

Processamento Natural de Linguagem

Aula 0

Introdução do curso



Semantix[®]

All about data

Quem sou eu?



Filipe Theodoro

Cientista de dados da Semantix
Instrutor do Semantix Academy

Contatos

filipe.theodoro@semantix.com.br
linkedin.com/in/filipe-theodoro



O professor

- Formado em Engenharia de Computação pela Universidade Federal de Ouro Preto – MG
- Projetos:
 - Autenticação Biométrica usando fonocardiograma
 - Detecção de Trincas em dormentes
 - Visão Computacional
 - Classificação de Texto
 - Recuperação de Informação
 - Sistema de Recomendação

Ementa

- Introdução
 - O que é?
 - História
 - Exemplos de aplicação
- Expressão regular e préprocessamento
 - Regex
 - Tokenização
 - Normalização
 - N-gram
 - Remover stopwords
 - Correção ortográfica
 - Exercício prático

Ementa

- Modelagem clássica
 - Bag of words (BOW)
 - TF-IDF
 - BM25
 - Modelagem de Tópicos
 - Exercício prático
- Modelagem redes neurais artificiais
 - Word embedding
 - Transformers – Attention is all you need



Semantix[®]

All about data

filipe.theodoro@semantix.com.br

www.semantix.com.br