

## 1. INTRODUCTION

Currently, education is always undergoing various changes and problems as well as improvements to continuously strive to enhance the quality of education itself. One of the efforts to improve the quality of education is by implementing all the procedures set by the accreditation body. With accreditation, certain criteria have been established as benchmarks to assess the quality of education [1]. Universities, as the pinnacle of the education system in Indonesia, design learning systems to equip students with the competencies needed in the workforce. To prepare for this, universities, especially faculties and study programs, prepare curriculum as a guide for the learning plans for students [2]. In the curriculum, PLOs (Program Learning Outcomes) is defined as a guide for the criteria for student graduation. To support PLOs, CLOs (Course Learning Outcomes) is defined as a guideline for conducting learning activities in a course. To ensure the fulfillment of PLOs and CLOs, an assessment mechanism is implemented. In the preparation of PLOs and CLOs, inconsistencies between the narratives of PLOs and CLOs are often encountered [1]. This misalignment can lead to inaccurate assessments, errors in curriculum design, failure to meet accreditation standards, and students being less prepared for the workforce or further studies [1].

Therefore, we developed an alignment analysis between PLOs and CLOs as an automatic mechanism to check the alignment between PLOs and CLOs. We examined the alignment between PLO and CLO using a text similarity-based method. Text similarity is an instrument in NLP that not only considers the semantic similarity between texts but also analyzes the semantic properties shared by two words, which are usually used in NLP tasks, one of which is document matching [3]. We chose tf-idf and fastText as text representation methods and cosine similarity as the text distance method. Tf-idf is used to create vectors that will be calculated using cosine similarity [1]. FastText as a word embedding method was chosen because in previous research [1], it was suggested that future studies could use text representation methods that pay more attention to the semantic relationships between words.

Based on previous research, fastText has proven that it has better performance compared to glove, lexvec, and conceptnet [4]. In the previous research [5], it also showed that fastText has better accuracy of 0.89 compared to tf-idf with PCA. Also, since the dataset in our study has a wide variety of words, fastText was chosen because, based on research [6] it is superior in handling the variety of words in sentences than tf-idf. In our study, from the results of the text representation using tf-idf and fastText on CLOs and PLOs, the results calculated using the cosine similarity method.