1. Introduction

Based on Ali Olow's research, [1]. SRS is a document that serves as a guide in making an application. This document is the result of discussions between developers and application users. In this SRS, various requirements must be approved by application developers and users. [2].

Unified Modeling Language (UML) is one of the most widely used standard modelling languages. The function of UML is to model a system design from the application to be made [3]. UML has several types of diagrams, such as use case, sequence, and class diagrams [4]. The class diagram is one of the essential parts of UML, because in the class diagram, there are components containing class names, attributes, and relations each class has. So this section is needed in describing the system design of the application [4].

This study uses SRS from the Penguin application. This application assists small and medium businesses in recording store finances and bookkeeping. In SRS Penguin, there is an artifact class diagram. However, the class diagram artifact components have not met the compatibility between artifacts, because there is no link between class diagram artifacts, such as use case diagrams and sequence diagrams, as well as the absence of a validation that tests whether the formation of the class diagram component is following the other artifacts. Therefore, a deeper understanding is needed for application developers when creating class diagrams in an SRS document.

The purpose of this research is to explain how class diagrams are formed. Using the text mining approach on use case diagrams and sequence diagrams. This method is used to collect information sourced from the Penguin's SRS application by calculating the suitability of class diagrams based on artifacts, use case diagrams, and sequence diagrams.

Several contributions will result from this research, namely :

- Extract using the text mining method on the SRS Documents of the Penguin application for artifact class diagrams, use cases, and sequence diagrams.

- Create class diagram artifacts based on text mining extraction results from SRS Penguin.

- Validate the class diagram results using Gwet's AC1 and Cohen Kappa.

- Producing the formation of class diagram artifact components based on the engineering results of the extraction process and similarity with text processing.

The rest of this paper will be divided into several parts. The related work section contains studies related to this research topic. The methodology chapter contains the processes carried out during the research. The results and discussion chapter includes the results obtained in the research process and their explanations. And the conclusion has the conclusions from the research results obtained.