



Contribution ID: 53

Type: **not specified**

## Disease Validation in Health Insurance Claims via Concept Extraction based on Natural Language Processing

Healthcare fraud is a significant issue that leads to substantial financial losses and compromises the quality of patient care. Traditional fraud detection methods often rely on rule-based systems and manual audits, which are inefficient and lack scalability. Machine learning methods have begun to be incorporated in the fraud detection systems of insurance companies; however these methods mainly focus on tabular data. However, each claim may also be accompanied by a wealth of unstructured textual data, such as hospitalization summaries, medical opinions, clinical notes, surgical operation reports and discharge papers. These provide a fertile ground for the application of Natural Language Processing (NLP) methods. Transformer architectures have been at the forefront of research in NLP and Deep Learning in general. Language models specific to the medical domain have been recently introduced, for example BioBERT and BioGPT. In this paper, we leverage traditional NLP techniques like Topic Modeling in addition to the newest advances in transformers to create a framework for detecting diseases concerning each case by extracting information from the available unstructured data. This framework can then be used as a fraud detection system by checking if the extracted information is consistent with the information contained in the claim.

### Special/ Invited session

Young Statistician

### Classification

Both methodology and application

### Keywords

topic modelling, fraud detection, health insurance

**Primary authors:** Mr MICHELAKIS, Charalampos-Panagiotis (University of Piraeus, Department of Business Administration); Mr PAPAGEORGIOU, Grigorios (Department of Civil Engineering, University of Patras)

**Co-authors:** Prof. ECONOMOU, Polychronis (Department of Civil Engineering, University of Patras); BERSIMIS, Sotiris (University of Piraeus, Greece)

**Presenter:** Mr MICHELAKIS, Charalampos-Panagiotis (University of Piraeus, Department of Business Administration)

**Track Classification:** Young