SGSN word embedder optimization

Lazar Jelić

April 23, 2020

Abstract

This note provides new insight on time and memory optimization techniques for skip-gram word embedding neural network.

1 Introduction

Blah blah blah...

$$p(w_{c,j} = w_{O,c}|w_I) = y_{c,j} = \frac{e^{u_{c,j}}}{\sum_{j'=1}^{V} e^{u}_{j'}}$$

$$u_{c,j} = u_j = \boldsymbol{v'}^{\top} \cdot \boldsymbol{h}$$

$$\boldsymbol{w'_{ij}} = \boldsymbol{w_{ij}} - \alpha * EI_j * \boldsymbol{h_i}$$

$$\boldsymbol{v'_{w_{ij}}} = \boldsymbol{v_{w_{ij}}} - \alpha * EH^{\top}$$

$$EI_j = \sum_{c=1}^{C} e_{c,j}$$

$$EH_i = \sum_{j=1}^{V} EI_j * w'_{ij}$$