**Assignment 2:**

Name Khlod Mohamed Ali

ID: 5695

1. **Word2vec:**
   1. **Sigmoid:**

The sigmoid function is given by Equation:



Code:

    s= 1/(1+np.exp(-x))

* 1.  **naiveSoftmaxLossAndGradient:**

loss**:**

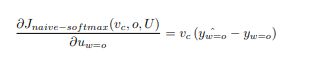


gradCenterVec**:**



y is a 1-hot vector with a 1 at word.

gradOutsideVec**:**



code:

 parameter = np.dot(outsideVectors,centerWordVec)*#shape (V,), (U.⊤ \* Vc)*

    y\_bar  = softmax(parameter) *# calculate P at all O and c*

    P\_o = y\_bar[outsideWordIdx]

    loss = -np.log(P\_o)

    y = np.zeros(outsideVectors.shape[0])

    y[outsideWordIdx] = 1

    gradCenterVec = np.dot(y\_bar - y, outsideVectors) *# shape(len of a vector,)*

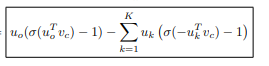
    gradOutsideVecs = np.outer(y\_bar-y, centerWordVec) *# shape(V,N)*

* 1. **negSamplingLossAndGradient:**

loss**:**



gradCenterVec**:**



gradOutsideVec**:**



Code:

    score = np.dot(outsideVectors[outsideWordIdx],centerWordVec) *#Uo.T\*Vc*

    term1 = -np.log(sigmoid(score))

    term2=np.sum(np.log(sigmoid(-np.dot(outsideVectors[negSampleWordIndices],centerWordVec))))

    loss=term1-term2

    gradCenterVec = (sigmoid(score)-1)\*outsideVectors[outsideWordIdx] *#shape (N,),derivative  of term1*

    gradOutsideVecs = np.zeros(outsideVectors.shape)

    gradOutsideVecs[outsideWordIdx] =(sigmoid(score)-1)\*centerWordVec *#derivative  of term1*

    for k in negSampleWordIndices: *#derivative  of term2*

        gradOutsideVecs[k] = gradOutsideVecs[k] + (1-sigmoid(-np.dot(outsideVectors[k],centerWordVec)))\*centerWordVec

        gradCenterVec += (1-sigmoid(-np.dot(outsideVectors[k],centerWordVec)))\*outsideVectors[k]

* 1. **skipgram:**

**code:**

    centerWordVec=centerWordVectors[word2Ind[currentCenterWord]]

    for outword in outsideWords:

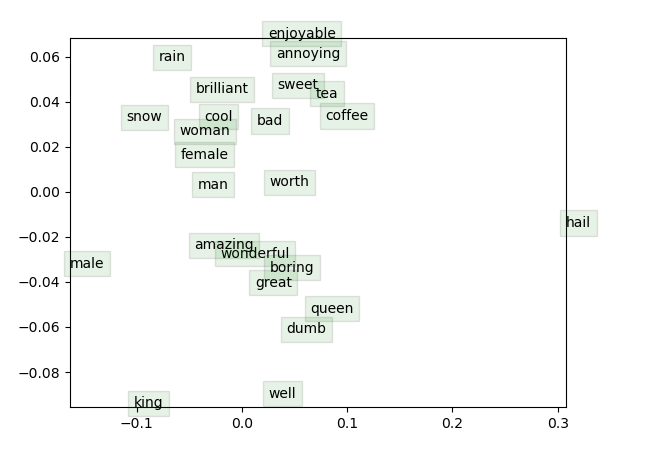
        loss\_,gradCenterVecs\_,gradOutsideVectors\_=word2vecLossAndGradient(centerWordVec,word2Ind[outword],outsideVectors,dataset)

        loss+=loss\_

        gradCenterVecs[word2Ind[currentCenterWord]]+=gradCenterVecs\_

        gradOutsideVectors+=gradOutsideVectors\_

1. **the result:**

It takes around 3 hours and 45 min.