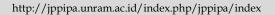
JPPIPA 9(5) (2023)



## Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education





## Corpus Development of Forest Fire using Manual Annotation

Sadam Andhika Putra Darmawan 1, Warih Maharani\*2

1,2 Informatics Department, Telkom University, Jl. Telekomunikasi No. 1, Terusan Buah Batu, Bandung, Jawa Barat, Indonesia

Received: May xx, 2023 Revised: May xx, 2023 Accepted: May xx, 2023 Published: May xx, 2023

Corresponding Author: Warih Maharani wmaharani@telkomuniversity.ac.id

DOI:

© 2023 The Authors. This open access article is distributed under a (CC-BY License)



Abstract: Forest fires and haze are very common, even annual occurrences in Indonesia. These forest fires are common during the dry season. The impact of forest fires is very bad in various aspects, disrupting health, economy, transportation, and even into the realm of politics, especially foreign politics, because the smoke from forest fires can spread to neighboring countries, disrupting the lives of people in these neighboring countries and becoming a topic of conversation on social media Twitter. Twitter is one of the most popular social media in the world, where users can send and read text-based messages called tweets. This makes Twitter a place to discuss various topics. Opinions expressed by users can be used to build a corpus for both Bahasa Indonesia and Bahasa Melayu. However, few studies still address Bahasa Indonesia and Bahasa Melayu, and manual annotation methodologies have not been widely used for both corpus. Therefore, this study focuses on words with the same vocabulary but different meanings in the current forest fires. Data was obtained from Twitter, so informal words were found and then randomly selected with the same vocabulary but have different meanings.

Keywords: Manual Annotation, Bahasa Indonesia, Bahasa Melayu, Corpus, Forest Fire