|  |
| --- |
| Machine Learning Project: Used car price prediction (Format 2) |
| 2023-04-19 |

# Overview

## Participants:

Julius Čereška, Simeon xd?

## Dataset:

[100,000 UK Used Car Data set](https://www.kaggle.com/datasets/adityadesai13/used-car-dataset-ford-and-mercedes)

Around 100k variables, 10 unique car brands with various models, 9 features.

## Project proposition

Create a model capable of predicting car resell price, based on the feature inputs.

## Project format

This project will be based on the 2nd format, focusing on exploring several machine learning algorithms for regression problems:

* Linear Regression (with/without regularization)
* Artificial Neural Networks
* Support Vector Machines
* Decision trees
* Random Forests

## Project implementation

Project will be implemented in several steps:

1. **Data exploration**  
   Dataset will be explored to gain insights about features and understand the problem better. This will also help with feature engineering and selection.
2. **Feature engineering and selection**  
   Features will be cleaned of outliers and normalized if needed, composite features created for better prediction power. This step is intertwined with the algorithm implementation step, since different features might be better for some algorithms and worse for others.
3. **Algorithm implementation**Different machine learning algorithms will be explored and implemented with different parameters in hopes of achieving best results.
4. **Results**Results of the work will be presented, model insights discussed and best model for problem solution proposed.