

Adapting Google Translat Word Embedding for Ara lingual Informat

1st Maryamah Maryamah

*Departement of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
maryamah.18051@mhs.its.ac.id*

2nd Agus Zaina

*Departement of In
Institut Teknologi Sepu
Surabaya, Indo
agusza@if.its*

4th Ahmad Makki Hasan

*Department of Arabic Language
Education
Universitas Islam Negeri Maulana
Malik Ibrahim
Malang, Indonesia
ahmadmakkih@pba.uin-malang.ac.id*

Abstract— The translation has an essential role in Cross-lingual Information Retrieval. Translation using a dictionary is reliable even though it has a limited vocabulary. Translation using google translate, in some cases, using different words used in document target words. The translation process causes word translation to be less accurate to get relevant documents. In this paper, we proposed a new translation approach by adapting google translate using a dictionary and word embedding in Arabic-Indonesian Cross-lingual Information Retrieval. The dictionary is the primary resource used for translation improved by Levenshtein distance and FastText for finding the correct word translation. Google translate is used