



Open Internship Position

Project Title

Smart Push - Adaptive Learning for Effective Recommendation

Project Description

Research topics:

Supervised and unsupervised learning, recommendation systems, decision systems and constraints, heterogeneous data sources, data mediation, distributed architecture, big data

Research project:

Motivation. SQUADATA is a small-sized company specialized in data processing and Web marketing technologies. The company proposes several customer solutions to its clients (e.g. advertisers, websites and/or brand owners) to manage their advertising campaigns or their user and profile databases. Among these solutions, we find : **retargeting**, i.e. identifying a user on a website and reaching him after a while through a specific channel (email, sms, push notification, ...); **targeting**, broadcasting an advertising campaign to a list of identified users (through a channel); **wakeupdata** : identifying potentially active users from a given list of customers; etc.

Based on a large amount of data, representing customer activities on the Web, we perform statistics and configure advertising and marketing campaigns, to broadcast advertiser's content. Configuring a campaign scenario requires lot of human input and a specific business knowledge. Moreover, there are several privacy and pressure constraints to respects before soliciting a user.

Today, all process are implemented using jobs and scripts, configured by human end-users. The objective is to migrate from this manually configured system into a learning based decision system.

Open Scientific Challenges.

- Composing from/with heterogeneous data sources (different formats, different logic, stream vs static, structured vs. unstructured)

- Handling large data volumes (storage, processing, latency, ...)
- Data standardization /normalisation
- Building and testing supervised and unsupervised learning models
- Evaluation

Research Objectives.

- Representing/aggregating data into a relevant format, in order to extract a sub-group of interesting features, focusing on valuable data, when also dealing with a possible large volume of noise
- Proposing an adaptive and learning approach able to make decision based on past known results

Methodology

The intern will have to carry out the following steps:

- Perform a study of the related work in the fields of learning or adaptive decision systems.
- Adopt a prototypical approach, providing recurrent testable models
- Propose novel protocols or framework for testing and evaluating their impact
- Assess the proposed solution both theoretically and practically through extensive experiments involving the deployment on real distributed platforms with real data (e.g., Grid 5000, Apache Spark).

Technologies in use by the company : php, javascript, lua, nginx, httpd, cloudera hadoop suite (hive, impala, spark, ...), standard data format (parquet, csv, json)...

References

- [1] Andrej Miklosik, Martin Kuchta, Nina Evans, Stefan Zak. Towards the Adoption of Machine Learning-Based Analytical Tools in Digital Marketing. doi.org/10.1109/ACCESS.2019.2924425
- [2] Xiaomei Bai, Mengyang Wang, Ivan Lee, Zhuo Yang, Xiangjie Kong, Feng Xia. Scientific Paper Recommendation: A Survey. IEEE Access, Vol. 7, 9324-9339, 2019.
- [3] Thiago Silveira, Min Zhang, Xiao Lin, Yiqun Liu, Shaoping Ma. How good your recommender system is? A survey on evaluations in recommendation. International Journal of Machine Learning & Cybernetics, 10(5), 813-831, 2019.

Advisors

Sara Bouchenak

Professor, INSA Lyon – LIRIS
Sara.Bouchenak@insa-lyon.fr
<https://perso.liris.cnrs.fr/sara.bouchenak/>

Short bio:

Sara Bouchenak is Professor in Computer Science at INSA Lyon. She is a member of the DRIM research group at LIRIS laboratory, where she conducts research on highly-available, dependable and manageable distributed computer systems. Prior to that, she was Associate Professor at University of Grenoble until 2014, a visiting professor at Universidad Politécnica de Madrid, Spain, in 2009/2010, and an associate researcher at EPFL, Switzerland, in 2003.

Sara Bouchenak is a member of ACM and IEEE; she was an officer of the French chapter of ACM-SIGOPS. She is an officer of CNRS GDR RSD. She received her HDR (Habilitation à Diriger les Recherches) from University of Grenoble I in 2010, and her PhD in computer science from Grenoble Institute of Technology in 2001.

Sara Bouchenak has coordinated and participated in a number of national and EU projects. She acts as expert for H2020 and FP7 projects, ANR projects, and for other european national research agencies. She served as general chair and program committee chair for several international and national conferences, and she is a member of the program committee of more than 50 conferences among which IEEE DSN, ICDCS, SRDS. She is the author or co-author of more than 90 publications.

Pierre De Vettor

PhD - R&D Engineer, Squadata
pdevettor@squadata.net
<https://www.linkedin.com/in/pdevettor/>

Short bio:

Pierre De Vettor is a computer science PhD who works as an R&D engineer at Squadata since 2017. His work relates to large data processings and composition/mediation of heterogeneous data sources. He participates in the study and design of data processing approach, relying on data workflows and query optimisation.

Pierre De Vettor obtained his PhD in 2016 after a 3 years industrial doctorate directed by Djamel Benslimane and Michael Mriassa, from the SOC research group at LIRIS laboratory. His thesis focus on data integration of raw data coming from heterogeneous and multi-origin data sources on the Web, providing a generic architecture able to analyze and combine this heterogeneous, informal, and sometimes meaningless data into a coherent smart data set.

Internship Location & Allowance

The internship has a monthly allowance of 1000 €.

The internship will take place at SQUADATA.

5, impasse Morel. 69003 Lyon. France

Phone: +33 4 78 54 45 45

www.squadata.net

Contact

For further information regarding the internship, please contact Pierre De Vettor (pdevettor@squadata.net) and Sara Bouchenak (sara.bouchenak@insa-lyon.fr).