

## **Biclusterização aplicada em Sistemas de Recomendação baseados em Conteúdo textual**

**Aluno:** Lucas Fernandes Brunialti, **Orientador:** Dr<sup>a</sup>. Prof<sup>a</sup>. Sarajane Marques Peres

### **Contexto**

Biclusterização é uma técnica com potencial para encontrar grupos de objetos similares considerando a correlação parcial existente entre os mesmos. Isto significa que tal técnica é capaz de detectar um subconjunto de características que melhor representa um determinado agrupamento de dados. Essa capacidade é potencialmente útil para Sistemas de Recomendação baseados em Conteúdo nas quais se faz necessária a sugestão de itens úteis, atrativos e principalmente diversificados.

### **Objetivo**

Analisar propriedades de algoritmos de biclusterização para verificar a aderência dessa estratégia como ferramenta de decisões sobre recomendações de conteúdo textual, e a partir de tal análise, propor adaptações aos algoritmos de forma a adequá-los ao contexto de recomendação de notícias.

### **Método**

Aplicação de algoritmos de biclusterização aplicados sobre uma representação vetorial de um conjunto de dados textual extraído de um portal de notícias. Como método de avaliação é planejado aplicar medidas clássicas de avaliação de clusterização e de recomendação, em conjunto com estratégias de avaliação baseadas em informações reais de comportamento de usuários no portal de notícias.

### **Resultados**

Dentro do contexto de Sistemas de Recomendação, tradicionalmente, é esperado que itens úteis e adequados às preferências do usuário sejam recomendados. Entretanto, um potencial problema no contexto de Sistemas de Recomendação baseados em Conteúdo é a recomendação de itens previsíveis, sem a potencialidade de trazer conteúdo novo (interessante) para o usuário. Esse problema é conhecido como serendipidade. Então, é esperado que a aplicação de algoritmos de biclusterização contribua para melhorar a capacidade de recomendar itens considerando a otimização da propriedade de serendipidade.

### **Conclusões**

Por meio da análise realizada em uma revisão sistemática, constatou-se que algoritmos de clusterização são pouco explorados no contexto de Sistemas de Recomendação baseados em Conteúdo textual. Contudo, técnicas de clusterização têm sido empregadas com sucesso para análise de conteúdo textual em outras áreas de aplicação. E, em face da necessidade do problema de serendipidade e da hipótese de que considerar similaridades parciais pode ajudar a otimizá-la, conclui-se que o estudo de adequabilidade de biclusterização no contexto de Sistemas de Recomendação baseados em Conteúdo textual é uma questão em aberto com potencial de produzir resultados interessantes e adequados para estudo neste projeto.