Abstract

Software requirement specification is a document that can be used as a guide for developers to develop applications. This study uses SRS from the Penguin application to help determine the development of class diagrams based on use case and sequence diagrams using the text mining method. The results of this process will be calculated for similarity, after which validation and testing will be carried out using Gwet's AC1 and Cohen Kappa. Based on the results and discussion, three artifacts were formed, namely actors from use case diagrams (AUC), objects from sequence diagrams (OSD), and class names from class diagrams (NCD). The three artifacts produce two comparisons in the formation of class diagrams. The first comparison is between AUC and NCD, with the highest cosine similarity score of 0.666. From this score, the resulting construction of class diagram component names is seller and customer. The first comparison also resulted in a score of 0.088 for Cohen Kappa and 0.756 for Gwet's AC1. Furthermore, for the second comparison, between OSD and NCD, two results were obtained with the same score, namely 0.9. This score resulted in the formation of class component names such as seller, transaction page, revenue page, expenditure page, and penguin app system. And the second comparison has a Cohen kappa score of 0.112 and 0.926 for Gwet's AC1 score. The results of the Cohen Kappa score, and Gwet's AC1 can be used as recommendations for improving class names that match the actors names in use case diagram, and object names in the sequence diagram.

Keywords—software requirement specification, class diagram, use case diagram, text mining, similarities, validation