LINKSFOUNDATION.COM

FONDAZIONE links PASSION FOR INNOVATION

Politecnico di Torino 1859

ellis European Laboratory for Learning and Intelligent Systems

Applied Data Science Project

L17 - Project communication tools II

Giuseppe Rizzo

Turin, October 14, 2021

Pillars

Design

Manage

Develop

Communicate

Report

Report type

- Paper
  - concise reporting
  - usually, the reader of a paper is a researcher

- Deliverable
  - it is a longer report than a paper
  - usually, the reader of a deliverable is an innovator with knowledge about the topic of the deliverable

Report type

- Paper
  - concise reporting
  - usually, the reader of a paper is a researcher
- Deliverable
  - it is a longer report than a paper
  - usually, the reader of a deliverable is an innovator with knowledge about the topic of the deliverable

Paper outline

Abstract → Introduction → Related Work → Method → Experiment → Conclusion

Abstract

The whole paper condensed in a few hundred words

1 sentence for each of the following:
- What are the objectives
- How does the work address the objectives
- What about the experiment
- What are the results
- Which conclusions can be derived

Introduction

The introduction sets the scene by:

- stating the objective and the background
- introducing the methodology
- providing the remainder of the whole paper

Related Work

This section lists the relevant research activities that have been used as inspiration sources

While listing each:

- 1 sentence to summarize the whole work taken as inspiration
- 1 sentence to underline one of the concepts and/or methodologies further worked out in your work

The take home message of this section is to have a fair and factual comparison with the state of the art meaning what others did before you

Method

This section contains:

- a problem statement described with words and with formulas
- a diagram to summarize the method
- a detailed description and formulas
- a complete list of technical tools and configurations

Experiment

A detailed description of the dataset

- why it is relevant for the work
- how it has been collected/found
- statistics of the dataset (no. of records, no. of features, …)

Experiment configuration

- whether there are assumptions to run the experiments
- how the experiment is performed

Evaluation

- how the evaluation is performed
- tables summarizing the performance
- take home message from the evaluation: what was successful, what not and why

Conclusion

This section

- restates the objective
- recaps the method and how it addresses the objective
- provides a short summary of the evaluation and whether it has been successful or not
- shares some lesson learned to be ported to other studies
- outlines future additions to this work

Length

Concise report

Length depends on the venue where the paper is presented:

- less than 10 pages if a conference
- from 10 to 25 pages for a journal (though there are exceptional cases of longer papers)

Hands on

Predicting Your Next Stop-over from Location-based Social Network Data with Recurrent Neural Networks

Enrico Palumbo
ISMB
Turin, Italy
EURECOM
Sophia Antipolis, France
palumbo@ismb.it

Giuseppe Rizzo
ISMB
Turin, Italy
giuseppe.rizzo@ismb.it

Raphaël Troncy
EURECOM
Sophia Antipolis, France
raphael.troncy@eurecom.fr

Elena Baralis
Politecnico di Torino
Torino, Italy
elena.baralis@polito.it

ABSTRACT

In the past years, Location-based Social Network (LBSN) data have strongly fostered a data-driven approach to the recommendation of Points of Interest (POIs) in the tourism domain. However, an important aspect that is often not taken into account by current approaches is the temporal correlations among POI categories in tourist paths. In this work, we collect data from Foursquare, we extract timed paths of POI categories from sequences of temporally neighboring check-ins and we use a Recurrent Neural Network (RNN) to learn to generate new paths by training it to predict observed paths. As a further step, we cluster the data considering users' demographics and learn separate models for each category of users. The evaluation shows the effectiveness of the proposed approach in predicting paths in terms of model perplexity on the test set.

KEYWORDS

Sequence learning, path recommendation, tourism, POI recommendation

1 INTRODUCTION

Location-based Social Networks (LBSN) allow users to *check-in* in a Point-of-Interest (POI)[1] and share their activities with friends, providing publicly available data about their behavior. One of the distinctive features of LBSN data with respect to traditional location prediction systems, which are mainly based on GPS data and focus on physical mobility [33], is the rich categorization of POIs in consistent taxonomies, which attribute an explicit semantic meaning to users' activities. The availability of venue categories has opened new research lines, such as statistical studies of venues peculiarities [17], automatic creation of representations of city neighborhoods and users [22, 25], definition of semantic similarities between cities [24]. Most importantly, venue categories play an important role in POI recommender systems, as they enable to model user interests and personalize the recommendations [18]. In the past years, little attention has been dedicated to the temporal correlations among venue categories in the exploration of a city, which is nonetheless a crucial factor in recommending POIs. Consider the example of a check-in in an *Irish Pub* at 8 PM: is the user more likely to continue her evening in a *Karaoke Bar* or in an *Opera House*? Better a *Chinese Restaurant* or an *Italian Restaurant* for dinner after a *City Park* in the morning and a *History Museum* in the afternoon? Note that predicting these sequences require an implicit modeling of at least two dimensions: 1) temporal, as certain types of venues are more temporally related than others (e.g. after an *Irish Pub*, people are more likely to go to *Karaoke* than to a *History Museum* 2) personal, as venue categories implicitly define a user profile, independently from their order (e.g. *Steakhouse* and *Vegetarian Restaurant* do not go frequently together). Most of existing studies attempt to model directly sequences of POIs rather than their categories to recommend the next POI to a user (see 'next POI prediction' in Sec. 2). In this work, we focus on modeling sequences of POI categories to enhance the generality and the portability of the obtained results. This can be considered as a first step in the next POI prediction problem, as the POI category can then be turned into a specific POI by querying a database of POIs according to a variety of parameters, such as the user context (e.g. position, weather) and/or specific POI features such as popularity, average prices and the like.

In order to address this problem, we first collect users' check-ins from Foursquare and extract their corresponding venue categories, segmenting them into a set of temporally neighboring activities, which we call *paths*. Then, we train a Recurrent Neural Network to learn to predict these paths in order to generate new ones, thanks to its architecture that is specifically meant to model temporal sequences without specifying a specific memory length. In the attempt to take into consideration the fact that the nature of the generated sequences is not universal, but it critically depends on the typology of user, we cluster users in groups and learn separate models for each of them. Differently from previous work [32], we cluster users based on their demographics rather than on their past activities, consistently with the intent of obtaining results that are portable to new data without a cold start problem.

The main scientific contributions of this paper are: (1) addressing the problem of next POI category using a machine learning

Palumbo et al. (2017) Predicting Your Next Stop-over from Location-based Social Network Data with Recurrent Neural Networks. In (RecSys) ACM RecSys Workshop on Recommenders in Tourism (RecTour), Como, Italy

App for editing

Overleaf
https://www.overleaf.com

Google Docs

Word

All applications allow to collaborative edit the making of a presentation

Overleaf offers an editing environment optimized for writing scientific articles

Consider to export your paper when complete to PDF to maximize the portability

FONDAZIONE links

PASSION FOR INNOVATION

Thank you for your attention.

Questions?

FONDAZIONE LINKS
PASSION FOR INNOVATION

CONTACTS

Giuseppe Rizzo

Team Leader

p. +39 011 2276244

e. giuseppe.rizzo@linksfoundation.com