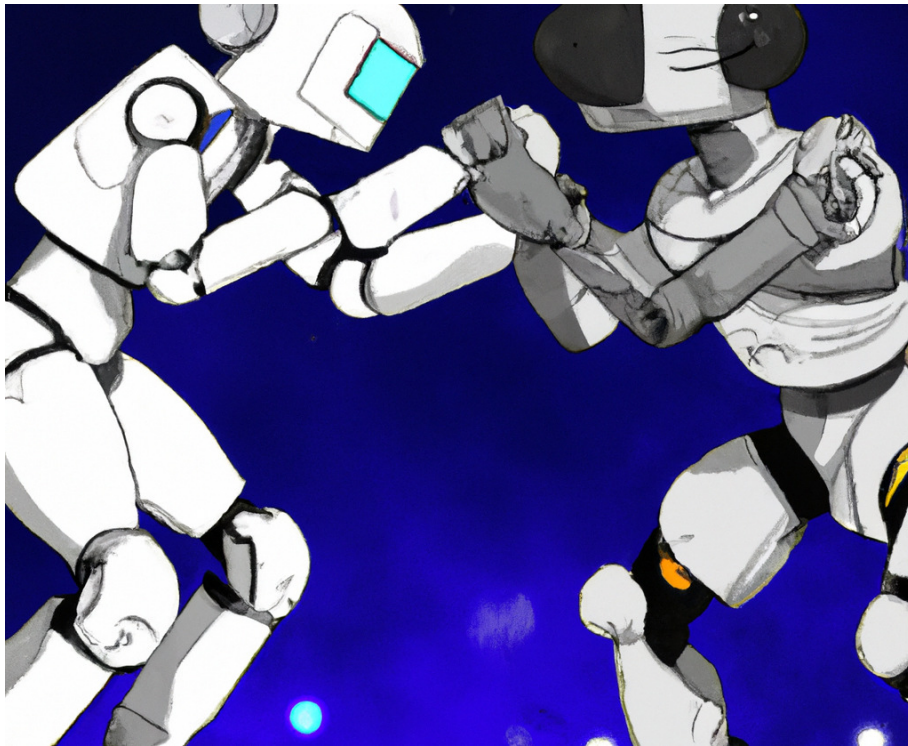


# Word Embedding

---



Palavras ou frases do vocabulário são mapeadas para vetores de números reais

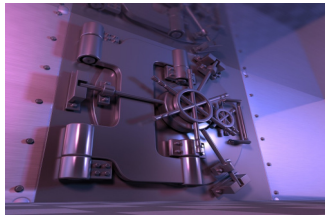
# Word Embedding

## ➤ Independentes de contexto:

- Word2Vec
- GloVe



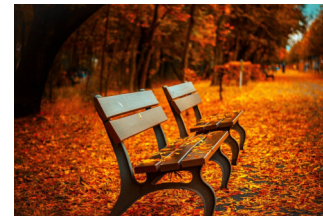
Dim1	Dim2	Dim3
0.1	0.3	0.7



Dim1	Dim2	Dim3
0.1	0.3	0.7

## ➤ Dependentes de contexto

- ELMo
- BERT
- GPT



Dim1	Dim2	Dim3
0.1	0.3	0.7



Dim1	Dim2	Dim3
0.45	0.8	0.3

# Word Embedding

---

Word	Dimension 1	Dimension 2	Dimension 3
Cat	0.1	0.3	0.7
Dog	0.15	0.35	0.68
Mouse	0.9	0.05	0.2
Cheese	0.85	0.05	0.22
Jump	0.45	0.8	0.3
Run	0.47	0.82	0.33

# Similaridade

---

Word	Dimension 1	Dimension 2	Dimension 3
Cat	0.1	0.3	0.7
Dog	0.15	0.35	0.68

# Analogia

---

Word	Dimension 1	Dimension 2	Dimension 3
King	0.6	0.8	0.4
Queen	0.4	0.8	0.4
Man	0.6	0.7	0.5
Woman	0.4	0.7	0.5



# Word Embedding

- **Propriedade Semânticas:** Vetor de Gato é semelhante ao de Cachorro
- **Propriedades sintáticas:** Função gramatical da palavra: Substantivo e Verbo
- **Propriedades Distributivas:** Palavras que aparecem em contexto semelhante



Word Embedding					
Gato	0.3	0.4	0.1	-1.3	0.0
Cachorro	0.4	0.2	0.7	-1.1	0.2
Muro	0.2	0.2	0	-0.8	0.1

Semelhança de cosseno			
	Gato	Cachorro	Muro
Gato	1	0.9	0.5
Cachorro	0.9	1	0.7
Muro	0.5	0.7	1

# Transformers

---

- Arquitetura de RNA utilizada para NLP
- Mecanismos “self-attention”
- Ótima em Tarefas de Contexto
- Pode utilizar word embeddings como entrada para a RNA

