O objetivo deste projeto é desenvolver um modelo preditivo para identificar a ocorrência de insuficiência cardíaca em pacientes com base nos dados clínicos existentes. A insuficiência cardíaca é uma condição crítica que requer diagnóstico precoce e intervenção apropriada para melhorar os resultados dos pacientes. Portanto, a capacidade de prever a probabilidade de insuficiência cardíaca com base nas informações demográficas e de saúde do paciente é fundamental para melhorar a qualidade dos cuidados médicos.

Para realizar essa análise foi necessário a preparação dos dados, isso incluiu lidar com valores ausentes, discrepantes e possíveis inconsistências nos dados.

(Falar das mudanças que foi realizado no SQL ou Excel (exemplo valores que foi mudado dos exames de valor número para caractere – Alto))

Para este projeto, optamos por usar o Weka, uma poderosa ferramenta de mineração de dados e aprendizado de máquina de código aberto, desenvolvida na Universidade de Waikato, na Nova Zelandia. O ambiente do Weka contém uma coleção de visualizações e ferramentas para análise e previsão de dados, com interfaces gráficas de fácil acesso:

(Imagem Weka.)

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

O Weka conta com uma comunidade robusta de usuários e desenvolvedores, bem como uma ampla documentação disponível online, o que facilita o suporte e o aprendizado. Devido à sua popularidade em ambientes educacionais e de pesquisa, o Weka é uma escolha comum entre estudantes, pesquisadores e profissionais que desejam explorar e aplicar técnicas de mineração de dados e aprendizado de máquina em diversos campos.

Optamos por utilizar o J48 (C4.5) do Weka para construir nosso modelo de árvore de decisão. O J48 é uma implementação bem estabelecida do renomado algoritmo C4.5, reconhecido por sua eficácia na construção de árvores de decisão. Esse algoritmo é valorizado por sua habilidade em lidar com variáveis categóricas e numéricas nos conjuntos de dados, além de sua facilidade de interpretação, o que torna as decisões do modelo mais transparentes e compreensíveis. O funcionamento do J48 envolve a construção de uma árvore de decisão de forma iterativa. Em cada etapa, o algoritmo seleciona o melhor atributo para dividir os dados com base em critérios específicos, como o ganho de informação. Essa divisão é feita de maneira recursiva, criando-nos e ramos na árvore até que todas as amostras tenham a mesma classe ou não possam ser divididas mais. A árvore resultante é utilizada para classificar novas instâncias, seguindo os ramos de acordo com os valores de seus atributos.