



ID: 71523047

30-09-2017

Investigador Filipe Rodrigues vence prémio com tese de doutoramento

Vida Académica Estudo na área de “machine learning” valeu prémio pela Associação Portuguesa de Reconhecimento de Padrões

A “Melhor Tese de Doutoramento” para a Associação Portuguesa de Reconhecimento de Padrões está na Faculdade de Ciências e Tecnologias da Universidade de Coimbra (FCTUC) e pertence a Filipe Rodrigues. O investigador do Departamento de Engenharia Informática da FCTUC desenvolveu um estudo em que propõe um conjunto de modelos probabilísticos para aprendizagem a partir de dados gerados pela multidão, tese que também já lhe valeu uma menção honrosa da Associação Portuguesa para a Inteligência Artificial.

Intitulada “Probabilistic Models for Learning from Crowdsourced Data”, a tese foi orientada pelos docentes Francisco Câmara Pereira e Bernardete Ribeiro e desenvolveu-se na área da machine learning, uma sub-área da inteligência artificial. «Têm-se verificado grandes desenvolvimentos na área de “machine learning”, permitindo ter, hoje em dia, sistemas inteligentes capazes de reco-



Filipe Rodrigues é o autor da tese premiada

nhecer a nossa voz e responder a perguntas, traduzir textos de forma automática, conduzir veículos de forma autónoma, etc. Por detrás deste recente progresso estão contribuições ao nível da teoria e novos algoritmos de aprendizagem, bem como a crescente disponibilidade de dados - vulgarmente designados de “Big Data”. Contudo, o sucesso dos algoritmos de machine learning está frequentemente de-

pendente da disponibilidade de grandes volumes de exemplos etiquetados», explica Bernardete Ribeiro.

Para se perceber melhor, a docente e investigadora da FCTUC exemplifica com um cenário: «imaginemos que queremos desenvolver um sistema capaz de distinguir imagens de peões de imagens de automóveis. Para isso é necessária a criação de um conjunto de dados etiquetados com

centenas, ou idealmente milhares ou milhões, de exemplos de imagens de peões e imagens de automóveis».

As grandes dimensões destes conjuntos de dados tornam impraticável o recurso à etiquetagem manual dos dados por um único anotador, sendo por isso prática comum o recurso a plataformas de “crowdsourcing”. No entanto, os diferentes níveis de perícia individual dos anotadores que colaboram nestas plataformas tornam necessário o desenvolvimento de abordagens específicas e direccionadas para este tipo de dados multi-anotador.

Considerando «o problema da heterogeneidade dos anotadores, na tese são propostos modelos probabilísticos capazes de diferenciar, de forma automática e totalmente não-supervisionada, os anotadores confiáveis dos menos confiáveis ou até identificar anotadores cujas respostas são dadas de forma aleatória ou pouco premeditada», explica Bernardete Ribeiro. ◀