

Hands-on W11: Assignment/Project 2 – Regression

Team Project (one team consists of 4 members –same team with the previous one–)

As the previous assignment, the objective of this project is to explore your capability as data analyst. You are challenged to implement data mining method starting from scratch, i.e., you must crawl dataset from a website. Then, you need to implement several regression methods on data you already crawled, analyze and evaluate.

Basic tasks you need to perform in this project: (Note: you can add additional tasks that you think necessary for this project)

1. Visit OLX website, find list of all second hand cars that have been sold in that website. Link : https://www.olx.co.id/mobil-bekas_c198
2. Crawl the data from that website. The data should include the price offered for the second hand cars and several related features, e.g.: brand, model, year, mileage, etc.
3. You should define how many data and features that you've crawled from OLX.
4. Build regression models to predict the price for second hand cars. This step should include EDA, pre-processing, creating model and evaluation. You **may** use several regression methods, do benchmarking, employ additional techniques, such as ensemble method, to improve the prediction accuracy, **etc.**
5. Please use the most appropriate evaluation metrics for this regression project.

Deadline: **03 November 2019, 23:55 WIB.**

Deliverables:

1. Technical report:
 - a. A4, Font Times New Roman, 11, single line spacing
 - b. Report of step 1 – step 5, highlight some obstacles (if any)
 - c. Report your conclusions

2. Presentation

Nb: In the end of slide and report, please write members' names and their own contribution during this project.