

Data Description

trainFeatures.csv & testFeatures.csv

age: The age of the individual; this attribute is continuous.

work-class: The type of the employer that the individual has, involving Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked; this attribute is nominal.

fnlwgt: Final weight, this is the number of people the census believes the entry represents; this attribute is continuous.

education: The highest level of education achieved for that individuals involving Preschool, 1st-4th, 5th-6th, 7th-8th, 9th, 10th, 11th, 12th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, Some-college, Bachelors, Masters, Doctorate; This attribute is nominal.

education-num: Highest level of education in numerical form; this attribute is continuous.

marital-status: Marital status of the individual, involving Divorced, Married-AF-spouse, Married-civ-spouse, Married-spouse-absent, Never-married, Separated, Widowed; This attribute is nominal.

occupation: The occupation of the individual, involving 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; This attribute is nominal.

relationship: The family relationship of the individual, involving Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried; This attribute is nominal.

race: The race of the individual, involving White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black; This attribute is nominal.

sex: Female, Male; This attribute is nominal.

capital-gain: capital gains recorded; This attribute is continuous.

capital-loss: capital losses recorded; This attribute is continuous.

hours-per-week: Hours worked per week; This attribute is continuous.

native-country: label person's country, involving 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.

trainLabels.csv

0: means a person makes no more than 50K a year, i.e. $\leq 50k$

1: means a person makes over 50K a year, i.e. $> 50K$