CENTRAL EUROPEAN UNIVERSITY DEPARTMENT OF ECONOMICS AND BUSINESS MS IN BUSINESS ANALYTICS PROGRAM

DAVID GABOR UTASSY CAPSTONE PROJECT SUMMARY Location-based Demand Forecasting @Hilflylabs

Faculty supervisor @CEU, mentor @Hiflylabs: *Eszter Somos Data Solution Advisor* Sponsor @Hiflylabs :

Szabolcs Biró Head of Advanced Analytics

Budapest, 2021

EXECUTIVE SUMMARY

The stakeholder of this project was Hiflylabs Zrt. which is providing business intelligence solutions, consulting, and mobile app development for their clients. The goal was to develop a platform/ code-base and methodology for Hiflylabs, to be able to offer location-based demand forecasting for future clients. It is important to highlight, that the point of this project is not to provide real analysis, but to deliver an applicable and reusable framework that supports Hiflylabs to satisfy their client's needs.

After some experimental data analysis and reading articles related to location-based demand forecasting, I defined the following keywords to follow during the project:

• General:

The framework and the codebase that is provided by this project are as general as possible. This means that is applicable in different industries (supermarket, petrol-station, drugstores, restaurants...), and various locations (different cities, countries, or even continents)

• Add explanatory variables

The first way the framework is providing support for future analysis is by providing additional location-based explanatory variables.

• Find popular places where the traffic is high

When we are talking about valuable location-based information from the point of demand forecasting, the movement of potential customers is the one that we need. There are ways, that can give us such information, but most of them cost a lot (for example buying such data from telecommunications companies). For this reason, my platform is providing some clever proxy variables that can be used to get such information.



After the initial experimentations and the definition of the project scope, the core part was to develop a "Location-based explanatory variable collector" framework which followed the following high-level structure:



The input of this framework is a table containing the shops of the client. While the client shop list is going through this framework, it extends it with various explanatory variables that can be used to create more accurate prediction models in the end. The following modules are provided by the platform that the user can call with parameterizable function calls to extend the analytical table. To summarize the findings and experience of this project the following table contains the pros and cons of the modules:

	PROs	CONs
Google Places API	 super general up to date connected with commercial brands convenient API great for competitor scraping 	 Google protects it hard limited information provided
GHLS	 super general cross-border free captures living area 	outdated (2015)(250 m resolution)
BKK Futár	 very detailed up to date free convenient API captures people's daily traveling 	only Budapest
OSM (public transit stops)	 super general up to date free open source captures people's daily traveling 	no transit stop schedules

In the technical documentation, each module is explained in detail.

As the core part of the project was to provide a code-base that can be used in the future, the biggest challenge was to make the platform reproducible. To make the use of this platform handy and convenient for the users (data scientists/ analysts) I made the following actions:

- Use GitHub repository
- Modular structure
- Detailed documentation, also in README.md
- Docstring documentation of functions
- Logging to give feedback to the user
- example_calls jupyter notebook, to show examples of use

At the end of the project, some experimental was analysis made on the explanatory variables that are collected by the platform (implemented during this project). From the results, we can see, that these variables have real value in them.

As the project is very broad and this project had to be ended somewhere, several further development possibilities are provided to Hiflylabs to give directions to them that are worth working on in the future.

