

Laboratório - 1 - Arquivo: jogar_volei.arff

Os dados são relativos a informações meteorológicas levantadas em dias de jogos de volei realizados ou não. São 14 dados sobre umidade, temperatura, vento, dentre outros, que refletem a atitude dos jogadores em jogar ou não a partida de acordo com o tempo.

Pergunta: Caso você acordasse pela manhã e desejasse jogar volei, qual seria sua decisão baseada nos dados informados.

Resposta: Minere os dados através do algoritmo de classificação **J48** e estabeleça uma decisão de acordo com os atributos disponibilizados.

Passo a Passo

Siga o Passo a Passo e análise os resultados:

Desenvolvido com o objetivo de gerar regras de classificação. É a implementação do algoritmo J48 desenvolvido por Ross Quinlan. Essa mineração de dados criará a estrutura de uma árvore de decisão e será interpretada através de uma árvore de decisão, transformando-a em regras.

Vamos minerar a seguinte base de dados:

Tabela condicoes: DECISÕES SOBRE JOGAR OU NÃO VOLEI

PREVISAO	TEMPERATURA	UMIDADE	VENTO	JOGAR
Chuvoso	Frio	Normal	Falso	Sim
Chuvoso	Agradável	Alta	Falso	Sim
Encoberto	Quente	Alta	Falso	Sim
Sol	Quente	Alta	Verdade	Nao
Sol	Quente	Alta	Falso	Nao
Chuvoso	Frio	Normal	Verdade	Nao
Encoberto	Frio	Normal	Verdade	Sim
Sol	Agradável	Alta	Falso	Nao
Sol	Frio	Normal	Falso	Sim
Chuvoso	Agradável	Normal	Falso	Sim
Sol	Agradável	Normal	Verdade	Sim
Encoberto	Agradável	Alta	Verdade	Sim
Encoberto	Quente	Normal	Falso	Sim
Chuvoso	Agradável	Alta	Verdade	nao

PASSO 1 - Minerando com o software WEKA

Primeiro transformaremos a tabela condicoes em um arquivo texto no formato ARFF, pois é o tipo de arquivo que o WEKA consegue interpretar.

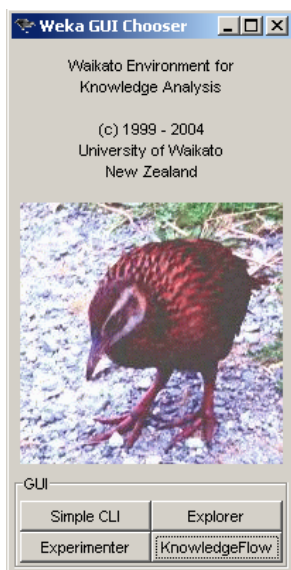
O resultado é o seguinte:

```
@attribute previsao {ensolarado, nublado, chuvoso}  
@attribute temperatura {quente, brando, frio}  
@attribute umidade{alta, normal}  
@attribute vento {ventando, nao_ventando}  
@attribute jogar {sim,nao}
```

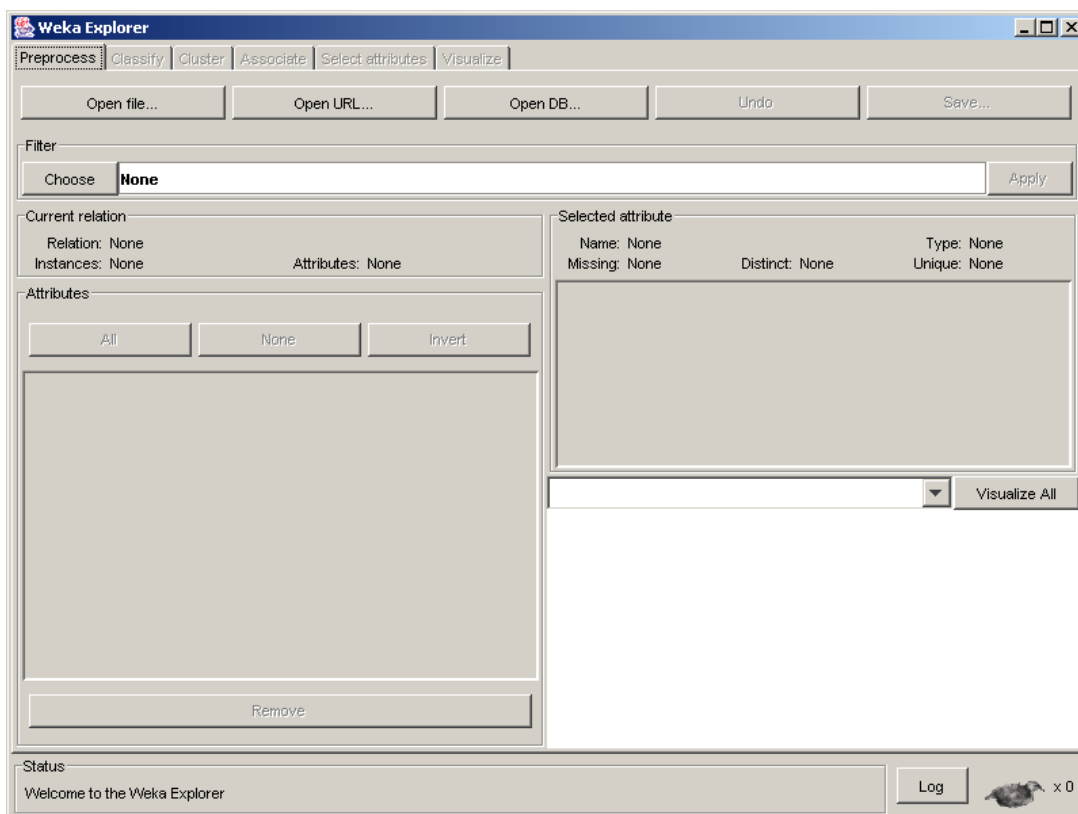
```
@data  
ensolarado,quente,alta,nao_ventando,nao  
ensolarado,quente,alta,ventando,nao  
nublado,quente,alta,nao_ventando,sim  
chuvoso,brando,alta,nao_ventando,sim  
chuvoso,frio,normal,nao_ventando,sim  
chuvoso,frio,normal,ventando,nao  
nublado,frio,normal,ventando,sim  
ensolarado,brando,alta,nao_ventando,nao  
ensolarado,frio,normal,nao_ventando,sim  
chuvoso,brando,normal,nao_ventando,sim  
ensolarado,brando,normal,ventando,sim  
nublado,brando,alta,ventando,sim  
nublado,quente,normal,nao_ventando,sim  
chuvoso,brando,alta,ventando,nao
```

Salvamos como arquivo de texto. Com o nome **condicoes.arff**.

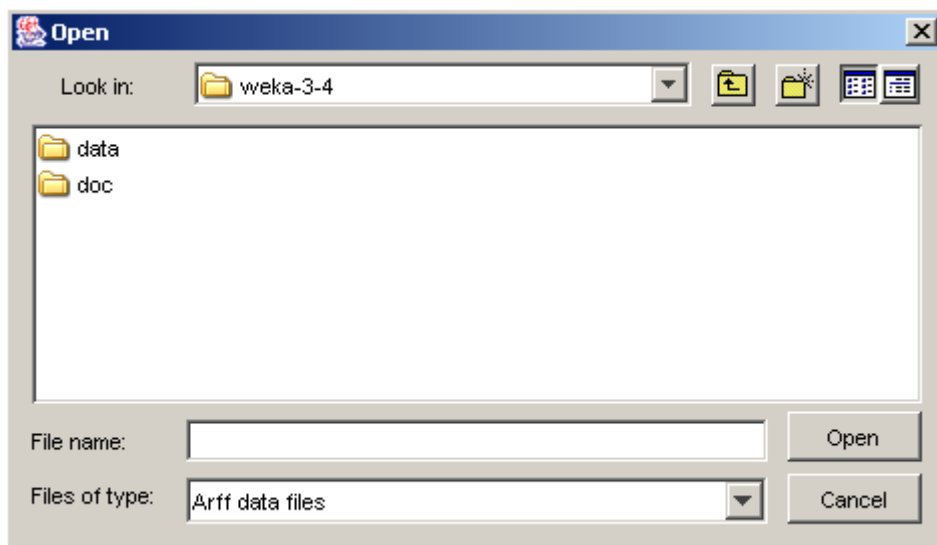
PASSO 2 - Abrimos o WEKA, apresentará a tela abaixo:



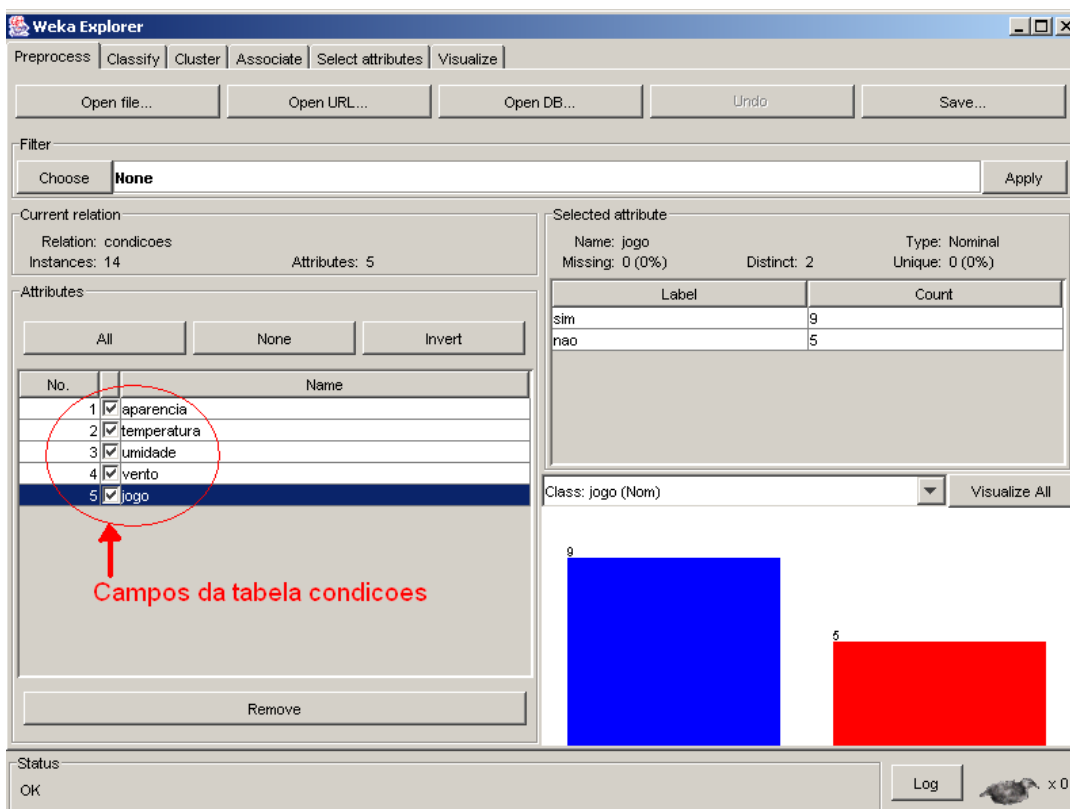
Clicamos no botão “Explorer”, logo será aberto a seguinte janela:



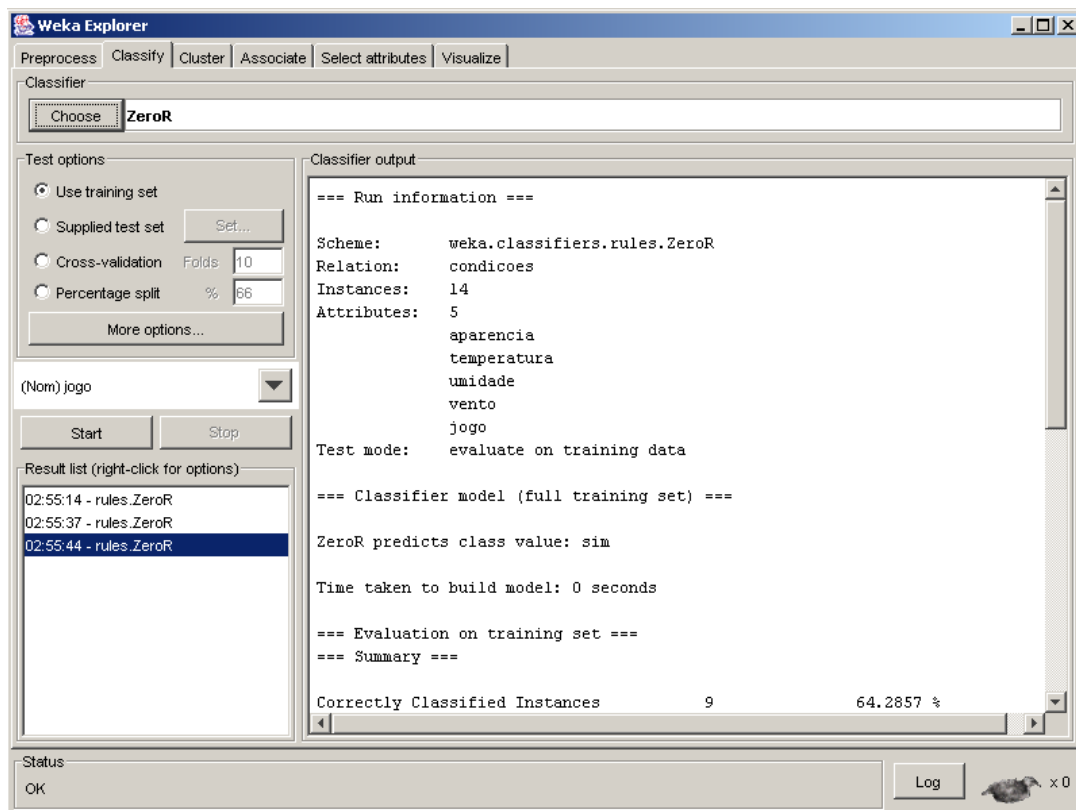
Clicamos no botão “Open file...”, abrirá uma janela de abertura de arquivo:



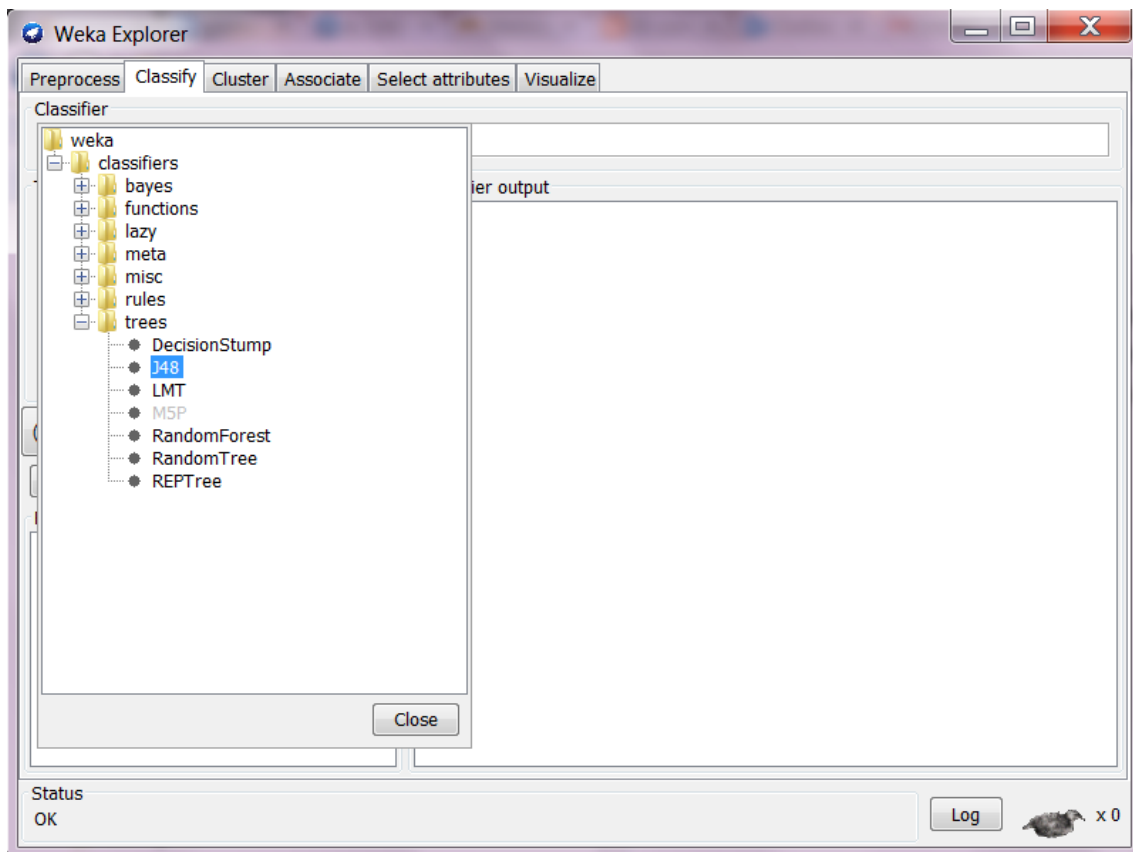
Colocamos a localização do arquivo condições.arff e clicamos no botão “open”, a janela WEKA Explorer mostra os campos, marcamos todos os campos para serem minerados:



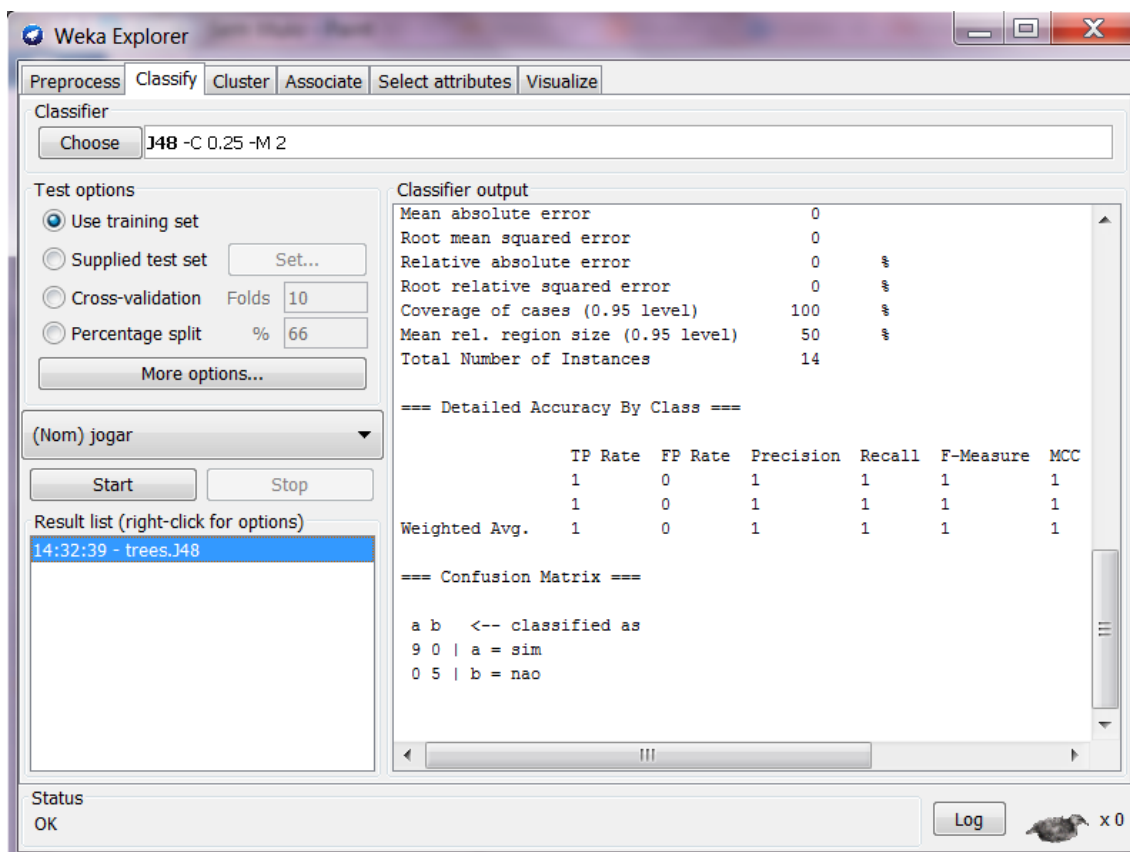
Logo após clicamos na guia classify, será mostrado a janela abaixo:



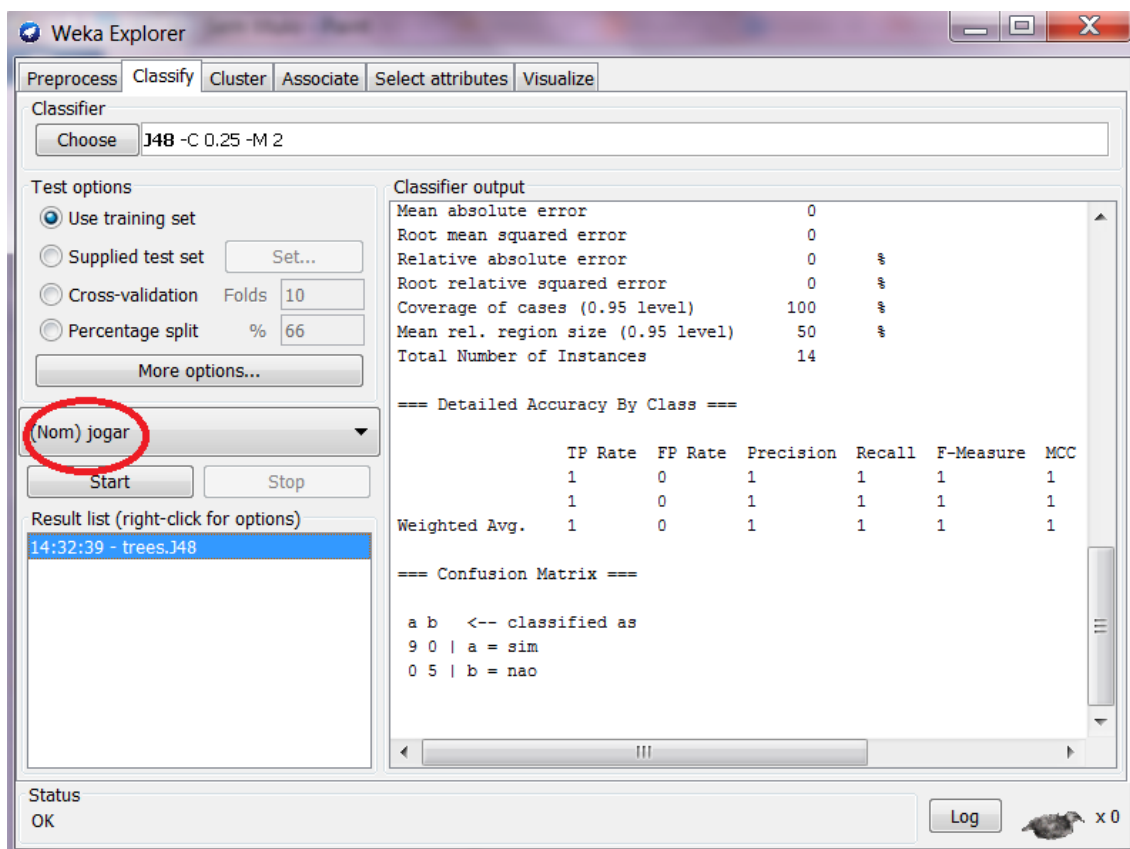
Para selecionar o tipo do algoritmo que foi implementado clicamos no botão “**choose**”, em nosso caso foi o algoritmo J48, para isso selecione dentro da pasta weka, a pasta **classifiers**, dentro dela **trees** e logo após J48. Como mostra a figura abaixo:



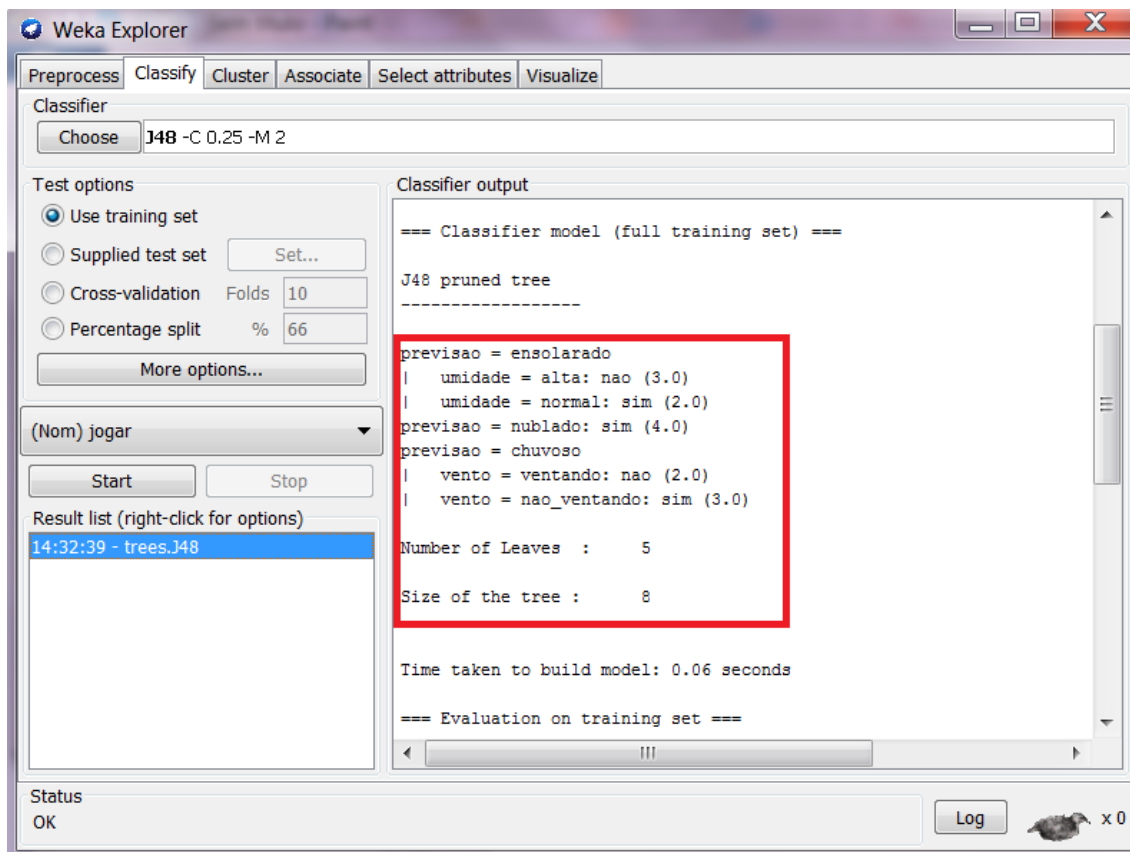
Agora que selecionamos o tipo do algoritmo, vamos clicar em “**use training test**”, para dizer que estamos na fase de treinamento do método de classificação, como mostra a figura:



Selecionaremos o nosso atributo **classe**, que em nosso caso é **jogar**, como mostra a figura abaixo:



logo após clicamos no botão “start”, o weka irá minerar mostrando na parte direita da janela as seguintes regras:



Para melhor visualização:

```

previsao = ensolarado
|  umidade = alta: nao (3.0)
|  umidade = normal: sim (2.0)
previsao = nublado: sim (4.0)
previsao = chuvoso
|  vento = ventando: nao (2.0)
|  vento = nao_ventando: sim (3.0)
  
```

Podemos observar que o weka não mostra regras, mas sim uma árvore que pode ser facilmente analisada. Por exemplo as três primeiras linhas dizem o seguinte:

Quando previsao = chuvoso e vento = nao_ventando então jogo = sim
Ou
Quando previsao = chuvoso e vento = ventando então jogo = não

Lembrando que deve ser analisado a estatística de **Kappa**, que informa se os dados estão bem classificados ou não, **Kappa > 0,80** indica uma boa classificação dos dados.