**Some information of our problem.**



**Source:**

Donor:   
Ronny Kohavi and Barry Becker   
Data Mining and Visualization   
Silicon Graphics.   
e-mail: ronnyk **'@'** sgi.com for questions.

**Data Set Information:**

Extraction was done by Barry Becker from the 1994 Census database. A set of reasonably clean records was extracted using the following conditions: ((AAGE>16) && (AGI>100) && (AFNLWGT>1)&& (HRSWK>0))   
  
Prediction task is to determine whether a person makes over 50K a year.

**Attribute Information:**

Listing of attributes:   
  
>50K, <=50K.

|  |  |
| --- | --- |
| age: | continuous. |
| workclass: | Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked. |
| fnlwgt: | continuous |
| education: | Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool |
| education-num: | continuous. |
| marital-status: | Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse. |
| occupation: | 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. |
| relationship: | Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried |
| race: | White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black. |
| sex: | Female, Male. |
| capital-gain: | continuous. |
| capital-loss: | continuous. |
| hours-per-week: | continuous. |
| native-country: | 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, Trinadad&Tobago, Peru, Hong, Holand-Netherlands. |

11th: Eleventh

9th: Ninth

Seventh\_Eighth

Twelfth

First\_Fourth

Tenth

Fith\_Sixth

Outlying-US

TrinadadTobago

**Relevant Papers:**

Ron Kohavi, "Scaling Up the Accuracy of Naive-Bayes Classifiers: a Decision-Tree Hybrid", Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, 1996   
[[Web Link]](http://rexa.info/paper/bf6cec50b7f7d48d105c8c649210cc3a42d3d71e)