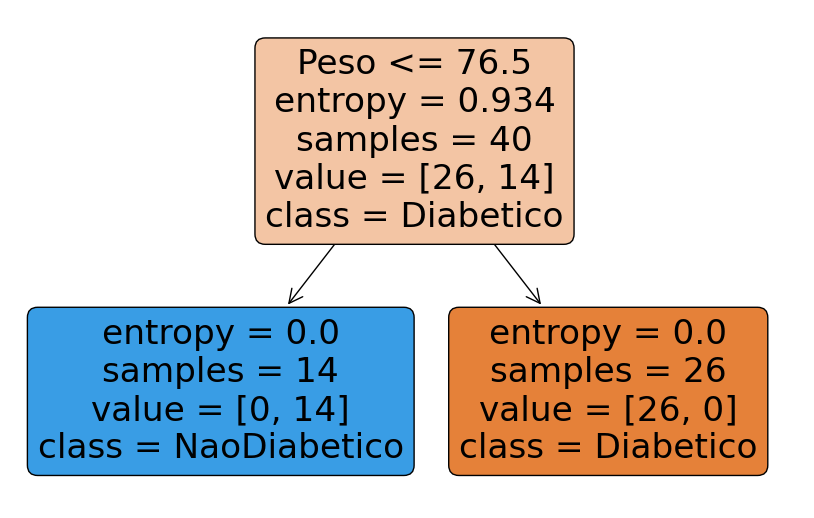


*Imagem 2- Matriz de confusão (CriterioProvas.arff)*

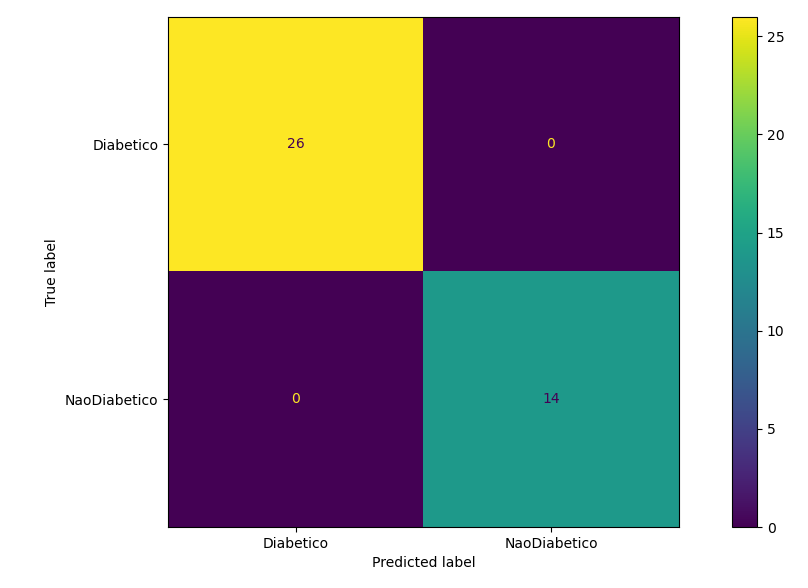


O grupo decidiu criar uma base de dados sobre diabetes, que inclui informações sobre glicose, peso e seu respectivo diagnóstico (Diabético ou Não Diabético). O critério adotado para o diagnóstico é o seguinte: um paciente será classificado como *Diabético* se a glicose estiver acima de 130 mg/dL e o peso acima de 60 kg. Caso contrário, o paciente será classificado como *Não Diabético*. A base de dados estará disponível em um arquivo no formato .arff, localizado junto ao presente documento na entrega.

*Imagem 1- Árvores encontradas (CriterioDiabetes.arff)*



*Imagem 2- Matriz de confusão (CriterioDiabetes.arff)*

**