

1. INTRODUCTION

The Merdeka Curriculum uses several teaching tools as learning media. These teaching tools include textbooks, teaching modules, and learning videos. However, textbooks must be used in the learning process[1]. In accordance with Regulation 22 of 2022 issued by the Minister of Education, Culture, Research, and Technology of the Republic of Indonesia, textbooks are required to adhere to specific quality standards. These standards encompass various aspects, including process standards, rules for acquiring manuscripts, and guidelines for book publishing.[2].

According to the regulation, text readability is one of the book quality standards[2]. Text readability is a determinant of text that readers can understand and read quickly[3]. The level of text readability in textbooks affects students' enthusiasm and interest in reading[3]. Reading is an important activity and a parameter for the success of the learning process[4]. It is necessary to have a textbook that aligns with the student's reading level regarding readability.

There are two ways to measure text readability; traditional approaches (readability formula) and the utilization of natural language processing. The readability formula is a traditional way to calculate text readability by using linguistic features. Besides, NLP can also be used to extract text features.[4].

Several studies have discussed text readability in several languages. Chakraborty et al.[5] analyzed text readability for Bengali text using a readability formula and deep learning model building. Tajed et al.[6] conducted an analysis on the complexity of Slovene text by employing both a readability formula and the construction of a machine learning model.

In this research, we analyze text readability for Indonesian text using a deep learning algorithm (Bi-LSTM) and readability formula. Bi-LSTM method produce text readability category, complex and simple, based on text language complexity. On the other hand, readability formula output is the reader/student grade level, which can be converted to reader's age. The output produced by the readability formula is the grade level converted to the reader's age. We follow readability formula used in [5] work, which utilizes linguistic features: word length, sentence length, word count, sentence count, syllable count, and count of difficult words.