Lyric translation from English to other language with meaningful sentences

Abstract

In the constantly connected world where more people are listening to music from different genre and mostly different languages. Some example was the K-Pop hit Gangum Style, which was great hit without people knowing what the lyrics are. What if they was a way to convert the lyrics which are in English to other languages and vice versa.

What are the applications of such a Machine Translation?

We currently have subtitles for English series that are watched all over the world, there is no translation for English songs. What if we could provide a translation from English to other language using Machine Translation, then people will know what the song means.

What are the feature that we would use?

We would use words as features, but also consider the sentence construction so that we don’t break the meaning during the translation.

Is there any research done in this?

There has been some research that are done on machine translation on the part of document translation. Some of research link has been done in <http://web.eecs.umich.edu/~mihalcea/downloads.html>. There has already an application online <http://lyricstranslate.com/> that manually translates language to other.

Is there any dataset available?

<http://lyricstranslate.com/> has about ~90000 translated songs from English to other languages. We could gather data from these to get the sentence construction and data analysis

Brief implementation steps?

There needs to be two approaches – Training the model and Translating the data.

Since we would create a patch to test the model and also verify the translated text in the output. It would be a generative model.

1. Split the text into sentence separated by punctuation such as [(),.] for both the languages
2. Extract the word from the sentences
3. Do the PoS tagging
4. Build the word tree and mapping from one language to another.
5. Estimate the probabilities of the words and return the data.