Water Treatment Plant Data Set

Download: Data Folder, Data Set Description

Abstract: Multiple classes predict plant state

Data Set Characteristics:	Multivariate	**Number of Instances:**	527	**Area:**	Physical
Attribute Characteristics:	Integer, Real	**Number of Attributes:**	38	**Date Donated**	1993-06-01
Associated Tasks:	Clustering	**Missing Values?**	N/A	**Number of Web Hits:**	108572

Source:

Creators:

Manel Poch (igte2 '@' cc.uab.es)
Unitat d'Enginyeria Quimica
Universitat Autonoma de Barcelona. Bellaterra. Barcelona; Spain

Donor:

Javier Bejar and Ulises Cortes (bejar '@' lsi.upc.es)
Dept. Llenguatges i Sistemes Informatics;
Universitat Politecnica de Catalunya. Barcelona; Spain

Data Set Information:

This dataset comes from the daily measures of sensors in a urban waste water treatment plant. The objective is to classify the operational state of the plant in order to predict faults through the state variables of the plant at each of the stages of the treatment process. This domain has been stated as an ill-structured domain.

Attribute Information:

All atrributes are numeric and continuous

N. Attrib.
1 Q-E (input flow to plant)
2 ZN-E (input Zinc to plant)
3 PH-E (input pH to plant)
4 DBO-E (input Biological demand of oxygen to plant)
5 DQO-E (input chemical demand of oxygen to plant)
6 SS-E (input suspended solids to plant)
7 SSV-E (input volatile supended solids to plant)
8 SED-E (input sediments to plant)

9 COND-E (input conductivity to plant)
10 PH-P (input pH to primary settler)
11 DBO-P (input Biological demand of oxygen to primary settler)
12 SS-P (input suspended solids to primary settler)
13 SSV-P (input volatile supended solids to primary settler)
14 SED-P (input sediments to primary settler)
15 COND-P (input conductivity to primary settler)
16 PH-D (input pH to secondary settler)
17 DBO-D (input Biological demand of oxygen to secondary settler)
18 DQO-D (input chemical demand of oxygen to secondary settler)
19 SS-D (input suspended solids to secondary settler)
20 SSV-D (input volatile supended solids to secondary settler)
21 SED-D (input sediments to secondary settler)
22 COND-D (input conductivity to secondary settler)
23 PH-S (output pH)
24 DBO-S (output Biological demand of oxygen)
25 DQO-S (output chemical demand of oxygen)
26 SS-S (output suspended solids)
27 SSV-S (output volatile supended solids)
28 SED-S (output sediments)
29 COND-S (output conductivity)
30 RD-DBO-P (performance input Biological demand of oxygen in primary settler)
31 RD-SS-P (performance input suspended solids to primary settler)
32 RD-SED-P (performance input sediments to primary settler)
33 RD-DBO-S (performance input Biological demand of oxygen to secondary settler)
34 RD-DQO-S (performance input chemical demand of oxygen to secondary settler)
35 RD-DBO-G (global performance input Biological demand of oxygen)
36 RD-DQO-G (global performance input chemical demand of oxygen)
37 RD-SS-G (global performance input suspended solids)
38 RD-SED-G (global performance input sediments)

Relevant Papers:

J. De Gracia. ``Avaluacio de tecniques de classificacio per a la gestio de Bioprocessos: Aplicacio a un reactor de fangs activats'' Master Thesis. Dept. de Quimica. Unitat d'Enginyeria Quimica. Universitat Autonoma de Barcelona. Bellaterra (Barcelona). 1993.

J. Bejar, U. Cort\'es and M. Poch. "LINNEO+: A Classification Methodology for Ill-structured Domains". Research report RT-93-10-R. Dept. Llenguatges i Sistemes Informatics. Barcelona. 1993.
[Web Link]

Ll. Belanche, U. Cortes and M. S\`anchez. "A knowledge-based system for the diagnosis of waste-water treatment plant". Proceedings of the 5th international conference of industrial and engineering applications of AI and Expert Systems IEA/AIE-92. Ed Springer-Verlag. Paderborn, Germany, June 92.
[Web Link]

Citation Request:

Please refer to the Machine Learning Repository's citation policy