Lecture: «Extracting Deep Learning Representations: The Tiramisu project»

Data: xoves, 3 decembro, 2015 -12:00 - 13:00

Lugar: Salón de actos

Poñente(s): Ulises Cortés (Universitat Politècnica de Catalunya)

Idioma: Castelán

Streaming: Segue este evento en directo

Ulises Cortés is a Full-Professor and Researcher of the Technical University of Catalonia (UPC)

Professor Cortés and his group have been applying their work in Artificial Intelligence to Environmental Sciences, in special to Wastewater Treatment Plants

In the last three years Deep Learning has become a prominent area of research within Artificial Intelligence. Its application to cognitive computing through natural language processing and image recognition has caught the attention of researchers from all over the world. Furthermore, IT companies such as Google, Twitter, Microsoft of IBM are investing heavily on the development of Deep Learning systems for their integration into commercial products.

By looking at the core of Deep Learning we find that these are representation building networks, as a result of learning and tuning millions of non-linear patterns on a large set of input data. The representations being learnt by these networks are bigger both in number and complexity to any previous system, which provides Deep Learning with unique knowledge representation capabilities.

In our work we explore the knowledge stored internally in Deep Learning networks, while trying to extract it for other reasoning purposes. Using convolutional neural networks (CNN), the original purpose of which is to discriminate (i.e., classify) among a set of known classes, we transform images into large and sparse vectors of features. The resultant vector space embedding is shown to have remarkable knowledge representation capabilities, identifying untaught abstract classes (e.g., what is a living thing?) and allowing vector operations consistent with visual semantics. This work sheds some light on the internal aspects of Deep Learning networks, and opens up a whole new field of application for their learnt models.