|  |  |
| --- | --- |
| Programa de Pós-Graduação em Psicologia Tópicos Especiais em Avaliação e Medidas Psicológicas II: **Inteligência Artificial e Processamento de Linguagem Natural aplicada à Psicometria e Avaliação**  Prof. Dr. Ricardo Primi  1o Semestre de 2022 - 2a feira – 9:00-12:00 hs |  |

## Objetivo da disciplina

Ao final da disciplina espera-se que o aluno tenha desenvolvido:

* Conhecimento atualizado sobre métodos recentes de inteligência artificial aplicados à psicometria e avaliação
* Conhecimento de métodos de Deep Learning para processamento de textos
* Conhecimento aplicado de como usar Keras para R e outros pacotes que implmentam Deep Learning no R.
* Autonomia para aplicar métodos de Deep Learning para situações práticas de análise de dados com textos envolvendo avaliação psicológica e educacional
* Conhecimento aprofundado da linguagem R

## Ementa e programa

* Introdução a aprendizagem de máquina, inteligência artificial e processamento de linguagem natural
* Fundamentos das redes neurais.
* Keras para R
* Processamento básico de textos
* Modelos representações vetoriais - word embeddings (bag of words, distributed representaions/word embeddings, desenho das matrizes, ponderação, redução de dimesnionalidade, comparação de similaridade de vetores)
* Deep learning para textos: RNN e LSTM, 1D CNN.
* Representações contextuais: Transformes e BERT

## Avaliação

Faremos quatro exercícios durante o semestre. Será solicitado que você faça parte das análises desses exercícios para entrega

1. Correção automática de criatividade usos alternativos
2. Análise fatorial de palavras com o BFI
3. Predição dos perfis do MBTI a partir de textos de posts no facebook
4. Análise de Tweets sobre Política com a Teoria dos Fundamentos Morais

Teremos algumas leituras com pequenos testes para verificar a compreensão do texto.

## Site da disciplina

Link da aula <https://meet.google.com/xwc-goot-gva>

Discussão do texto: Employing digital footprints in psychometrics

Processamento básico de textos

## Cronograma 2022

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Dia** | **Conteúdo** | **Leituras / Exercícios** |
| 1 | 7/03 | Aula inaugural do Programa de Pós |  |
| 2 | 14/03 | Introdução a inteligência artificial e Processamento de linguagem natural  Gravação da aula  <https://drive.google.com/file/d/1IKZC1hqeB6RT_1tHASfAFv7d3iHAeasP/view?usp=sharing> |  |
| 3 | 21/03 | Discussão do texto: *Employing digital footprints in psychometrics*  Processamento básico de textos  Gravação da aula  <https://drive.google.com/file/d/1ywithRi1F08sXz9HaOVzauAMyFGCcyAY/view?usp=sharing> | Textos  *Education DataScience- Past, Present, Future*  *Employing digital footprints in psychometrics* |
| 4 | 28/03 | Discussão do texto: *Education DataScience- Past, Present, Future*  Principal components e Singular Value Decomposition no R  Gravação da aula  <https://drive.google.com/file/d/14nHQkPGzlXxS53-4L98xQXzUVTeNcjc2/view?usp=sharing> |  |
| 5 | 4/04 | Bancas coletivas de mestrado |  |
| 6 | 11/04 | Pré-processamento básico de textos  Fundamentos das redes neurais  Gravação da aula  <https://drive.google.com/file/d/1671BLaBzE1fDW-h4R_rGcasnIA1UdeIr/view?usp=sharing> |  |
| 7 | 18/04 | Fundamentos das redes neurais  Keras para R  Gravação da aula  <https://drive.google.com/file/d/1O0axJSX81U1JMppHbhLrb9H5LFY6t47n/view?usp=sharing> | Textos  *The Semantic Scale Network: An Online Tool to Detect Semantic Overlap of Psychological Scales and Prevent Scale Redundancies*  *Mining Big Data to Extract Patterns and Predict Real-Life Outcomes* |
| 8 | 25/04 | Exercício 1. Pré-processamento básico de textos  Fundamentos das redes neurais  Introdução ao Keras para R  Gravação da aula  <https://drive.google.com/file/d/1BAUWnD8ZxIzQFgsv3jUszKN8MqnqadMf/view?usp=sharing> | Exercício 1. Correção automática de criatividade usos alternativos |
| 9 | 2/05 | Introdução ao Keras para R  Aprimorando o pré-processamento de textos: steaming, TFIDF e Visualização de tópicos com LDAvis  <https://drive.google.com/file/d/1i1Eb10PJ-yWHXK380V3Wq7NWlFjBx4jp/view?usp=sharing> |  |
| 10 | 9/05 | Discussão dos textos  Gravação da aula  <https://drive.google.com/file/d/1sQX9XS6WeS_qa5hcMaWHU6-j98YONrir/view?usp=sharing> | Textos  *Latent human traits in the language of social media: An open-vocabulary approach*  *Automatic Personality Assessment Through Social Media Language* |
| 11 | 16/05 | Modelos representações vetoriais  Gravação da aula  <https://drive.google.com/file/d/1eUB8JMuFqzzUaH2TAlylygMJmyXnG5TK/view?usp=sharing> |  |
| 12 | 23/05 | Deep learning para textos: RNN e LSTM, 1D CNN  Gravação da aula  <https://drive.google.com/file/d/1ML3R9BcUJPnXIMeofxNcbxNRm38l3gVf/view?usp=sharing> |  |
| 13 | 30/05 | Deep learning para textos: RNN e LSTM, 1D CNN  Métricas de julgamento  Como treinar um embedding  Gravação da aula  <https://drive.google.com/file/d/1tuBrYb3aiHDkXq6IuC8UW0j1alykT3Kj/view?usp=sharing>  Palestra do Prof. Roberto Lotufo  <https://www.youtube.com/watch?v=hEmpPfkP3TE> |  |
| 14 | 6/06 | Corrigindo o teste de criatividade com redes LSTM, 1D CNN  Representações contextuais: Transformes e BERT  Gravação da aula  <https://drive.google.com/file/d/147W8DkV9FgiloIxVOSbvckns5SBy8mzl/view?usp=sharing> |  |
| 15 | 13/06 | Representações contextuais: Transformes e BERT |  |
|  |  |  | Exercício 4. Análise de Tweets sobre política com a Teoria dos Fundamentos Morais (MFT)  Textos  *Mapping the Moral Domain*  *Compassionate Liberals and Polite Conservatives: Associations of Agreeableness With Political Ideology and Moral Values*  Exercício 2. Análise fatorial de palavras com o BFI  Exercício 3. Predição dos perfis do MBTI a partir de textos de posts no facebook |
|  |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Dia** | **Conteúdo** | **Leituras / Exercícios** |
| 1 | 7/03 | Aula inaugural do Programa de Pós |  |
| 2 | 14/03 | Introdução a inteligência artificial e Processamento de linguagem natural |  |
| 3 | 21/03 | Fundamentos das redes neurais. | Textos |
| 4 | 28/03 | Keras para R |  |
| 5 | 4/04 | Processamento básico de textos |  |
| 6 | 11/04 | Processamento básico de textos | Textos  Exercício 1. |
| 7 | 18/04 | Modelos representações vetoriais |  |
| 8 | 25/04 | Modelos representações vetoriais |  |
| 9 | 2/05 | Modelos representações vetoriais | Textos  Exercício 2. |
| 10 | 9/05 | Deep learning para textos: RNN e LSTM, 1D CNN |  |
| 11 | 16/05 | Deep learning para textos: RNN e LSTM, 1D CNN | Exercício 3. |
| 12 | 23/05 | Representações contextuais: Transformes e BERT | Textos |
| 13 | 30/05 | Representações contextuais: Transformes e BERT |  |
| 14 | 6/06 | Representações contextuais: Transformes e BERT |  |
| 15 | 13/06 | Representações contextuais: Transformes e BERT | Exercício 4. |
|  |  |  |  |
|  |  |  |  |

## Referências

Chollet, F. & Allaire, J. J. (2018). *Deep Learning with R*. Manning: Shelter Island, NY

Hvitfeldt, H. & Silge, J (2021). *Supervised Machine Learning for Text Analysis in R*. CRC Press (<https://smltar.com>)

Kabacoff, R. (2015). *R in action: data analysis and graphics with R*. Manning Publications Co.

Silge, J & Robinson, D. (2017). *Text Mining with R*. CA: O’Reilly Media inc. (<https://www.tidytextmining.com>, <https://github.com/juliasilge/tidytext> ).

Wickham, H., & Grolemund, G. (2017). *R for data science*. Sebastopol, CA: O’Reilly (<http://r4ds.had.co.nz/index.html>

## Sites

<https://r4ds.had.co.nz>

<http://www.cookbook-r.com>

## Anexo: Links das aulas de 2020

|  |  |  |
| --- | --- | --- |
|  | **Dia** | **Conteúdo / Atividades / Exercícios** |
| 1 | 17/8 | Introdução: data science, machine learning e deep learnig  <https://drive.google.com/file/d/1Iy7S4HUu6T8oSynO0Y19rHwIXaXvIkQW/view?usp=sharing> |
| 2 | 24/8 | Instalação R e R-studio  R-markdown  Análise exploratória gráfica: ggplot  Exercício 1: R-markdown  Parte 1  <https://drive.google.com/file/d/1o6j5yk8v-5qaTgCgX8GBHZNEZMHfgfyc/view?usp=sharing>  Parte 2:  <https://drive.google.com/file/d/1BjNHIzpU08iVKgGkkHyPWpukggT3oqGM/view?usp=sharing> |
| 3 | 31/8 | Análise exploratória gráfica: ggplot  Exercício 2: Análise Exploratória de dados  Gravação  Parte 1  <https://drive.google.com/file/d/1vNO5URWQIFunCfFcxkDM76YEAZGzwGz8/view?usp=sharing>  Parte 2  <https://drive.google.com/file/d/1eokl1_ptLk8TvUN5U5YEboQU7aO1S3ga/view?usp=sharing> |
|  | 7/9 | Independência |
| 4 | 14/9 | Análise exploratória gráfica: ggplot (continuação)  Organização transformação de dados: dplyr  Gravação  Parte 1  <https://drive.google.com/file/d/1NUVaYd_kPlAhunuebSlGnnEFHFrs9xWK/view?usp=sharing>  Parte 2  <https://drive.google.com/file/d/1zUBhqEaPrvNSgmhBzfjpxS3Q3fWYQpnp/view?usp=sharing> |
| 5 | 21/9 | Organização transformação e manipulação de dados: dplyr e tidyr  Gravação  Parte 1  <https://drive.google.com/file/d/10WWK-sjdjBk36h6HmbV-tPxUOrWgsc-n/view?usp=sharing>  Parte 2  <https://drive.google.com/file/d/1VHJtm-ENeGmu4nK9At7TjRya7-SY1Kv4/view?usp=sharing>  Parte 3  <https://drive.google.com/file/d/1tEC6z0llsqbUuzE7uA9bkNaacFhgQSdO/view?usp=sharing> |
| 6 | 28/9 | Organização transformação e manipulação de dados:  dplyr e tidyr  Parte 1  <https://drive.google.com/file/d/1567W-86Dx25osezawGcJGBFb1_nlEGJp/view?usp=sharing>  Parte 2  <https://drive.google.com/file/d/1rjmWWHBQSMDdCrQn7zwa__etb6FBCo48/view?usp=sharing>  Exercício 2: Análise exploratória de dados |
| 7 | 5/10 | Continuação dplyr e tidyr  Tipos de dados  Programação (interação e funções): purrr  https://github.com/rprimi/acqu\_mirt  Parte 1  <https://drive.google.com/file/d/1gSKX_1Y8gZS2JHSI3CQx84kmUxw47JLy/view?usp=sharing>  Parte 2  <https://drive.google.com/file/d/1hZfzWKthrSevWg4dtyDEoV-_INgwqDdC/view?usp=sharing> |
|  | 12/10 | Dia da criança |
| 8 | 19/10 | Entrega Exerccício 3  Qualificações coletivas |
| 9 | 26/10 | Modelagem: modelr, broom, keras  List columns  Parte 1  <https://drive.google.com/file/d/1KlHczR0uuG2TprCF7B5HwUHwTh7BVT97/view?usp=sharing>  Parte 2  <https://drive.google.com/file/d/1UqXVDQbnQ-pVhSV76Vbm1A5z_y5NUx34/view?usp=sharing> |
|  | 2/11 | Finados |
| 10 | 9/11 | Modelagem: modelr, broom, keras  List columns  Mineração de textos: tidytext |
| 11 | 16/11 | Mineração de textos: tidytext |
| 12 | 23/11 | Mineração de textos: tidytext, quanteda  Modelagem avançada com Deep Learning com Keras para R  Intro  <https://drive.google.com/file/d/17ip1l_JJkZIJbOzr_5wviQd4xtQoeNWA/view?usp=sharing>  Parte 1  <https://drive.google.com/file/d/1ujE9PR2MzMkSkuCdSSFW2i3AIeGnO4s3/view?usp=sharing>  Parte 2  <https://drive.google.com/file/d/1sW7OH_BG4hStwb5Aca2N03xudWvC0uh-/view?usp=sharing>  Deep Learning  <https://www.youtube.com/playlist?list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi>  Ver Chapter 1, 2 e 3 |
| 13 | 30/11 | Modelagem com Deep Learning com Keras para R  Parte 1  <https://drive.google.com/file/d/1HHFOGe2SgaK7P6WaUN9vyZALA8hXd1XO/view?usp=sharing>  Parte 2  <https://drive.google.com/file/d/1UkCTJfMRU-8UyxWxoeTNHooqEZF-KbCY/view?usp=sharing> |
| 14 | 7/12 | Modelagem com Deep Learning com Keras para R  Parte 1  <https://drive.google.com/file/d/1MuTrrhJ1O662V75biyOpG-WkEtTsYE7D/view?usp=sharing>  Parte 2  <https://drive.google.com/file/d/1KmoZu0SdmdE7u02lAwBN279ns29MG1Iu/view?usp=sharing> |
|  |  |  |