

To what extent predicting variation of treatment score is affected by demographic variables

Team Project Proposal

January 8, 2021

1 Problem

The project will use data from tinnitus patients that visit a certain clinical center and are offered treatments. At admission to the center, patients fill out questionnaires that assess the severity of their symptoms. These questionnaires are also filled out after finalizing the treatment.

The target of the project is the analysis of how demographic variables impact the predictions of the treatment outcome (measured by the variation of the questionnaire score). Splitting the data into certain groups may or not impact the performance of the prediction models.

Firstly, sampling methods should be studied and implemented. After, simple regression methods (linear regression, lasso, ridge, ...) should be tested per group and the performance of the learning models should be compared with the scenario without any grouping.

2 Requirements

In order to ensure the success of the team project, the students should meet the following requirements:

- familiar with the ML pipeline and be able to implement it
- understanding of regression methods
- knowledge of feature encoding techniques
- fluency in Python (preferable) or R
- familiar with evaluation techniques
- familiar with sampling methods

3 Duration & Supervision

This project is planned for the Summer term 2021.

First meeting should be arranged, latest, for the week of the 7th April.

Interested students are invited to email: clara[dot]puga[at]ovgu.de