In this research, extraction is carried out to change the text on an artifact to be processed through Text Mining to process and analyze the text whose results are validated, using the coefficient equation formula. There is a discrepancy between the flow in the Activity Diagram and the steps performed in the Use Case Description. Based on these problems, this study aims to extract an artifact to find similarities between the Activity Diagram and the actions taken in the Use Case Description and perform text analysis using the Text Mining method to obtain similarities, then validation and reliability are carried out. The extraction process is broken down into several documents according to the artifacts used, for Use Case Description based on Step Performed, namely: UC01, UC02, UC03, UC04, and UC05. As for the Activity Diagram, namely: AD1, AD2, AD3, AD4, and AD5. The results obtained from this study are the highest similarity between activity diagrams and use case descriptions in sentences UC04 and AD4 with a total similarity of 0.51394086. The highest similarity between words is located in words "request" and "application" with a complete similarity of 0.9231. The Kappa score with the AC1 Gwet formula using the Python programming language is -008583 which means "Less than Opportunity Deal", but it is different from the Kappa Score value using a questionnaire filled out by the Expert which is 0.92246 which means "Almost perfect".

Keywords—Activity Diagram, Step Performed, Extraction, Similarity, Text Mining, Validity, Reliability.