Data Warehousing and Data Mining
Mini Project

Jawaharlal Nehru Technological University Anantapur, Ananthapuramu

in partial fulfillment of the requirements for
the award of the degree of

BACHELOR OF TECHNOLOGY
IN
INFORMATION TECHNOLOGY

Submitted by

20121A1207	***B. Ganesh***
20121A1208	***B. Vyshnavi***
20121A1209	***B. Venkata Bhavana***
20121A1210	***B. Bindhu Siva Sree***
20121A1211	***B. Lalith Kumar***

Under the Supervision of

Ms. CH. Prathima, *M.Tech.(ph.d)*
Professor
Department of Information Technology

Department of Information Technology
SREE VIDYANIKETHAN ENGINEERING COLLEGE
(AUTONOMOUS)
(Affiliated to JNTUA, Ananthapuramu, Approved by AICTE, Accredited by NBA & NAAC)
Sree Sainath Nagar, Tirupati – 517 102, A.P., INDIA

2022-2023

DATA SETS

Adult Data Set

Abstract: Predict whether income exceeds $50K/yr based on census data. Also known as "Census Income" dataset.

Data Set

Data Set Characteristics:	Multivariate	**Number of Instances:**	48842	**Area:**	Social
Attribute Characteristics:	Categorical, Integer	**Number of Attributes:**	14	**Date Donated**	1996-05-01
Associated Tasks:	Classification	**Missing Values?**	Yes	**Number of Web Hits:**	2662217

Source:

Donor:

Ronny Kohavi and Barry Becker

Data Mining and Visualization

Silicon Graphics.

e-mail: ronnyk '@' live.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.

work class: 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.

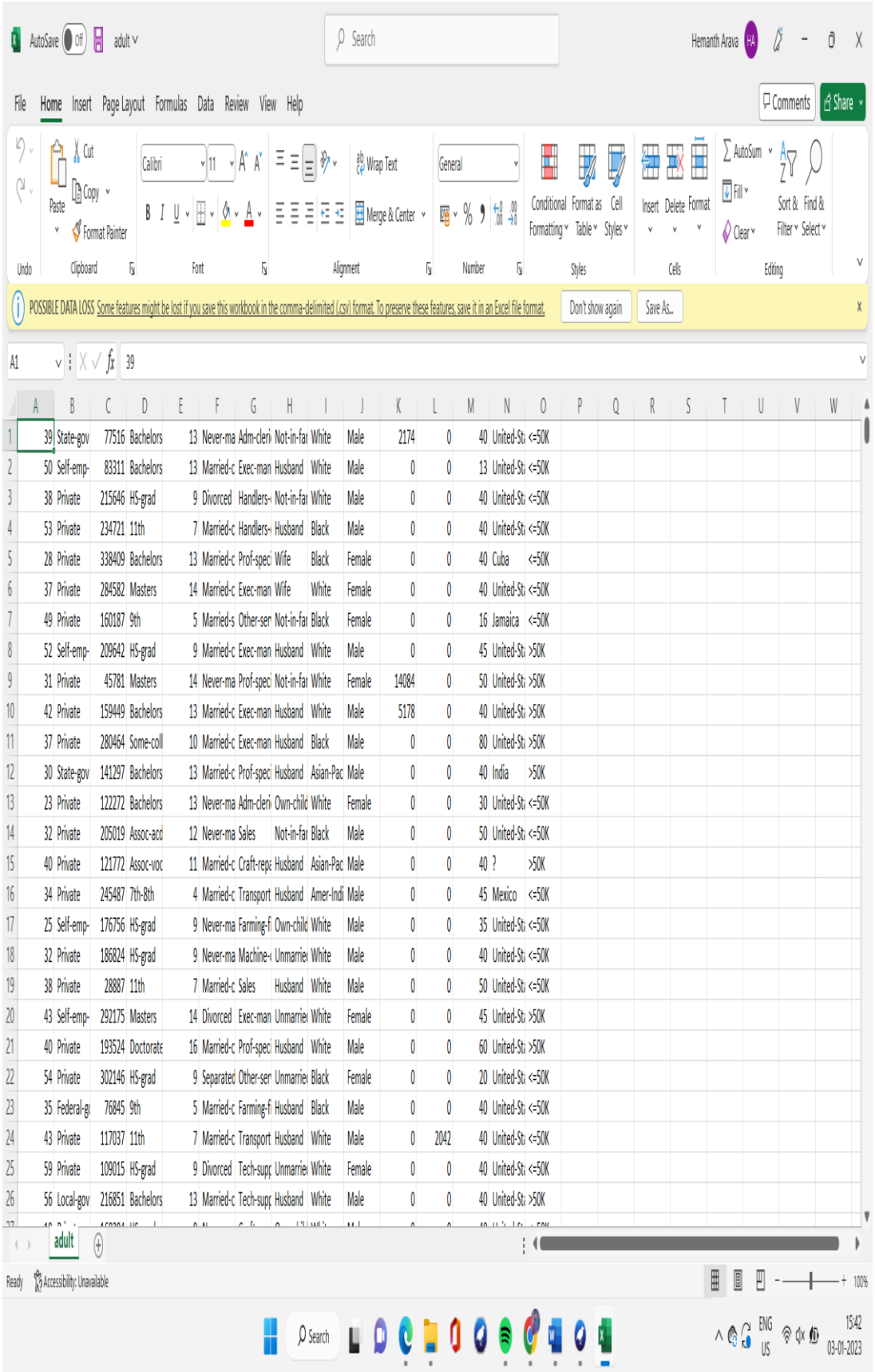
sex: Female, Male.

capital-gain: continuous.

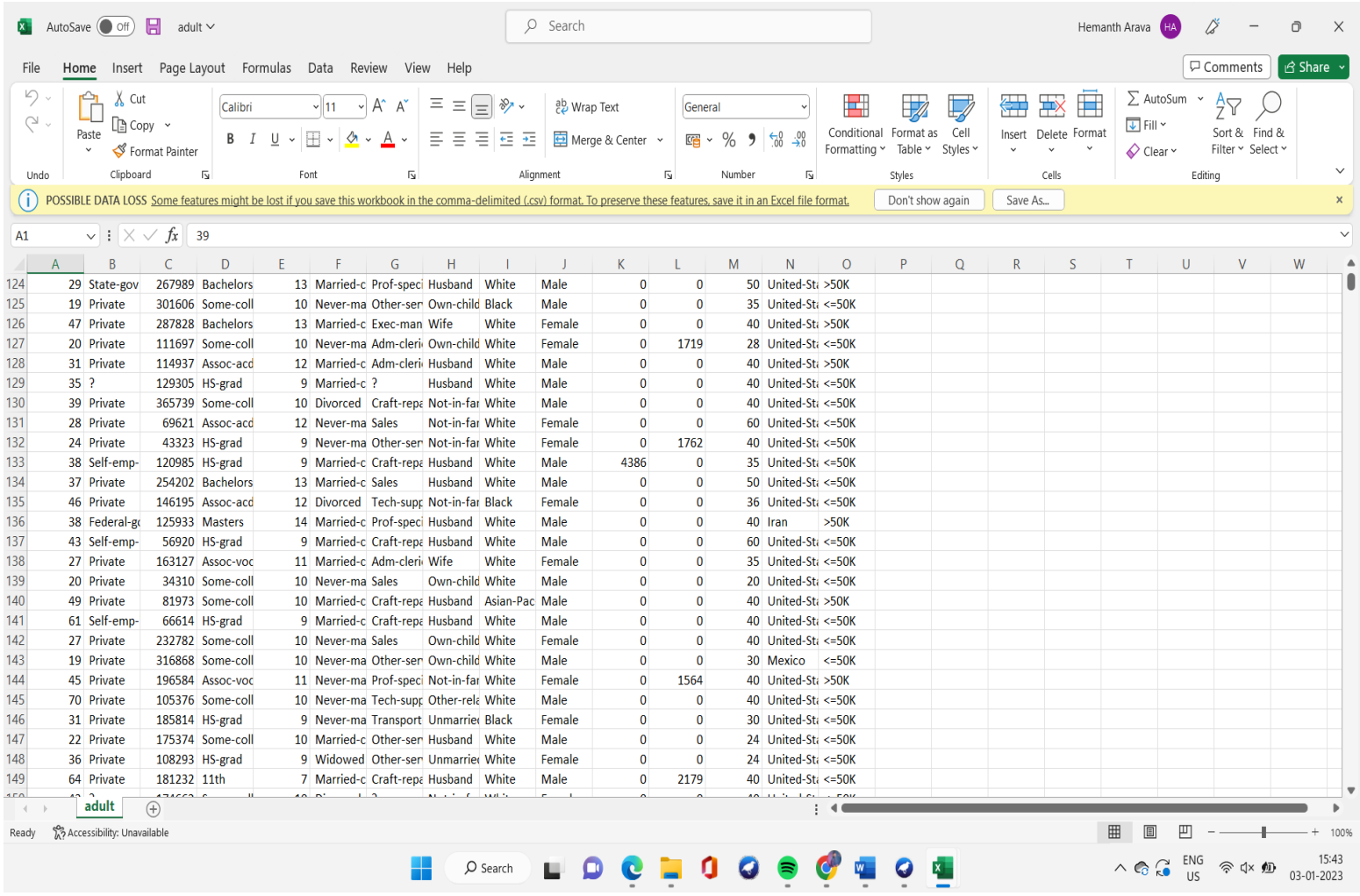
capital-loss: continuous.

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.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	39	State-gov	77516	Bachelors	13	Never-ma	Adm-clerl	Not-in-fa	White	Male	2174	0	40	United-St	<=50K
2	50	Self-emp-	83311	Bachelors	13	Married-c	Exec-man	Husband	White	Male	0	0	13	United-St	<=50K
3	38	Private	215646	HS-grad	9	Divorced	Handlers-	Not-in-fa	White	Male	0	0	40	United-St	<=50K
4	53	Private	234721	11th	7	Married-c	Handlers-	Husband	Black	Male	0	0	40	United-St	<=50K
5	28	Private	338409	Bachelors	13	Married-c	Prof-speci	Wife	Black	Female	0	0	40	Cuba	<=50K
6	37	Private	284582	Masters	14	Married-c	Exec-man	Wife	White	Female	0	0	40	United-St	<=50K
7	49	Private	160187	9th	5	Married-s	Other-ser	Not-in-fa	Black	Female	0	0	16	Jamaica	<=50K
8	52	Self-emp-	209642	HS-grad	9	Married-c	Exec-man	Husband	White	Male	0	0	45	United-St	>50K
9	31	Private	45781	Masters	14	Never-ma	Prof-speci	Not-in-fa	White	Female	14084	0	50	United-St	>50K
10	42	Private	159449	Bachelors	13	Married-c	Exec-man	Husband	White	Male	5178	0	40	United-St	>50K
11	37	Private	280464	Some-coll	10	Married-c	Exec-man	Husband	Black	Male	0	0	80	United-St	>50K
12	30	State-gov	141297	Bachelors	13	Married-c	Prof-speci	Husband	Asian-Pac	Male	0	0	40	India	>50K
13	23	Private	122272	Bachelors	13	Never-ma	Adm-clerl	Own-child	White	Female	0	0	30	United-St	<=50K
14	32	Private	205019	Assoc-acd	12	Never-ma	Sales	Not-in-fa	Black	Male	0	0	50	United-St	<=50K
15	40	Private	121772	Assoc-voc	11	Married-c	Craft-repa	Husband	Asian-Pac	Male	0	0	40	?	>50K
16	34	Private	245487	7th-8th	4	Married-c	Transport	Husband	Amer-Indi	Male	0	0	45	Mexico	<=50K
17	25	Self-emp-	176756	HS-grad	9	Never-ma	Farming-fi	Own-child	White	Male	0	0	35	United-St	<=50K
18	32	Private	186824	HS-grad	9	Never-ma	Machine-i	Unmarried	White	Male	0	0	40	United-St	<=50K
19	38	Private	28887	11th	7	Married-c	Sales	Husband	White	Male	0	0	50	United-St	<=50K
20	43	Self-emp-	292175	Masters	14	Divorced	Exec-man	Unmarried	White	Female	0	0	45	United-St	>50K
21	40	Private	193524	Doctorate	16	Married-c	Prof-speci	Husband	White	Male	0	0	60	United-St	>50K
22	54	Private	302146	HS-grad	9	Separated	Other-ser	Unmarried	Black	Female	0	0	20	United-St	<=50K
23	35	Federal-g	76845	9th	5	Married-c	Farming-fi	Husband	Black	Male	0	0	40	United-St	<=50K
24	43	Private	117037	11th	7	Married-c	Transport	Husband	White	Male	0	2042	40	United-St	<=50K
25	59	Private	109015	HS-grad	9	Divorced	Tech-supp	Unmarried	White	Female	0	0	40	United-St	<=50K
26	56	Local-gov	216851	Bachelors	13	Married-c	Tech-supp	Husband	White	Male	0	0	40	United-St	>50K



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
124	29	State-gov	267989	Bachelors	13	Married-c	Prof-speci	Husband	White	Male	0	0	50	United-Sta	>50K
125	19	Private	301606	Some-coll	10	Never-ma	Other-ser	Own-child	Black	Male	0	0	35	United-Sta	<=50K
126	47	Private	287828	Bachelors	13	Married-c	Exec-man	Wife	White	Female	0	0	40	United-Sta	>50K
127	20	Private	111697	Some-coll	10	Never-ma	Adm-cleri	Own-child	White	Female	0	1719	28	United-Sta	<=50K
128	31	Private	114937	Assoc-acd	12	Married-c	Adm-cleri	Husband	White	Male	0	0	40	United-Sta	>50K
129	35	?	129305	HS-grad	9	Married-c	?	Husband	White	Male	0	0	40	United-Sta	<=50K
130	39	Private	365739	Some-coll	10	Divorced	Craft-repa	Not-in-fa	White	Male	0	0	40	United-Sta	<=50K
131	28	Private	69621	Assoc-acd	12	Never-ma	Sales	Not-in-fa	White	Female	0	0	60	United-Sta	<=50K
132	24	Private	43323	HS-grad	9	Never-ma	Other-ser	Not-in-fa	White	Female	0	1762	40	United-Sta	<=50K
133	38	Self-emp-	120985	HS-grad	9	Married-c	Craft-repa	Husband	White	Male	4386	0	35	United-Sta	<=50K
134	37	Private	254202	Bachelors	13	Married-c	Sales	Husband	White	Male	0	0	50	United-Sta	<=50K
135	46	Private	146195	Assoc-acd	12	Divorced	Tech-supp	Not-in-fa	Black	Female	0	0	36	United-Sta	<=50K
136	38	Federal-g	125933	Masters	14	Married-c	Prof-speci	Husband	White	Male	0	0	40	Iran	>50K
137	43	Self-emp-	56920	HS-grad	9	Married-c	Craft-repa	Husband	White	Male	0	0	60	United-Sta	<=50K
138	27	Private	163127	Assoc-voc	11	Married-c	Adm-cleri	Wife	White	Female	0	0	35	United-Sta	<=50K
139	20	Private	34310	Some-coll	10	Never-ma	Sales	Own-child	White	Male	0	0	20	United-Sta	<=50K
140	49	Private	81973	Some-coll	10	Married-c	Craft-repa	Husband	Asian-Pac	Male	0	0	40	United-Sta	>50K
141	61	Self-emp-	66614	HS-grad	9	Married-c	Craft-repa	Husband	White	Male	0	0	40	United-Sta	<=50K
142	27	Private	232782	Some-coll	10	Never-ma	Sales	Own-child	White	Female	0	0	40	United-Sta	<=50K
143	19	Private	316868	Some-coll	10	Never-ma	Other-ser	Own-child	White	Male	0	0	30	Mexico	<=50K
144	45	Private	196584	Assoc-voc	11	Never-ma	Prof-speci	Not-in-fa	White	Female	0	1564	40	United-Sta	>50K
145	70	Private	105376	Some-coll	10	Never-ma	Tech-supp	Other-rela	White	Male	0	0	40	United-Sta	<=50K
146	31	Private	185814	HS-grad	9	Never-ma	Transport	Unmarrie	Black	Female	0	0	30	United-Sta	<=50K
147	22	Private	175374	Some-coll	10	Married-c	Other-ser	Husband	White	Male	0	0	24	United-Sta	<=50K
148	36	Private	108293	HS-grad	9	Widowed	Other-ser	Unmarrie	White	Female	0	0	24	United-Sta	<=50K
149	64	Private	181232	11th	7	Married-c	Craft-repa	Husband	White	Male	0	2179	40	United-Sta	<=50K

Anuran Calls(MFCCs)

Abstract: Acoustic features extracted from syllables of anuran (frogs) calls, including the mily, the genus, and the species labels (multilabel).

Data Set Characteristic	Multivariate	**Number of Instances:**	7195	**Area:**	Life
Attribute Characteristics:	Real	**Number of Attributes:**	22	**Date Donated**	2017-02-24
Associated Tasks:	Classification, Clustering	**Missing Values?**	N/A	**Number of Web Hits:**	71989

Source:

Eng. Juan Gabriel Colonna <juancolonna '@' icomp.ufam.edu.br>, Prof. Eduardo Freire Nakamura <nakamura '@' icomp.ufam.edu.br>, Prof. Marco A. P. Cristo <marco.cristo '@' gmail.com>, Biologist and collaborator Prof. Marcelo Gordo <mgordo '@' ufam.edu.br> Universidade Federal do Amazonas, Av. General Rodrigo Octavio JordÃ£o Ramos, 1200 - Coroado I, Manaus - AM, 69067-005, Brasil.

Data Set Information:

This dataset was used in several classifications tasks related to the challenge of anuran species recognition through their calls. It is a multilabel dataset with three columns of labels. This dataset was created segmenting 60 audio records belonging to 4 different families, 8 genus, and 10 species. Each audio corresponds to one specimen (an individual frog), the record ID is also included as an extra column. We used the spectral entropy and a binary cluster method to detect audio frames belonging to each syllable. The segmentation and feature extraction were carried out in Matlab. After the segmentation we got 7195 syllables, which became instances for train and test the classifier. These records were collected in situ under real noise conditions (the background sound). Some species are from the campus of Federal University of Amazonas, Manaus, others from Mata AtlÃ¢ntica, Brazil, and one of them from CÃ³rdoba, Argentina. The recordings were stored in wav format with 44 triangular filters. These coefficients were normalized between -1 â‰¤ mfcc â‰¤ 1. The amount of instances per class

Families:

Bufonidae 68

Dendrobatidae 542

Hylidae 2165

Leptodactylidae 4420

Genus:

Adenomera 4150

Ameerega 542

Dendropsophus 310

Hypsiboas 1593

Leptodactylus 270

Osteocephalus 114

Rhinella 68

Scinax 148

Species:

AdenomeraAndre 672

AdenomeraHylaedactâ€¦ 3478

Ameeregatrivittata 542

HylaMinuta 310

HypsiboasCinerascens 472

HypsiboasCordobae 1121

LeptodactylusFuscus 270

OsteocephalusOophaâ€¦ 114

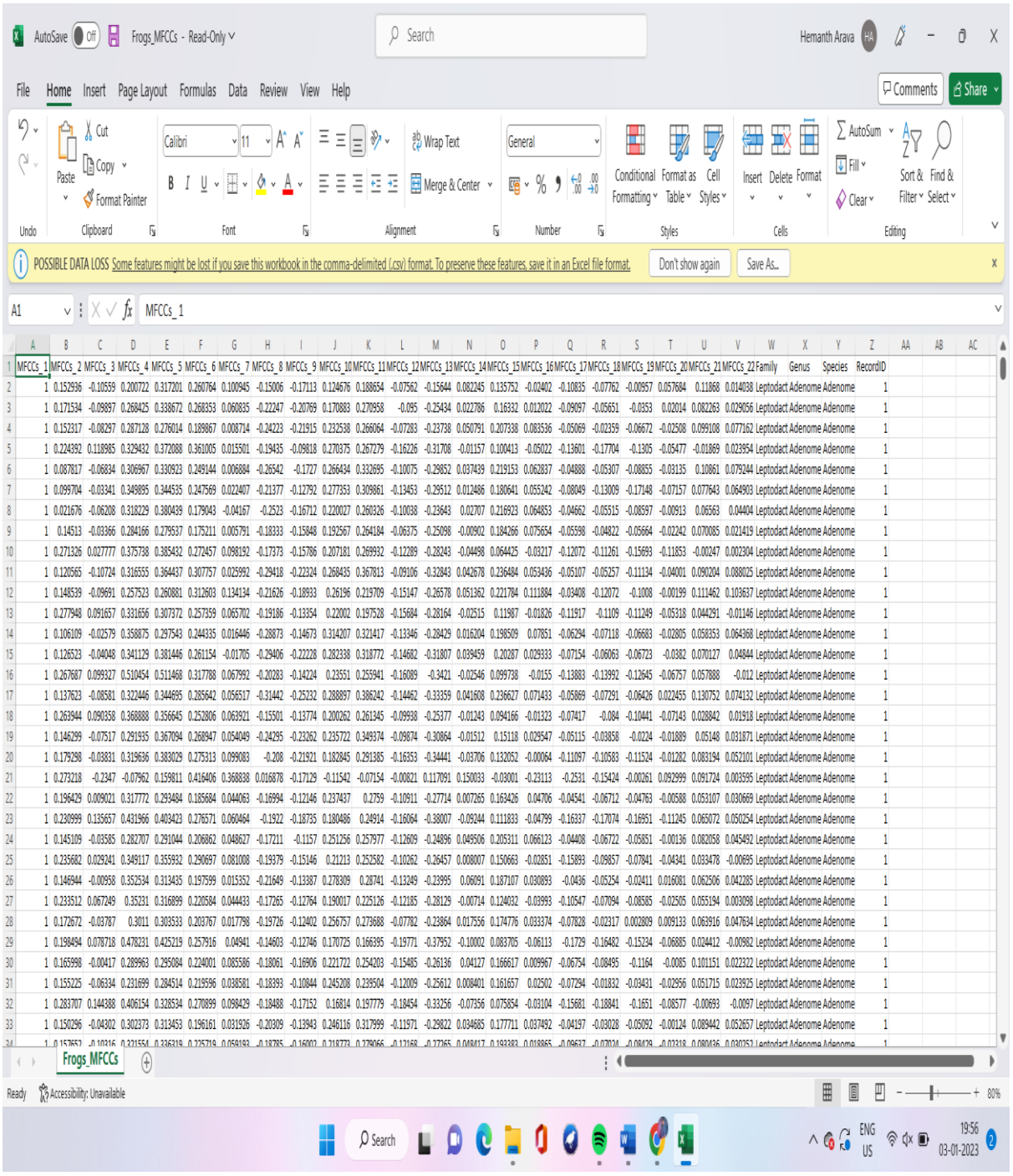
Rhinellagranulosa 68

ScinaxRuber 148

Attribute Information:

Mel-frequency cepstral coefficients (MFCCs) are coefficients that collectively make up an mel-frequency cepstrum (MFC). Due to each syllable has different length, every row (i) was normalized acording to MFCCs_i/(max(abs(MFCCs_i))).

DATASET



Accelerometer

Abstract: Accelerometer data from vibrations of a cooler fan with weights on its blades. It can be used for predictions, classification and other tasks that require vibration analysis, specially in engines.

Data Set Characteristics:	Multivariate	**Number of Instances:**	153000	**Area:**	Computer
Attribute Characteristics:	Integer, Real	**Number of Attributes:**	5	**Date Donated**	2021-05-02
Associated Tasks:	Classification Regression	**Missing Values?**	N/A	**Number of Web Hits:**	52287

Source:

Gustavo Scalabrini Sampaio
gustavo.sampaio '@' mackenzista.com.br
Postgraduate Program in Electrical Engineering and Computing, Mackenzie Presbyterian University, SÃ£o Paulo, Brazil.
Arnaldo Rabello de Aguiar Vallim Filho
arnaldo.aguiar '@' mackenzie.br
Computer Science Dept., Mackenzie Presbyterian University, SÃ£o Paulo, Brazil.
Leilton Santos da Silva
leilton '@' emae.com.br
EMAEâ€ ” Metropolitan Company of Water & Energy, SÃ£o Paulo, Brazil.
Leandro Augusto da Silva
leandroaugusto.silva '@' mackenzie.br
Postgraduate Program in Electrical Engineering and Computing, Mackenzie Presbyterian University, SÃ£o Paulo, Brazil.

Donator:

Gustavo Scalabrini Sampaio

Data Set Information:

This dataset was generated for use on 'Prediction of Motor Failure Time Using An Artificial Neural Network' project (DOI: 10.3390/s19194342). A cooler fan with weights on its blades

was used to generate vibrations. To this fan cooler was attached an accelerometer to collect the vibration data. With this data, motor failure time predictions were made, using an artificial neural networks. To generate three distinct vibration scenarios, the weights were distributed in three different ways: 1) 'red' – normal configuration: two weight pieces positioned on neighboring blades; 2) 'blue' - perpendicular configuration: two weight pieces positioned on blades forming a 90Â angle; 3) 'green' - opposite configuration: two weight pieces positioned on opposite blades. A schematic diagram can be seen in figure 3 of the paper.

Devices used:

Akasa AK-FN059 12cm Viper cooling fan (Generate the vibrations)
MMA8452Q accelerometer (Measure vibration)

Data collection method:

17 rotation speeds were set up, ranging from 20% to 100% of the cooler maximum speed at 5% intervals; for the three weight distribution configurations in the cooler blades. Note that the Akasa AK-FN059 cooler has 1900 rpm of max rotation speed. The vibration measurements were collected at a frequency of 20 ms for 1 min for each percentage, generating 3000 records per speed. Thus, in total, 153,000 vibration records were collected from the simulation model.

Attribute Information:

There are 5 attributes in the dataset: wconfid,pctid,x,y and z.
wconfid: Weight Configuration ID (1 - 'red' - normal configuration; 2 - 'blue' - perpendicular configuration; 3 - 'green' - opposite configuration) pctid: Cooler Fan RPM Speed Percentage ID (20 means 20%, and so on).
x: Accelerometer x value.
y: Accelerometer y value.
z: Accelerometer z value.

Data set

wconfid	pcd	x	y	z	
1	20	1.004	0.09	-0.125	
1	20	1.004	-0.043	-0.125	
1	20	0.969	0.09	-0.121	
1	20	0.973	-0.012	-0.137	
1	20	1	-0.016	-0.121	
1	20	0.961	0.082	-0.121	
1	20	0.973	-0.055	-0.109	
1	20	1	0.012	-0.133	
1	20	0.969	-0.102	-0.141	
1	20	0.973	-0.059	-0.125	
1	20	1.012	0.043	-0.133	
1	20	0.996	-0.109	-0.148	
1	20	0.988	-0.02	-0.125	
1	20	1.012	0.043	-0.129	
1	20	0.996	-0.09	-0.152	
1	20	0.965	-0.102	-0.117	
1	20	1.004	0.055	-0.121	
1	20	0.988	-0.059	-0.141	
1	20	0.969	-0.086	-0.117	
1	20	1.039	0.094	-0.117	
1	20	0.984	0.113	-0.148	
1	20	1.008	0.012	-0.141	
1	20	0.996	0.035	-0.141	
1	20	0.988	-0.066	-0.125	
1	20	0.965	-0.156	-0.113	
1	20	0.992	0.059	-0.129	
1	20	0.992	-0.031	-0.125	
1	20	0.973	-0.133	-0.148	
1	20	1.031	0.102	-0.145	
1	20	0.984	-0.035	-0.125	
1	20	0.984	-0.004	-0.137	
1	20	1	0.098	-0.125	
1	20	1.031	-0.012	-0.129	
1	20	0.965	-0.148	-0.152	
1	20	1.012	0.102	-0.129	
1	20	1.008	-0.008	-0.133	
1	20	0.965	-0.129	-0.164	
1	20	1.004	0.145	-0.121	
1	20	1.008	0.004	-0.121	
1	20	0.984	0.027	-0.121	
1	20	1.031	0.07	-0.117	
1	20	0.977	-0.051	-0.152	
1	20	0.98	-0.094	-0.117	

1	20	1.02	0.082	-0.125	
1	20	0.98	-0.012	-0.137	
1	20	0.977	-0.09	-0.125	
1	20	1.023	0.109	-0.109	
1	20	0.992	0.008	-0.137	
1	20	1.004	0.004	-0.133	
1	20	0.996	0.078	-0.133	
1	20	1	-0.039	-0.125	
1	20	0.965	0.113	-0.117	
1	20	0.973	-0.027	-0.117	
1	20	1.008	-0.02	-0.121	
1	20	0.965	0.09	-0.121	
1	20	0.957	-0.047	-0.113	
1	20	1.012	-0.004	-0.125	
1	20	0.969	-0.051	-0.133	
1	20	1	0.121	-0.125	
1	20	1.016	0.023	-0.121	
1	20	0.988	-0.121	-0.156	
1	20	0.992	-0.035	-0.137	
1	20	1	0.039	-0.125	
1	20	0.992	-0.094	-0.156	
1	20	0.969	-0.086	-0.129	
1	20	1	0.051	-0.121	
1	20	0.984	0.082	-0.148	
1	20	1.031	0.102	-0.121	
1	20	0.996	0	-0.145	
1	20	0.973	-0.09	-0.125	
1	20	0.961	-0.023	-0.133	
1	20	0.988	0.023	-0.137	
1	20	0.984	-0.082	-0.121	
1	20	0.965	-0.152	-0.148	
1	20	0.992	0.008	-0.129	
1	20	0.996	-0.051	-0.121	
1	20	0.969	-0.141	-0.148	
1	20	0.996	0.094	-0.141	
1	20	1.027	-0.031	-0.133	
1	20	0.977	-0.09	-0.145	
1	20	0.996	0.109	-0.133	
1	20	1.047	-0.016	-0.113	
1	20	0.973	-0.133	-0.148	
1	20	0.988	0.055	-0.129	
1	20	1.02	0.008	-0.129	
1	20	1.004	0.148	-0.121	
1	20	1.023	0.078	-0.121	
1	20	0.996	-0.031	-0.137	

Classification

Data set used → ADULT DATA SET

Selecting a Classifier:

At the top of the classify section is the Classifier box. This box has a text field that gives the name of the currently selected classifier, and its options. Clicking on the text box brings up a GenericObjectEditor dialog box, just the same as for filters, that you can use to configure the options of the current classifier. The Choose button allows you to choose one of the classifiers that are available in WEKA.

Test Options:

The result of applying the chosen classifier will be tested according to the options that are set by clicking in the Test options box. There are four test modes:

1. Use training set. The classifier is evaluated on how well it predicts the class of the instances it was trained on.
2. Supplied test set. The classifier is evaluated on how well it predicts the class of a set of instances loaded from a file. Clicking the Set... button brings up a dialog allowing you to choose the file to test on.
3. Cross-validation. The classifier is evaluated by cross-validation, using the number of folds that are entered in the Folds text field.
4. Percentage split. The classifier is evaluated on how well it predicts a certain percentage of the data which is held out for testing. The amount of data held out depends on the value entered in the % field.

The Class Attribute:

The classifiers in WEKA are designed to be trained to predict a single ‘class’ attribute, which is the target for prediction. Some classifiers can only learn nominal classes; others can only learn numeric classes (regression problems); still others can learn both. By default, the class is taken to be the last attribute in the data. you want to train a classifier to predict a different attribute, click on the box below the Test options box to bring up a drop-down list of attributes to choose from.

Training a Classifier:

Once the classifier, test options and class have all been set, the learning process is started by clicking on the Start button. While the classifier is busy being trained, the little bird moves around. You can stop the training process at any time by clicking on the Stop button. When training is complete, several things happen. The Classifier is output area to the right of display is filled with text describing the results of training and testing. A new entry appears in the Result list box. We look at the result list below; but first we investigate the text that has been output.

The Classifier Output Text:

The text in the Classifier output area has scroll bars allowing you to browse the results. Of course, you can also resize the Explorer window to get a larger display area.

The output is split into several sections:

1. Run information. A list of information giving the learning scheme options, relation name, instances, attributes and test mode that were involved in the process.
2. Classifier model (full training set). A textual representation of the classification model that was produced on the full training data.
3. The results of the chosen test mode are broken down thus
4. Summary. A list of statistics summarizing how accurately the classifier was able to predict the true class of the instances under the chosen test mode.
5. Detailed Accuracy By Class. A more detailed per-class break down of the classifier's prediction accuracy.
6. Confusion Matrix. Shows how many instances have been assigned to each class. lements show the number of test examples whose actual class is the row and whose predicted class is the column.

Implementation in Weka:

1) Start □ Programs □ Weka
2) Click on explorer.
3) Click on open file.
4) Select dataset file and click on open.
5) Click on edit button which shows weather table on weka

Procedure for Constructing Decision Tree:

1) Open Start Programs Weka

2) Open explorer.

3) Click on open file and select adult dataset

4) Select Classifier option on the top of the Menu bar.

5) Select Choose button and click on Tree option.

6) Click on J48.

7) Click on Start button and output will be displayed on the right side of the window.

8) Select the result list and right click on result list and select Visualize Tree option.

9) Then Decision Tree will be displayed on new window.

Result

Algorithm—J-48

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: adult

Instances: 32560

Attributes: 15

39

State-gov

77516

Bachelors

13

Never-married

Adm-clerical

Not-in-family

White

Male

2174

0

40

United-States

<=50K

Test mode: evaluate on training data

=== Classifier model (full training set) ===

J48 pruned tree

```
 2174 <= 6849
|   Never-married =  Married-civ-spouse
|   |   0 <= 1762
|   |   |   13 <= 12
|   |   |   |   13 <= 8
|   |   |   |   |   2174 <= 3942:  <=50K (1606.0/149.0)
|   |   |   |   |   2174 > 3942
|   |   |   |   |   |   2174 <= 5060:  <=50K (15.0/3.0)
|   |   |   |   |   |   2174 > 5060:  >50K (11.0)
|   |   |   |   13 > 8
|   |   |   |   |   2174 <= 5060
|   |   |   |   |   |   40 <= 34:  <=50K (777.0/111.0)
|   |   |   |   |   |   40 > 34
|   |   |   |   |   |   |  39 <= 35:  <=50K (2505.0/532.0)
|   |   |   |   |   |   |  39 > 35
|   |   |   |   |   |   |   |   0 <= 1504
|   |   |   |   |   |   |   |   |   Adm-clerical =  Exec-managerial
|   |   |   |   |   |   |   |   |   |   White =  White
|   |   |   |   |   |   |   |   |   |   |   State-gov =  Self-emp-not-inc:  <=50K (89.0/29.0)
|   |   |   |   |   |   |   |   |   |   |   State-gov =  Private:  >50K (331.0/125.0)
|   |   |   |   |   |   |   |   |   |   |   State-gov =  State-gov:  <=50K (17.0/6.0)
|   |   |   |   |   |   |   |   |   |   |   State-gov =  Federal-gov:  >50K (16.0/3.0)
|   |   |   |   |   |   |   |   |   |   |   State-gov =  Local-gov:  >50K (42.0/20.0)
|   |   |   |   |   |   |   |   |   |   |   State-gov =  ?:  >50K (0.0)
|   |   |   |   |   |   |   |   |   |   |   State-gov =  Self-emp-inc:  >50K (98.0/35.0)
|   |   |   |   |   |   |   |   |   |   |   State-gov =  Without-pay:  >50K (0.0)
|   |   |   |   |   |   |   |   |   |   |   State-gov =  Never-worked:  >50K (0.0)
|   |   |   |   |   |   |   |   |   |   White =  Black
|   |   |   |   |   |   |   |   |   |   |   77516 <= 209236:  >50K (10.0)
|   |   |   |   |   |   |   |   |   |   |   77516 > 209236
|   |   |   |   |   |   |   |   |   |   |   |  39 <= 45
|   |   |   |   |   |   |   |   |   |   |   |   |   77516 <= 275703:  <=50K (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   77516 > 275703:  >50K (5.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |  39 > 45:  <=50K (5.0)
|   |   |   |   |   |   |   |   |   |   White =  Asian-Pac-Islander
|   |   |   |   |   |   |   |   |   |   |   40 <= 47:  >50K (6.0/2.0)
|   |   |   |   |   |   |   |   |   |   |   40 > 47:  <=50K (9.0/1.0)
```

```
|  |  |  |  |  |  |  |  |  |  White =  Amer-Indian-Eskimo:  <=50K (6.0)
|  |  |  |  |  |  |  |  |  |  White =  Other:  >50K (0.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Handlers-cleaners:  <=50K (125.0/29.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Prof-specialty
|  |  |  |  |  |  |  |  |  |  White =  White
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-not-inc
|  |  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Husband:  <=50K (11.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Not-in-family:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Wife:  >50K (5.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Own-child:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Unmarried:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Other-relative:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Private
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Bachelors:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  HS-grad:  <=50K (30.0/14.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  11th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Masters:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  9th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Some-college
|  |  |  |  |  |  |  |  |  |  |  |  |  77516 <= 111283:  <=50K (7.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  77516 > 111283:  >50K (56.0/23.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-acdm:  >50K (14.0/5.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-voc
|  |  |  |  |  |  |  |  |  |  |  |  |  77516 <= 249585:  >50K (14.0/3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  77516 > 249585:  <=50K (7.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  7th-8th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Doctorate:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Prof-school:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  5th-6th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  10th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  1st-4th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Preschool:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  12th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  State-gov
|  |  |  |  |  |  |  |  |  |  |  |  77516 <= 140854:  <=50K (3.0)
|  |  |  |  |  |  |  |  |  |  |  |  77516 > 140854:  >50K (12.0/3.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Federal-gov:  >50K (14.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Local-gov:  <=50K (5.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  ?:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-inc
|  |  |  |  |  |  |  |  |  |  |  |  39 <= 53:  >50K (5.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  39 > 53:  <=50K (2.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Without-pay:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Never-worked:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  White =  Black
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Husband:  >50K (5.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Not-in-family:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Wife:  <=50K (4.0)
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Own-child:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Unmarried:  <=50K (0.0)
```

```
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Other-relative:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  White =  Asian-Pac-Islander:  >50K (4.0/1.0)
|  |  |  |  |  |  |  |  |  |  White =  Amer-Indian-Eskimo:  <=50K (4.0/1.0)
|  |  |  |  |  |  |  |  |  |  White =  Other:  >50K (1.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Other-service:  <=50K (212.0/37.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Adm-clerical
|  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-not-inc:  >50K (4.0/1.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Private
|  |  |  |  |  |  |  |  |  |  |  Male =  Male
|  |  |  |  |  |  |  |  |  |  |  |  40 <= 53
|  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 43:  <=50K (106.0/31.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  40 > 43
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Bachelors:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  HS-grad
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  77516 <= 163847:  <=50K (10.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  77516 > 163847:  >50K (7.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  11th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Masters:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  9th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Some-college
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  39 <= 39:  <=50K (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  39 > 39:  >50K (8.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-acdm:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-voc:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  7th-8th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Doctorate:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Prof-school:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  5th-6th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  10th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  1st-4th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Preschool:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  12th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  40 > 53:  >50K (8.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  Male =  Female
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Bachelors:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  HS-grad
|  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 52
|  |  |  |  |  |  |  |  |  |  |  |  |  |  39 <= 50:  >50K (37.0/12.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  39 > 50:  <=50K (15.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  40 > 52:  <=50K (4.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  11th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Masters:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  9th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Some-college
|  |  |  |  |  |  |  |  |  |  |  |  |  77516 <= 156526
|  |  |  |  |  |  |  |  |  |  |  |  |  |  39 <= 44:  <=50K (5.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  39 > 44
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  39 <= 46:  >50K (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  39 > 46:  <=50K (7.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  77516 > 156526:  >50K (21.0/4.0)
```

```
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-acdm:  >50K (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-voc:  >50K (3.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  7th-8th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Doctorate:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Prof-school:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  5th-6th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  10th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  1st-4th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Preschool:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  12th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  State-gov
|  |  |  |  |  |  |  |  |  |  |  |  39 <= 45:  <=50K (13.0/3.0)
|  |  |  |  |  |  |  |  |  |  |  |  39 > 45
|  |  |  |  |  |  |  |  |  |  |  |  |  39 <= 49:  >50K (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  39 > 49:  <=50K (5.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Federal-gov
|  |  |  |  |  |  |  |  |  |  |  |  40 <= 39:  <=50K (4.0)
|  |  |  |  |  |  |  |  |  |  |  |  40 > 39:  >50K (81.0/28.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Local-gov:  <=50K (27.0/11.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  ?:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-inc:  >50K (4.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Without-pay:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Never-worked:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  Adm-clerical =  Sales
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-not-inc
|  |  |  |  |  |  |  |  |  |  |  |  13 <= 10
|  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 54:  <=50K (60.0/10.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  40 > 54
|  |  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 70
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  77516 <= 126513:  >50K (7.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  77516 > 126513:  <=50K (18.0/5.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  40 > 70:  >50K (6.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  13 > 10
|  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 55:  >50K (4.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  40 > 55:  <=50K (3.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  State-gov =  Private
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Bachelors:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  HS-grad:  <=50K (175.0/67.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  11th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Masters:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  9th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Some-college
|  |  |  |  |  |  |  |  |  |  |  |  |  2174 <= 1409:  >50K (146.0/63.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  2174 > 1409
|  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 <= 3103:  >50K (4.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 > 3103:  <=50K (5.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-acdm
|  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 53
|  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 <= 1506:  <=50K (11.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 > 1506:  >50K (3.0/1.0)
```

```
|  |  |  |  |  |  |  |  |  |  |  |  40 > 53:  >50K (5.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-voc:  <=50K (19.0/9.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  7th-8th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Doctorate:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Prof-school:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  5th-6th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  10th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  1st-4th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Preschool:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  12th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  State-gov:  <=50K (2.0/1.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Federal-gov:  >50K (1.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Local-gov:  <=50K (1.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  ?:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-inc:  >50K (71.0/28.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Without-pay:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Never-worked:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Craft-repair
|  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-not-inc
|  |  |  |  |  |  |  |  |  |  |  39 <= 56:  <=50K (101.0/19.0)
|  |  |  |  |  |  |  |  |  |  |  39 > 56
|  |  |  |  |  |  |  |  |  |  |  |  39 <= 62:  >50K (12.0/4.0)
|  |  |  |  |  |  |  |  |  |  |  |  39 > 62:  <=50K (4.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Private
|  |  |  |  |  |  |  |  |  |  |  13 <= 9:  <=50K (515.0/182.0)
|  |  |  |  |  |  |  |  |  |  |  13 > 9
|  |  |  |  |  |  |  |  |  |  |  |  39 <= 48
|  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 47
|  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 <= 1506:  <=50K (170.0/57.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 > 1506
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  39 <= 42:  >50K (4.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  39 > 42
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 <= 4101:  <=50K (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 > 4101:  >50K (3.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  40 > 47:  >50K (54.0/20.0)
|  |  |  |  |  |  |  |  |  |  |  |  39 > 48
|  |  |  |  |  |  |  |  |  |  |  |  |  39 <= 59
|  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 <= 3942
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 52:  >50K (80.0/23.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  40 > 52:  <=50K (4.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  2174 > 3942:  <=50K (5.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  39 > 59:  <=50K (13.0/2.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  State-gov:  <=50K (22.0/6.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Federal-gov
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Bachelors:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  HS-grad:  >50K (12.0/3.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  11th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Masters:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  9th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Some-college:  <=50K (11.0/2.0)
```

```
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Assoc-acdm: >50K (2.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Assoc-voc: >50K (5.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 7th-8th: >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Doctorate: >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Prof-school: >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 5th-6th: >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 10th: >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 1st-4th: >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Preschool: >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 12th: >50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov = Local-gov: <=50K (47.0/18.0)
|  |  |  |  |  |  |  |  |  |  State-gov = ?: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov = Self-emp-inc
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Bachelors: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = HS-grad
|  |  |  |  |  |  |  |  |  |  |  |  77516 <= 116165: <=50K (7.0)
|  |  |  |  |  |  |  |  |  |  |  |  77516 > 116165
|  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 49
|  |  |  |  |  |  |  |  |  |  |  |  |  |  39 <= 51: <=50K (6.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  39 > 51: >50K (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  40 > 49: >50K (4.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 11th: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Masters: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 9th: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Some-college
|  |  |  |  |  |  |  |  |  |  |  |  40 <= 42: <=50K (4.0)
|  |  |  |  |  |  |  |  |  |  |  |  40 > 42: >50K (7.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Assoc-acdm: <=50K (2.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Assoc-voc: <=50K (2.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 7th-8th: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Doctorate: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Prof-school: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 5th-6th: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 10th: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 1st-4th: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = Preschool: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors = 12th: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov = Without-pay: <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov = Never-worked: <=50K (0.0)
|  |  |  |  |  |  |  |  |  Adm-clerical = Transport-moving
|  |  |  |  |  |  |  |  |  |  State-gov = Self-emp-not-inc: <=50K (40.0/14.0)
|  |  |  |  |  |  |  |  |  |  State-gov = Private: <=50K (289.0/97.0)
|  |  |  |  |  |  |  |  |  |  State-gov = State-gov
|  |  |  |  |  |  |  |  |  |  |  40 <= 39: >50K (2.0)
|  |  |  |  |  |  |  |  |  |  |  40 > 39: <=50K (8.0)
|  |  |  |  |  |  |  |  |  |  State-gov = Federal-gov
|  |  |  |  |  |  |  |  |  |  |  77516 <= 177499: <=50K (5.0)
|  |  |  |  |  |  |  |  |  |  |  77516 > 177499: >50K (7.0/1.0)
|  |  |  |  |  |  |  |  |  |  State-gov = Local-gov: <=50K (37.0/6.0)
|  |  |  |  |  |  |  |  |  |  State-gov = ?: <=50K (0.0)
```

```
|  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-inc
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Bachelors:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  HS-grad:  <=50K (5.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  11th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Masters:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  9th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Some-college:  >50K (5.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-acdm:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-voc:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  7th-8th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Doctorate:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Prof-school:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  5th-6th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  10th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  1st-4th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Preschool:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  12th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Without-pay:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Never-worked:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Farming-fishing:  <=50K (224.0/39.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Machine-op-inspct
|  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-not-inc
|  |  |  |  |  |  |  |  |  |  |  13 <= 9:  <=50K (3.0)
|  |  |  |  |  |  |  |  |  |  |  13 > 9:  >50K (4.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Private
|  |  |  |  |  |  |  |  |  |  |  2174 <= 2829:  <=50K (336.0/103.0)
|  |  |  |  |  |  |  |  |  |  |  2174 > 2829
|  |  |  |  |  |  |  |  |  |  |  |  2174 <= 3908:  <=50K (19.0/5.0)
|  |  |  |  |  |  |  |  |  |  |  |  2174 > 3908:  >50K (8.0/1.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  State-gov:  >50K (4.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Federal-gov:  <=50K (4.0/1.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Local-gov
|  |  |  |  |  |  |  |  |  |  |  77516 <= 205706:  <=50K (3.0)
|  |  |  |  |  |  |  |  |  |  |  77516 > 205706:  >50K (3.0/1.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  ?:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-inc:  <=50K (3.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Without-pay:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Never-worked:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Tech-support
|  |  |  |  |  |  |  |  |  |  77516 <= 161857
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Bachelors:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  HS-grad
|  |  |  |  |  |  |  |  |  |  |  |  39 <= 49:  <=50K (13.0/3.0)
|  |  |  |  |  |  |  |  |  |  |  |  39 > 49:  >50K (8.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  11th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Masters:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  9th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Some-college:  >50K (31.0/11.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-acdm:  <=50K (11.0/4.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-voc
```

```
|  |  |  |  |  |  |  |  |  |  |  |  77516 <= 72844:  <=50K (3.0)
|  |  |  |  |  |  |  |  |  |  |  |  77516 > 72844:  >50K (10.0/4.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  7th-8th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Doctorate:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Prof-school:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  5th-6th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  10th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  1st-4th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  Preschool:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Bachelors =  12th:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  77516 > 161857:  >50K (73.0/18.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  ?
|  |  |  |  |  |  |  |  |  |  13 <= 10
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Husband
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Bachelors:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  HS-grad:  <=50K (98.0/27.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  11th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Masters:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  9th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Some-college
|  |  |  |  |  |  |  |  |  |  |  |  |  40 <= 49
|  |  |  |  |  |  |  |  |  |  |  |  |  |  77516 <= 110445:  <=50K (6.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  77516 > 110445
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  77516 <= 263024:  >50K (7.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  77516 > 263024:  <=50K (5.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  40 > 49:  <=50K (7.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-acdm:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Assoc-voc:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  7th-8th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Doctorate:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Prof-school:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  5th-6th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  10th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  1st-4th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  Preschool:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  |  Bachelors =  12th:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Not-in-family:  <=50K (1.0)
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Wife
|  |  |  |  |  |  |  |  |  |  |  |  77516 <= 179973:  <=50K (5.0)
|  |  |  |  |  |  |  |  |  |  |  |  77516 > 179973:  >50K (6.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Own-child:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Unmarried:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  Not-in-family =  Other-relative:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  13 > 10:  >50K (11.0/3.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Protective-serv
|  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-not-inc:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Private
|  |  |  |  |  |  |  |  |  |  |  White =  White:  <=50K (27.0/7.0)
|  |  |  |  |  |  |  |  |  |  |  White =  Black
|  |  |  |  |  |  |  |  |  |  |  |  77516 <= 98361:  <=50K (2.0)
```

```
|  |  |  |  |  |  |  |  |  |  |  |  77516 > 98361:  >50K (7.0/2.0)
|  |  |  |  |  |  |  |  |  |  |  White =  Asian-Pac-Islander:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  White =  Amer-Indian-Eskimo:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  White =  Other:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  State-gov
|  |  |  |  |  |  |  |  |  |  |  White =  White:  <=50K (28.0/7.0)
|  |  |  |  |  |  |  |  |  |  |  White =  Black:  >50K (3.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  White =  Asian-Pac-Islander:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  |  White =  Amer-Indian-Eskimo:  >50K (1.0)
|  |  |  |  |  |  |  |  |  |  |  White =  Other:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Federal-gov:  <=50K (4.0/2.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Local-gov
|  |  |  |  |  |  |  |  |  |  |  39 <= 56:  >50K (73.0/21.0)
|  |  |  |  |  |  |  |  |  |  |  39 > 56:  <=50K (5.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  ?:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Self-emp-inc:  <=50K (2.0/1.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Without-pay:  >50K (0.0)
|  |  |  |  |  |  |  |  |  |  State-gov =  Never-worked:  >50K (0.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Armed-Forces:  <=50K (0.0)
|  |  |  |  |  |  |  |  |  Adm-clerical =  Priv-house-serv:  <=50K (5.0)
|  |  |  |  |  |  |  |  0 > 1504:  <=50K (51.0)
|  |  |  |  |  2174 > 5060
|  |  |  |  |  |  2174 <= 6514:  >50K (66.0)
|  |  |  |  |  |  2174 > 6514:  <=50K (4.0)
|  |  |  13 > 12
|  |  |  |  40 <= 31
|  |  |  |  |  Not-in-family =  Husband:  <=50K (232.0/70.0)
|  |  |  |  |  Not-in-family =  Not-in-family:  <=50K (1.0)
|  |  |  |  |  Not-in-family =  Wife
|  |  |  |  |  |  United-States =  United-States:  >50K (67.0/24.0)
|  |  |  |  |  |  United-States =  Cuba:  >50K (1.0)
|  |  |  |  |  |  United-States =  Jamaica:  >50K (0.0)
|  |  |  |  |  |  United-States =  India:  >50K (0.0)
|  |  |  |  |  |  United-States =  ?:  <=50K (2.0)
|  |  |  |  |  |  United-States =  Mexico:  >50K (0.0)
|  |  |  |  |  |  United-States =  South:  >50K (0.0)
|  |  |  |  |  |  United-States =  Puerto-Rico:  >50K (0.0)
|  |  |  |  |  |  United-States =  Honduras:  >50K (0.0)
|  |  |  |  |  |  United-States =  England:  >50K (0.0)
|  |  |  |  |  |  United-States =  Canada:  <=50K (1.0)
|  |  |  |  |  |  United-States =  Germany:  >50K (0.0)
|  |  |  |  |  |  United-States =  Iran:  >50K (0.0)
|  |  |  |  |  |  United-States =  Philippines:  >50K (0.0)
|  |  |  |  |  |  United-States =  Italy:  >50K (0.0)
|  |  |  |  |  |  United-States =  Poland:  <=50K (1.0)
|  |  |  |  |  |  United-States =  Columbia:  >50K (0.0)
|  |  |  |  |  |  United-States =  Cambodia:  >50K (0.0)
|  |  |  |  |  |  United-States =  Thailand:  >50K (0.0)
|  |  |  |  |  |  United-States =  Ecuador:  >50K (0.0)
|  |  |  |  |  |  United-States =  Laos:  >50K (0.0)
```

```
|  |  |  |  |  |  United-States = Taiwan: >50K (1.0)
|  |  |  |  |  |  United-States = Haiti: >50K (0.0)
|  |  |  |  |  |  United-States = Portugal: >50K (0.0)
|  |  |  |  |  |  United-States = Dominican-Republic: >50K (0.0)
|  |  |  |  |  |  United-States = El-Salvador: >50K (0.0)
|  |  |  |  |  |  United-States = France: >50K (0.0)
|  |  |  |  |  |  United-States = Guatemala: >50K (0.0)
|  |  |  |  |  |  United-States = China: >50K (0.0)
|  |  |  |  |  |  United-States = Japan: >50K (0.0)
|  |  |  |  |  |  United-States = Yugoslavia: >50K (0.0)
|  |  |  |  |  |  United-States = Peru: >50K (0.0)
|  |  |  |  |  |  United-States = Outlying-US(Guam-USVI-etc): >50K (0.0)
|  |  |  |  |  |  United-States = Scotland: >50K (0.0)
|  |  |  |  |  |  United-States = Trinadad&Tobago: >50K (0.0)
|  |  |  |  |  |  United-States = Greece: >50K (0.0)
|  |  |  |  |  |  United-States = Nicaragua: >50K (0.0)
|  |  |  |  |  |  United-States = Vietnam: >50K (0.0)
|  |  |  |  |  |  United-States = Hong: >50K (0.0)
|  |  |  |  |  |  United-States = Ireland: >50K (0.0)
|  |  |  |  |  |  United-States = Hungary: >50K (0.0)
|  |  |  |  |  |  United-States = Holand-Netherlands: >50K (0.0)
|  |  |  |  |  Not-in-family = Own-child: <=50K (3.0/1.0)
|  |  |  |  |  Not-in-family = Unmarried: <=50K (0.0)
|  |  |  |  |  Not-in-family = Other-relative: <=50K (2.0)
|  |  |  |  40 > 31
|  |  |  |  |  39 <= 28
|  |  |  |  |  |  39 <= 25: <=50K (68.0/17.0)
|  |  |  |  |  |  39 > 25
|  |  |  |  |  |  |  Male = Male: <=50K (117.0/53.0)
|  |  |  |  |  |  |  Male = Female: >50K (30.0/7.0)
|  |  |  |  |  39 > 28
|  |  |  |  |  |  Adm-clerical = Exec-managerial: >50K (827.0/188.0)
|  |  |  |  |  |  Adm-clerical = Handlers-cleaners
|  |  |  |  |  |  |  2174 <= 2176: <=50K (18.0/6.0)
|  |  |  |  |  |  |  2174 > 2176: >50K (2.0)
|  |  |  |  |  |  Adm-clerical = Prof-specialty
|  |  |  |  |  |  |  Not-in-family = Husband: >50K (962.0/262.0)
|  |  |  |  |  |  |  Not-in-family = Not-in-family: <=50K (2.0)
|  |  |  |  |  |  |  Not-in-family = Wife: >50K (121.0/29.0)
|  |  |  |  |  |  |  Not-in-family = Own-child: <=50K (2.0)
|  |  |  |  |  |  |  Not-in-family = Unmarried: >50K (0.0)
|  |  |  |  |  |  |  Not-in-family = Other-relative: <=50K (4.0)
|  |  |  |  |  |  Adm-clerical = Other-service
|  |  |  |  |  |  |  40 <= 52: <=50K (33.0/6.0)
|  |  |  |  |  |  |  40 > 52: >50K (3.0/1.0)
|  |  |  |  |  |  Adm-clerical = Adm-clerical
|  |  |  |  |  |  |  United-States = United-States
|  |  |  |  |  |  |  |  39 <= 61: >50K (104.0/41.0)
|  |  |  |  |  |  |  |  39 > 61: <=50K (6.0/1.0)
|  |  |  |  |  |  |  United-States = Cuba: >50K (0.0)
```

```
|  |  |  |  |  |  |  United-States = Jamaica: >50K (0.0)
|  |  |  |  |  |  |  United-States = India: <=50K (3.0)
|  |  |  |  |  |  |  United-States = ?: >50K (5.0/1.0)
|  |  |  |  |  |  |  United-States = Mexico: <=50K (3.0)
|  |  |  |  |  |  |  United-States = South: <=50K (2.0/1.0)
|  |  |  |  |  |  |  United-States = Puerto-Rico: >50K (0.0)
|  |  |  |  |  |  |  United-States = Honduras: >50K (0.0)
|  |  |  |  |  |  |  United-States = England: <=50K (1.0)
|  |  |  |  |  |  |  United-States = Canada: >50K (0.0)
|  |  |  |  |  |  |  United-States = Germany: >50K (2.0)
|  |  |  |  |  |  |  United-States = Iran: >50K (1.0)
|  |  |  |  |  |  |  United-States = Philippines
|  |  |  |  |  |  |  |  Bachelors = Bachelors
|  |  |  |  |  |  |  |  |  39 <= 44: >50K (9.0/1.0)
|  |  |  |  |  |  |  |  |  39 > 44: <=50K (3.0)
|  |  |  |  |  |  |  |  Bachelors = HS-grad: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 11th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Masters: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 9th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Some-college: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Assoc-acdm: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Assoc-voc: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 7th-8th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Doctorate: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Prof-school: <=50K (3.0/1.0)
|  |  |  |  |  |  |  |  Bachelors = 5th-6th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 10th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 1st-4th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Preschool: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 12th: >50K (0.0)
|  |  |  |  |  |  |  United-States = Italy: >50K (2.0)
|  |  |  |  |  |  |  United-States = Poland: >50K (0.0)
|  |  |  |  |  |  |  United-States = Columbia: >50K (0.0)
|  |  |  |  |  |  |  United-States = Cambodia: >50K (0.0)
|  |  |  |  |  |  |  United-States = Thailand: >50K (0.0)
|  |  |  |  |  |  |  United-States = Ecuador: >50K (0.0)
|  |  |  |  |  |  |  United-States = Laos: >50K (1.0)
|  |  |  |  |  |  |  United-States = Taiwan: >50K (0.0)
|  |  |  |  |  |  |  United-States = Haiti: >50K (0.0)
|  |  |  |  |  |  |  United-States = Portugal: >50K (0.0)
|  |  |  |  |  |  |  United-States = Dominican-Republic: >50K (0.0)
|  |  |  |  |  |  |  United-States = El-Salvador: >50K (0.0)
|  |  |  |  |  |  |  United-States = France: >50K (0.0)
|  |  |  |  |  |  |  United-States = Guatemala: >50K (0.0)
|  |  |  |  |  |  |  United-States = China: <=50K (1.0)
|  |  |  |  |  |  |  United-States = Japan: >50K (0.0)
|  |  |  |  |  |  |  United-States = Yugoslavia: >50K (0.0)
|  |  |  |  |  |  |  United-States = Peru: >50K (0.0)
|  |  |  |  |  |  |  United-States = Outlying-US(Guam-USVI-etc): >50K (0.0)
|  |  |  |  |  |  |  United-States = Scotland: >50K (0.0)
```

```
|  |  |  |  |  |  |  United-States = Trinadad&Tobago: >50K (0.0)
|  |  |  |  |  |  |  United-States = Greece: >50K (0.0)
|  |  |  |  |  |  |  United-States = Nicaragua: >50K (0.0)
|  |  |  |  |  |  |  United-States = Vietnam: <=50K (1.0)
|  |  |  |  |  |  |  United-States = Hong: >50K (0.0)
|  |  |  |  |  |  |  United-States = Ireland: >50K (0.0)
|  |  |  |  |  |  |  United-States = Hungary: >50K (0.0)
|  |  |  |  |  |  |  United-States = Holand-Netherlands: >50K (0.0)
|  |  |  |  |  |  Adm-clerical = Sales
|  |  |  |  |  |  |  State-gov = Self-emp-not-inc
|  |  |  |  |  |  |  |  40 <= 46: >50K (23.0/7.0)
|  |  |  |  |  |  |  |  40 > 46: <=50K (33.0/10.0)
|  |  |  |  |  |  |  State-gov = Private
|  |  |  |  |  |  |  |  Bachelors = Bachelors: >50K (209.0/68.0)
|  |  |  |  |  |  |  |  Bachelors = HS-grad: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 11th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Masters
|  |  |  |  |  |  |  |  |  39 <= 56
|  |  |  |  |  |  |  |  |  |  40 <= 47
|  |  |  |  |  |  |  |  |  |  |  39 <= 51
|  |  |  |  |  |  |  |  |  |  |  |  77516 <= 163516: <=50K (5.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  77516 > 163516: >50K (9.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  39 > 51: <=50K (2.0)
|  |  |  |  |  |  |  |  |  |  40 > 47: >50K (9.0)
|  |  |  |  |  |  |  |  |  39 > 56: <=50K (5.0/1.0)
|  |  |  |  |  |  |  |  Bachelors = 9th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Some-college: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Assoc-acdm: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Assoc-voc: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 7th-8th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Doctorate: <=50K (2.0)
|  |  |  |  |  |  |  |  Bachelors = Prof-school: <=50K (6.0/2.0)
|  |  |  |  |  |  |  |  Bachelors = 5th-6th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 10th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 1st-4th: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Preschool: >50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 12th: >50K (0.0)
|  |  |  |  |  |  |  State-gov = State-gov: <=50K (1.0)
|  |  |  |  |  |  |  State-gov = Federal-gov: <=50K (3.0/1.0)
|  |  |  |  |  |  |  State-gov = Local-gov: >50K (1.0)
|  |  |  |  |  |  |  State-gov = ?: >50K (0.0)
|  |  |  |  |  |  |  State-gov = Self-emp-inc: >50K (62.0/16.0)
|  |  |  |  |  |  |  State-gov = Without-pay: >50K (0.0)
|  |  |  |  |  |  |  State-gov = Never-worked: >50K (0.0)
|  |  |  |  |  |  Adm-clerical = Craft-repair
|  |  |  |  |  |  |  State-gov = Self-emp-not-inc
|  |  |  |  |  |  |  |  Bachelors = Bachelors
|  |  |  |  |  |  |  |  |  40 <= 52: <=50K (16.0/4.0)
|  |  |  |  |  |  |  |  |  40 > 52: >50K (2.0)
|  |  |  |  |  |  |  Bachelors = HS-grad: <=50K (0.0)
```

```
|  |  |  |  |  |  |  |  Bachelors = 11th: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Masters: >50K (5.0/1.0)
|  |  |  |  |  |  |  |  Bachelors = 9th: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Some-college: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Assoc-acdm: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Assoc-voc: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 7th-8th: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Doctorate: <=50K (1.0)
|  |  |  |  |  |  |  |  Bachelors = Prof-school: <=50K (2.0)
|  |  |  |  |  |  |  |  Bachelors = 5th-6th: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 10th: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 1st-4th: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = Preschool: <=50K (0.0)
|  |  |  |  |  |  |  |  Bachelors = 12th: <=50K (0.0)
|  |  |  |  |  |  |  State-gov = Private: >50K (83.0/34.0)
|  |  |  |  |  |  |  State-gov = State-gov: >50K (1.0)
|  |  |  |  |  |  |  State-gov = Federal-gov: <=50K (3.0/1.0)
|  |  |  |  |  |  |  State-gov = Local-gov
|  |  |  |  |  |  |  |  39 <= 44: <=50K (6.0)
|  |  |  |  |  |  |  |  39 > 44: >50K (2.0)
|  |  |  |  |  |  |  State-gov = ?: >50K (0.0)
|  |  |  |  |  |  |  State-gov = Self-emp-inc: <=50K (10.0/3.0)
|  |  |  |  |  |  |  State-gov = Without-pay: >50K (0.0)
|  |  |  |  |  |  |  State-gov = Never-worked: >50K (0.0)
|  |  |  |  |  |  Adm-clerical = Transport-moving
|  |  |  |  |  |  |  White = White
|  |  |  |  |  |  |  |  39 <= 44: <=50K (14.0/3.0)
|  |  |  |  |  |  |  |  39 > 44
|  |  |  |  |  |  |  |  |  39 <= 52: >50K (8.0/1.0)
|  |  |  |  |  |  |  |  |  39 > 52: <=50K (5.0)
|  |  |  |  |  |  |  White = Black: <=50K (3.0)
|  |  |  |  |  |  |  White = Asian-Pac-Islander: >50K (1.0)
|  |  |  |  |  |  |  White = Amer-Indian-Eskimo: <=50K (0.0)
|  |  |  |  |  |  |  White = Other: <=50K (0.0)
|  |  |  |  |  |  Adm-clerical = Farming-fishing
|  |  |  |  |  |  |  77516 <= 34574: >50K (8.0/1.0)
|  |  |  |  |  |  |  77516 > 34574: <=50K (40.0/10.0)
|  |  |  |  |  |  Adm-clerical = Machine-op-inspct: <=50K (32.0/11.0)
|  |  |  |  |  |  Adm-clerical = Tech-support
|  |  |  |  |  |  |  United-States = United-States: >50K (66.0/18.0)
|  |  |  |  |  |  |  United-States = Cuba: >50K (0.0)
|  |  |  |  |  |  |  United-States = Jamaica: >50K (0.0)
|  |  |  |  |  |  |  United-States = India: <=50K (1.0)
|  |  |  |  |  |  |  United-States = ?: <=50K (1.0)
|  |  |  |  |  |  |  United-States = Mexico: >50K (0.0)
|  |  |  |  |  |  |  United-States = South: >50K (0.0)
|  |  |  |  |  |  |  United-States = Puerto-Rico: >50K (0.0)
|  |  |  |  |  |  |  United-States = Honduras: >50K (0.0)
|  |  |  |  |  |  |  United-States = England: <=50K (1.0)
|  |  |  |  |  |  |  United-States = Canada: >50K (0.0)
```

```
|  |  |  |  |  |  |  United-States = Germany: >50K (0.0)
|  |  |  |  |  |  |  United-States = Iran: <=50K (1.0)
|  |  |  |  |  |  |  United-States = Philippines: >50K (2.0)
|  |  |  |  |  |  |  United-States = Italy: >50K (0.0)
|  |  |  |  |  |  |  United-States = Poland: >50K (0.0)
|  |  |  |  |  |  |  United-States = Columbia: >50K (0.0)
|  |  |  |  |  |  |  United-States = Cambodia: >50K (0.0)
|  |  |  |  |  |  |  United-States = Thailand: >50K (0.0)
|  |  |  |  |  |  |  United-States = Ecuador: >50K (0.0)
|  |  |  |  |  |  |  United-States = Laos: >50K (0.0)
|  |  |  |  |  |  |  United-States = Taiwan: >50K (0.0)
|  |  |  |  |  |  |  United-States = Haiti: >50K (0.0)
|  |  |  |  |  |  |  United-States = Portugal: >50K (0.0)
|  |  |  |  |  |  |  United-States = Dominican-Republic: >50K (0.0)
|  |  |  |  |  |  |  United-States = El-Salvador: >50K (0.0)
|  |  |  |  |  |  |  United-States = France: >50K (0.0)
|  |  |  |  |  |  |  United-States = Guatemala: >50K (0.0)
|  |  |  |  |  |  |  United-States = China: >50K (0.0)
|  |  |  |  |  |  |  United-States = Japan: >50K (0.0)
|  |  |  |  |  |  |  United-States = Yugoslavia: >50K (0.0)
|  |  |  |  |  |  |  United-States = Peru: >50K (0.0)
|  |  |  |  |  |  |  United-States = Outlying-US(Guam-USVI-etc): >50K (0.0)
|  |  |  |  |  |  |  United-States = Scotland: >50K (0.0)
|  |  |  |  |  |  |  United-States = Trinadad&Tobago: >50K (0.0)
|  |  |  |  |  |  |  United-States = Greece: >50K (0.0)
|  |  |  |  |  |  |  United-States = Nicaragua: >50K (0.0)
|  |  |  |  |  |  |  United-States = Vietnam: >50K (0.0)
|  |  |  |  |  |  |  United-States = Hong: >50K (0.0)
|  |  |  |  |  |  |  United-States = Ireland: >50K (0.0)
|  |  |  |  |  |  |  United-States = Hungary: >50K (0.0)
|  |  |  |  |  |  |  United-States = Holand-Netherlands: >50K (0.0)
|  |  |  |  |  |  Adm-clerical = ?
|  |  |  |  |  |  |  40 <= 43
|  |  |  |  |  |  |  |  40 <= 38: >50K (2.0)
|  |  |  |  |  |  |  |  40 > 38
|  |  |  |  |  |  |  |  |  77516 <= 369909: <=50K (27.0/7.0)
|  |  |  |  |  |  |  |  |  77516 > 369909: >50K (3.0)
|  |  |  |  |  |  |  40 > 43: >50K (19.0/7.0)
|  |  |  |  |  |  Adm-clerical = Protective-serv: >50K (47.0/12.0)
|  |  |  |  |  |  Adm-clerical = Armed-Forces: >50K (0.0)
|  |  |  |  |  |  Adm-clerical = Priv-house-serv: <=50K (1.0)
|  |  0 > 1762
|  |  |  0 <= 1980: >50K (585.0/14.0)
|  |  |  0 > 1980
|  |  |  |  0 <= 2163: <=50K (63.0)
|  |  |  |  0 > 2163
|  |  |  |  |  13 <= 12
|  |  |  |  |  |  0 <= 2174: >50K (5.0)
|  |  |  |  |  |  0 > 2174
|  |  |  |  |  |  |  0 <= 2206: <=50K (14.0)
```

```
|  |  |  |  |  |  |  0 > 2206
|  |  |  |  |  |  |  |  0 <= 2415
|  |  |  |  |  |  |  |  |  0 <= 2392
|  |  |  |  |  |  |  |  |  |  0 <= 2246:  >50K (2.0)
|  |  |  |  |  |  |  |  |  |  0 > 2246
|  |  |  |  |  |  |  |  |  |  |  39 <= 67:  <=50K (13.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  39 > 67:  >50K (4.0/1.0)
|  |  |  |  |  |  |  |  |  0 > 2392:  >50K (10.0)
|  |  |  |  |  |  |  |  0 > 2415:  <=50K (10.0)
|  |  |  |  |  13 > 12:  >50K (62.0/2.0)
|  Never-married =  Divorced
|  |  13 <= 12:  <=50K (3526.0/167.0)
|  |  13 > 12
|  |  |  40 <= 43
|  |  |  |  0 <= 1668:  <=50K (455.0/52.0)
|  |  |  |  0 > 1668
|  |  |  |  |  0 <= 2231:  <=50K (12.0/1.0)
|  |  |  |  |  0 > 2231:  >50K (9.0/2.0)
|  |  |  40 > 43
|  |  |  |  Adm-clerical =  Exec-managerial
|  |  |  |  |  0 <= 880
|  |  |  |  |  |  Bachelors =  Bachelors:  <=50K (63.0/28.0)
|  |  |  |  |  |  Bachelors =  HS-grad:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  11th:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  Masters
|  |  |  |  |  |  |  0 <= 0
|  |  |  |  |  |  |  |  77516 <= 118462:  <=50K (5.0/1.0)
|  |  |  |  |  |  |  |  77516 > 118462:  >50K (22.0/6.0)
|  |  |  |  |  |  |  0 > 0:  <=50K (2.0)
|  |  |  |  |  |  Bachelors =  9th:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  Some-college:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  Assoc-acdm:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  Assoc-voc:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  7th-8th:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  Doctorate:  >50K (2.0)
|  |  |  |  |  |  Bachelors =  Prof-school:  <=50K (2.0/1.0)
|  |  |  |  |  |  Bachelors =  5th-6th:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  10th:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  1st-4th:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  Preschool:  <=50K (0.0)
|  |  |  |  |  |  Bachelors =  12th:  <=50K (0.0)
|  |  |  |  |  |  0 > 880:  >50K (8.0)
|  |  |  |  Adm-clerical =  Handlers-cleaners:  >50K (1.0)
|  |  |  |  Adm-clerical =  Prof-specialty
|  |  |  |  |  State-gov =  Self-emp-not-inc
|  |  |  |  |  |  13 <= 14
|  |  |  |  |  |  |  77516 <= 90338:  >50K (3.0)
|  |  |  |  |  |  |  77516 > 90338:  <=50K (6.0/1.0)
|  |  |  |  |  |  13 > 14:  >50K (4.0)
|  |  |  |  |  State-gov =  Private:  <=50K (49.0/16.0)
```

```
|  |  |  |  |  State-gov =  State-gov
|  |  |  |  |  |  77516 <= 230657:  <=50K (7.0/1.0)
|  |  |  |  |  |  77516 > 230657:  >50K (3.0)
|  |  |  |  |  State-gov =  Federal-gov:  >50K (4.0/1.0)
|  |  |  |  |  State-gov =  Local-gov:  <=50K (40.0/2.0)
|  |  |  |  |  State-gov =  ?:  <=50K (0.0)
|  |  |  |  |  State-gov =  Self-emp-inc:  <=50K (4.0/1.0)
|  |  |  |  |  State-gov =  Without-pay:  <=50K (0.0)
|  |  |  |  |  State-gov =  Never-worked:  <=50K (0.0)
|  |  |  |  Adm-clerical =  Other-service:  >50K (7.0/2.0)
|  |  |  |  Adm-clerical =  Adm-clerical:  <=50K (9.0/1.0)
|  |  |  |  Adm-clerical =  Sales:  <=50K (49.0/17.0)
|  |  |  |  Adm-clerical =  Craft-repair
|  |  |  |  |  77516 <= 143833:  >50K (3.0/1.0)
|  |  |  |  |  77516 > 143833:  <=50K (5.0)
|  |  |  |  Adm-clerical =  Transport-moving:  <=50K (3.0)
|  |  |  |  Adm-clerical =  Farming-fishing:  <=50K (3.0)
|  |  |  |  Adm-clerical =  Machine-op-inspct:  <=50K (0.0)
|  |  |  |  Adm-clerical =  Tech-support:  <=50K (0.0)
|  |  |  |  Adm-clerical =  ?:  <=50K (7.0)
|  |  |  |  Adm-clerical =  Protective-serv
|  |  |  |  |  White =  White:  <=50K (6.0/2.0)
|  |  |  |  |  White =  Black:  >50K (3.0)
|  |  |  |  |  White =  Asian-Pac-Islander:  >50K (0.0)
|  |  |  |  |  White =  Amer-Indian-Eskimo:  >50K (0.0)
|  |  |  |  |  White =  Other:  >50K (0.0)
|  |  |  |  Adm-clerical =  Armed-Forces:  <=50K (0.0)
|  |  |  |  Adm-clerical =  Priv-house-serv:  <=50K (1.0)
|  Never-married =  Married-spouse-absent:  <=50K (410.0/26.0)
|  Never-married =  Never-married
|  |  13 <= 12
|  |  |  0 <= 2080:  <=50K (8238.0/88.0)
|  |  |  0 > 2080
|  |  |  |  0 <= 2377:  <=50K (18.0/2.0)
|  |  |  |  0 > 2377:  >50K (9.0)
|  |  13 > 12
|  |  |  13 <= 14:  <=50K (2120.0/189.0)
|  |  |  13 > 14
|  |  |  |  39 <= 32:  <=50K (63.0/6.0)
|  |  |  |  39 > 32
|  |  |  |  |  0 <= 653
|  |  |  |  |  |  Adm-clerical =  Exec-managerial:  >50K (11.0/2.0)
|  |  |  |  |  |  Adm-clerical =  Handlers-cleaners:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Prof-specialty
|  |  |  |  |  |  |  State-gov =  Self-emp-not-inc:  >50K (13.0/4.0)
|  |  |  |  |  |  |  State-gov =  Private
|  |  |  |  |  |  |  |  39 <= 52
|  |  |  |  |  |  |  |  |  40 <= 42:  >50K (12.0/2.0)
|  |  |  |  |  |  |  |  |  40 > 42
|  |  |  |  |  |  |  |  |  |  Male =  Male:  <=50K (6.0/1.0)
```

```
|  |  |  |  |  |  |  |  |  |  Male =  Female:  >50K (5.0)
|  |  |  |  |  |  |  |  |  39 > 52:  <=50K (9.0/1.0)
|  |  |  |  |  |  |  |  State-gov =  State-gov:  <=50K (6.0/1.0)
|  |  |  |  |  |  |  |  State-gov =  Federal-gov:  >50K (5.0/1.0)
|  |  |  |  |  |  |  |  State-gov =  Local-gov:  <=50K (4.0/1.0)
|  |  |  |  |  |  |  |  State-gov =  ?:  >50K (0.0)
|  |  |  |  |  |  |  |  State-gov =  Self-emp-inc:  <=50K (1.0)
|  |  |  |  |  |  |  |  State-gov =  Without-pay:  >50K (0.0)
|  |  |  |  |  |  |  |  State-gov =  Never-worked:  >50K (0.0)
|  |  |  |  |  |  |  Adm-clerical =  Other-service:  >50K (1.0)
|  |  |  |  |  |  |  Adm-clerical =  Adm-clerical:  >50K (0.0)
|  |  |  |  |  |  |  Adm-clerical =  Sales:  >50K (0.0)
|  |  |  |  |  |  |  Adm-clerical =  Craft-repair:  >50K (1.0)
|  |  |  |  |  |  |  Adm-clerical =  Transport-moving:  >50K (0.0)
|  |  |  |  |  |  |  Adm-clerical =  Farming-fishing:  >50K (0.0)
|  |  |  |  |  |  |  Adm-clerical =  Machine-op-inspct:  >50K (1.0)
|  |  |  |  |  |  |  Adm-clerical =  Tech-support:  >50K (0.0)
|  |  |  |  |  |  |  Adm-clerical =  ?:  <=50K (2.0/1.0)
|  |  |  |  |  |  |  Adm-clerical =  Protective-serv:  >50K (0.0)
|  |  |  |  |  |  |  Adm-clerical =  Armed-Forces:  >50K (0.0)
|  |  |  |  |  |  |  Adm-clerical =  Priv-house-serv:  >50K (0.0)
|  |  |  |  |  |  0 > 653:  >50K (11.0)
|  |  Never-married =  Separated
|  |  |  13 <= 12:  <=50K (885.0/22.0)
|  |  |  13 > 12
|  |  |  |  Bachelors =  Bachelors
|  |  |  |  |  2174 <= 4687:  <=50K (85.0/10.0)
|  |  |  |  |  2174 > 4687:  >50K (3.0/1.0)
|  |  |  |  Bachelors =  HS-grad:  <=50K (0.0)
|  |  |  |  Bachelors =  11th:  <=50K (0.0)
|  |  |  |  Bachelors =  Masters
|  |  |  |  |  Male =  Male
|  |  |  |  |  |  Adm-clerical =  Exec-managerial:  >50K (4.0)
|  |  |  |  |  |  Adm-clerical =  Handlers-cleaners:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Prof-specialty:  <=50K (3.0/1.0)
|  |  |  |  |  |  Adm-clerical =  Other-service:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Adm-clerical:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Sales:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Craft-repair:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Transport-moving:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Farming-fishing:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Machine-op-inspct:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Tech-support:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  ?:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Protective-serv:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Armed-Forces:  >50K (0.0)
|  |  |  |  |  |  Adm-clerical =  Priv-house-serv:  >50K (0.0)
|  |  |  |  |  Male =  Female:  <=50K (14.0/2.0)
|  |  |  |  Bachelors =  9th:  <=50K (0.0)
|  |  |  |  Bachelors =  Some-college:  <=50K (0.0)
```

```
|  |  |  Bachelors = Assoc-acdm: <=50K (0.0)
|  |  |  Bachelors = Assoc-voc: <=50K (0.0)
|  |  |  Bachelors = 7th-8th: <=50K (0.0)
|  |  |  Bachelors = Doctorate
|  |  |  |  77516 <= 192286: <=50K (3.0)
|  |  |  |  77516 > 192286: >50K (4.0)
|  |  |  Bachelors = Prof-school
|  |  |  |  77516 <= 129246: <=50K (2.0)
|  |  |  |  77516 > 129246: >50K (6.0)
|  |  |  Bachelors = 5th-6th: <=50K (0.0)
|  |  |  Bachelors = 10th: <=50K (0.0)
|  |  |  Bachelors = 1st-4th: <=50K (0.0)
|  |  |  Bachelors = Preschool: <=50K (0.0)
|  |  |  Bachelors = 12th: <=50K (0.0)
|  Never-married = Married-AF-spouse
|  |  Bachelors = Bachelors: >50K (3.0)
|  |  Bachelors = HS-grad: <=50K (13.0/3.0)
|  |  Bachelors = 11th: <=50K (0.0)
|  |  Bachelors = Masters: <=50K (0.0)
|  |  Bachelors = 9th: <=50K (0.0)
|  |  Bachelors = Some-college: >50K (3.0/1.0)
|  |  Bachelors = Assoc-acdm: <=50K (2.0)
|  |  Bachelors = Assoc-voc: >50K (1.0)
|  |  Bachelors = 7th-8th: <=50K (0.0)
|  |  Bachelors = Doctorate: <=50K (0.0)
|  |  Bachelors = Prof-school: <=50K (0.0)
|  |  Bachelors = 5th-6th: <=50K (0.0)
|  |  Bachelors = 10th: <=50K (0.0)
|  |  Bachelors = 1st-4th: <=50K (0.0)
|  |  Bachelors = Preschool: <=50K (0.0)
|  |  Bachelors = 12th: <=50K (0.0)
|  Never-married = Widowed: <=50K (973.0/65.0)
2174 > 6849
|  13 <= 10
|  |  40 <= 35
|  |  |  39 <= 27: <=50K (5.0)
|  |  |  39 > 27
|  |  |  |  40 <= 34: >50K (29.0)
|  |  |  |  40 > 34
|  |  |  |  |  Bachelors = Bachelors: >50K (0.0)
|  |  |  |  |  Bachelors = HS-grad: >50K (12.0/2.0)
|  |  |  |  |  Bachelors = 11th: >50K (0.0)
|  |  |  |  |  Bachelors = Masters: >50K (0.0)
|  |  |  |  |  Bachelors = 9th: >50K (0.0)
|  |  |  |  |  Bachelors = Some-college: <=50K (2.0)
|  |  |  |  |  Bachelors = Assoc-acdm: >50K (0.0)
|  |  |  |  |  Bachelors = Assoc-voc: >50K (0.0)
|  |  |  |  |  Bachelors = 7th-8th: >50K (0.0)
|  |  |  |  |  Bachelors = Doctorate: >50K (0.0)
|  |  |  |  |  Bachelors = Prof-school: >50K (0.0)
```

```
|  |  |  |  |  Bachelors =  5th-6th:  >50K (0.0)
|  |  |  |  |  Bachelors =  10th:  >50K (0.0)
|  |  |  |  |  Bachelors =  1st-4th:  >50K (0.0)
|  |  |  |  |  Bachelors =  Preschool:  >50K (0.0)
|  |  |  |  |  Bachelors =  12th:  >50K (0.0)
|  |  40 > 35
|  |  |  39 <= 60:  >50K (384.0/4.0)
|  |  |  39 > 60
|  |  |  |  2174 <= 9562:  >50K (20.0)
|  |  |  |  2174 > 9562
|  |  |  |  |  2174 <= 10566:  <=50K (4.0)
|  |  |  |  |  2174 > 10566:  >50K (18.0/1.0)
|  13 > 10:  >50K (925.0/2.0)

Number of Leaves  :     719

Size of the tree :      901


Time taken to build model: 1.43 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0.17 seconds

=== Summary ===

Correctly Classified Instances       28696               88.1327 %
Incorrectly Classified Instances     3864                11.8673 %
Kappa statistic                          0.6554
Mean absolute error                      0.1794
Root mean squared error                  0.2995
Relative absolute error                 49.0749 %
Root relative squared error             70.0543 %
Total Number of Instances            32560
```

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.949	0.331	0.900	0.949	0.924	0.660	0.914	0.964	<=50K
0.669	0.051	0.806	0.669	0.731	0.660	0.914	0.811	>50K
Weighted Avg. 0.881	0.264	0.877	0.881	0.877	0.660	0.914	0.927	

=== Confusion Matrix ===

a	b <-- classified as
23454	1265 | a = <=50K
2599	5242 | b = <=50K

Output

Fig 1.1: Loading data into weka

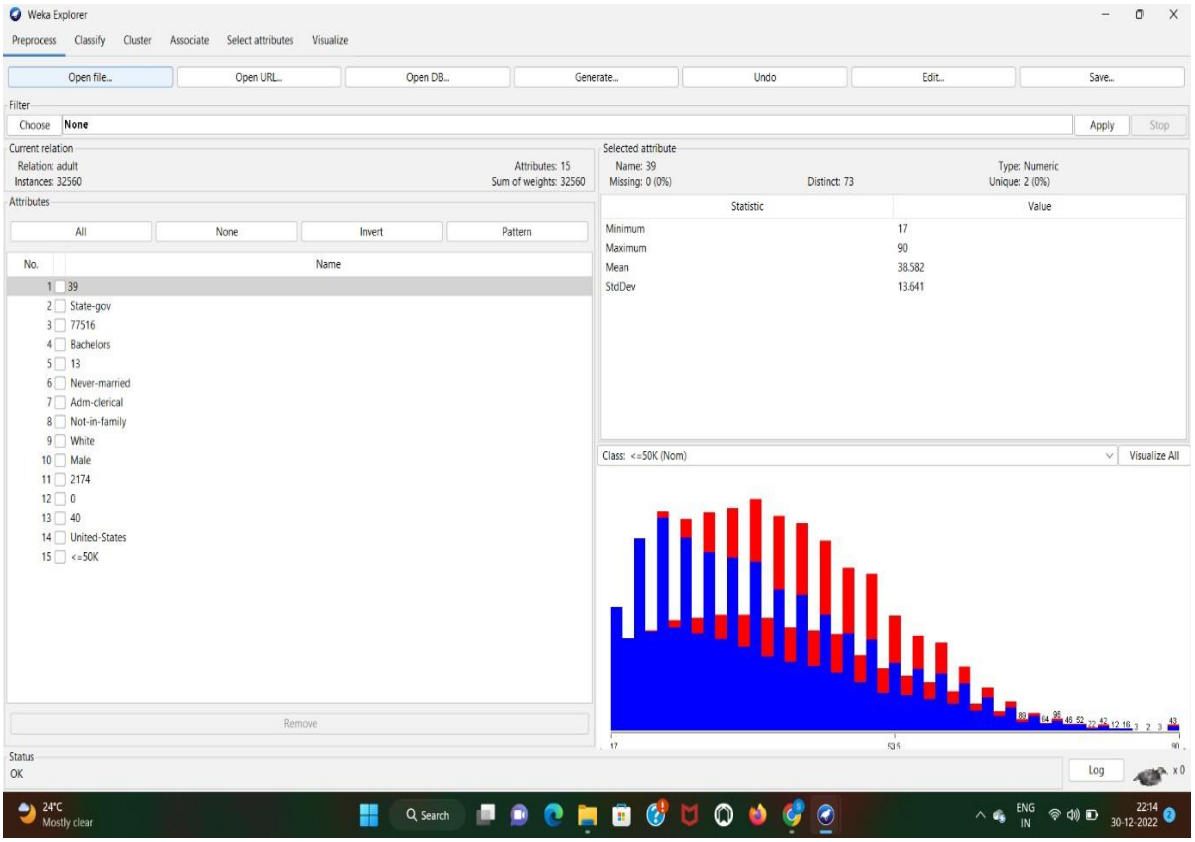


FIG 1.2: classification

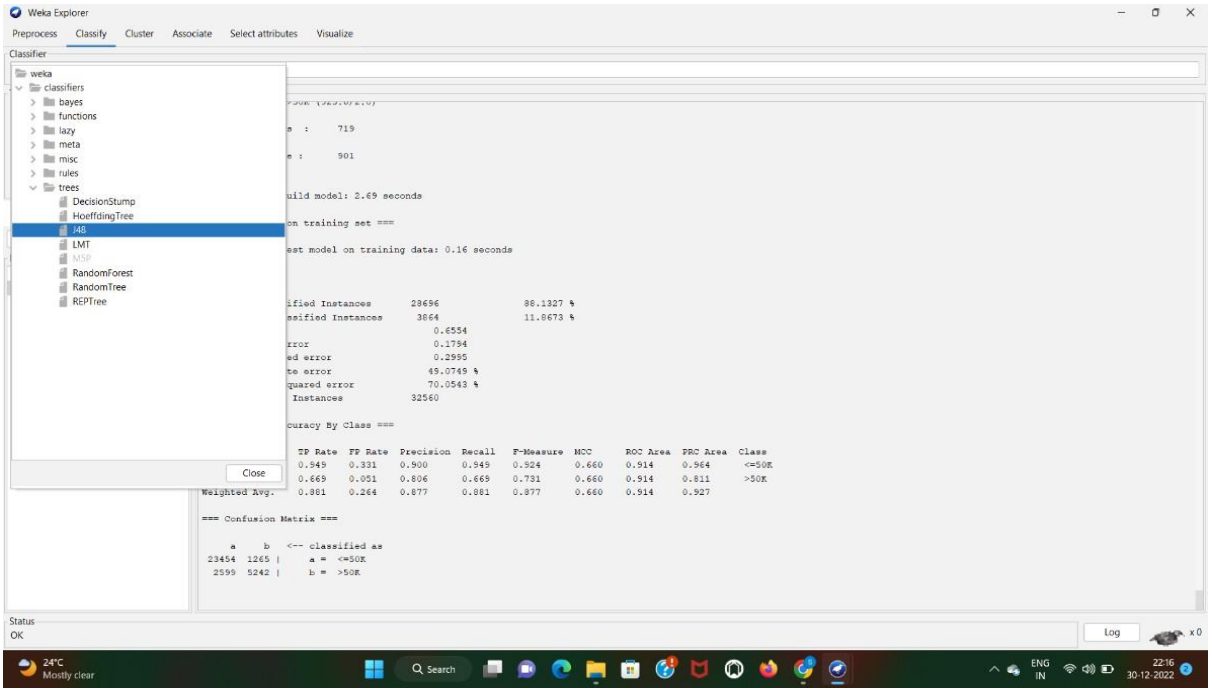
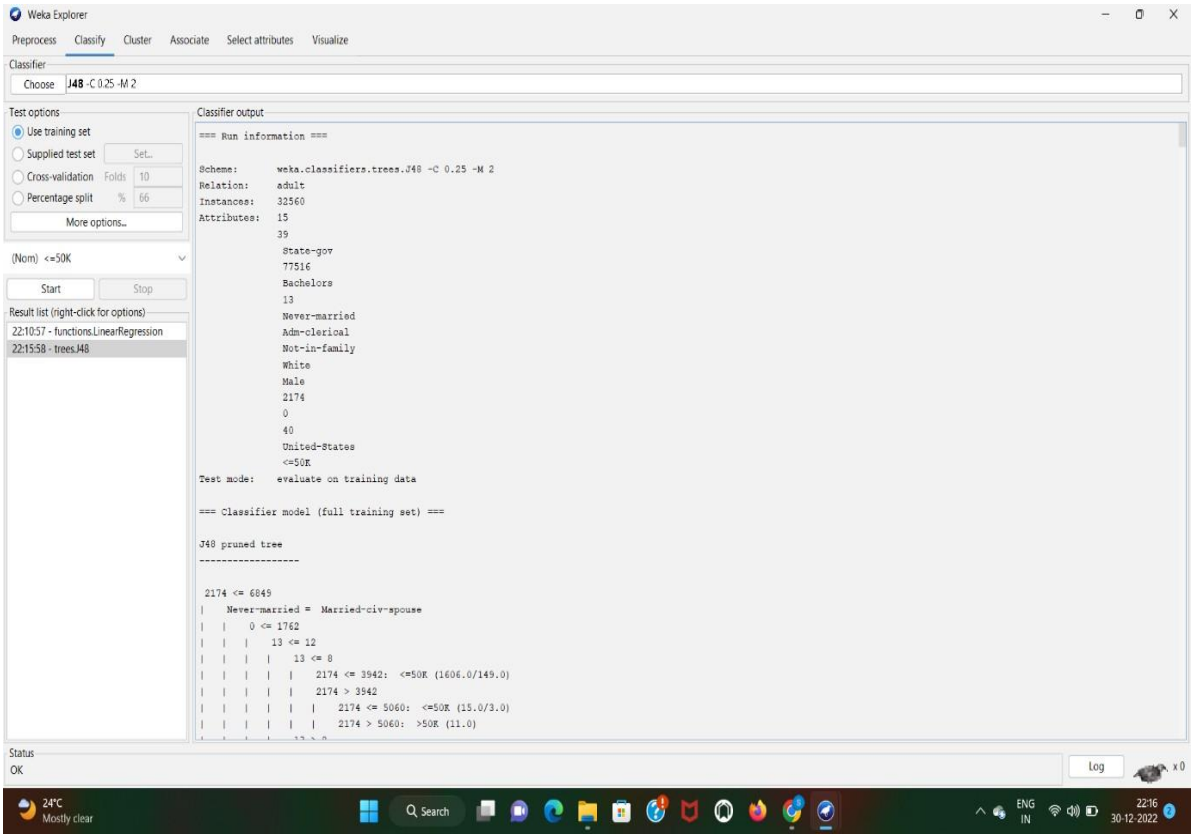
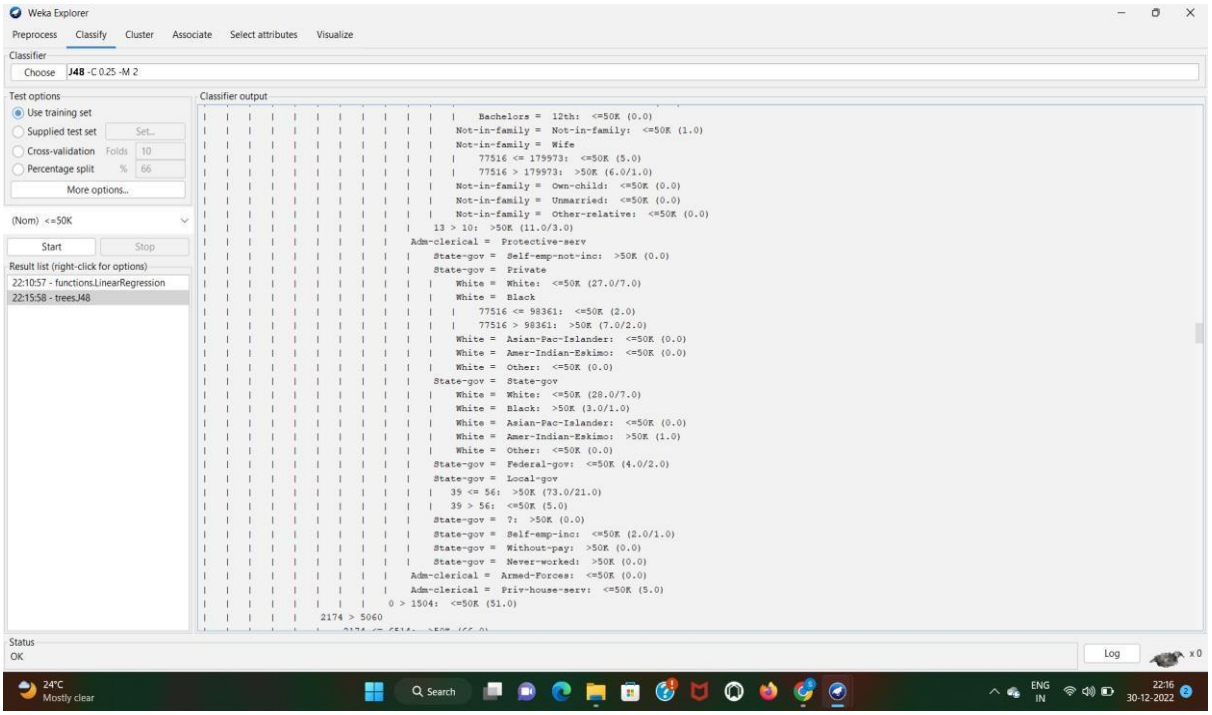
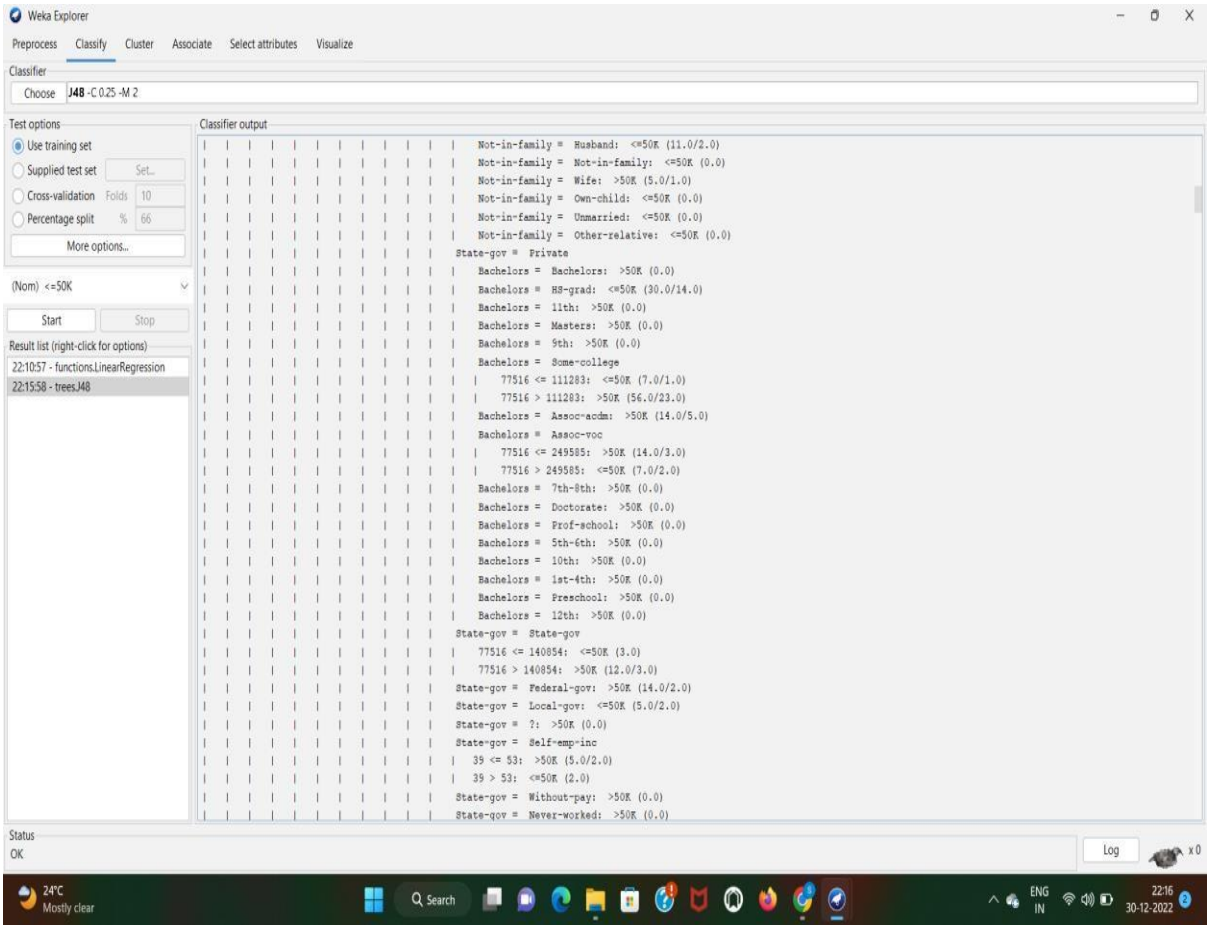
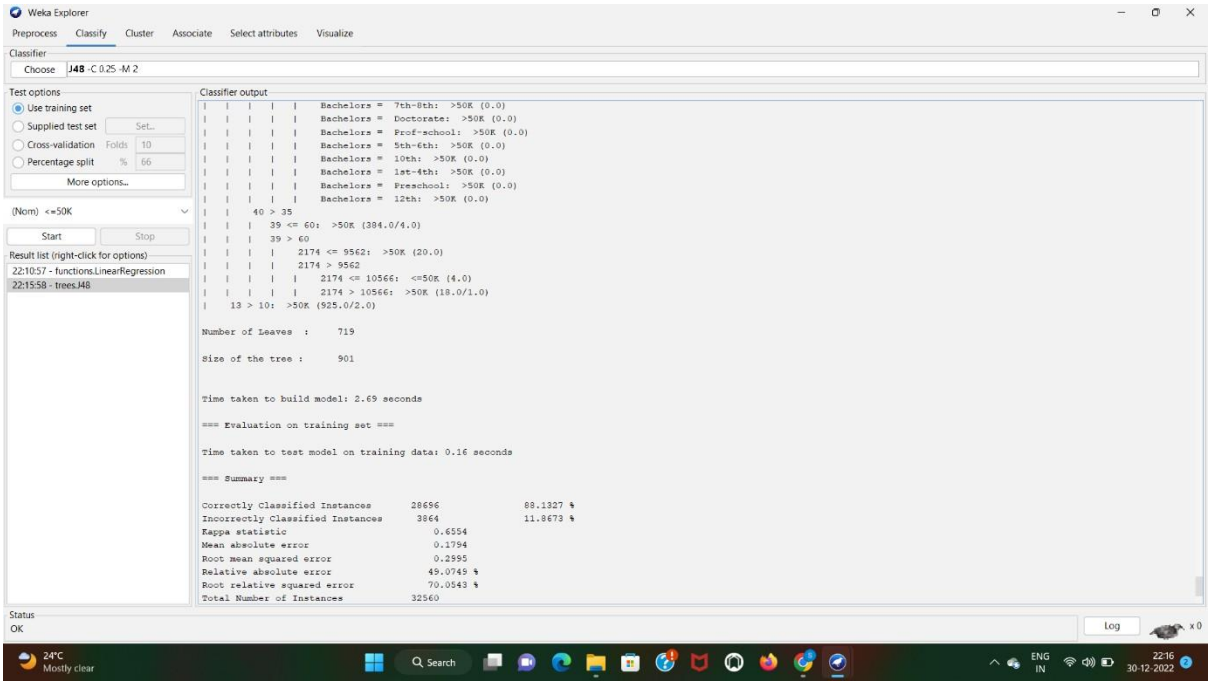


Fig 1.3,1.4,1.5,1.6: output




Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) <=50K
Start
Stop
Result list (right-click for options)
22:10:57 - functions.LinearRegression
22:15:58 - trees.J48
Classifier output
Bachelors = 12th: <=50K (0.0)
Not-in-family = Not-in-family: <=50K (1.0)
Not-in-family = Wife
77516 <= 179973: <=50K (5.0)
77516 > 179973: >50K (6.0/1.0)
Not-in-family = Own-child: <=50K (0.0)
Not-in-family = Unmarried: <=50K (0.0)
Not-in-family = Other-relative: <=50K (0.0)
13 > 10: >50K (11.0/3.0)
Adm-clerical = Protective-serv
State-gov = Self-emp-not-inc: >50K (0.0)
State-gov = Private
White = White: <=50K (27.0/7.0)
White = Black
77516 <= 98361: <=50K (2.0)
77516 > 98361: >50K (7.0/2.0)
White = Asian-Pac-Islander: <=50K (0.0)
White = Amer-Indian-Eskimo: <=50K (0.0)
White = Other: <=50K (0.0)
State-gov = State-gov
White = White: <=50K (28.0/7.0)
White = Black: >50K (3.0/1.0)
White = Asian-Pac-Islander: <=50K (0.0)
White = Amer-Indian-Eskimo: >50K (1.0)
White = Other: <=50K (0.0)
State-gov = Federal-gov: <=50K (4.0/2.0)
State-gov = Local-gov
39 <= 56: >50K (73.0/21.0)
39 > 56: <=50K (5.0)
State-gov = ?: >50K (0.0)
State-gov = Self-emp-inc: <=50K (2.0/1.0)
State-gov = Without-pay: >50K (0.0)
State-gov = Never-worked: >50K (0.0)
Adm-clerical = Armed-Forces: <=50K (0.0)
Adm-clerical = Priv-house-serv: <=50K (5.0)
0 > 1504: <=50K (51.0)
2174 > 5060
Status
OK
Log
x 0
24°C
Mostly clear
Search
ENG
IN
22:16
30-12-2022


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) <=50K
Start
Stop
Result list (right-click for options)
22:10:57 - functions.LinearRegression
22:15:58 - trees.J48
Classifier output
Not-in-family = Husband: <=50K (11.0/2.0)
Not-in-family = Not-in-family: <=50K (0.0)
Not-in-family = Wife: >50K (5.0/1.0)
Not-in-family = Own-child: <=50K (0.0)
Not-in-family = Unmarried: <=50K (0.0)
Not-in-family = Other-relative: <=50K (0.0)
State-gov = Private
Bachelors = Bachelors: >50K (0.0)
Bachelors = HS-grad: <=50K (30.0/14.0)
Bachelors = 11th: >50K (0.0)
Bachelors = Masters: >50K (0.0)
Bachelors = 9th: >50K (0.0)
Bachelors = Some-college
77516 <= 111283: <=50K (7.0/1.0)
77516 > 111283: >50K (56.0/23.0)
Bachelors = Assoc-acdm: >50K (14.0/5.0)
Bachelors = Assoc-voc
77516 <= 249585: >50K (14.0/3.0)
77516 > 249585: <=50K (7.0/2.0)
Bachelors = 7th-8th: >50K (0.0)
Bachelors = Doctorate: >50K (0.0)
Bachelors = Prof-school: >50K (0.0)
Bachelors = 5th-6th: >50K (0.0)
Bachelors = 10th: >50K (0.0)
Bachelors = 1st-4th: >50K (0.0)
Bachelors = Preschool: >50K (0.0)
Bachelors = 12th: >50K (0.0)
State-gov = State-gov
77516 <= 140854: <=50K (3.0)
77516 > 140854: >50K (12.0/3.0)
State-gov = Federal-gov: >50K (14.0/2.0)
State-gov = Local-gov: <=50K (5.0/2.0)
State-gov = ?: >50K (0.0)
State-gov = Self-emp-inc
39 <= 53: >50K (5.0/2.0)
39 > 53: <=50K (2.0)
State-gov = Without-pay: >50K (0.0)
State-gov = Never-worked: >50K (0.0)
Status
OK
Log
x 0
24°C
Mostly clear
Search
ENG
IN
22:16
30-12-2022


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) <=50K
Start
Stop
Result list (right-click for options)
22:10:57 - functions.LinearRegression
22:15:58 - trees.J48
Classifier output
| | | | | | Bachelors = 7th-8th: >50K (0.0)
| | | | | | Bachelors = Doctorate: >50K (0.0)
| | | | | | Bachelors = Prof-school: >50K (0.0)
| | | | | | Bachelors = 5th-6th: >50K (0.0)
| | | | | | Bachelors = 10th: >50K (0.0)
| | | | | | Bachelors = 1st-4th: >50K (0.0)
| | | | | | Bachelors = Preschool: >50K (0.0)
| | | | | | Bachelors = 12th: >50K (0.0)
| | | 40 > 35
| | | 39 <= 60: >50K (384.0/4.0)
| | | 39 > 60
| | | | 2174 <= 9562: >50K (20.0)
| | | | 2174 > 9562
| | | | | 2174 <= 10566: <=50K (4.0)
| | | | | 2174 > 10566: >50K (18.0/1.0)
| 13 > 10: >50K (925.0/2.0)
Number of Leaves : 719
Size of the tree : 901
Time taken to build model: 2.69 seconds
=== Evaluation on training set ===
Time taken to test model on training data: 0.16 seconds
=== Summary ===
Correctly Classified Instances 28696 88.1327 %
Incorrectly Classified Instances 3864 11.8673 %
Kappa statistic 0.6554
Mean absolute error 0.1794
Root mean squared error 0.2995
Relative absolute error 49.0749 %
Root relative squared error 70.0543 %
Total Number of Instances 32560
Status
OK
Log
x 0
24°C
Mostly clear
Search
ENG
IN
22:16
30-12-2022

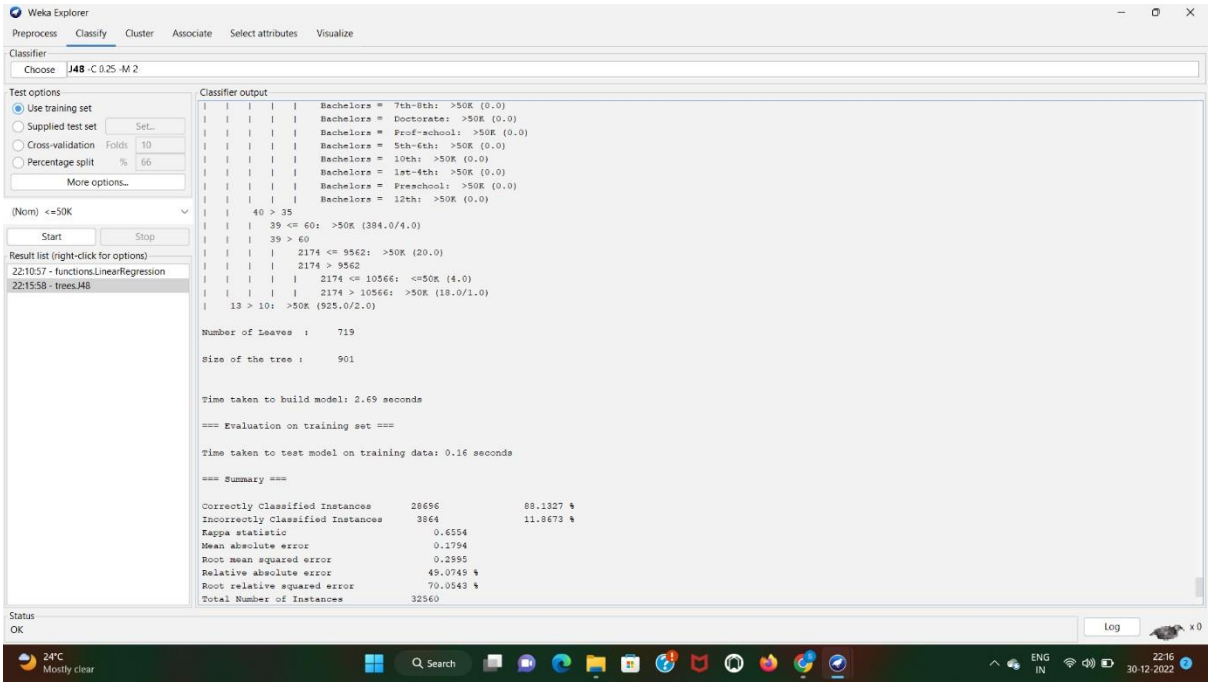

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) <=50K
Start
Stop
Result list (right-click for options)
22:10:57 - functions.LinearRegression
22:15:58 - trees.J48
Classifier output
| | | | | | Bachelors = 7th-8th: >50K (0.0)
| | | | | | Bachelors = Doctorate: >50K (0.0)
| | | | | | Bachelors = Prof-school: >50K (0.0)
| | | | | | Bachelors = 5th-6th: >50K (0.0)
| | | | | | Bachelors = 10th: >50K (0.0)
| | | | | | Bachelors = 1st-4th: >50K (0.0)
| | | | | | Bachelors = Preschool: >50K (0.0)
| | | | | | Bachelors = 12th: >50K (0.0)
| | | 40 > 35
| | | 39 <= 60: >50K (384.0/4.0)
| | | 39 > 60
| | | | 2174 <= 9562: >50K (20.0)
| | | | 2174 > 9562
| | | | | 2174 <= 10566: <=50K (4.0)
| | | | | 2174 > 10566: >50K (18.0/1.0)
| 13 > 10: >50K (925.0/2.0)
Number of Leaves : 719
Size of the tree : 901
Time taken to build model: 2.69 seconds
=== Evaluation on training set ===
Time taken to test model on training data: 0.16 seconds
=== Summary ===
Correctly Classified Instances 28696 88.1327 %
Incorrectly Classified Instances 3864 11.8673 %
Kappa statistic 0.6554
Mean absolute error 0.1794
Root mean squared error 0.2995
Relative absolute error 49.0749 %
Root relative squared error 70.0543 %
Total Number of Instances 32560
Status
OK
Log
x 0
24°C
Mostly clear
Search
ENG
IN
22:16
30-12-2022

Fig 1.7: Tree visualization

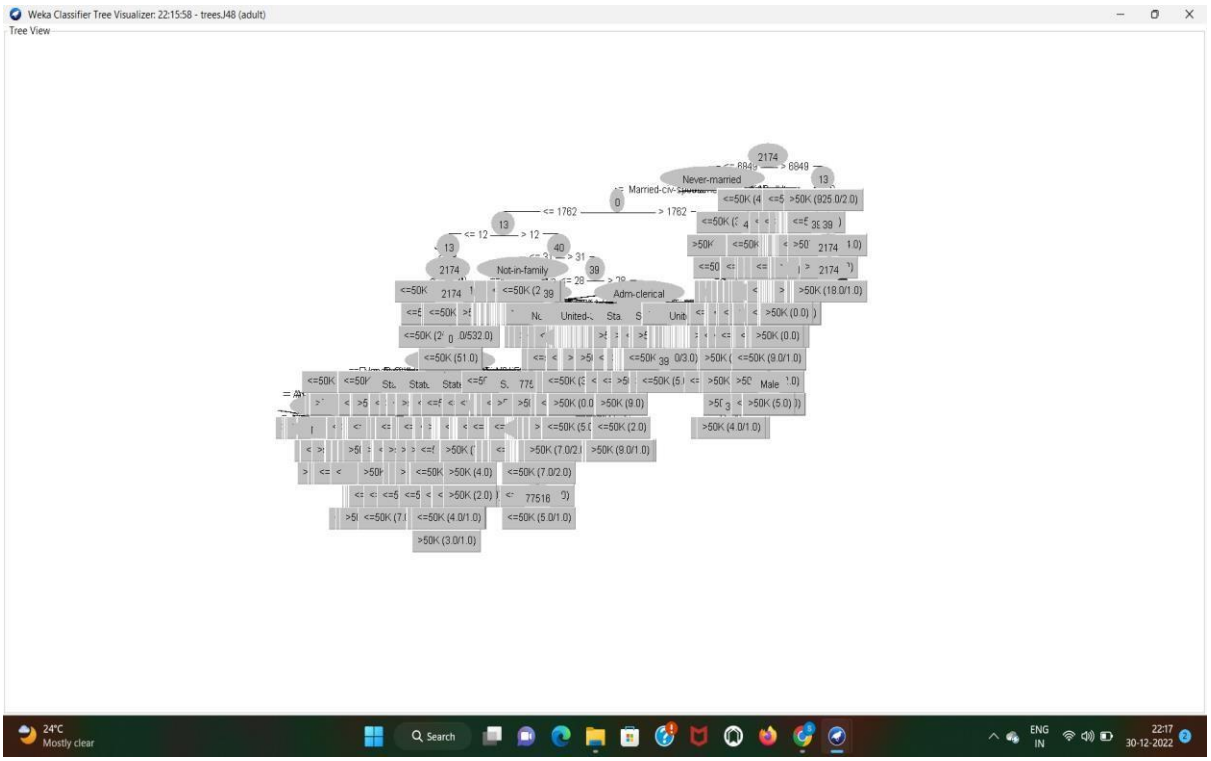


Fig 1.8: curve visualization

Naive Bayes

Bayesian Learning provides a probabilistic approach to inference.

It is based on the assumption that the quantities of interest are governed by probability distributions and that optimal decisions can be made by reasoning about these probabilities together with observed data. Bayesian learning algorithms calculate explicit probabilities for hypotheses. Each observed training example can incrementally decrease or increase the estimated probability that a hypothesis is correct. This provides a more flexible approach to learning than algorithms that completely eliminate a hypothesis if it is found to be inconsistent with any single example. Prior knowledge can be combined with observed data to determine the final probability of a hypothesis. In Bayesian learning, prior knowledge is provided by asserting

1) a prior probability for each candidate hypothesis, and

2) a probability distribution over observed data for each possible hypothesis. Bayesian methods can accommodate hypotheses that make probabilistic predictions New instances can be classified by combining the predictions of multiple hypotheses, weighted by their probabilities.

They require initial knowledge of many probabilities. When these probabilities are not known in advance they are often estimated based on background knowledge previously available data, and assumptions about the form of the underlying distributions. They require significant computational cost. They can provide a standard of optimal decision making against which other practical methods can be measured.

Implementation in Weka:

Open Start -> Programs ->Weka

Open **explorer**.

Click on **open file** and select **Adult.csv**

Select **Classify option** on the top of the Menu bar.

Select **Choose button** and click on **weka->classifier->bayes->NaiveBayes**.

Click on **Start button** and output will be displayed on the **right side** of the window.

Result

=== Run information ===

Scheme: weka.classifiers.bayes.NaiveBayes
Relation: adult
Instances: 32560
Attributes: 15
39
State-gov
77516
Bachelors
13
Never-married
Adm-clerical
Not-in-family
White
Male
2174
0
40
United-States
<=50K
Test mode: evaluate on training data

=== Classifier model (full training set) ===

Naive Bayes Classifier

	Class	
Attribute	<=50K	>50K
	(0.76)	(0.24)
39		
mean	36.8489	44.105
std. dev.	13.8031	10.364
weight sum	24719	7841
precision	1.0139	1.0139
State-gov		
Self-emp-not-inc	1818.0	725.0
Private	17734.0	4964.0
State-gov	945.0	354.0
Federal-gov	90.0	372.0

Local-gov	1477.0	618.0
?	1646.0	192.0
Self-emp-inc	495.0	623.0
Without-pay	15.0	1.0
Never-worked	8.0	1.0
[total]	24728.0	7850.0

77516

mean	190345.459	188004.5351
std. dev.	106479.7029	102535.5228
weight sum	24719	7841
precision	68.0227	68.0227

Bachelors

Bachelors	3134.0	2222.0
HS-grad	8827.0	1676.0
11th	1116.0	61.0
Masters	765.0	960.0
9th	488.0	28.0
Some-college	5905.0	1388.0
Assoc-acdm	803.0	266.0
Assoc-voc	1022.0	362.0
7th-8th	607.0	41.0
Doctorate	108.0	307.0
Prof-school	154.0	424.0
5th-6th	318.0	17.0

Priv-house-serv	149	2.0
[total]	24734.0	7856.0

Not-in-family

Husband	7276.0	5919.0
Not-in-family	7449.0	857.0

Wife	824.0	746.0
Own-child	5002.0	68.0
Unmarried	3229.0	219.0
Other-relative	945.0	38.0
[total]	24725.0	7847.0

White

White	20699.0	7118.0
Black	2738.0	388.0
Asian-Pac-Islander	764.0	277.0
Amer-Indian-Eskimo	276.0	37.0
Other	247.0	26.0
[total]	24724.0	7846.0

Male

Male	15128.0	6663.0
Female	9593.0	1180.0
[total]	24721.0	7843.0

2174

mean	149.6467	4029.7337
std. dev.	965.09	14582.8927
weight sum	24719	7841
precision	847.4492	847.4492

0

mean	53.0231	194.5982
std. dev.	310.1925	594.0944
weight sum	24719	7841
precision	47.8681	47.8681

40

mean	38.8175	45.4355

std. dev.	12.3269	10.9806
weight sum	24719	7841
precision	1.0538	1.0538

United-States

United-States	21999.0	7172.0
Cuba	71.0	26.0
Jamaica	72.0	11.0
India	61.0	41.0
?	438.0	147.0
Mexico	611.0	34.0
South	65.0	17.0
Puerto-Rico	103.0	13.0
Honduras	13.0	2.0
England	61.0	31.0
Canada	83.0	40.0
Germany	94.0	45.0
Iran	26.0	19.0
Philippines	138.0	62.0
Italy	49.0	26.0
Poland	49.0	13.0
Columbia	58.0	3.0
Cambodia	13.0	8.0
Thailand	16.0	4.0
Ecuador	25.0	5.0
Laos	17.0	3.0
Taiwan	32.0	21.0
Haiti	41.0	5.0
Portugal	34.0	5.0
Dominican-Republic	69.0	3.0
El-Salvador	98.0	10.0
France	18.0	13.0

Guatemala	62.0	4.0

China	56.0	21.0
Japan	39.0	25.0
Yugoslavia	11.0	7.0
Peru	30.0	3.0
Outlying-US(Guam-USVI-etc)	15.0	1.0
Scotland	10.0	4.0
Trinadad&Tobago	18.0	3.0
Greece	22.0	9.0
Nicaragua	33.0	3.0
Vietnam	63.0	6.0
Hong	15.0	7.0
Ireland	20.0	6.0
Hungary	11.0	4.0
Holand-Netherlands	2.0	1.0
[total]	24761.0	7883.0

Time taken to build model: 0.11 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0.35 seconds

=== Summary ===

Correctly Classified Instances	27193	83.5166 %
Incorrectly Classified Instances	5367	16.4834 %
Kappa statistic	0.5032	
Mean absolute error	0.1728	
Root mean squared error	0.3713	
Relative absolute error	47.2607 %	
Root relative squared error	86.8377 %	
Total Number of Instances	32560	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.934	0.478	0.860	0.934	0.896	0.513	0.893	0.964	<=50K
	0.522	0.066	0.716	0.522	0.604	0.513	0.893	0.728	>50K
Weighted Avg.	0.835	0.379	0.826	0.835	0.826	0.513	0.893	0.907	

=== Confusion Matrix ===

a	b	<-- classified as
23099	1620	a = <=50K
3747	4094	b = >50K

Output

Fig 2.1: naviebayes dataset view

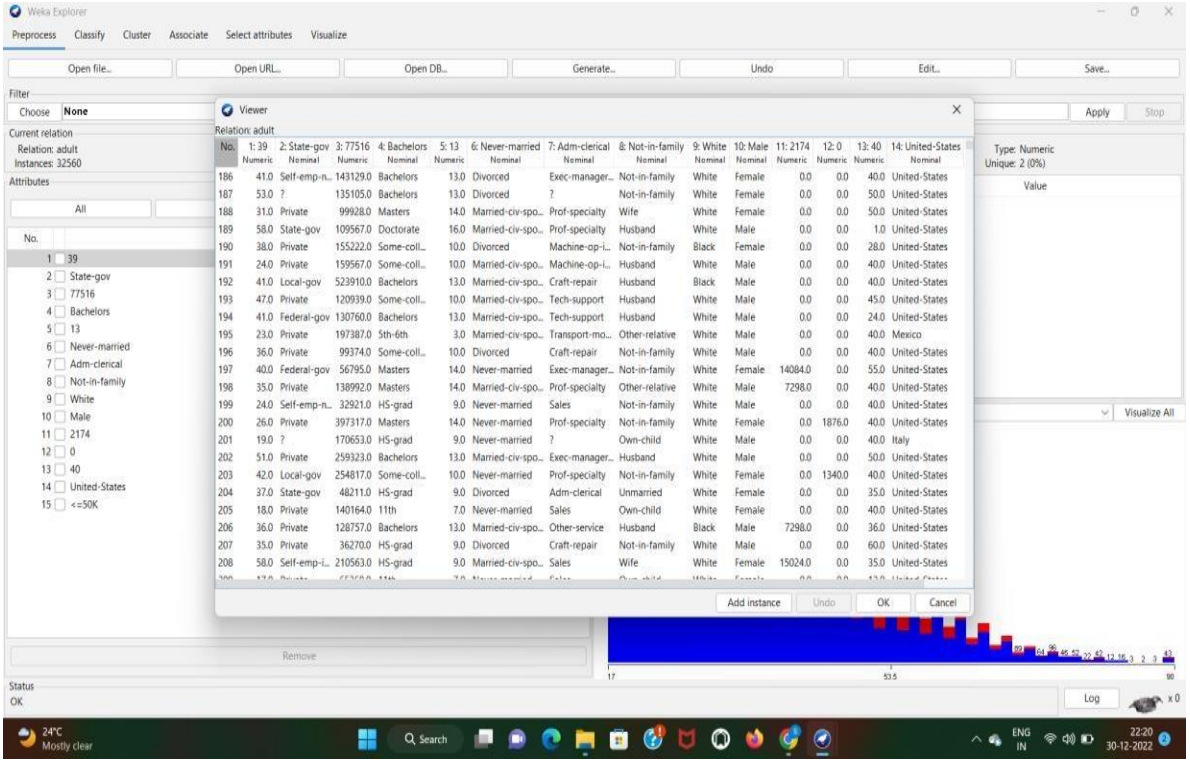


Fig 2.2: naivebayes view

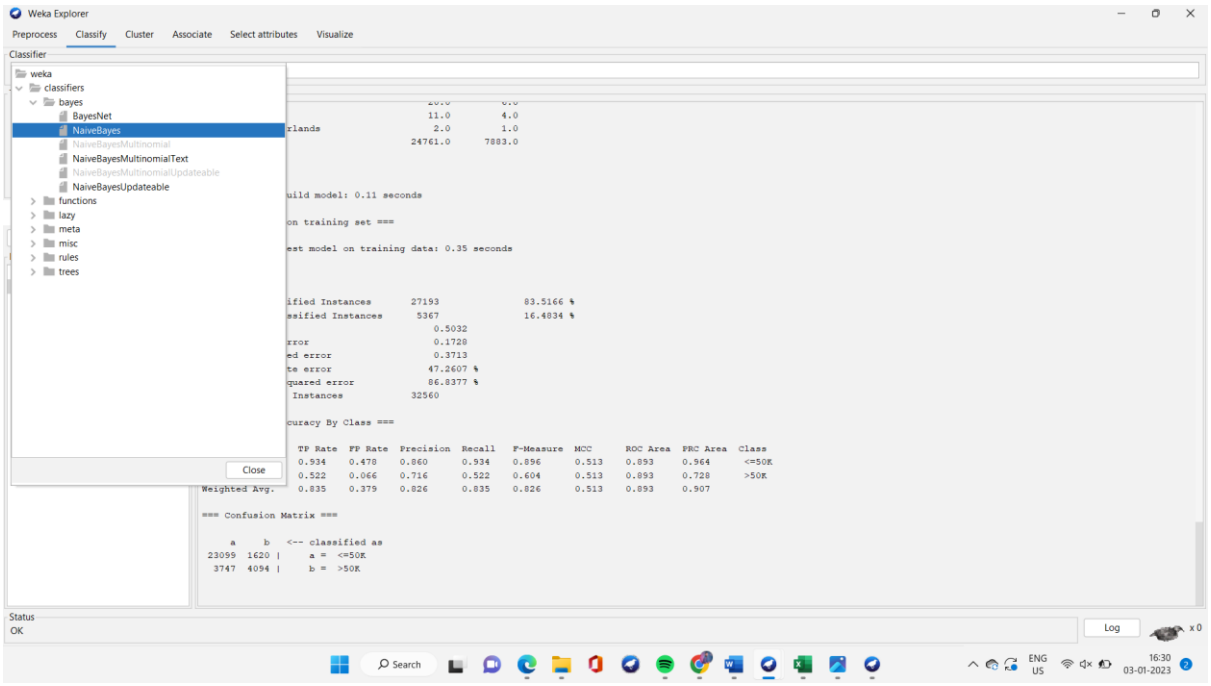
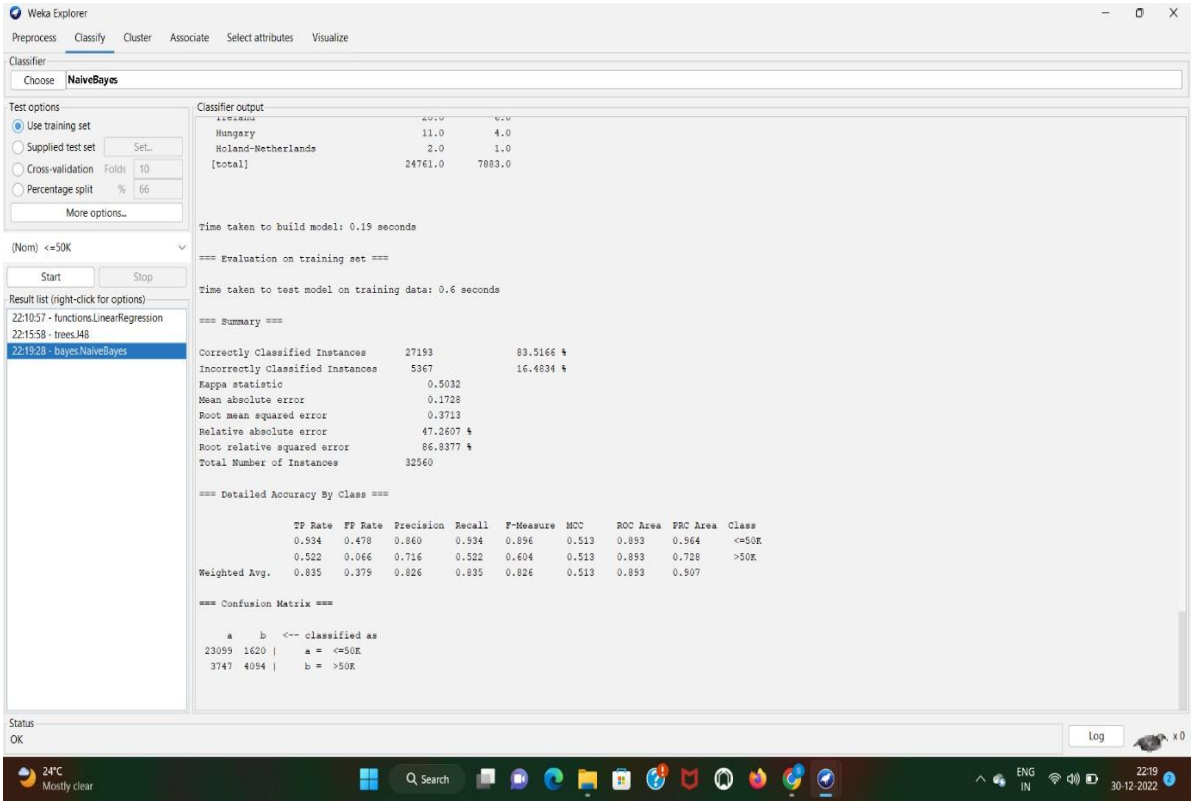
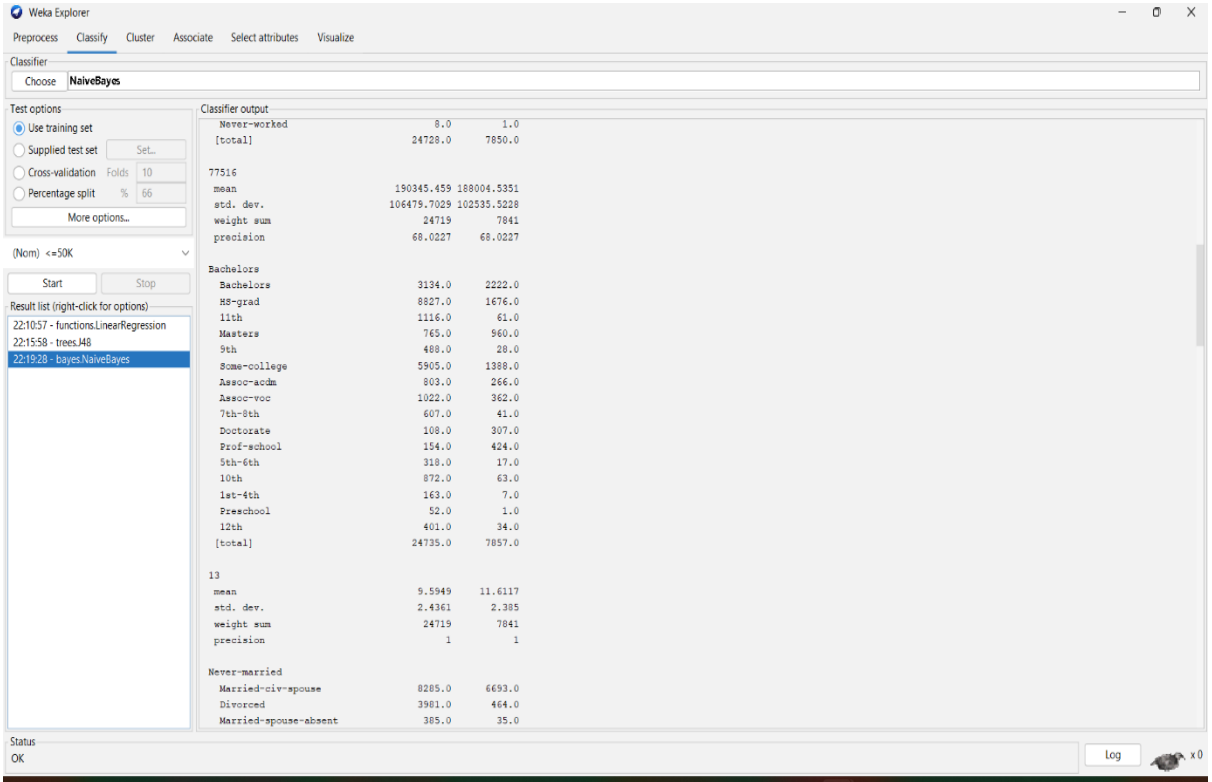
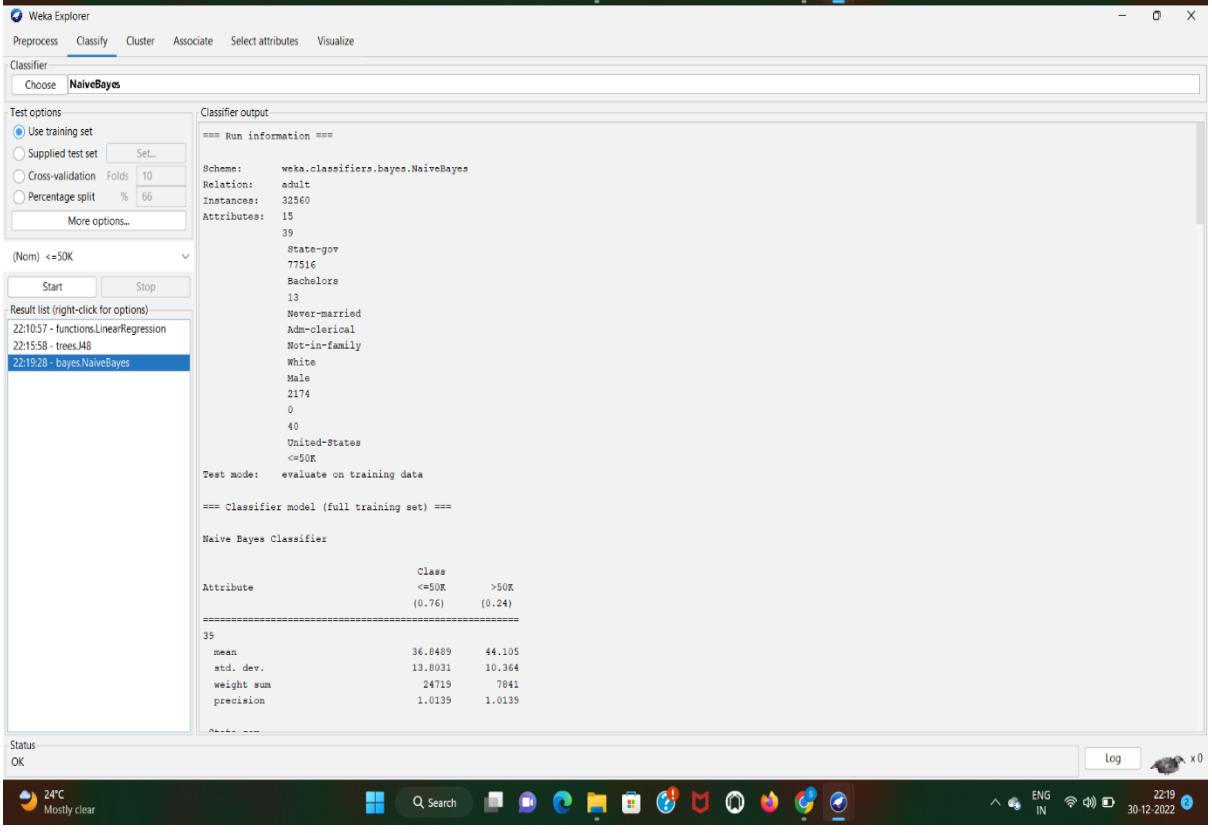


Fig 2.3,2.4,2.5: Result




Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
NaiveBayes
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) <=50K
Start
Stop
Result list (right-click for options)
22:10:57 - functions.LinearRegression
22:15:58 - trees.J48
22:19:28 - bayes.NaiveBayes
Classifier output
Never-worked 8.0 1.0
[total] 24728.0 7850.0
77516
mean 190345.459 188004.5351
std. dev. 106479.7029 102535.5228
weight sum 24719 7841
precision 68.0227 68.0227
Bachelors
Bachelors 3134.0 2222.0
HS-grad 8827.0 1676.0
11th 1116.0 61.0
Masters 765.0 960.0
9th 488.0 28.0
Some-college 5905.0 1388.0
Assoc-acdm 803.0 266.0
Assoc-voc 1022.0 362.0
7th-8th 607.0 41.0
Doctorate 108.0 307.0
Prof-school 154.0 424.0
5th-6th 318.0 17.0
10th 872.0 63.0
1st-4th 163.0 7.0
Preschool 52.0 1.0
12th 401.0 34.0
[total] 24735.0 7857.0
13
mean 9.5949 11.6117
std. dev. 2.4361 2.385
weight sum 24719 7841
precision 1 1
Never-married
Married-civ-spouse 8285.0 6693.0
Divorced 3981.0 464.0
Married-spouse-absent 385.0 35.0
Status
OK
Log
x 0
24°C
Mostly clear
Search
ENG
IN
22:19
30-12-2022


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
NaiveBayes
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) <=50K
Start
Stop
Result list (right-click for options)
22:10:57 - functions.LinearRegression
22:15:58 - trees.J48
22:19:28 - bayes.NaiveBayes
Classifier output
=== Run information ===
Scheme: weka.classifiers.bayes.NaiveBayes
Relation: adult
Instances: 32560
Attributes: 15
39
State-gov
77516
Bachelors
13
Never-married
Adm-clerical
Not-in-family
White
Male
2174
0
40
United-States
<=50K
Test mode: evaluate on training data
=== Classifier model (full training set) ===
Naive Bayes Classifier
Class
Attribute <=50K >50K
(0.76) (0.24)
39
mean 36.8489 44.105
std. dev. 13.8031 10.364
weight sum 24719 7841
precision 1.0139 1.0139
Status
OK
Log
x 0
24°C
Mostly clear
Search
ENG
IN
22:19
30-12-2022

Clustering

Clustering is the process of grouping a set of data objects into multiple groups or *clusters* o that objects within a cluster have high similarity, but are very dissimilar to objects in other clusters. Dissimilarities and similarities are assessed based on the attribute values describe escribing the objects and often involve distance measures. Clustering as a data mining tool has its roots in many application areas such as biology, security, business intelligence, and Web search.

Algorithm: *k*-means. The *k*-means algorithm for partitioning, where each cluster scenter is represented by the mean value of the objects in the cluster.

Input: *k*: the number of clusters, *D*: a data set containing *n* objects.

Output: A set of *k* clusters.

Method:

Arbitrarily choose *k* objects from *D* as the initial cluster centers; **Repeat** (Re) assign each object to the cluster to which the object is the most similar, based on the mean value of the objects in the cluster; update the cluster mew means, that is, calculate the mean value of the objects for each cluster; **until** no change;

Implementation in Weka

Steps for run K-mean Clustering algorithms in WEKA

1. Open WEKA Tool.
2. Click on WEKA Explorer.
3. Click on Preprocessing tab button.
4. Click on open file button.
5. Choose iris data set and open file.
6. Click on cluster tab and Choose k-mean and select use training set test.
7. Click on start button.

Result

=== Run information ===

Scheme:
weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10
Relation: Acmii
Instances: 7195
Attributes: 26

MFCCs_ 1
MFCCs_ 2
MFCCs_ 3
MFCCs_ 4
MFCCs_ 5
MFCCs_ 6
MFCCs_ 7
MFCCs_ 8
MFCCs_ 9
MFCCs_10
MFCCs_11
MFCCs_12
MFCCs_13
MFCCs_14
MFCCs_15
MFCCs_16
MFCCs_17
MFCCs_18
MFCCs_19
MFCCs_20
MFCCs_21
MFCCs_22
Family
Genus
Species

RecordID

Test mode: evaluate on training data

=== Clustering model (full training set) ===

kMeans

======

Number of iterations: 15

Within cluster sum of squared errors: 6667.792973183623

Initial starting points (random):

Cluster 0: 1,0.766848,0.761217,0.454927,-0.008135,0.228177,0.022602,0.142533,0.290178 0.225708,0.009719,0.457463,-0.199074,-0.295366,0.289808,0.159816,-0.117387,-0.052753 33 675 , 0.074759,-0.029766,-0.218756,Hylidae,Hypsiboas,HypsiboasCinerascens,3633214

Cluster: ,0.051892,- .011486,0.471796,0.252316,0.140863,- .073924 .021872,0.215224,0.114 761,-0.258898,-0.030776,0.268448,-0.046265 0.256193,0.000348,0.226305,0.123726,- lidae, Adenomera,AdenomeraHylaedactylus,21

Missing values globally replaced with mean/mode

Final cluster centroids:

	Cluster#		
Attribute	Full Data (7195.0)	0 (2466.0)	1 (4729.0)
MFCCs_ 1	0.9899	0.9762	0.997
MFCCs_ 2	0.3236	0.3733	0.2977
MFCCs_ 3	0.3112	0.4467	0.2406
MFCCs_ 4	0.446	0.3451	0.4986
MFCCs_ 5	0.127	0.0768	0.1532
MFCCs_ 6	0.0979	0.1667	0.0621

MFCCs_ 7	-0.0014	0.0674	-0.0373
MFCCs_ 8	-0.0004	-0.0263	0.0132
MFCCs_ 9	0.1282	0.0449	0.1716
MFCCs_10	0.056	0.0303	0.0694
MFCCs_11	-0.1157	0.0017	-0.1769
MFCCs_12	0.0434	0.0617	0.0338
MFCCs_13	0.1509	0.0138	0.2225
MFCCs_14	-0.0392	-0.024	-0.0472
MFCCs_15	-0.1017	0.0203	-0.1654
MFCCs_16	0.0421	0.0142	0.0566
MFCCs_17	0.0887	-0.0095	0.1399
MFCCs_18	0.0078	0.0155	0.0037
MFCCs_19	-0.0495	0.0159	-0.0836
MFCCs_20	-0.0532	0.0069	-0.0846
MFCCs_21	0.0373	-0.0066	0.0602
MFCCs_22	0.0876	-0.0124	0.1397
Family	Leptodactylidae	Hylidae	Leptodactylidae
Genus	Adenomera	Hypsiboas	Adenomera
Species	AdenomeraHylaedactylus	HypsiboasCor	AdenomeraHylaedactylus
RecordID	25.22	37.8751	18.6209

Time taken to build model (full training data) : 0.22 seconds

=== Model and evaluation on training set ===

Clustered Instances

0 2466(34%)

1 4729 (66%)

Output

Fig 3.1: cluster view

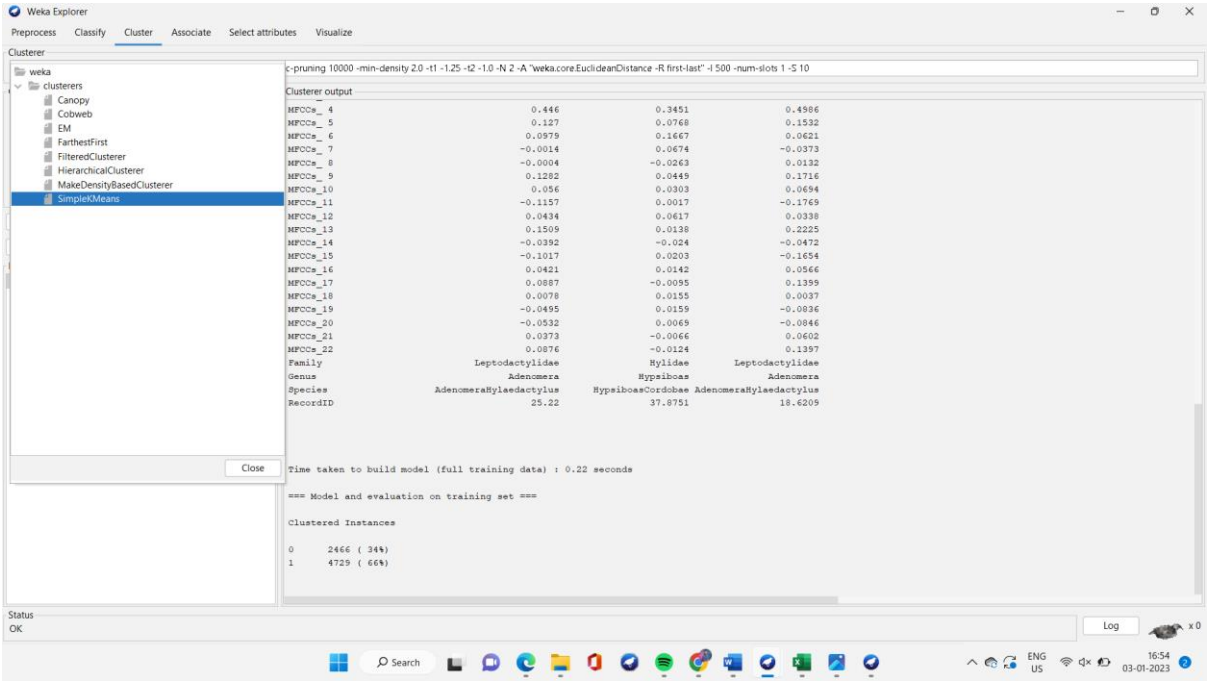
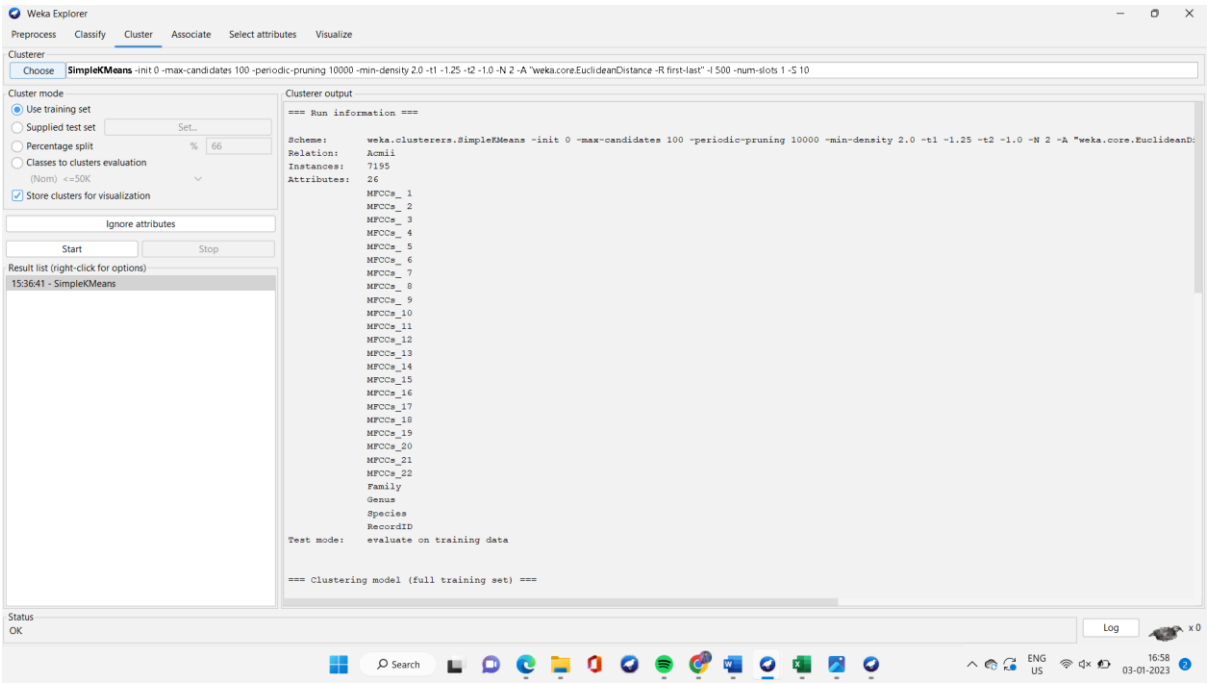


Fig 3.2,3.3: Result



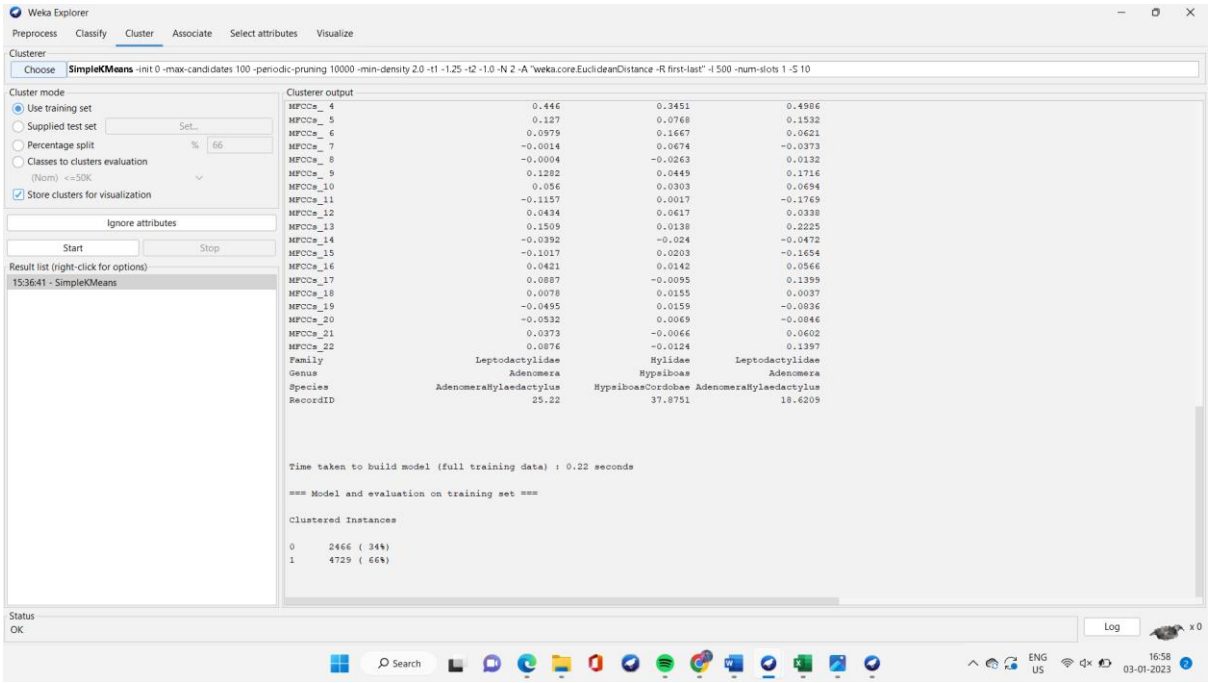
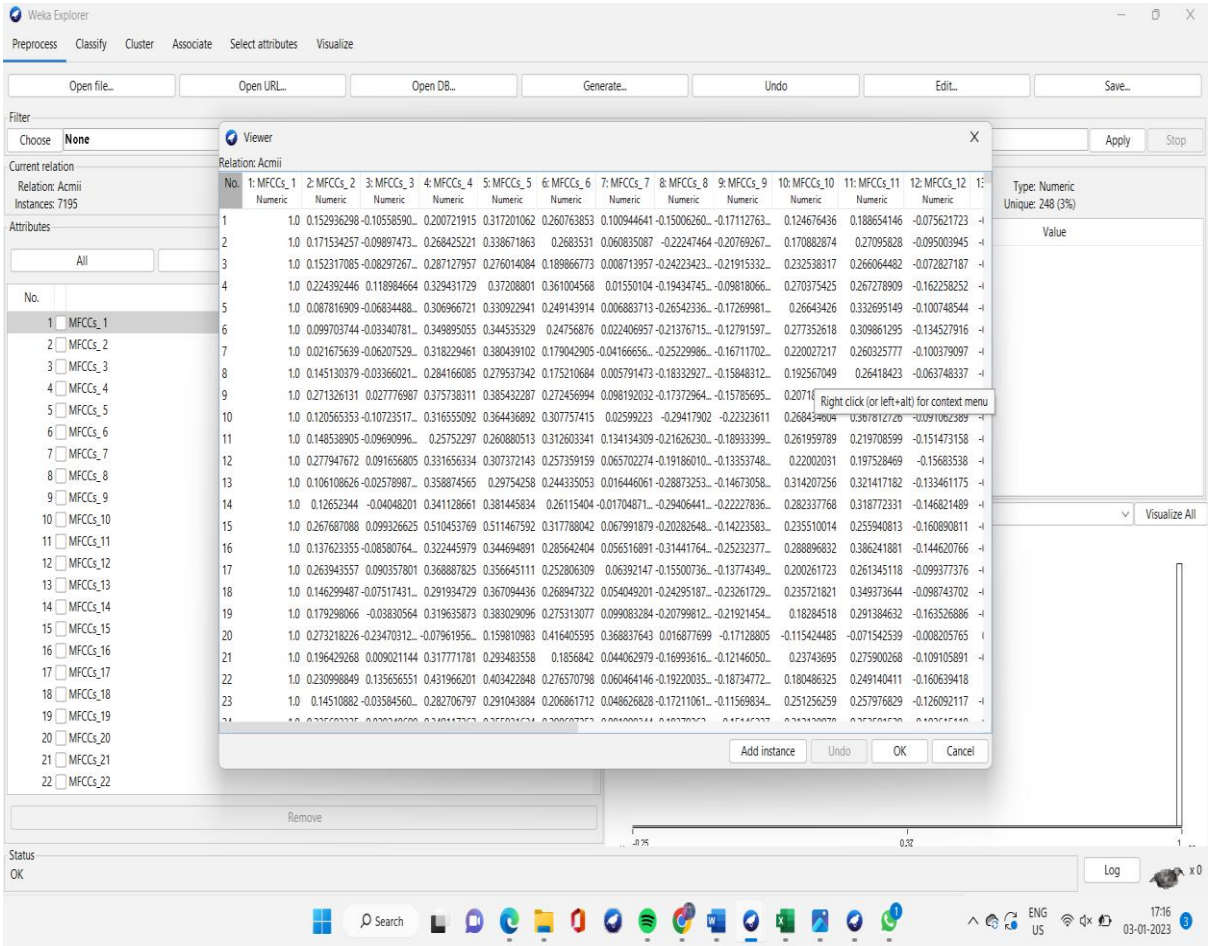


Fig 3.4,3.5 : data view



Linear regression

Numeric prediction is the task of predicting continuous values for given input. example, it required to predict the salary of employees with 10 years of experience, or tomorrow's tmperature. The most widely used approach for numeric prediction is regression. Regression analysis can be used to model the relationship between a set of predictor variables and a response variable (which is continuous-valued). The response variable is also referred to as the predicted attribute.

Regression analysis is a good choice when all of the predictor variables are continuous valued as well. Many problems can be solved by linear regression, and even more can be tackled by, applying transformations to the variables so that a nonlinear problem can be converted to a linear one. Several software packages exist to solve regression problems. Examples include SAS, SPSS, and S-Plus. Simple Linear regression analysis involves a response variable, y, an a single redictor variable, x. It is the simplest form of regression, and models y as a linear function of x. That is

$$y = b + wx$$

where the variance of y is assumed to be constant, and b and w are regression coefficients specifying the Y-intercept and slope of the line, respectively.

Implementation in Weka

Procedure:

1) Open Start -> Programs -> Weka
2) Open explorer.
3) Click on open file and select Dataset
4) Select Classify option on the top of the Menu bar.
5) Select Choose button and click on weka->classifiers->functions->Linear Regression.
6) Click on Start button and output will be displayed on the right side of the window.

Result

=== Run information ===

Scheme: weka.classifiers.functions.SimpleLinearRegression
Relation: accelerometer
Instances: 153000
Attributes: 5
wconfid
pctid
x
y
z
Test mode: evaluate on training data

=== Classifier model (full training set) ===

Linear regression on x

-0.06 * x - 0.06

Predicting 0 if attribute value is missing.

Time taken to build model: 0.12 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0.3 seconds

=== Summary ===

Correlation coefficient 0.0912

Mean absolute error	0.2432
Root mean squared error	0.5149
Relative absolute error	99.5363 %
Root relative squared error	99.5836 %
Total Number of Instances	153000

Output

Fig 4.1:Loading data

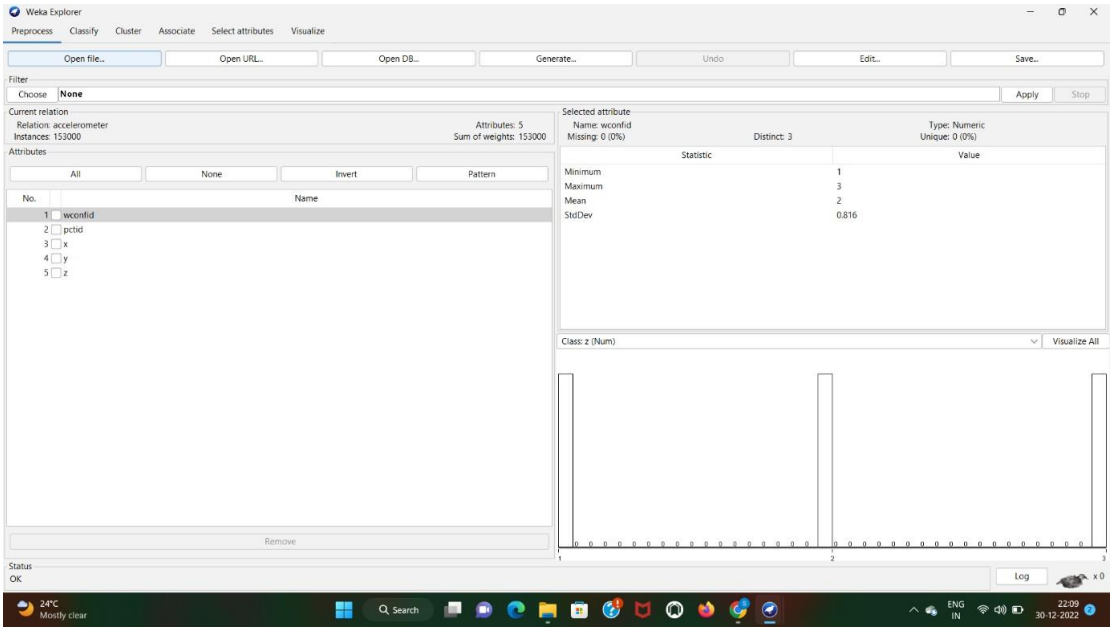


Fig 4.2 : linear regression

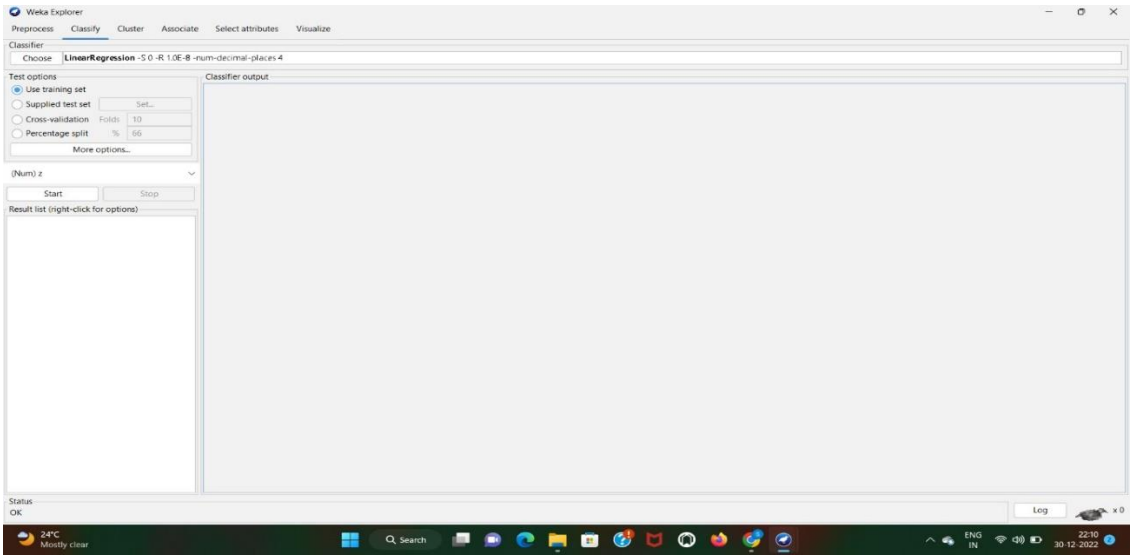


Fig 4.3 : Result

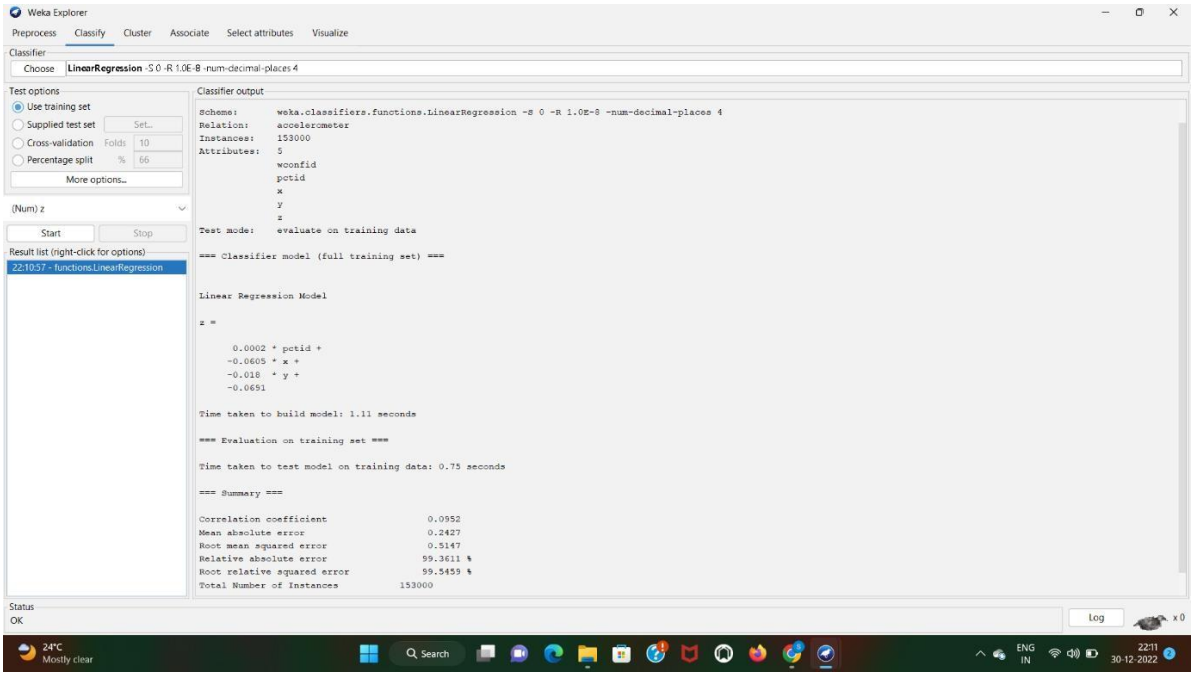


Fig 4.4: Data View