Machine Learning Engineer Nanodegree

Capstone Proposal

Rodrigo Moreira Fagundes December 14th, 2018

Proposal

Domain Background

In such a large developing country as Brazil, with so many different economic and cultural setups, tackling modern slavery efficiently is a huge challenge. According the Observatório Digital do Trabalho Escravo (https://observatorioescravo.mpt.mp.br/), from 2003 to 2018, 44,229 people were rescued from degrading working conditions in 3,318 inspections (13.33 rescues per inspaction). 2,006 successful diligences took place in 766 of the 5,570 brazilian municipalities (a coverage of 13.75%). If we add the 1,847 inspections with no rescues, the coverage raises to a 37.9%, with 2,112 locations, but it reveals a success rate of 52.06%. Since 2015, the number of diligences have benn falling, returning to the yearly inspections frequency seen in 2003-2007.

Some regions are not accessible, taking too long to mobilize an inspection - allowing perpetrators to move their operations or just hide their illegal aspects during the audit. Traditionally, the inspections are mobilized based on a denounce and evidences that support it. The most vulnerable people, though, lack the opportunity to reach government agencies.

In opposition, urban centers with high population density have a large number of enterprises to be verified, rendering a traditional coverage goal unrealistic - both due to cost and manpower. Modern slavery practices in urban centers tend to be disguised as just poor labor standards or practices, facing long disputes before estabilishing the perpetration.

Problem Statement

Given the scenarios, government agencies have to craft a way to concentrate its resources in targets that reach more vulnerable people. One way to create such a prioritization can be trying to identify, using classification, the locations that would most likely to result in more rescues per diligence. Some inputs are very promising for that task, such as municipalities profiles - based on census data provided by IBGE (Instituto Nacional de Geografia e Estatística) - and the record of previous inspections - from MTE (Ministério do Trabalho e Emprego). The output of the classification will indicate municipalities in which inspections may result in maximum rescue rates.

Datasets and Inputs

For this study, we're using three datasets. The first is a collection of information on municipality's census, collected by IBGE (Instituto Nacional de Geografia e Estatística), available to the public. The comprehensive dataset, with no relevant missing data so far, is a suitable source for identifying profiles and similarities between locations. Is is also a reliable dataset, once the census conducted by IBGE goes through a rigorous methodology.

The second and third datasets will be the disidentificated registers of operations and inspections. They contain information on the municipalities where inspections took place, how many people were rescued from degrading work conditions, their origin and where they claimed to reside at the moment. In the current study, they'll be used as a base for risk rating, which in turn will become the label for classification.

Solution Statement

One solution to the problem can be resource optimization by defining high priorities munuicipalities based on statistical inference. By using municipalities similarities and previous diligences data, it's reasonable to focus on locations that are most likely to result in a more effective action, rescuing more people in a single inspection, for instance. By prioritizing municipalities according to the distribution of rescues per inspection, the model can be repeated, hopefully with decreasing numbers of perpetrations.

Benchmark Model

There's no benchmark model available but the actual rating of rescues per inspection exclusively using the traditional resource placing strategies. Therefore, we'll split the data in two equal parts: a subset for training and validating, that will be used as the benchmark model, and another to test the model.

Evaluation Metrics

To measure the success of the solution, the municipalities that have no previous records flagged by the solution should be split in control and test subjects. The testing group should be subject to a task force and the result should be compared to the control (disclosed at the end of the evaluation, to avoid bias). A number of false positives should be expected, and compared to the metrics revealed in the modeling phase.

If the rescues per inspections in the testing group is higher than the control group (error margin considered), the model is proven effective. The result should also be compared to the overall rating and to the previous records - it can reveal a migration of modern slavery practices.

For the purpose of the current study, we'll use the inspections dataset to train, test and validate the model, comparing the results using precision and recall metrics.

Project Design

First, we'll estabilish three menace rating (LOW, MEDIUM, HIGH), using the register of previous inspections using the distribution of rescues per inspection. The levels will be, LOW, MEDIUM and HIGH, from the terciles of the selected indicator.

Since there's no previous definition of such rating and we're establishing the levels according to a normal distribution, the labeling class should be balanced (each tercile holds a third of the data). After that, we'll label the municipalities that actually had inspections according to these ratings and the average of the rescues per inspection in them. This rating will be added to the census data as the label for the classification later on.

The census data holds a high dimensionality. In order to keep the explainability of the resulting model, we'll use random forests to understand how the dimensions in census data contribute to the model. Beforehand, all municipalities will be disidentificated for the purpose of this study.

On the first run, no outliers will be removed and we expect to identify features that contribute little (or not at all) to the outcome - we expect to keep at least 95% of explainability and a minimum of 10 features). Then, with the dimensionality reduced, we'll revisit the sliced dataset, remove the extreme outliers (above Q1 - 3*IQ, where Q1 is the first quartile and IQ is the interquartile) and normalize the data.

The secund run will consist in the evaluation model for this study. The dataset will be split in two. The first segment will be used for training and testing, using a random forest with 10-fold cross validation. The generated model will be used to classify the second part of the dataset. The results will be compared using precision and recall metrics.

Future work

Since the results of this study is expected to be included in the Observatório Digital do Trabalho Escravo, the model will be built using the full inspections dataset. Then, it will be applied for labeling municipalities with no inspection record. Those with a label HIGH will be flagged, as the ones already identified in the inspections records. The final output is a brazilian map of municipalities colored by rating level, revealing which areas should be subject to a more thorough investigation.