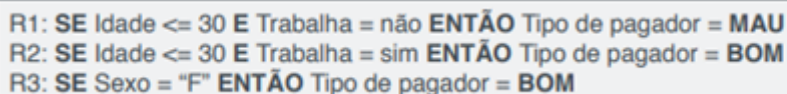


3 | TAREFA DE CLASSIFICAÇÃO

Na tarefa de classificação, o objetivo é descrever ou prever as características de um atributo especial chamado de “atributo-classe” ou “classe”, podendo ter dois ou mais valores possíveis. Por exemplo, considere um atributo-classe chamado de “Tipo de Pagador”, podendo ter dois valores possíveis: sim (representa um bom pagador) e não (representa um mau pagador). Assim, a partir dos atributos existentes no conjunto de dados, o algoritmo de classificação identificará características que representem pessoas que tenham um perfil de bom e mau pagador. A Figura 2 ilustra um conjunto de 03 regras do tipo Se ... , então (R1, R2 e R3) que identificam situações que diferencia nas duas classes citadas.



R1: SE Idade <= 30 E Trabalha = não ENTÃO Tipo de pagador = MAU
R2: SE Idade <= 30 E Trabalha = sim ENTÃO Tipo de pagador = BOM
R3: SE Sexo = "F" ENTÃO Tipo de pagador = BOM

Figura 2. Exemplo de regras geradas pela tarefa de classificação.

Fonte: autoria própria.

Neste exemplo, a classificação serviu para identificar e diferenciar perfil de bons e maus pagadores. A regra R1 informa que se uma pessoa tiver idade menor ou igual a 30 e se não trabalha, então será um mau pagador. Por outro lado, na regra R2, se uma pessoa tiver idade menor ou igual a 30 e trabalhar, então é um bom pagador. A regra R3, por sua vez, informa que pessoas do sexo feminino são boas pagadoras.

A tarefa de classificação tem como objetivo classificar/descrever os grupos existentes observando as características comuns em um conjunto de dados. Na tarefa A tarefa de classificação tem como objetivo classificar/descrever os grupos existentes observando as características comuns em um conjunto de dados. Na tarefa de classificação, a forma de representação do conhecimento (padrões) pode ser com regras do tipo (SE...ENTÃO), (conforme exemplo da Figura 2) ou árvores de decisão (forma gráfica e hierárquica na apresentação das características dos grupos). Cada tarefa pode ainda ter várias implementações através de diferentes algoritmos. Na literatura, há uma infinidade de algoritmos de classificação disponíveis. No caso da ferramenta de mineração de dados Weka, Universidade de Waikato na Nova Zelândia, existem os seguintes algoritmos de classificação: Part, OneR, J48 e randomtree (FRANK et al., 2016).