

Trabalho - Etapa 4

Resumo

O objetivo é desenvolver uma aplicação para avaliar e prever o risco de envolvimento de um usuário com drogas ilícitas a partir de seus dados de perfil, traços de personalidade obtidos através de testes psicológicos (NEO-FFI-R¹, BIS-11² e ImpSS³) e histórico de consumo de drogas lícitas e ilícitas. O resultado da avaliação de risco será um valor em porcentagem da chance estimada de o usuário se envolver com qualquer droga ilícita, assim como uma porcentagem específica para cada droga ilícita de um conjunto pré-determinado. Para tal, utiliza-se técnicas de Machine Learning e uma base de dados de 1885 instâncias resultados de uma pesquisa de consumo de drogas.

Adequação de Modelo para XQuery

Na etapa 2, desenvolvemos um modelo conceitual e lógico que foi usado na construção de um banco de dados relacional. Nesta etapa, temos como objetivo a criação de um banco de dados em XML para fazer consultas com XQuery. Para isso, alteramos o modelo da etapa 2:

- Consumo era uma entidade da qual as demais drogas herdavam. Como cada droga para cada usuário possui apenas um atributo de frequência, essa herança foi desfeita. Consumo passa a ser uma entidade com vários atributos: a frequência de uso de cada droga.
- Para a composição do XML, estabelece-se a hierarquia em que o usuário (ou "pessoa") contém seus atributos pessoais (idade, etnia, etc.), o atributo "consumo", que concentra dentro de si os atributos das drogas, e o atributo "personalidade", que concentra dentro de si os atributos de traços de personalidade.

Convertemos o arquivo ".csv" da base original de dados em um arquivo .xml com a ferramenta online CSV to XML Converter (disponível [aqui](#)). O código utilizado na ferramenta

¹ Este exame avalia os seguintes traços de personalidade:

- O Neuroticismo (N) é uma tendência de longo prazo de experimentar emoções negativas, como nervosismo, tensão, ansiedade e depressão;
- A Extroversão (E) manifesta-se em extrovertida, calorosa, ativa, assertiva, falante, alegre e em busca de características de estimulação;
- A abertura à experiência (O) é uma apreciação geral da arte, de ideias incomuns e de interesses imaginativos, criativos, não convencionais e amplos;
- A Amabilidade (A) é uma dimensão das relações interpessoais, caracterizada pelo altruísmo, confiança, modéstia, bondade, compaixão e cooperação;
- A Consciência (C) é uma tendência da pessoa ser organizada e confiável, obstinada, persistente, confiável e eficiente.

² Este exame avalia impulsivity.

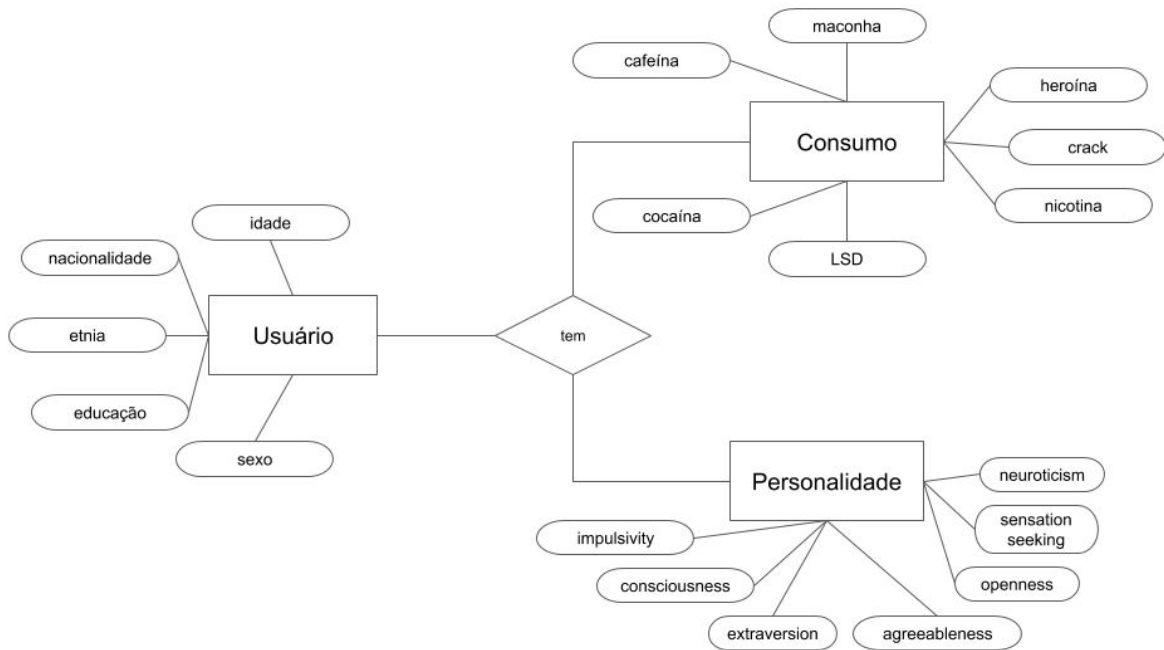
³ Este exame avalia sensation seeking.

consiste em especificar qual coluna (número da coluna) fornece o dado que preenche o valor do atributo do XML no formato especificado, como segue abaixo:

```
< Pessoa id="##1##">
  < age>##2##</age>
  < gender>##3##</gender>
  < education>##4##</education>
  < country>##5##</country>
  < ethnicity>##6##</ethnicity>
  < personalidade>
    < nscore>##7##</nscore>
    < escore>##8##</escore>
    < oscore>##9##</oscore>
    < ascore>##10##</ascore>
    < cscore>##11##</cscore>
    < impulsive>##12##</impulsive>
    < SS>##13##</SS>
  </personalidade>
  < consumo>
    < alcohol>##14##</alcohol>
    < amphet>##15##</amphet>
    < amyl>##16##</amyl>
    < benzos>##17##</benzos>
    < caff>##18##</caff>
    < cannabis>##19##</cannabis>
    < choc>##20##</choc>
    < coke>##21##</coke>
    < crack>##22##</crack>
    < ecstasy>##23##</ecstasy>
    < heroin>##24##</heroin>
    < ketamine>##25##</ketamine>
    < legalh>##26##</legalh>
    < lsd>##27##</lsd>
    < meth>##28##</meth>
    < mushrooms>##29##</mushrooms>
    < nicotine>##30##</nicotine>
    < semer>##31##</semer>
    < vsa>##32##</vsa>
  </consumo>
</Pessoa>
```

Modelo Conceitual

- ER:



Modelo Lógico

