

TOPIX: AI-BASED SOLUTION ALLOWING TO SUMMARIZE MASSIVE DATABASES INVOLVING TEXT WRITTEN BY USERS ON PRODUCTS

Can be applied in a large variety of situations in order to easily interpret user opinions/feelings/expectations regarding objects.

ERG\NEO

L'AVENIR EST FAIT D'AUDACE

PRESENTATION

The AI-based solution produces a simultaneous clustering (co-clustering) of users and products, based on the key emerging topics from the reviews and by the underlying model. The typical use case consists in an e-commerce company interested in understanding the relationship between its users and the sold products thanks to the analysis of user comments. While existing solutions focus on analyzing the fact that users have bought specific products and conversely, other solutions can analyze huge amount of text data to extract useful information, TOPIX is a unique technology that can do both by simultaneously analyzing the links between users and products and the associated texts such as reviews, comments or opinions.



Artificial Intelligence - Co-clustering
Latent Topic Block Model (LTBM) - Variational inference algorithm
Textual data analysis - E-commerce

APPLICATIONS

- E-commerce
- Marketing
- Healthcare
- Science watch
- Social network, e-journalism

INTELLECTUAL PROPERTY

Filing of the software

CONTACT

+33 (0)1 44 23 21 50
industriels@erganeo.com
Ref. project : 417

COMPETITIVE ADVANTAGES

- Simultaneous and automatic clustering (co-clustering) of rows (ex.: users), columns (ex.: products) based on the key topics (identified from the words analyzed in the texts)
- Produces key topics from the reviews and by the underlying model
- Massive and possibly extremely sparse datasets
- Unsupervised AI
- Image pre-processing routine allowing TOPIX to deal with images (in addition to text)
- Four types of graphs proposed to the user for the outcomes

DEVELOPMENT PHASE

- ✓ Operational Software coded in C++
- ✓ Demonstrator available with a user-friendly web interface at <https://topix.mi.parisdescartes.fr>. Typical use cases proposed. The web platform allows users to upload their own dataset to test the performance of the algorithm

PUBLICATIONS

L. Bergé, C. Bouveyron, M. Corneli, P. Latouche *The latent topic block model for the co-clustering of textual interaction data*. Computational Statistics and Data Analysis, vol. 137, pp.247-270, Elsevier (2019)