Analysis of the combination of a machine learning model with a rule engine

With the development of the digital world, the techniques used in Natural Language Processing (NLP) are constantly evolving. While machine learning is the main technique used today, it may be reaching its limits, especially in terms of accuracy. This is why we analyze the results of combining a machine-learning model with a rule-based classifier.

Introduction

Whether to use rules or a machine learning model depends on the use cases, the objective of the task to be solved, but also the data we have at our disposal. Each one achieves different objectives and performs more or less well depending on the application. However, each one has its own limitations and we need to look for more efficient solutions. This is why, in this work, we propose the combination of a model with a rule-based approach. We explore the concept of combining rules and machine learning models based on a real-world use case.

Context

Our use case is a Swiss start-up that supports medical and legal professionals in their administrative workflow on accidents and disease diagnosis. Their application allows to upload PDF documents and offers many tools to reduce the time spent searching for information in these files. One of the features is the classification of documents based on a complex and custom rule engine that is entirely developed in-house. The main challenge of this feature is to further improve the results of this classifier.

Objectives

This is why we propose a combination solution with a model. Our goal is to explore the potential benefit of combining the outputs of a machine learning model with the start-up's rule engine.

We explore two main approaches regarding the combination of a rule and model based method. In a first set of experiments, we investigate on how to combine both outputs. To do so, we analyze whether the model can be used to validate or correct the rule engine's predictions. Furthermore, we try to define the best method for each class and if the objective is directed towards precision or recall. In a second step, we suggest other approaches to integrate the model into the start-up's classification system.

Results

The results of the validation or correction of the outputs from the rule-based approach with those from the model are rather poor. Only a subset of the data can be considered as reliable, with a small margin of error. Then, if we group the results by label, we observe a notable difference in quality between the two methods. Therefore, we can suggest which classifier performs better for certain labels. Moreover, when the focus is on precision or recall, we also suggest one of the two approaches. This varies by class or if we look at the results as a whole.

As the fusion of the rule engine with the model is very limited, we also propose three possible applications of the model that are tools for the developer working on the rule engine. In addition to helping writing the rules, the model can make the labeling of data more efficient. Finally, it can be associated with a neural model interpretability tool, in order to better understand the keywords of a prediction.

Yael Tatjana Iseli