Recruitment test – Volvo Group Connected Solutions (VGCS)

For the roles of Data Scientist and Senior Data Scientist

Background

The Volvo Group has a multitude of connected vehicles and assets delivering data to VGCS using telematics solutions. This information is a valuable source to better understand how customers utilize Volvo products for their daily operations, as well as pin-pointing and revealing issues related to Volvo components and software.

In this invented scenario, our contacts at Volvo Trucks (VT) have become aware of the amount of data we manage and raised two questions. Firstly, they would like to better understand the potential in the telematics data for developing their business, and secondly, they would like to use the data to identify the root cause of a severe quality issue that is currently bugging around one fifth of the vehicle population.

As a data scientist, it is your task to help them with a first assessment. We have a workshop scheduled with the stakeholders at Volvo Trucks and we need to come prepared. Data has just been extracted from our systems and the legal department has approved GDPR compliancy for our analysis. We are ready to go.

Task

Your task is to get to grips with the dataset (see next section) and assess it in view of the two purposes mentioned by our VT contacts. Feel free to go about this in any way you like. Examples of the type of questions that could be of interest are:

- What information can be extracted from the vehicle GPS traces? Any informative visualization that can improve the understanding?
- How would you go about building an interactive view of this data and its possibilities?
- Are you missing some vital information in the data? What should be added to make your life as a data scientist easier, and what new possibilities could this generate?
- A hot topic is understanding how goods flow in the European transport market. Can this data be used to estimate transport flows on the roads?
- Are there clear symptoms indicating whether a vehicle has quality issues or not?
- Some vehicles have unknown issue status (file "_cx.csv" below). If we treat them as a test data set, could we already now, based on the patterns from c1 and c2, indicate if they are likely to have issues or not?
- What recommendations would you give our stakeholders for further analysis and next steps, given the information you have found in the data?

Information on provided data

Three files have been extracted by our data engineers. The data is divided into two groups depending on if a vehicle had issues or not. Vehicles for which the issue status is still unclear have been separated into a third group.

- trackdata_c1.csv – first group of vehicles
- trackdata_c2.csv – second group of vehicles
- trackdata_cx.csv – vehicles with unknown issue status

The files contain GPS traces of the vehicles, with each vehicle associated with one trip (representing its typical usage pattern). The format is similar in all files and the columns are:

- uid – unique vehicle identifier
- row_number – counter for the data points that make up a trip (i.e. a "time coordinate" for the trip)
- lat – latitude position in WGS84
- lon – longitude position in WGS84
- ele – elevation above sea level in meters

Presentation

The results of this case study should be compiled and presented during a meeting at VGCS (beamer with HDMI connection will be available), with our Volvo Trucks contacts as your target audience. You will have maximum 20 minutes for the presentation followed by 20 minutes for questions and discussion. We value the time you have invested and will try our best to provide constructive feedback and recommendations to help further develop your skills and abilities. During the presentation, we will evaluate the following aspects:

- Data science abilities in terms of how the questions are approached and analyzed, as well as the choice of techniques and their implementation/execution
- Business mindset and understanding of stakeholder needs
- Presentation and communication skills

We appreciate if any reports or analysis/scripts are shared with us for evaluation. You are free to choose any tools, techniques or methods you like to address the questions.

For questions, please contact:
Fredrik Moeschlin fredrik.moeschlin@volvo.com
Adam Stahl adam.stahl@volvo.com