Setul de date contine informatii personale, educationale si profesionale ale diferitor angajati. Obiectivul setului de date este clasificarea BINARA a angajatilor in categoriile de a castiga peste sau sub $50K per an (<=50K / >50K).

Atributele exemplelor sunt urmatoarele:

|  |  |  |
| --- | --- | --- |
| Nume atribut | Tip de date | Semnificatie |
| fnl | numeric | Caracteristica socio-economica a populatiei din care provine individul |
| hpw | numeric | Numar de ore de munca pe saptamana |
| relation | categoric | Tipul de relatie in care este implicat individul |
| gain | numeric | Castigul de capital |
| country | categoric | Tara de origine |
| job | categoric | Meseria individului |
| edu\_int | numeric | Numarul de ani de studiu |
| years | numeric | Varsta individului |
| loss | numeric | Pierderea de capital |
| work\_type | categoric | Tipul de meserie |
| partner | categoric | Tipul de partener pe care il are individul |
| edu | categoric | Tipul de educatie al individului |
| gender | categoric | Genul individului |
| race | categoric | Rasa individului |
| prod | numeric | Producerea de capital |
| gtype | categoric | Tipul contractului de munca |