DATA SCIENCE AND FRAUD DETECTION IN INSURANCE

Academic partner(s) : Laboratoire d'Economie d'Orléans (Université d'Orléans)

Financial partner(s) : Thélem Assurances

Scientific director(s) : Denisa Banulescu-Radu

Website : available HERE

RESEARCH PROGRAM DESCRIPTION

The research initiative "Data Science and Fraud Detection in Insurance" is a research project bringing together specialists in econometrics, data science and economics, whose objective is to develop new prevention and insurance fraud detection techniques. The project thus contributes to the development and dissemination (in the academic and professional fields) of recent advances in Data science applied to the field of insurance fraud detection.

Automatic fraud detection is a specific area of statistical modeling: unlike the detection of anomalies in the industrial field, for example, it aims to detect fraudulent transactions resulting from the rational behavior of agents. Therefore, detecting cases of fraud requires not only skills in data science and econometrics (knowledge of statistical models and their properties), but also economic and legal skills to understand the motivations and the strategic behavior of fraudsters. The main motivations of the project are mainly related to (i) the modeling of the strategic behavior of fraudsters in the field of insurance; (ii) the need to implement effective detection systems given the huge financial losses associated with fraud; (iii) the use of new databases to identify the mechanisms of insurance fraud.

For each research action, the work conducted within the chair considers the primary challenges in fraud modeling, encountered both in the academic domain and by Thélem Assurances. In this context, several major issues have been identified: i) the development of new technical tools for the real-time prevention/detection of various types of fraud; 2) the utilization of appropriate methods to handle imbalanced databases; 3) the imperative for evaluating fraud detection models; 4) the demand for interpretability in detection models and the exploration of novel databases.

The project is organized around several actions:

- Active academic research in the field of insurance fraud detection;
- Funding of high-quality research proposals;
- Organization of regular plenary conferences and a program of visiting researchers;
- Training of data scientists and academics on the specific issues of fraud detection;
- Promotion and dissemination of research;
- Recruitment of a PhD student.

RESEARCH TEAM

Researchers

- *Denisa BANULESCU RADU, University of Orléans*
- *Alexis DIRER, University of Orléans*
- *Matthieu PICAULT, University of Orléans*

PhD Students

- *Yannick KOUGBLENOU, University of Orléans*

Interns

- *Severine PRIN, Thélem Assurances*
- *Pascal FELIOT, Thélem Assurances*
- *Céline PICHARD, Thélem Assurances*
- *Damien THOMAS, Thélem Assurances*
- *Julie AZOUAOU GENERMONT, Thélem Assurances*

Other people involved in the program (e.g. engineers)

- *Alexandra AMANI, University of Orléans (internship, 2 months)*

PUBLICATIONS OF THE YEAR DIRECTLY RELATED TO THE RESEARCH PROGRAM

Published

- Banulescu-Radu, D., Kougblenou, Y., (2023). Data science for insurance fraud detection: a review. Forthcoming in Handbook of Insurance, Springer

- Banulescu-Radu, D., Yankol-Schalck, M., (2023). Practical guideline to efficiently detect insurance fraud in the era of machine learning: a household insurance case. Journal of Risk and Insurance. Online version, November 2023

Under review

- Banulescu-Radu, D., Hansen, P.R., Huang, Z., Matei, M., (2023). Volatility during the financial crisis through the lens of high frequency data: a Realized GARCH approach. R&R in Journal of Financial Econometrics.

Working papers

- Baesens, B., Banulescu-Radu, D., Hurlin, C., Kougblenou, Y., Verdonck, T., (2023). Benchmarking state-of-the-art resampling techniques for classification models: do optimal ratios exist? Working paper.
- Banulescu-Radu, G. D., Bénoît, S., Hurlin, C., (2023). Shortfall in Tax Revenue: Evaluating the Social Security Contribution Fraud. Working paper.
- Kougblenou, Y., (2023). Omission errors in binary classification tasks: impact and mitigation. Working paper.
- Banulescu-Radu, D., Kougblenou, Y., (2023). Machine Learning and cost sensitive learning for insurance fraud detection. Ongoing project
- Coté, O., Coté, MP., Charpentier, A., (2023). A Fair price to pay: exploiting causal graphs for fairness in insurance. Working paper (winner of the 1st Call for papers).
- Pei, J., Lu, Y., (2024). Forecasting natural disaster frequencies using nonstationary count time series models. Working paper (winner of the 2nd Call for papers).
- Dotta, M., Milhaud, X., Pommeret, D., (2024). Copulas based fraud detection. Ongoing project (winner of the 2nd Call for papers).

Reports, books, press articles, etc.

MAJOR COMMUNICATIONS RELATED TO THE RESEARCH PROGRAM

Major academic conferences, invited speaker, etc.

Three plenary conferences have been organized so far during 2023-2024, as shown in the table below. Two other plenary conferences are scheduled by the end of 2024.

	Date	Theme	Presenter
Conference n°1 (see Poster)	Feb 3, 2023	Quelles perspectives pour le système de retraite après la réforme de 2023 ?	Prof. Anne Lavigne (Professor at University of Orléans)
Conference n°2 (see Poster)	Oct 21, 2023	Energie, climat : décider par temps de crise ?	Prof. Patrice Geoffron (Professor at Paris Dauphine University)
Conference n°3 (see Poster)	Jan 19, 2024	Éthique et équité en tarification	Prof. Arthur Charpentier (Professor at University of Quebec)

Conference n°8	May 31, 2024	TBA	Prof. Florence Jusot (Professor at Paris Dauphine University)
Conference n°9	Fall, 2024	TBA	Dr. Gonéri le Cozannet (Senior researcher at BRGM)

The project research findings have been disseminated through presentations at various conferences in 2023, including:

- IFABS, International Finance and Banking Society Conference, Oxford, 24-26 July 2023
- CaLiBank Workshop, LAPE, Limoges, 7 June 2023
- AFSE, 71st Congress of the French Economic Association, Paris, 14-16 June 2023
- FEBS, 12th International Conference of the Financial Engineering and Banking Society, Crete, Greece, 1-4 June 2023

Events organized by the program

1. The inaugural **summer school** hosted by the chair took place from June 26 to 28, 2023, welcoming doctoral students, researchers, and employees of Thélem Assurances. The primary focus revolved around the fundamental mechanisms of Machine Learning algorithms, as outlined in the curriculum detailed in the Training plan:

   Part 1: Introduction to supervised and semi-supervised Machine Learning techniques and applications: general principle, use cases, challenges.

   Part 2: Interpretability of Machine Learning methods.

2. The second edition of the **summer school**, on *Advanced Machine Learning,* will be held from June 26 to 28, 2024, at Hôtel Dupanloup.
3. A half-day **training session**, focusing on the principles and applications of Natural Language Processing (NLP), is scheduled for April 8. This training, structured in accordance with formal academic standards, seeks to provide Thélem employees with essential insights and skills in this dynamic field.

OTHER HIGHLIGHTS

Awards, scientific recognition, organization of calls for projects, involvement in master's courses, PhD program visiting researchers, etc...

Organisation of calls for projects

A call for papers (available here) was launched in February 2023, inviting researchers to submit first-class theoretical and/or empirical research papers related to the following topics:

Topic # 1 Insurance

- Impact of new technologies on the insurance offer
- Natural disaster risk insurance – Climate risk and insurance

- The impact of inflation on the insurance sector
- Data Science for insurance fraud detection
- Pricing in insurance and fairness
- Governance of AI algorithms in the insurance sector
- Interpretability of AI algorithms in the insurance sector

Topic # 2 Finance

- Detection of financial fraud (banking, insurance, social security contributions, etc.)
- Fight against money laundering and terrorism financing

The Steering Committee selected the following projects:

1. *Forecasting natural disaster frequencies using nonstationary count time series models,* submitted by Yang Lu (Assistant professor, Concordia University, Quebec, Canada).
2. *Copula-based fraud detection,* submitted by Maxime Dotta, Xavier Milhaud, Denys Pommeret (Aix-Marseille University, Ecole Centrale).

Publication of results: September 2023.

The academic articles associated to the selected projects are:

1. Pei, J., Lu, Y., (2024). Forecasting natural disaster frequencies using nonstationary count time series models. Working paper (link)
2. Dotta, M., Milhaud, X., Pommeret, D., (2024). Copulas based fraud detection. Ongoing project.

PhD program visiting researchers & training

Y. Kougblenou, the PhD student recruited within the project, completed a research stay at KU Leuven from June 3 to June 25, 2023. Another visiting is scheduled for June 2024.

Additionally, he participated in a training day on conformal inference (the 7th Mathematical Statistics Day - January 18, 2024) on January 18, 2024.