



data

*cf2vec: Collaborative
Filtering algorithm
selection using graph
distributed
representations*

Juan Valdivia

Advanced Analytics Prod

juan.valdivia@itau-unibanco.com.br

11/10/2021

Agenda

- 1 Introdução
 - Motivação
- 2 *Metalearning*
 - *No-free lunch theorem e viés*
- 3 *Recommender Systems*
- 4 *Representational Learning*
- 5 *cf2vec*

O que o paper tenta resolver?

- Sistemas de recomendação (*RecSys*) ajudam usuários (clientes) a descobrir os itens (produtos) mais relevantes tornando a experiência do usuário mais agradável. *RecSys* ajuda as empresas a entenderem melhor o comportamento de seus usuários, e assim aumentarem suas receitas.
- Collaborative filtering (*CF*): abordagem de *RecSys* que usa as preferências de um grupo de usuários para fazer recomendações ou previsões das preferências desconhecidas por outros usuários.
- Desafio: falta de orientação sobre qual algoritmo de *Machine Learning* seria mais adequado para uma determinada tarefa (por exemplo: *Rating Prediction*).

Possíveis soluções

- Primeira solução: avaliar o desempenho de vários algoritmos e depois escolher o mais adequado. Abordagem de tentativa e erro, sendo custosa e subjetiva.
- Segunda solução: usar *Metalearning* (*MtL*) que é uma abordagem de *ML* que gera metaconhecimento mapeando *metafeatures* para *metatargets*, formando assim um *metadataset*.
- As *metafeatures* são características extraídas do conjunto de dados que podem influenciar o desempenho dos algoritmos de *ML*.
- As *metatargets* mostram o desempenho de um grupo de algoritmos num conjunto de dados, por exemplo, um ranking de algoritmos. Assim, *MtL* pode ser usado na tarefa de seleção de algoritmos.

Objetivo do paper

- Objetivo: Selecionar algoritmos de *collaborative filtering* usando *Metalearning*.
- Necessário: extrair *metafeatures* de conjuntos de dados de *CF*.
- Trabalhos anteriores: as *metafeatures* extraídas dependiam do conhecimento prévio do cientista de dados (*Rating Matrix metafeatures*, *Subsampling landmarks*, *Graph metafeatures* e *Comprehensive metafeatures*)
- Solução: extrair *metafeatures* usando o mínimo de intervenção humana. Para isso, é usado *Representational Learning*.

no-free lunch theorem e Viés

no-free lunch theorem

Existe um espaço de hipóteses H (conjunto infinito de possíveis funções h). O *no-free lunch theorem* diz que não existe um único classificador $h \in H$ que seja adequado para todas as situações, pois cada algoritmo faz suposições sobre os dados.

Viés

Preferência na escolha da função (hipótese) que deve ser usada. Essa preferência é muito importante para garantir o aprendizado, pois ela garante que se tenha um conjunto finito de funções a serem testadas.

O que seria *Metalearning*?

- *Metalearning* (*MtL*) consiste basicamente em extrair metaconhecimento de tarefas resolvidas com sucesso por Aprendizado de Máquina para usá-lo na solução de novas tarefas [?].
- Abordagem trabalhosa por exigir caracterizar conjuntos de dados, mas bastante objetiva.
- Ao construir um meta-modelo não há a necessidade de treinar múltiplos classificadores para um novo conjunto de dados.
- Além disso, o uso de *MtL* pode evitar o problema do *overfitting* e facilitar a reprodutibilidade de experimentos.
- Dessa forma, *MtL* pode ser usado para selecionar os melhores algoritmos para uma determinada tarefa.

Multiple Columns

Heading

- ① Statement
- ② Explanation
- ③ Example

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer lectus nisl, ultricies in feugiat rutrum, porttitor sit amet augue. Aliquam ut tortor mauris. Sed volutpat ante purus, quis accumsan dolor.

Table

Treatments	Response 1	Response 2
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

Table: Table caption

Theorem

Theorem (Mass–energy equivalence)

$$E = mc^2$$

Example (Theorem Slide Code)

```
\begin{frame}  
\frametitle{Theorem}  
\begin{theorem}[Mass--energy equivalence]  
$E = mc^2$  
\end{theorem}  
\end{frame}
```

Figure

Uncomment the code on this slide to include your own image from the same directory as the template .TeX file.

Figure

Uncomment the code on this slide to include your own image from the same directory as the template .TeX file.

Figure

Uncomment the code on this slide to include your own image from the same directory as the template .TeX file.

Citation

An example of the `\cite` command to cite within the presentation:

This statement requires citation [Smith, 2012].

References



John Smith (2012)

Title of the publication

Journal Name 12(3), 45 – 678.

The End