

Ass 3 – Challenge

Dorin Keshales – 313298424

User name: keshald

- **Can the two languages be distinguished using a bag-of-words approach? Explain why.**

Answer:

The BOW approach it's an encoding in which we sum the words' vectors, but then we lose word order information. In our case, the order is the only thing that distinguishes between the two languages. So, using BOW approach wouldn't work for this task.

- **Can the two languages be distinguished using a bigram or trigram based approach? Explain why.**

Answer:

Bigram and Trigram approaches support only in local order, but in this task, both local order and global order are necessary. So those approaches wouldn't work, as well.

- **Can the two languages be distinguished using a convolutional neural network? Explain why**

Answer:

CNN is an architecture that handles sequences of varying lengths while retaining local information. But CNN is n-gram based, and so just like bigrams and trigrams it wouldn't suffice.