

Benemérita Universidad Autónoma de Puebla  
Facultad de Ciencias de la Computación

# Máquinas de aprendizaje

## Reporte: Ejemplo de árboles de decisión



# BUAP

**Docente: Abraham Sánchez López**

**Alumno**

Taisen Romero Bañuelos

**Matrícula**

202055209



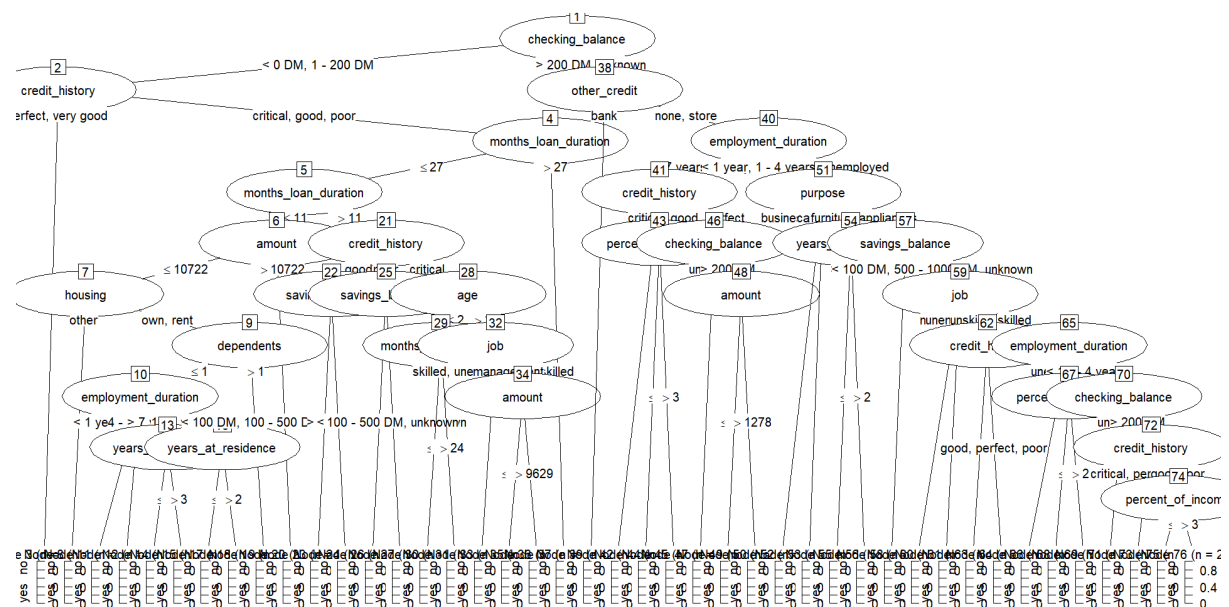
Segunda rama principal:

```

checking_balance in {> 200 DM,unknown}:
...other_credit = bank: yes (50/35)
  other_credit in {none,store}:
...employment_duration in {> 7 years,4 - 7 years}:
  ...credit_history = very good: no (0)
  : credit_history = poor:
  :
  : ...percent_of_income <= 3: no (6)
  : : percent_of_income > 3: yes (7/4)
  : credit_history in {critical,good,perfect}:
  : ...checking_balance = unknown: no (139/3)
  : checking_balance = > 200 DM:
  : : ...amount <= 1278: yes (3/1)
  : : amount > 1278: no (14)
employment_duration in {< 1 year,1 - 4 years,unemployed}:
...purpose in {business,renovations}: yes (24/16)
  purpose in {car,car0}: no (64/5)
  purpose = education:
  ...years_at_residence <= 2: yes (6/3)
  : years_at_residence > 2: no (7)
  purpose = furniture/appliances:
  ...savings_balance in {> 1000 DM,100 - 500 DM}: no (14)
  savings_balance in {< 100 DM,500 - 1000 DM,unknown}:
  ...job = management: yes (7/3)
  job = unemployed: no (1)
  job = unskilled:
  ...credit_history = critical: no (5)
  : credit_history in {good,perfect,poor,
  : : very good}: yes (11/6)
  job = skilled:
  ...employment_duration = unemployed: yes (2/1)
  employment_duration = < 1 year:
  ...percent_of_income <= 2: yes (6/2)
  : percent_of_income > 2: no (7)
  employment_duration = 1 - 4 years:
  ...checking_balance = unknown: no (33)
  checking_balance = > 200 DM:
  ...credit_history in {critical,perfect,poor,
  : : very good}: yes (1)
  : credit_history = good:
  ...percent_of_income <= 3: no (6)
  : percent_of_income > 3: yes (2/1)

```

## Ploteo:



Aunque no es el ploteo más elegante que se podría tener, es de ayuda para visualizar mejor los datos que aparecen en consola (aunque realmente por sí mismos ya son entendibles).

Ahora, hay un punto muy importante del que hablar y que no se mencionó en el PDF, ¿Cómo definimos con precisión los costos que se le deben asignar a cada variable?, digo, resulta evidente que una variable va a tener más costo que otra, pero cómo determinamos ese costo/proporción con relación a las otras variables. Supongo que ese será un tema para preguntar o investigar. Como es algo sencillo lo investigué y obtuve la siguiente respuesta: “calcular el costo variable unitario y multiplicarlo por la cantidad producida o vendida”.