ESTIMATING BOLT USER PENETRATION IN CITY DISTRICTS

Members:
Rasul Agharzayev, Orkhan Jamaladdinov, Taavi Pipar, Kristjan Ševerev

Project owner:
Polina Shevchenko
PROBLEM STATEMENT

- **Why?**
  - Better resource distribution between city districts in Tallinn
  - Improve our skills by solving real life problem with machine learning

- **How?**
  - Bolt data and data from external sources
  - Machine Learning model (regression)
APPROACH

- Official statistics
- Land use
- Social media usage
- Bolt usage
DATA

Spatial data

- Grid 1x1 km

Descriptives

- Time unit
- Population
- Land usage
- Official statistics
<table>
<thead>
<tr>
<th>Model</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ Time unit 1h</td>
<td>▪ RMSE: 10.3164</td>
</tr>
<tr>
<td>▪ Gradient Boosting Regressor</td>
<td>▪ Score: 0.9116</td>
</tr>
<tr>
<td>'n_estimators': 450,</td>
<td></td>
</tr>
<tr>
<td>'max_depth': 4,</td>
<td></td>
</tr>
<tr>
<td>'min_samples_split': 22,</td>
<td></td>
</tr>
<tr>
<td>'learning_rate': 0.2</td>
<td></td>
</tr>
</tbody>
</table>
In order to predict penetration, we need to know where people are located in a particular area and time.

The smaller the area and time, the fewer factors affect it and the easier it is to give accurate results.
THANK YOU/ÄITÄH