

Lista de Exercícios - AM2 – Aula 13

Assunto: RNA em textos

Professor: Murilo Naldi

- 1 – São as principais diferenças entre processamento de linguagem natural (PLN) e processamento de textos?
- 2 – Explique a diferença entre linguagens formais e linguagens naturais. Como modelar as duas?
- 3 – Como funciona a técnica *Bag-of-Words* (*BoW*) e como ela representa documentos textuais?
- 4- Quais são os diferentes métodos para se definir o valor de um atributo relacionado a uma palavra em um conjunto de dados documento quando usamos *BoW*?
- 5 – A representação *BoW* é conhecida por sua alta dimensionalidade. Quais procedimentos podemos aplicar para reduzir o número de palavras utilizadas na representação?
- 6 – O que é a técnica de *word embedding*? Explique porque ela se diferencia do *BoW*?
- 7 – Explique duas arquiteturas que utilizam modelagem Word2Vec.
- 8 – No que o *embedding* ELMo se diferencia do Word2Vec? O que são LSTM e como ELMo se aproveita desse tipo de rede?
- 9 – Explique como utilizar *auto-encoders* para textos e como eles podem ser aplicados? Cite duas tarefas.
- 10 – O que são *transformers* e porque eles se diferenciam de outros tipos de *auto-encoders*? Que tipo de informação suas redes “*multi-head attention*” conseguem processar?