

Cracking the Code of Sick Absence: Using Data and AI to Predict the Future

Marcus Lindell & Casper Schwerin

The matter of sick absence is something all companies must handle. Despite this, they lack the tools to forecast how the absence will look in the nearby future. Statistical and machine learning approaches enable the sick absence to be modeled and predicted. By using appropriate additional information, more accurate predictions can be made. In the future, such models might be tailored to individual companies.

Sick absence is a very costly problem, accounting for more than 2% of the GDP of countries in the EU. Studies, both recent and historic, have shown connections between economical development and work absence. The discussion about the subject have also been found to affect future absence levels. Recent matters such as the COVID-19 pandemic have forced workplaces to take action and renew policies in order to maintain operational. In an ever-digitized world, MedHelp Care, an e-health company, offers companies a platform for registering and handling workplace absence. This has lead to a surge in available data, which could be explored in order to deepen the understanding concerning the topic.

MedHelp's data was transformed from a collection of datapoints into a time series so that it could be modeled. By modeling the data, it is possible to find reoccurring patterns in the sick absence and in doing so, form educated guesses, predictions, regarding how it will vary in the future. Such information could be very useful for companies when it comes to planning and staffing. The project used both statistical and machine learning models. Models using

sick absence in combination with additional information from either the Swedish stock market or tweets concerning COVID-19 was also created.

The results show that statistical models can predict weeks and months quite well (both with less than 7% average error) and that machine learning methods are better at predicting quarters of a year. Adding information from the stock market improves weekly prediction accuracy, but does not help with longer predictions. Using Twitter worsens performance all-round. The models may also be used to predict monthly averages of sick absence.

One can conclude that the models prove the possibility of predicting daily absence and could, in theory, inform companies about the future. However, due to their complexity, they are slow and require a lot of work to produce good predictions. As of now, the models can predict the larger population but are not useful for predicting individual companies. Further research is needed to develop new methods for prediction and to investigate the usefulness of factors such as economics and social media attention.