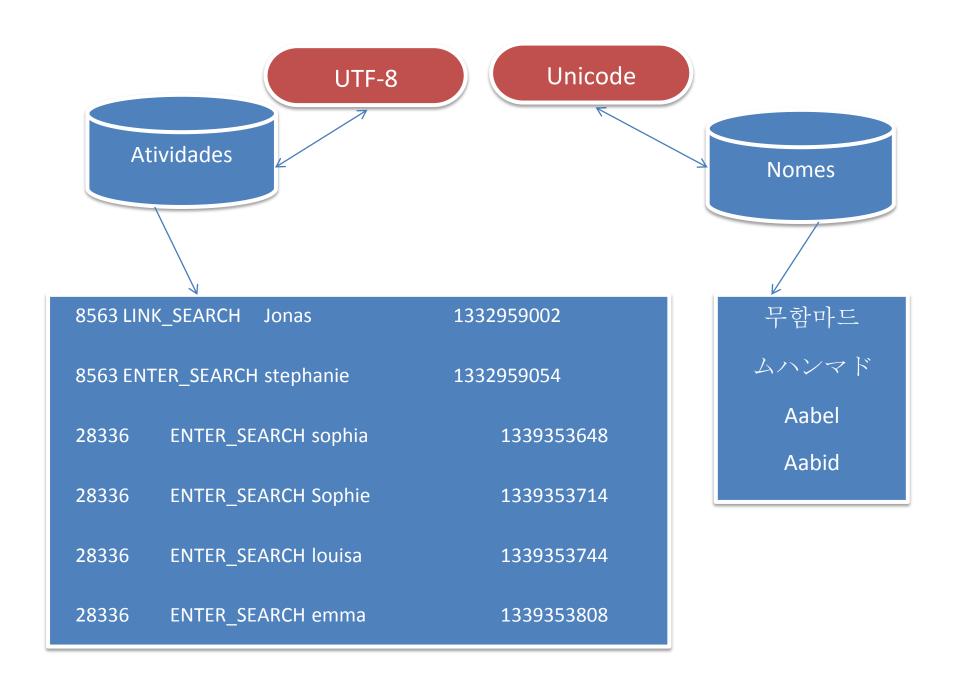
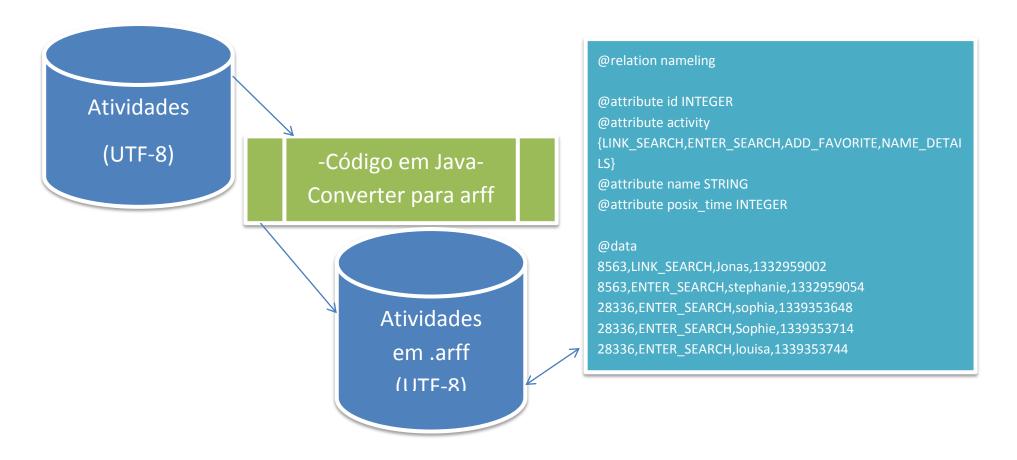
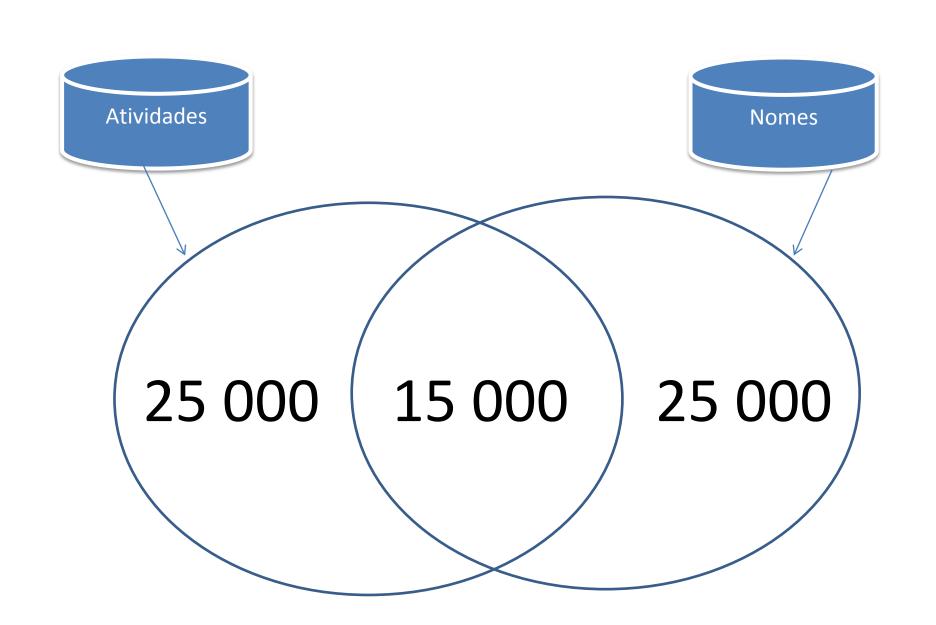
Desafio dos Nomes

Érick Oliveira





→ Os algoritmos de conversão para arquivos de texto de codificação diferentes da ANSI não funcionam bem na Weka. Ou não funcionam definitivamente. Possibilidade descartada.



Atividades

Processar base retirando todas as tuplas que contém nomes que não estão na base de nomes

Base de Atividades (com aprox 15 mil nomes distintos)

Base de Atividades (com aprox 15 mil nomes distintos) (aprox 315 000 tuplas) 28336,ENTER_SEARCH,sophia,1339353648 28336,ENTER_SEARCH,Sophie,1339353714 28336,ENTER_SEARCH,louisa,1339353744

Por usuário, calcular nomes que apareçam conjuntamente

Sophia<->Sophie,1
Sophie<->Sophia,1
Sophia<->Louisa,1
Louisa<->Sophia,1
Sophie<->Louisa,1
Louisa<->Sophie,1

2,ENTER_SEARCH,sophia,1339353648

Sophia<->Sophie,1

Sophie<->Sophia,1

Sophia<->Louisa,1

Louisa<->Sophia,1

Sophie<->Louisa,1

Louisa<->Sophie,1

Louisa<->Ana,2

Ana<->Maria,10

Ana<->Emma,12

Sophie, Louisa, Ana, Emma, Maria...

- Tempo de processamento
 - → Segunda abordagem