## HTTP DATASET CSIC 2010

**Dataset files**

* [Normal Traffic (Training)](http://www.isi.csic.es/dataset/normalTrafficTraining.rar)
* [Normal Traffic (Test)](http://www.isi.csic.es/dataset/normalTrafficTest.rar)
* [Anomalous Traffic (Test)](http://www.isi.csic.es/dataset/anomalousTrafficTest.rar)

**Dataset entry description**

Each entry is a web requests, for e.g.

POST http://localhost:8080/tienda1/publico/entrar.jsp HTTP/1.1

User-Agent: Mozilla/5.0 (compatible; Konqueror/3.5; Linux) KHTML/3.5.8 (like Gecko)

Pragma: no-cache

Cache-control: no-cache

Accept: text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,text/plain;q=0.8,image/png,\*/\*;q=0.5

Accept-Encoding: x-gzip, x-deflate, gzip, deflate

Accept-Charset: utf-8, utf-8;q=0.5, \*;q=0.5

Accept-Language: en

Host: localhost:8080

Cookie: JSESSIONID=EE6A06EE8174FC2B05432AA5A0FCE998

Content-Type: application/x-www-form-urlencoded

Connection: close

Content-Length: 33

**Labelling field**

No labelling field as we will only use the normal dataset for training phase. Anything not belonging to this set is anomalous.

**Questions to answer using ML**

1. A given web request is normal or anomalous?