



Adult

Donated on 4/30/1996

Predict whether annual income of an individual exceeds \$50K/yr based on census data. Also known as "Census Income" dataset.

Dataset Characteristics

Multivariate

Subject Area

Social Science

Associated Tasks

Classification

Feature Type

Categorical, Integer

Instances

48842

Features

14

Dataset Information



Additional 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))...

SHOW MORE

Has Missing Values?

Yes

Variables Table



Variable Name	Role	Type	Demographic	Description	Units	Missing Value
age	Feature	Integer	Age	N/A		no
workclass	Feature	Categorical	Income	Private, Self-emp-not-inc, Self-emp-inc,		yes

Variable Name	Role	Type	Demographic	Description	Units	Missing Value
				Federal-gov, Local-gov, State-gov, Without-pay, Never-worked.		
fnlwgt	Feature	Integer				no
education	Feature	Categorical	Education Level	Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool.		no
education-num	Feature	Integer	Education Level			no
marital-status	Feature	Categorical	Other	Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse.		no
occupation	Feature	Categorical	Other	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.		yes

Variable Name	Role	Type	Demographic	Description	Units	Missing Value
relationship	Feature	Categorical	Other	Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried.		no
race	Feature	Categorical	Race	White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black.		no
sex	Feature	Binary	Sex	Female, Male.		no
capital-gain	Feature	Integer				no
capital-loss	Feature	Integer				no
hours-per-week	Feature	Integer				no
native-country	Feature	Categorical	Other	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,		yes

Variable Name	Role	Type	Demographic	Description	Units	Missing Value
				Peru, Hong, Holand-Netherlands.		
income	Target	Binary	Income	>50K, <=50K.		no

Rows per page 25 0 to 15 of 15 < >

Additional Variable Information

Listing of attributes:

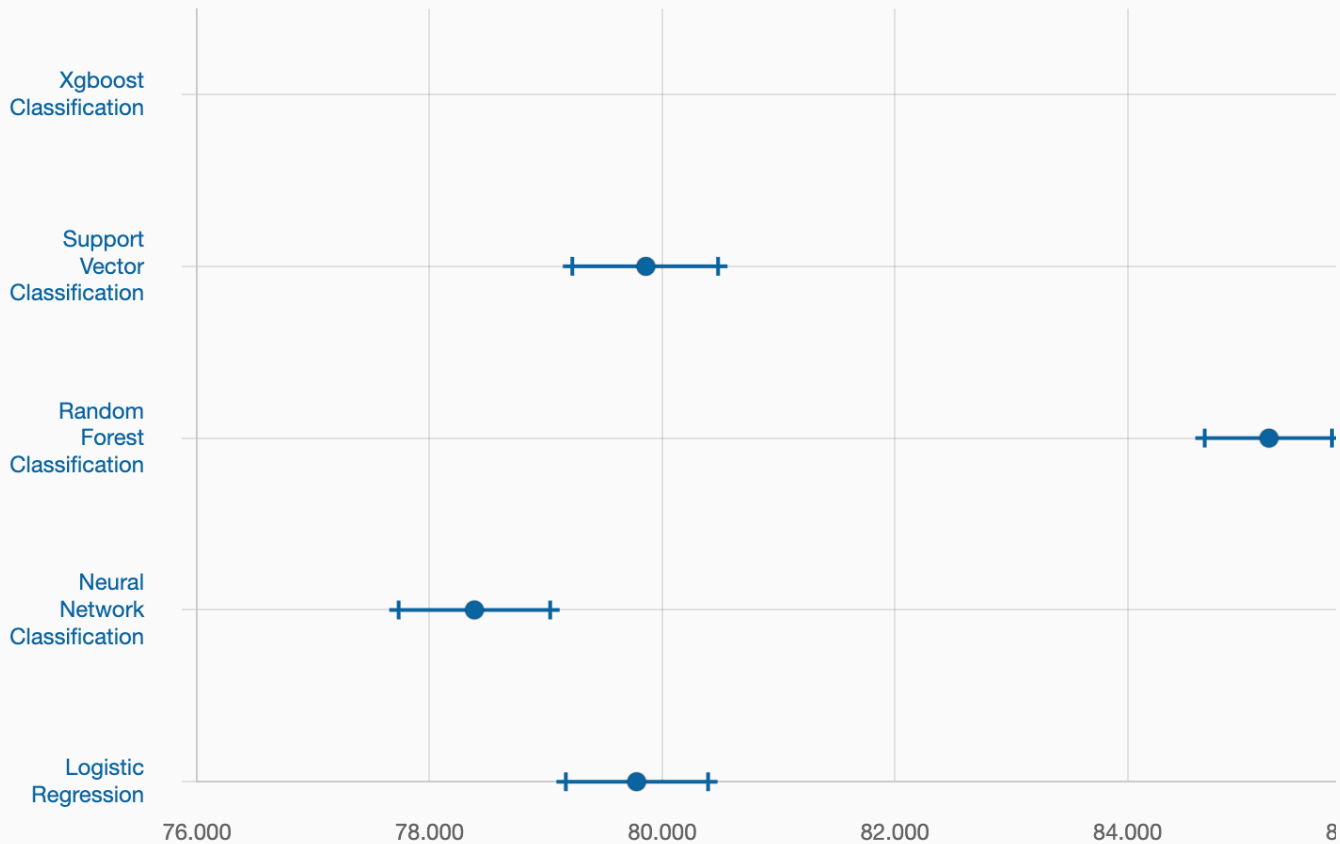
>50K, <=50K....

SHOW MORE

Baseline Model Performance

Accuracy

Precision



Dataset Files

File	Size
adult.data	3.8 MB
adult.test	1.9 MB
adult.names	5.1 KB
old.adult.names	4.2 KB
Index	140 Bytes

Papers Citing this Dataset

≡ SORT BY YEAR, DESC

[Integrating Association Rules with Decision Trees in Object-Relational Databases](#)

By Maruthi Ayyagari. 2019

Published in International Journal of Engineering Trends and Technology 67.3 (2019): 102-108.

[Outis: Crypto-Assisted Differential Privacy on Untrusted Servers](#)

By Amrita Chowdhury, Chenghong Wang, Xi He, Ashwin Machanavajjhala, Somesh Jha. 2019

Published in ArXiv.

[Refining the Structure of Neural Networks Using Matrix Conditioning](#)

By Roozbeh Yousefzadeh, Dianne O'Leary. 2019

Published in ArXiv.

[Fair k-Center Clustering for Data Summarization](#)

By Matthias Kleindessner, Pranjal Awasthi, Jamie Morgenstern. 2019

Published in ArXiv.

[Effectiveness of Equalized Odds for Fair Classification under Imperfect Group Information](#)

By Pranjal Awasthi, Matthias Kleindessner, Jamie Morgenstern. 2019

Published in ArXiv.

Rows per page

5



0 to 5 of 257



Reviews



2.1      (4 ratings)

[LOGIN TO WRITE A REVIEW](#)

po Kappa



trash dataset. income column contains extra "."

3/26/2025

robert jones



very cool!

12/19/2024

mr JoZef



very bad dataset. errors with parsing columns.

11/9/2024

Haji Qasim



The name of columns are not showing as mentioned here.

9/13/2024

 **DOWNLOAD** (605.7 KB)



IMPORT IN PYTHON

CITE

 **257 citations**

 **419177 views**

Keywords

census

Creators



Barry Becker

Silicon Graphics



Ronny Kohavi

ronnyk@live.com

Consultant/Stanford

DOI

10.24432/C5XW20

License

This dataset is licensed under a [Creative Commons Attribution 4.0 International](#) (CC BY 4.0) license.

This allows for the sharing and adaptation of the datasets for any purpose, provided that the appropriate credit is given.



THE PROJECT

[About Us](#)

[CML](#)

[National Science Foundation](#)

NAVIGATION

[Home](#)

[View Datasets](#)

[Donate a Dataset](#)

LOGISTICS

[Contact](#)

[Privacy Notice](#)

[Feature Request or Bug Report](#)