

Universidade de São Paulo
Instituto de Ciências Matemáticas e Computação de São Carlos
SCC0530 - Inteligência Artificial

Trabalho prático 2 – Parte 1

Docente:
Solange Oliveira Rezende

Estagiário PAE:
Renan de Padua

Alunos:
Luan Gustavo Orlandi (8531532)
Maiser José Alves Oliva (6309717)
Hamilton Gonçalves Junior (1053763)
Wesley Tiozzo (8077925)

Abril de 2016



Levantamento da base de dados

A base de dados obtida trata sobre o “Consumo de álcool por estudantes das disciplinas de matemática e idioma português do ensino médio (secondary school ou high school)”. A base foi obtida no seguinte repositório:

<http://archive.ics.uci.edu/ml/datasets/STUDENT+ALCOHOL+CONSUMPTION>

Descrição dos campos e dados

São 32 campos (atributos) relacionados na base de dados. A descrição de cada campo e os valores que eles podem receber estão em inglês para não perdermos a referência do conteúdo da base de dados, mas, serão efetivamente discutidos em português durante o decorrer das partes do trabalho.

Nº Nome - Descrição (Valores possíveis)

- 1 school - student's school (binary: 'GP' - Gabriel Pereira or 'MS' - Mousinho da Silveira)
- 2 sex - student's sex (binary: 'F' - female or 'M' - male)
- 3 age - student's age (numeric: from 15 to 22)
- 4 address - student's home address type (binary: 'U' - urban or 'R' - rural)
- 5 famsize - family size (binary: 'LE3' - less or equal to 3 or 'GT3' - greater than 3)
- 6 Pstatus - parent's cohabitation status (binary: 'T' - living together or 'A' - apart)
- 7 Medu - mother's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education or 4 - higher education)
- 8 Fedu - father's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education or 4 - higher education)
- 9 Mjob - mother's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
- 10 Fjob - father's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
- 11 reason - reason to choose this school (nominal: close to 'home', school 'reputation', 'course' preference or 'other')
- 12 guardian - student's guardian (nominal: 'mother', 'father' or 'other')
- 13 traveltime - home to school travel time (numeric: 1 - <15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - >1 hour)
- 14 studytime - weekly study time (numeric: 1 - <2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - >10 hours)
- 15 failures - number of past class failures (numeric: n if $1 \leq n < 3$, else 4)
- 16 schoolsup - extra educational support (binary: yes or no)
- 17 famsup - family educational support (binary: yes or no)
- 18 paid - extra paid classes within the course subject (Math or Portuguese) (binary: yes or no)
- 19 activities - extra-curricular activities (binary: yes or no)
- 20 nursery - attended nursery school (binary: yes or no)
- 21 higher - wants to take higher education (binary: yes or no)
- 22 internet - Internet access at home (binary: yes or no)
- 23 romantic - with a romantic relationship (binary: yes or no)

24 famrel - quality of family relationships (numeric: from 1 - very bad to 5 - excellent)
25 freetime - free time after school (numeric: from 1 - very low to 5 - very high)
26 goout - going out with friends (numeric: from 1 - very low to 5 - very high)
27 Dalc - workday alcohol consumption (numeric: from 1 - very low to 5 - very high)
28 Walc - weekend alcohol consumption (numeric: from 1 - very low to 5 - very high)
29 health - current health status (numeric: from 1 - very bad to 5 - very good)
30 absences - number of school absences (numeric: from 0 to 93)

these grades are related with the course subject, Math or Portuguese:

31 G1 - first period grade (numeric: from 0 to 20)
31 G2 - second period grade (numeric: from 0 to 20)
32 G3 - final grade (numeric: from 0 to 20, output target)

Possíveis conhecimentos úteis a serem retirados da base

Qual a relação entre o número de horas de estudo e o consumo de álcool pelos estudantes?

A escolaridade dos pais tem influência no consumo de álcool pelos estudantes?

O nível de frequência de saídas com os amigos influencia no consumo de álcool?

O tempo livre, ou seja, sem aula, tem impacto no consumo de álcool?

A qualidade do relacionamento com a família tem impacto no consumo de álcool?

O número de faltas (absences) na escola tem a ver com o consumo de álcool?

O tempo pra chegar até a escola tem influência no consumo de álcool?

Estudantes da área urbana consomem mais álcool do que estudantes da área rural?

Motivação

Será que o consumo de álcool por estudantes do Ensino Médio (Secondary School) tem influência em seu desempenho escolar e na vida social?

Referência:

Using Data Mining To Predict Secondary School Student Alcohol Consumption.

Fabio Pagnotta, Hossain Mohammad Amran

Department of Computer Science, University of Camerino