**Internet firewall**

this data set was collected from the internet traffic records on a university's firewall.

Dataset Characteristics

Multivariate

Subject Area

Computer Science

Associated Tasks

Classification

Feature Type

# Features

12

# Instances

65532

Dataset Information

Additional Information

There are 12 features in total. Action feature is used as a class. There are 4 classes in total. These are allow, action, drop and reset-both classes.

Has Missing Values?

no

Variable Information

Source Port,Destination Port,NAT Source Port,NAT Destination Port,Action,Bytes,Bytes Sent,Bytes Received,Packets,Elapsed Time (sec),pkts\_sent,pkts\_received

Dataset Files

File Size

log2.csv 2.9MBSM (

C CSM (Conventional and

12 features categorized as conventional and social media features. Both conventional features, collected from movies databases on Web as well as social media features(YouTube,Twitter). Dataset Characteristics

Multivariate

Subject Area

Computer Science

Associated Tasks

Classification, Regression

Feature Type

Integer

# Instances

217

# Features

12

Dataset Information

Additional Information

Year:2014 and 2015 Source: Twitter,YouTube,IMDB

Has Missing Values?

Yes

Dataset Files

File Size

2014 and 2015 CSM dataset.xlsx 34KB

**Power Consumption of Tetouan City**

Power Consumption of Tetouan CityThis dataset is related to power consumption of three different distribution networks of Tetouan city which is located in north Morocco

Dataset Characteristics

Multivariate

Time-SeriesSubject Area

Social Science

Associated Tasks

Regression

# Instances

52417

Feature Type

Real, Integer

# Features

6

Dataset Information

Has Missing Values?

No

Introductory Paper

Comparison of Machine Learning Algorithms for the Power Consumption Prediction : - Case Study of Tetouan city –

By A. Salam, A. E. Hibaoui. 2018

Published in International Renewable and Sustainable Energy Conference Power Consumption of Tetouan City

This dataset is related to power consumption of three different distribution networks of Tetouan city which is located in north Morocco

Dataset Characteristics

Multivariate

Time-Series

Subject Area

Social Science

Associated Tasks

Regression

# Instances

52417

Feature Type

Real, Integer

# Features

6

Dataset Information

Has Missing Values?

No

Introductory Paper

Comparison of Machine Learning Algorithms for the Power Consumption Prediction : - Case Study of Tetouan city –

By A. Salam, A. E. Hibaoui. 2018

Published in International Renewable and Sustainable Energy Conference

Variables Table

| Variable Name | Role | Type | Description | Units | Missing Values |
| --- | --- | --- | --- | --- | --- |
| DateTime | Feature | Date | Each ten minutes |  | no |
| Temperature | Feature | Continuous | Weather Temperature of Tetouan city |  | no |
| Humidity | Feature | Continuous | Weather Humidity of Tetouan city |  | no |
| Wind Speed | Feature | Continuous | Wind speed of Tetouan city |  | no |
| general diffuse flows | Feature | Continuous | general diffuse flows |  | no |
| diffuse flows | Feature | Continuous | diffuse flows |  | no |
| Zone 1 Power Consumption | Target | Continuous | power consumption of zone 1 of Tetouan city |  | no |
| Zone 2 Power Consumption | Target | Continuous | power consumption of zone 2 of Tetouan city |  | no |
| Zone 3 Power Consumption | Target | Continuous | power consumption of zone 3 of Tetouan city |  | no |

Facebook Metrics

Facebook performance metrics of a renowned cosmetic's brand Facebook page.

Dataset Characteristics

Multivariate

Subject Area

Bussiness

Associated Tasks

Regression

Feature Type

Integer

# Instances

500

# Features

18 Dataset Information

Additional Information

The data is related to posts' published during the year of 2014 on the Facebook's page of a renowned cosmetics brand.

This dataset contains 500 of the 790 rows and part of the features analyzed by Moro et al. (2016). The remaining were omitted due to confidentiality issues.

Has Missing Values?

No

**Variables Table**

| **Variable Name** | **Role** | **Type** | **Description** | **Units** | **Missing Values** |
| --- | --- | --- | --- | --- | --- |
| Page total likes | Feature | Integer |  |  | no |
| Type | Feature | Categorical |  |  | no |
| Category | Feature | Integer |  |  | no |
| Post Month | Feature | Integer |  |  | no |
| Post Weekday | Feature | Integer |  |  | no |
| Post Hour | Feature | Integer |  |  | no |
| Paid | Feature | Continuous |  |  | no |
| Lifetime Post Total Reach | Feature | Integer |  |  | no |
| Lifetime Post Total Impressions | Feature | Integer |  |  | no |
| Lifetime Engaged Users | Feature | Integer |  |  | no |

**Variables Table**

| **Variable Name** | **Role** | **Type** | **Description** | **Units** | **Missing Values** |
| --- | --- | --- | --- | --- | --- |
| Lifetime Post Consumers | Feature | Integer |  |  | no |
| Lifetime Post Consumptions | Feature | Integer |  |  | no |
| Lifetime Post Impressions by people who have liked your Page | Feature | Integer |  |  | no |
| Lifetime Post reach by people who like your Page | Feature | Integer |  |  | no |
| Lifetime People who have liked your Page and engaged with your post | Feature | Integer |  |  | no |
| comment | Feature | Integer |  |  | no |
| like | Feature | Integer |  |  | no |
| share | Feature | Integer |  |  | no |
| Total Interactions | Target | Integer |  |  | no |

**Additional Variable Information**

It includes 7 features known prior to post publication and 12 features for evaluating post impact (see Tables 2 and 3 from Moro et al., 2016 - complete reference in the 'Citation Request')

Steel Industry Energy Consumption

The data is collected from a smart small-scale steel industry in South Korea.

Dataset Characteristics

Multivariate

Subject Area

Physics and Chemistry

Associated Tasks

Regression

Feature Type

Real, Categorical

# Instances

35040

# Features

9

Dataset Information

Additional Information

The information gathered is from the DAEWOO Steel Co. Ltd in Gwangyang, South Korea. It produces several types of coils, steel plates, and iron plates. The information on electricity consumption is held in a cloud-based system. The information on energy consumption of the industry is stored on the website of the Korea Electric Power Corporation (pccs.kepco.go.kr), and the perspectives on daily, monthly, and annual data are calculated and shown.

Has Missing Values?

No

**Variables Table**

| **Variable Name** | **Role** | **Type** | **Description** | **Units** | **Missing Values** |
| --- | --- | --- | --- | --- | --- |
| date | Other | Date |  |  | no |
| Usage\_kWh | Feature | Continuous | Industry Energy Consumption | kWh | no |
| Lagging\_Current\_Reactive.Power\_kVarh | Feature | Continuous |  | kVarh | no |
| Leading\_Current\_Reactive\_Power\_kVarh | Feature | Continuous |  | kVarh | no |
| CO2(tCO2) | Feature | Continuous |  | ppm | no |
| Lagging\_Current\_Power\_Factor | Feature | Continuous |  | % | no |
| Leading\_Current\_Power\_Factor | Feature | Continuous |  | % | no |
| NSM | Feature | Integer |  | s | no |
| WeekStatus | Feature | Categorical | Weekend (0) or a Weekday(1) |  | no |
| Day\_of\_week | Feature | Categorical | Sunday, Monday, ..., Saturday |  | no |

Additional Variable Information

Data Variables Type Measurement

Industry Energy Consumption Continuous kWh

Lagging Current reactive power Continuous kVarh

Leading Current reactive power Continuous kVarh

tCO2(CO2) Continuous ppm

Lagging Current power factor Continuous %

Leading Current Power factor Continuous %

Number of Seconds from midnight Continuous S

Week status Categorical (Weekend (0) or a Weekday(1))

Day of week Categorical Sunday, Monday â€¦. Saturday

Load Type Categorical Light Load, Medium Load, Maximum Load

Class Labels

Light Load, Medium Load, Maximum Load