Aula8 – Regras de Associação (Subgroupos).

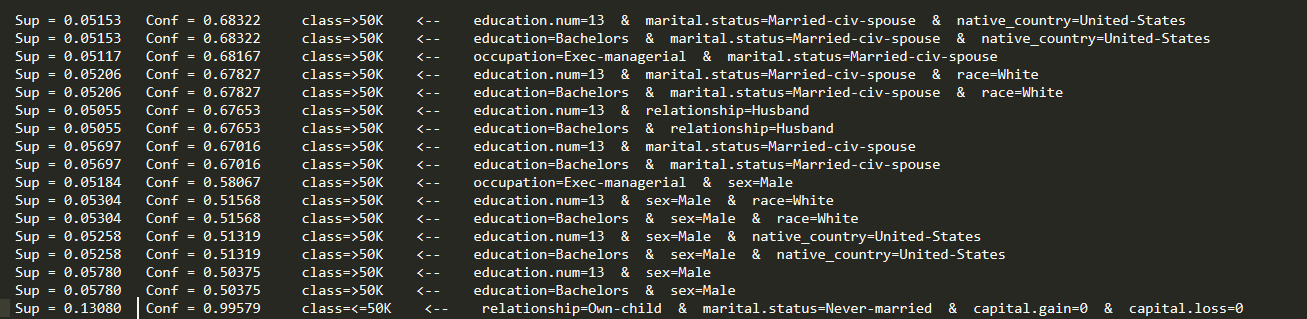
1. **Considere o dataset adult (no blackboard). Usando o Caren, apresente regras que caracterize a pessoas com rendimentos anuais superiores a 50000 USD.**

***Java caren adult.data 0.05 0.5 -Att -Hclass -s, -fisher -null?***

***-chi para teste qui quadrado***

***-imp0.000001 para importância***

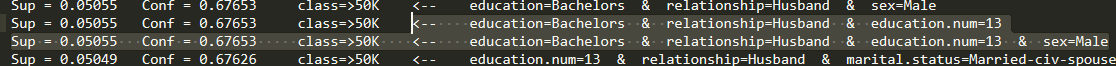
***-fisher para aplicar Fisher…***



1. **Apresente exemplos de regras redundantes, produtivas e significativas para rendimentos <= 50K dólares.**



Adicionar o atributo *education.num* não aumentou a confiança da regra gerada.



A nível de contexto, se o atributo *relationship* toma o valor de Marido, é redundante ter uma regra com igual confiança que apenas acrescenta que o atributo *sex* toma o valor Homem.

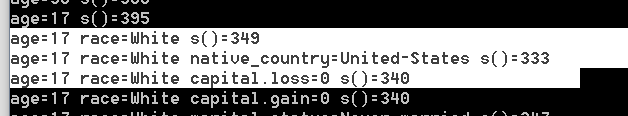
1. **Usando o switch –is do Caren para gerar a contagem dos termos frequentes, apresente exemplos de “closed patterns” e de “non-closed”.**

***Java caren adult.data 0.01 0.5 -s, -ovrt -is -Att -Hclass***

***more item.txt***



O conjunto {capital.gain} é não fechado, porque existe o conjunto {capital.gain, marital.status} com o mesmo número de ocorrências. Ou seja, existe um super conjunto imediato com o mesmo suporte que o conjunto inicial.



O conjunto {age, race} é fechado, porque não existe nenhum super conjunto com o mesmo número de ocorrências s() que ele.

Isto é, ao especificar/afunilar o conjunto {age, race} com mais atributos, o número de ocorrências desses conjuntos tem suporte inferior ao conjunto inicial.

1. **Apresente dois “subgrupos” diferentes que descrevam subpopulações com rendimento superior a 50K. Analise a sua incidência e sua interpretação.**