

1 Adult dataset

1.1 Apresentao dos Dados

Nome: Adult

Descrição: The Adult data set was extracted in 1994 from census data of the United States. It contains continuous and nominal attributes, describing some social information (age, race, sex, marital status, ...) about the citizens registered. The task is to predict whether the citizens income exceeds fifty thousand dollars a year.

Objetivo:

Instâncias: 45222 (48842)

Número de atributos: 14

Número de classes: 2

Atributo - Age: dado numérico, possíveis valores float de 17.0 a 90.0

Atributo - Workclass: dado categórico, possíveis valores: (Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked)

Atributo - Fnlwgt: dado numérico, possíveis valores float de 12285.0 a 1490400.0

Atributo - Education: dado categórico, possíveis valores: (Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool)

Atributo - Education-num: dado numérico, possíveis valores float de 1.0 a 16.0

Atributo - Marital-status: dado categórico, possíveis valores: (Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse)

Atributo - Occupation: dado categórico, possíveis valores: (Tech-support, Craft-repair, Other-service, Sales, Exec-managerial, Prof-specialty, Handlers-cleaners, Machine-op-inspct, Adm-clerical, Farming-fishing, Transport-moving, Priv-house-serv, Protective-serv, Armed-Forces)

Atributo - Relationship: dado categórico, possíveis valores: (Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried)

Atributo - Race: dado categórico, possíveis valores: (White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black)

Atributo - Sex: dado categórico, possíveis valores: (Female, Male)

Atributo - Capital-gain: dado numérico, possíveis valores float de 0.0 a 99999.0

Atributo - Capital-loss: dado numérico, possíveis valores float de 0.0 a 4356.0

Atributo - Hours-per-week: dado numérico, possíveis valores float de 1.0 a 99.0

Atributo - Native-country: dado categórico, possíveis valores: (United-States, Cambodia, England, Puerto-Rico, Canada, Germany, Outlying-US(Guam-USVI-etc), India, Japan, Greece, South, China, Cuba, Iran, Honduras, Philippines, Italy, Poland, Jamaica, Vietnam, Mexico, Portugal, Ireland, France, Dominican-Republic, Laos, Ecuador, Taiwan, Haiti, Columbia, Hungary, Guatemala, Nicaragua, Scotland, Thailand, Yugoslavia, El-Salvador, Trinidad&Tobago, Peru, Hong, Holand-Netherlands)

Atributo - Class: dado categórico, possíveis valores: (>50K, <=50K)