ABSTRACT

DESIGNING DATA PREPROCESSING BASED ON WEBSITE USING WEB SCRAPING TECHNIQUE TO PRODUCE *DATASET* VISUALIZATION IN SOCIAL MEDIA INSTAGRAM

By

AJENG RAKHMAWATI
NIM : 1202162144

In decision making and mining processes, data is one of the most important factors. For this reason, data is considered valuable in this regard. Before it becomes valuable data that can be used, it needs to be processed first. It takes several processes to process data so that it becomes high-value data. At the data mining stage, data processing becomes an important step that needs to be done. Data collected from a variety of different sources still have a different format or can be said as raw data that is still not feasible to be analyzed. With preprocessing data, raw data obtained from various sources with different formats will be processed to produce a dataset. The resulting dataset can be visualized to produce insights that can help in the decision making process. Output generated in preprocessing data that has been validated or is ready for processing can be further processed into the data mining stage. This research focuses on designing web-based preprocessing data using web erosion techniques to produce datasets as output. The resulting dataset will then be visualized using tools to generate insights in supporting the decision making process.

Keywords: *dataset, data collecting*, *data preprocessing*, *social media* instagram, web scraping.