Task: Predicting the Expected Travel Time of a vessel

PBL in Software Development

Short Description:

- Machine learning becomes more and more attention in many application domains, “be prepared” for decision making
- In maritime logistics AIS data can be used to support optimization, safety, …
- A concrete task is to predict the expected travel time (ETT) of a vessel for organizational tasks in a port.

Task Description:

- Harbors, locks want to optimize their daily business
- A tool must be built to predict the ETT of a vessel based on historical data
- The provided AIS data are nor reliable. Faulty data, missing data.
  Data cleaning is necessary to create valid models. You will get two routes.
- Some of the attributes are useless for prediction. Data selection and creation is a necessary task
- There are several learning methods that can be used for the ETT prediction. Experiments are necessary to find out which method or combination of methods is the best for the task.
- Beside the prediction, a graphical use interface should support the analytical view on historical data and actual data
- Prediction services can be provided by different “agents”. A distributed decision service is necessary that integrates the overall prediction of different providers.

Subtasks: An independent task is to ensure that the software components are always tested well.

ST01: Development infrastructure

GitLab of TUHH is used for the complete development process to support future maintenance issues. The project is at source level. At the end, the product owner will take over the infrastructure. As the product owner, I want to have access to all development stages. Setup the infrastructure so that at the end of each Sprint the product owner can download the resulting product including scripts to start the application.

ST02: Data preparation

The raw AIS data comes from an AIS data provider. The data has to be cleaned, prepared for the later stages. An initial set of features/attributes must be selected for machine learning. All decisions must be made transparent for me (the product owner), e.g., selection of features, preparation of features, trips that are used/not used, … .

This means, graphics, documentation and explanations of decisions.

ST03: Selecting machine learning methods for ETT and AIS data

As the product owner, I want the best ML method for the prediction of ETT - lowest error. This dependent on the methods, parameters chosen and data preparation.

Test at least 4 candidate methods.

All experiments must be executable by me, the product owner.

An explanatory report must be delivered to document the results of the experiments.

ST04: Selecting appropriate tools for the implementation

As the product owner, I want to understand the overall architecture of the application and the decisions for it.

The implementation environment must be described. The installation of the final application must not dependent on specific OS. A docker based implementation is preferred.

ST05: Creating the environment for the basic ETT setting

Build the ETT application for a short tour with only one predictor agent. Creating a GUI for problem visualization. The GUI app should show historical data of vessel movement and speed. On request of an ETT, the actual data must also be shown.

ST06: Building an infrastructure for distributed decision making

A broker agent can serve requests of clients that want to have an ETT prediction of a vessel. The broker agent learns the weightings for all service providers and integrates the prediction of all ETT predictor agents.

ST07: how to integrate this in the overall development process (GitLab, CI)

- You should integrate your testing environment with CI so that tests can be executed when new code is deployed.