

NLP and Word **Embeddings**

Word2Vec



Skip-grams for leavy word embeddigs

■ 网易云课堂

I want a glass of orange juice to go along with my cereal.



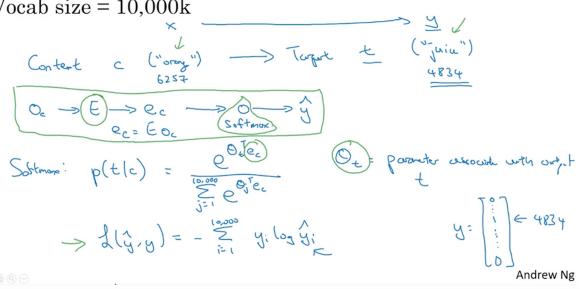




Model

■ 网易云课堂

Vocab size = 10,000k



Problems with softmax classification

$$p(t|c) = \frac{e^{\theta_t^T e_c}}{\sum_{j=1}^{10,000} e^{\theta_j^T e_c}}$$

$$\frac{10,000}{\sum_{j=1}^{10,000} e^{\theta_j^T e_c}}$$

$$\frac{10,000}{\sum_{j=1}^{10,000} e^{\theta_j^T e_c}}$$

How to sample the context c?

Andrew Ng