



Title: UCI – Machine Learning Repository**URL:** <http://archive.ics.uci.edu/ml/datasets.html>**Description:**

The above link contains multiple Data-sets which can be used for creating a Use-case.

Name	Data Types	Default Task	Attribute Types	# Instances	# Attributes	Year
 Abalone	Multivariate	Classification	Categorical, Integer, Real	4177	8	1995
 Adult	Multivariate	Classification	Categorical, Integer	48842	14	1996
 Annealing	Multivariate	Classification	Categorical, Integer, Real	798	38	
 Anonymous Microsoft Web Data		Recommender-Systems	Categorical	37711	294	1998
 Arrhythmia	Multivariate	Classification	Categorical, Integer, Real	452	279	1998
 Artificial Characters	Multivariate	Classification	Categorical, Integer, Real	6000	7	1992
 Audiology (Original)	Multivariate	Classification	Categorical	226		1987
 Audiology (Standardized)	Multivariate	Classification	Categorical	226	69	1992

Select one of them.

For instance you clicked Badges Data Set, then the below page will open. Open the highlighted link i.e Data Folder

Badges Data Set

Download: [Data Folder](#) [Data Set Description](#)

Abstract: Badges labeled with a "+" or "-" as a function of a person's name

Data Set Characteristics:	Univariate, Text	Number of Instances:	294	Area:	N/A
Attribute Characteristics:	N/A	Number of Attributes:	1	Date Donated	1994-09-01
Associated Tasks:	Classification	Missing Values?	No	Number of Web Hits:	27411

Source:

Creator:

Haym Hirsh, after an idea by Rob Schapire

Donor:





Haym Hirsh (hirsh.1@cs.nyu.edu)

Data Set Information:

Part of the problem in using an automated program to discover the unknown target function is to decide how to encode names such that the program can be used. The data below are presented in the form of a +/- label followed by the person's name. It is up to the learning-system user to decide how to convert this data into something usable by the system (e.g., what attributes to use if your favorite learner requires feature-vector data).

Data Folder : <http://archive.ics.uci.edu/ml/machine-learning-databases/badges/>

Index of /ml/machine-learning-databases/badges

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 Parent Directory			-
 Index	03-Dec-1996 04:07	110	
 badges.data	16-Nov-1994 22:44	4.9K	
 badges.info	16-Nov-1994 22:44	1.3K	

Apache/2.2.15 (CentOS) Server at archive.ics.uci.edu Port 80

Click on .info file of the Data-set. It will open the description of the Data-set.

<http://archive.ics.uci.edu/ml/machine-learning-databases/badges/badges.info>

1. Title: ML94/COLT94 Badge Problem
2. Source Information
 - Creator: Haym Hirsh, after an idea by Rob Schapire
 - Donor: Haym Hirsh (hirsh@cs.rutgers.edu)
 - Date: September, 1994
3. Past Usage:

Every pre-registered attendee at the 1994 Machine Learning Conference and 1994 Computational Learning Theory Conference received a badge labeled with a "+" or "-". The labeling was due to some function known only to the badge generator (Haym Hirsh), and it depended only on the attendee's name. The goal for conference attendees was to identify the unknown function used to generate the +/- labeling.
4. Relevant Information:

Part of the problem in using an automated program to discover the unknown target function is to decide how to encode names such that the program can be used. The data below are presented in the form of a +/- label followed by the person's name. It is up to the learning-system user to decide how to convert this data into something usable by the system (e.g., what attributes to use if your favorite learner requires feature-vector data).
5. Number of Instances: 294
6. Number of Attributes: N/A
7. Attribute Information: N/A
8. Missing Attribute Values: N/A
9. Class Distribution: 210 positives, 84 negatives

Click on .data file. It will open the dataset.

<http://archive.ics.uci.edu/ml/machine-learning-databases/badges/badges.data>

```
+ Naoki Abe  
- Myriam Abramson  
+ David W. Aha  
+ Kamal M. Ali  
- Eric Allender  
+ Dana Angluin  
- Chidanand Apte  
+ Minoru Asada  
+ Lars Asker  
+ Javed Aslam  
+ Haralabos Athanassiou  
+ Jose L. Balcazar  
+ Timothy P. Barber  
+ Michael W. Barley  
- Cristina Baroglio  
+ Peter Bartlett  
- Eric Baum  
+ Welton Becket  
.. . . .
```

Similarly there are multiple Data-sets available, you can follow the same steps to find the data and its description.