

Zadanie. Klasyfikacja - scoring kredytowy w oparciu o algorytm CHAID

Charakterystyka danych w arkuszu *tree_credit.sav*:

| | |
|----------------------|---|
| Credit_rating | Credit rating: 0=Bad, 1=Good, 9=missing values (zm. docelowa/predykcyjna) |
| Age | Age in years |
| Income | Income level: 1=Low, 2=Medium, 3=High |
| Credit_cards | Number of credit cards held: 1=Less than five, 2=Five or more |
| Education | Level of education: 1=High school, 2=College |
| Car_loans | Number of car loans taken out: 1=None or one, 2=More than two |

1. Wstaw odpowiedni węzeł do strumienia
2. Ustaw odpowiednio zmienną predykcyjną
3. Wstaw węzeł modelujący CHAID
 - w zakładce "Zmienne" powinno być wybrane ustawienie "Użyj wstępnie zdefiniowanych ról"
 - w zakładce "Budowanie":
 - "Zbudować pojedyncze drzewo" -> "Utwórz gotowy model"
 - w zakładce "Kryteria zatrzymania":
 - "Wartość bezwzględna" ("Minimum rekordów w gałęzi nadrzędnej"=400, "Minimum rekordów w gałęzi podrzędnej"=200)
 - uruchom węzeł
4. Oglądnij model i utwórz dokument Word'a w którym wkleisz zrzut ekranu otrzymanego drzewa i odpowiesz na poniższe pytania:
 1. w postaci regułowej: *które predyktory są najsilniejsze? jakie są różnice m. nimi? które nie mają znaczenia?*
 2. w postaci drzewa:
 - *jak wygląda węzeł 2 (low income) i czy wszystko jest w porządku?*
 - *podobnie węzeł 1 (high income)? jak zmienia się dokładność po dalszym podziale tego węzła?*
 - *co z dokładnością w węźle 3 (medium income)? co zmienia dalszy podział tego węzła?*
5. Zapisz strumień, zapisz plik z odpowiedziami, spakuj i umieść jako odpowiedź na zadanie ZAD_CHAID w Moodle.