

Title

How to make use of Unstructured Data – Critical Illness Claims Classification via Natural Language Processing

Speaker

Antoine Ly, SCOR

Abstract

(Re)insurers have a wealth of data at their disposal. Until recently, actuaries concentrated on exploiting structured data, leaving text data largely ignored even though it is collected and stored.

With the digital transformation of the insurance industry, however, exploiting unstructured data is becoming essential. The techniques offered by NLP mean that actuaries can explore new territories and have the opportunity to add value throughout the insurance value chain, from product design to claims management.

Based on the practical example of Critical Illness claims classification in the Chinese language, we will elaborate on modelling techniques and drivers for expected prediction qualities. After looking at and interpreting some of the results of this case, we will present an outlook for other languages.

Biography

Antoine Ly is Head of Data Science at SCOR Global Life. Doctor in mathematics on the application of machine learning to insurance, he is also graduated from ENSAE Paris and

Master's in science "Data Learning and Knowledge" of Sorbonne University. He teaches in the Data Science program given by the French Institute of Actuaries in collaboration with Pr. Elie and Pr. Charpentier. Antoine also teaches "Machine Learning with python" and "

Distributed computing" at ENSAE Paris. Antoine is a certified Actuary, member of the French Institute of Actuaries and contribute to the EAA taskforce in Artificial Intelligence and Data Science.
