Benemérita Universidad Autónoma de Puebla Facultad de Ciencias de la Computación

Máquinas de aprendizaje

Reporte: Ejemplo de árboles de decisión



Docente: Abraham Sánchez López

Alumno Matrícula

Ejemplo de árboles de decisión

La piedra angular de esta práctica fue el boosting, pero no el primero que se presentó, si no el que consideraba los costos. Aunque no se obtengan métricas de precisión generales tan buenas, creo que es importante no confundir el medio con la meta y por ende, si lo que se quiere reducir son las pérdidas, veo acertado hacer ese "sacrificio" de la precisión general por una mejor precisión en los valores que queremos considerar.

Como dato curioso, el último árbol generado (el que considera los costos) tiene un ploteo más entendible a la vista, todavía se ve encimado pero complementando con el comando summary() se puede tener un muy buen entendimiento del grafo.

Primera rama principal del árbol:

```
Decision tree:
checking_balance in {< 0 DM,1 - 200 DM}:
  ...credit_history in {perfect,very good}: yes (59/16) credit_history in {critical,good,poor}: ...months_loan_duration > 27: yes (102/40) months_loan_duration <= 27:
            :...savings_balance in {< 100 DM,100 - 500 DM,500 - 1000 DM,
                     unknown}: yes (
: savings_balance = > 1000 DM: no (6)
credit_history = poor:
                                                              unknown}: yes (162/91)
                        :...savings_balance in {< 100 DM,> 1000 DM,

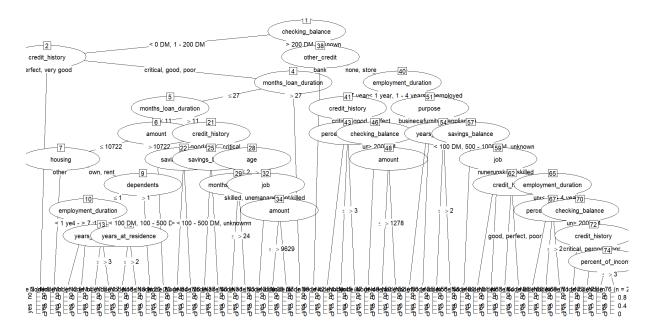
:... 500 - 1000 DM}: yes (13/7)

: savings_balance in {100 - 500 DM,unknown}: no (9)
                       credit_history = critical:
                        :...age <= 28:
:...months_loan_duration <= 24: no (16)
                                    months_loan_duration > 24: yes (1)
                              :...job in {skilled,unemployed,unskilled}: yes (35/21)
                                    job = management:
:...amount <= 9629: no (9)
amount > 9629: yes (1)
                  months loan duration <= 11:
                  :...amount > 10722: yes (2)
amount <= 10722:
                        amount <= 10/22:
:...housing = other: yes (4/2)
housing in {own,rent}:
:...dependents > 1: no (12)
dependents <= 1:
                                     :...employment_duration in {< 1 year,
                                           : unemployed}: yes (14/10)
employment_duration = 4 - 7 years: no (13)
employment_duration = > 7 years:
...vears at resident
                                            :...years_at_residence <= 3: yes (3/2)
                                                  years_at_residence > 3: no (9)
                                           employment_duration = 1 - 4 years:
:...years_at_residence <= 2: no (8)</pre>
                                                  years_at_residence > 2: yes (7/4)
```

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}:
               employment_duration in {< 1 year,1 - 4 years,unemployed}:
...purpose in {business,renovations}: yes (24/16)
    purpose in {car,car0}: no (64/5)</pre>
                     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".