LINK: <https://www.kaggle.com/datasets/fedesoriano/hepatitis-c-dataset>

**About Dataset**

**Context**

The data set contains laboratory values of blood donors and Hepatitis C patients and demographic values like age. The data was obtained from UCI Machine Learning Repository: <https://archive.ics.uci.edu/ml/datasets/HCV+data>

**Content**

All attributes except Category and Sex are numerical.  
Attributes 1 to 4 refer to the data of the patient:  
1) X (Patient ID/No.)  
2) Category (diagnosis) (values: '0=Blood Donor', '0s=suspect Blood Donor', '1=Hepatitis', '2=Fibrosis', '3=Cirrhosis')  
3) Age (in years)  
4) Sex (f,m)  
Attributes 5 to 14 refer to laboratory data:  
5) ALB  
6) ALP  
7) ALT  
8) AST  
9) BIL  
10) CHE  
11) CHOL  
12) CREA  
13) GGT  
14) PROT

The target attribute for classification is Category (2): blood donors vs. Hepatitis C patients (including its progress ('just' Hepatitis C, Fibrosis, Cirrhosis).