

The background is a dark, textured surface with faint, light-colored sketches of various scientific and mathematical concepts. These include a globe in the upper left, a telescope on the left side, a stack of books at the bottom left, a large percentage sign (%) at the bottom right, and various geometric shapes and symbols scattered throughout. The overall theme is academic and research-oriented.

Uso do método apriori na descoberta de regras de associação entre indicadores de propriedades rurais no RS

Autor: Luciano Moraes Da Luz Brum

Objetivo

- Aplicar o método apriori para a descoberta de regras de associação entre diferentes indicadores de dimensões:
 - Social;
 - Econômica;
 - Ambiental;
 - Produtiva;

Metodologia Adotada

Etapas de pré-processamento dos dados (Limpeza, Integração, Seleção e Transformação).

Limpeza e organização manual da planilha .csv

Carga dos dados da planilha para a ferramenta R.

Pré-processamento dos dados (Strings em n°s)

Separação dos dados em categorias por frequência.

Etapa de Mineração.

Uso do método Apriori

Passo a Passo

- Foi utilizada a ferramenta R para o processo de limpeza, categorização e mineração de dados (regras de associação).
- Das instruções utilizadas no R:

```
library("arules");  
library("arulesViz");  
x = read.csv("teste3.csv", stringsAsFactors=FALSE)  
x2=x  
  
for(i in 1:nrow(x2)){  
  for(j in 2:38){  
    x2[i,j] = as.numeric(gsub("[,]", ".", x[i,j]))  
    x2[i,j] = as.numeric(paste(x2[i,j]))  
  }  
}
```

Passo a Passo

- `for (i in 2:38){x2[,i]<-discretize(as.numeric(x2[,i]),categories=10)}`
- `x3 = as(x2[,2:38], "transactions");`
- `rules = apriori(x3, parameter=list(minlen=2, maxlen=2, support=0.01, confidence=0.6),appearance = list(lhs=c("MELHORAMENTO.GENÉTICO =[0.815,0.850]"), default="rhs"))`
- `inspect(head(sort(rules,by="confidence"),21))`
- `rules = apriori(x3, parameter=list(minlen=2, maxlen=2, support=0.01, confidence=0.6),appearance = list(lhs=c("GESTÃO.ECONÔMICA=[0.9,1.0]"), default="rhs"))`
- `inspect(head(sort(rules,by="confidence"),21))`

Conclusões: 1º caso: Produtores que têm alto nível de Gestão econômica [0,9 –1,0]:

- Têm Alta renda familiar per capita [0,9 –1,0] (14%, 100%);
- Têm Facilidade na atividade [0,9 –1,0] (14%, 100%);
- Têm Bom intervalo entre partos [0,9 –1,0] (14%, 100%);
- Têm Alta produção diária de leite [0,925 –1,0] (12%, 83%);
- Têm Diversificação de forrageiras [0,925 –1,000] (12%, 83%);
- Possuem bom nível de Sanidade do Rebanho [0,9 –1,0] (9,5%, 67%);
- Têm bom nível de Precocidade das novilhas [0,97 –1,00] (9,5%, 67%);
- Fazem Manejo da Pastagem [0,9 –1,0] (9,5%, 67%);
- Fazem Conservação do Solo [0,9 –1,0] (9,5%, 67%);

Conclusões: 2º caso: Produtores que têm alto nível de Melhoramento Genético [0,815 –0,850]:

- Têm baixo nível em Sanidade do Rebanho [0,0 –0,1] (38%, 100%);
- Têm baixo nível em Controle Leiteiro [0,0 –0,1] (38%, 100%);
- Têm Bom intervalo entre partos [0,9 –1,0] (38%, 100%);
- Têm Facilidade na atividade [0,9 –1,0] (36%, 94%);
- Têm baixo nível de Gestão Econômica [0,0 –0,1] (33%, 88%);
- Têm Alta renda familiar per capita [0,9 –1,0] (29%, 75%);
- Fazem bom Manejo da Pastagem [0,9 –1,0] (29%, 75%);
- Têm boa Diversificação de forrageiras [0,925 –1,000] (29%, 75%);
- Têm Alta produção diária de leite [0,925 –1,0] (26%, 69%);

Referências Bibliográficas

- AGRAWAL, R; SRIKANT, R. Fast algorithms for mining association rules. *In The 20th International Conference on Very Large Databases*, pages 487–499, 1994.
- BODON, F. A fast Apriori implementation. *In Proceedings of the IEEE ICDM Workshop on Frequent Itemset Mining Implementations*, Melbourne, Fl. B. CEUR Workshop Proceedings, vol. 90, 2003.
- CARDOSO, O. N. P; MACHADO, R. T. M. Gestão do conhecimento usando data mining: estudo de caso na Universidade Federal de Lavras. *Revista Administração Pública*, v. 42, n. 3, p. 495-528, 2008.
- FAYYAD, U. M.; PIATETSKY-SHAPIO, G.; SMYTH, P.; UTHURUSAMY, R. Advances in Knowledge Discovery and Data Mining. *AI Magazine*, v. 17 n° 3, 1996.
- HAN, J.; KAMBER, M. *Data Mining: Concepts and Techniques*. 2° ed. Morgan Kaufmann Publishers, p. 5–7, 2006.