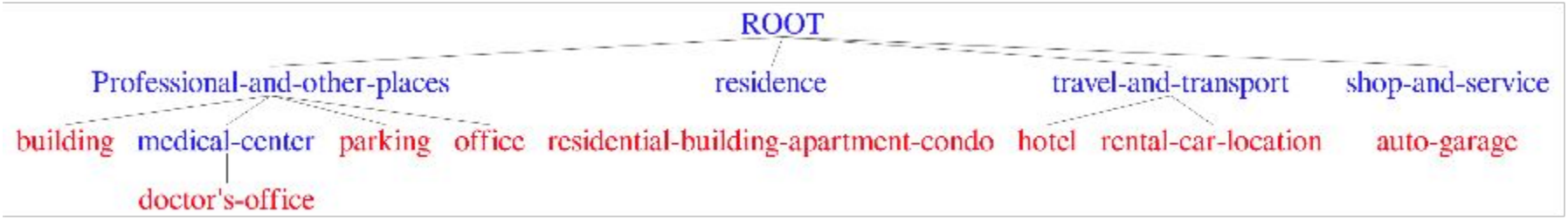
ANLPIR-2018

Projects proposal

Contact person: Gianni Barlacchi (barlacchi@fbk.eu)

Semantic Tree Kernel for Urban Area Analysis

Location-based Social Networks (e.g. Foursquare) provide information about each places in a region. Thus, the most straightforward way to represent it is through the use of these POIs data. Every venue is hierarchically categorized (e.g., Professional and Other Places → Medical Center → Doctor's office) and the categories are used to produce a tree representation of the area.



The goal of this project is to exploit the use of the semantic tree kernel, applied on these structures, for urban area classification (e.g. commercial, residential etc.).

Who are you? Personality prediction from social media texts.

- A person can be characterized in terms of 5 personality traits: Openness to experience, Conscientiousness, Extraversion, Agreeableness, Neuroticism.
- Personality has been shown to be related to many human characteristic, including writing style.

The goal of this project is to predict the user's personality using social media texts. In particular, the idea is based on the use of syntactic information that can be represented with tree kernel. We will provide a unique Facebook dataset composed by:

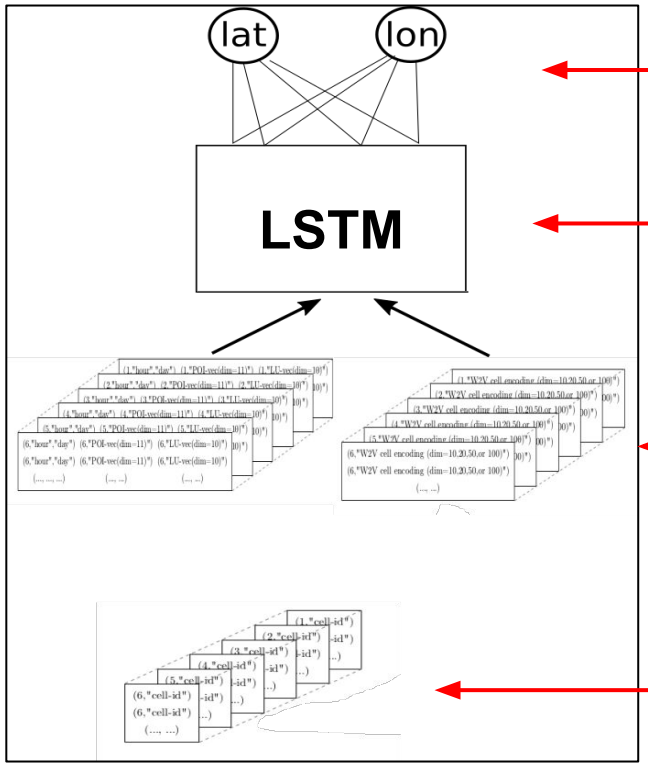
- Users with its personality traits: 115.874
- Posts: 19.181.625

References:

Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach
http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0073791

Predicting human mobility using using spatial embeddings

Starting from POI categories, we learn the area representation combining the semantic meaning of POIs with their spatial relationship. Then, we use this semantic information about places to predict the movements of our users.



Increment on latitude and longitude (in meters)

Neural Network

Enriched input with textual feature

Time and location representation

References:

DeepTransport: Prediction and Simulation of Human Mobility and Transportation Mode at a Citywide Level

What is happening there? Understanding crowds in a city from text and mobility data

Event detection by combining data from location-based social networks (e.g. Foursquare) and social media text (e.g. Tweets).

- Exploring twitter dataset for the city of New York
- Implements an event detector based on common statistical measures
- Use tweets text to classify the category of the event

References:

Event Detection using Twitter: A Spatio-Temporal Approach
http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0097807