CSC 691X Big Data

Project 4

Shuowen Wei

(1)

Not including the first 100 words, there are 271 words appear in at least 1/25 (4%) of all the lyrics.

(2)

Not including the first 100 words, there are 11 pairs of words appear in at least 1/25 (4%) of all the lyrics, they are:

[('101', '141'), ('102', '117'), ('106', '126'), ('101', '115'), ('103', '115'), ('101', '103'), ('101', '113'), ('115', '141'), ('103', '106'), ('101', '135'), ('101', '110')]

(3)

Not including the first 100 words, there is no triples appear in at least 1/25 (4%) of all the lyrics.