Forest fires data set (forestfires.csv)

Source:
P. Cortez and A. Morais. A Data Mining Approach to Predict Forest Fires using Meteorological Data. In J. Neves, M. F. Santos and J. Machado Eds., New Trends in Artificial Intelligence, Proceedings of the 13th EPIA 2007 - Portuguese Conference on Artificial Intelligence, December, Guimaraes, Portugal, pp. 512-523, 2007. APPIA, ISBN-13 978-989-95618-0-9.

This data considers forest fire data from the Montesinho natural park, from the Tras-os-Montes northeast region of Portugal. This park contains a high flora and fauna diversity. Inserted within a supra-Mediterranean climate, the average annual temperature is within the range 8 to 12C. The data was collected from January 2000 to December 2003 and it was built using two sources. The first database was collected by the inspector that was responsible for the Montesinho fire occurrences. At a daily basis, every time a forest fire occurred, several features were registered, such as the time, date, the type of vegetation involved, the six components of the FWI system and the total burned area. The second database was collected by the Braganc¸a Polytechnic Institute, containing several weather observations (e.g. wind speed) that were recorded with a 30 minute period by a meteorological station located in the center of the Montesinho Park. The two databases were stored in tens of individual spreadsheets, under distinct formats, and a substantial manual effort was performed to integrate them into a single dataset with a total of 517 entries.

Variables:

month - month of the year: "jan" to "dec"
day - day of the week: "mon" to "sun"
FFMC - FFMC index from the FWI system
DMC - DMC index from the FWI system
DC - DC index from the FWI system
ISI - ISI index from the FWI system
temp - temperature in Celsius degrees
RH - relative humidity in %:
wind - wind speed in km/h
rain - outside rain in mm/m^2
area - the burned area of the forest (in hectares)

The forest Fire Weather Index (FWI) is the Canadian system for rating fire danger and it includes four components here: Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), and Initial Spread Index (ISI). The first three are related to fuel codes: the FFMC denotes the moisture content surface litter and influences ignition and fire spread, while the DMC and DC represent the moisture content of shallow and deep organic layers, which affect fire intensity. The ISI is a score that correlates with fire velocity spread. Although different scales are used for each of the FWI elements, high values suggest more severe burning conditions.