

Early detection of customer projects in a retail setting

Supervisors: Tassadit Bouadi, Peggy Cellier, Alexandre Termier

Mail: tassadit.bouadi@irisa.fr, peggy.cellier@irisa.fr, alexandre.termier@irisa.fr

Keywords: data science, data mining, machine learning, pattern mining, prediction, retail

French customers especially enjoy DIY store (ex: Castorama, Leroy-Merlin, Bricomarché...) where they can buy numerous items for decorating, repairing or improving their homes. Usually there are two types of visits to such shops:

- The customer has a very specific and one-time need (ex: buy a new lightbulb) ;
- The customer is starting a large project, often at the time of a change in life (first apartment, moving to a new home, wedding...). This project requires planning and financing, and will lead to a large volume of purchases.

For the DIY store, it would be of utmost importance to detect early customers preparing for large projects, and to ensure that they will not go to a competitor. However, it is difficult to detect them, as they of course do not tell the store about the project, and there may be very few “preparatory purchases” before the project.

The goal of this internship, conducted with a large French DIY store chain, is to determine if it is possible to detect “life changes” of customers from their purchases, correlated with the start of an important project. We will have access to anonymized data from the DIY store, and possibly to a classic grocery store of the same group.

The research work of the internship will thus feature:

- An exploratory Data Science step over the data, in order to answer at least the following questions:
 - o Are there precursors of life changes leading to large purchases in the DIY historical data?
 - o Is there a correlation between the grocery data and the DIY data? Can it be used to infer life changes and large purchases?
- A methodological contribution to propose a proper analysis method for early detection of life changes leading to large purchases.

The methods used for the analysis need to give human-understandable results, hence methods from the pattern mining field will be preferred, for example discriminant pattern mining. Machine Learning methods of classification and next action prediction will also be reviewed.

The intern must have a deep interest for Data Science and for working on a real retail application, in collaboration with a retail company. There is an opportunity to pursue a CIFRE PhD after this internship.

References:

- Guozhu Dong: Exploiting the Power of Group Differences: Using Patterns to Solve Data Analysis Problems. Synthesis Lectures on Data Mining and Knowledge Discovery, Morgan & Claypool Publishers 2019
- Guozhu Dong, Jinyan Li: Efficient Mining of Emerging Patterns: Discovering Trends and Differences. KDD 1999: 43-52
- Arnaud Soulet, Chedy Raïssi, Marc Plantevit, Bruno Crémilleux: Mining Dominant Patterns in the Sky. ICDM 2011: 655-664
- Clément Gautrais, René Quiniou, Peggy Cellier, Thomas Guyet, Alexandre Termier: Purchase Signatures of Retail Customers. PAKDD (1) 2017: 110-121