Advanced NLP Assignment 1

Training parameters:

Skipgram:

Number of epochs: 5
Vocab size: 70,000 words from the original corpus or 7152 words in the vocabulary

CBoW:

Number of epochs: 50
Vocab size: 100,00 words form the original corpus or 8829 words in the vocabulary

The top 10 similar words to camera in the Skipgram model were:

```
unlocking : 0.9874593044949269
solely : 0.9744712499124768
iphone : 0.9711839244828665
kinda : 0.9707389163570181
conferance : 0.97039312363663
wm : 0.9699836060465211
securely : 0.9649907534299326
solved : 0.9647791575434149
character : 0.9626604126971192
detected : 0.9611651936516322
```

The top 10 similar words to camera in the CBoW model were:

```
critical : 0.9541734990988948
mv : 0.9514685179150645
throughout : 0.9381732043587981
rounded : 0.9374172566205393
inevitable : 0.9359312429018694
lean : 0.9314232157237866
vhsonly : 0.931418436089073
dealership : 0.9312713089903359
spit : 0.9290044500234483
east : 0.928113716372299
```

The top 10 similar words to camera in the Skipgram model were:

cameras : 0.813185233511313
camcorder : 0.800015616860111
tripod : 0.731128328745217
Nikon : 0.719561351922441
Trucker : 0.700374713254210
Usb : 0.700302453154287
Camera_lens : 0.68451281513211
nikon_D : 0.678133096961023
Canon_digital_slr : 0.647425293922243

The top 10 similar words to camera in the CBoW model were:

cameras : 0.813185233511313
camcorder : 0.800015616860111
tripod : 0.731128328745217
Nikon : 0.719561351922441
Trucker : 0.700374713254210
Usb : 0.700302453154287
Camera_lens : 0.68451281513211
nikon_D : 0.678133096961023
Canon_digital_slr : 0.647425293922243

Skipgram:

Camera: `[1.58324590e-02 5.48189239e-03 1.55487489e-02 ... 4.72645341e-05`
`4.48262438e-05 6.45275033e-05]`

Great: `[4.84651551e-02 1.28742394e-02 2.52897338e-02 ... 4.75477267e-05`
`2.72553561e-05 3.35519070e-05]`

Trucker: `[2.13710274e-02 5.88833156e-03 1.90667475e-02 ... 4.75755026e-05`
`5.82023025e-05 2.84846419e-05]`

Working: `[1.93926965e-02 7.77952580e-03 9.61556897e-03 ... 5.15534149e-05`
`5.44087681e-05 7.67880575e-05]`

CBoW:

Camera: [1.24382297e-02 5.72471078e-03 1.01766099e-02 ... 1.01306914e-04
 7.77968583e-05 1.56020869e-04]

Great: [7.24102047e-02 3.18431696e-02 7.39904180e-02 ... 5.91173085e-05
 2.86631007e-05 3.44806230e-05]

Trucker: [3.17227234e-02 8.81960736e-03 1.50652967e-02 ... 1.06406784e-04
 1.07675031e-04 6.41930600e-05]

Working: [2.63096362e-02 8.47869120e-03 1.24628694e-02 ... 9.93741701e-05
 8.23964290e-05 1.55582609e-04]