

PROCESSAMENTO DE LINGUAGEM NATURAL

Mestrado em Inteligência Artificial Aplicada

Código: 28513

Área Científica Predominante: Ciências da Computação

Docente: Óscar Rafael da Silva Ferreira Ribeiro

Idioma de Instrução: Português

Regime: S2

Carga Letiva: 30h Carga Trabalho: 130h

ECTS: 6,0

Objetivos

Esta unidade curricular permite aos estudantes um primeiro contacto com a área de processamento de linguagem natural, os seus objetivos e técnicas, sejam na sua vertente simbólica/algébrica ou na sua abordagem de aprendizagem automática.

Resultados da Aprendizagem

Os estudantes que tenham aprovação a esta unidade curricular devem ser capazes de perceber a estrutura típica de uma aplicação na área de processamento de linguagem natural, bem como os seus principais desafios. Nesse sentido, devem ser capazes de perceber a relevância das tarefas básicas de PLN, que incluem a atomização e segmentação, análise sintática e morfológica. Para além disso, devem conhecer as diferentes áreas de aplicação do PLN e serem capazes de desenvolver a investigação necessária para a sua implementação.

Conteúdos Programáticos

1. Introdução ao PLN
 - Áreas de atuação
 - Recursos (Léxicos e Corpora)
 - Pilha de NLP, das palavras à semântica
2. Análise Morfossintática
 - Análise Morfológica (Regras de Stemming, Afixos)
 - Modelos de linguagem
 - Análise sintática
3. Análise Semântica
 - Representação de significado
 - Semântica Lexical e Recursos Lexicais
 - Ontologias e Raciocínio
 - Semântica de vetores e embeddings de palavras

Bibliografia Recomendada

- Jurafsky, D. and Martin, J.H. (2020) Speech and Language Processing (3rd edition draft, available online)
- Indurkha, N. and Damerau, F. (2010). Handbook of Natural Language Processing (2nd edition). Chapman and Hall/CRC.
- Bird, S., Klein, E. and Loper, E. (2019). Natural Language Processing with Python. (updated edition, from O'Reilly first edition. available online).

Métodos de Ensino e de Aprendizagem

A estrutura curricular apresenta a área de processamento de linguagem natural de um ponto de vista amplo, focando na vertente lexical e semântica, usando técnicas simbólicas/algébricas e técnicas de aprendizagem automática.

Métodos de Avaliação

avaliação da unidade curricular é dividida em três componentes:

- Preparação prática de um protótipo de aplicação/solução de PLN (40%)
- Escrita de um pequeno artigo científico (em inglês) sobre o trabalho realizado (35%)
- Apresentação e discussão oral (em Inglês) do trabalho realizado (25%)