

UMass · CS685 | Advanced Natural Language Processing (2020)

CS685 (2020) · 课程资料包 @ShowMeAI



视频

中英双语字幕



课件

一键打包下载



笔记

官方笔记翻译



代码

作业项目解析



视频 · B 站 [扫码或点击链接]

<https://www.bilibili.com/video/BV1BL411t7RV>



课件 & 代码 · 博客 [扫码或点击链接]

<http://blog.showmeai.tech/umass-cs685>

NLP

语言模型 问答系统 文本生成 BERT

语义解析

模型蒸馏

GPT-3

注意力机制

迁移学习

知识推理

transformer

Awesome AI Courses Notes Cheatsheets 是 [ShowMeAI](#) 资料库的分支系列, 覆盖最具知名度的 **TOP50+** 门 AI 课程, 旨在为读者和学习者提供一整套高品质中文学习笔记和速查表。

点击课程名称, 跳转至课程**资料包**页面, **一键下载**课程全部资料!

| 机器学习 | 深度学习 | 自然语言处理 | 计算机视觉 |
|--|-------------------|--------------------|-------------------|
| Stanford · CS229 | Stanford · CS230 | Stanford · CS224n | Stanford · CS231n |
| # Awesome AI Courses Notes Cheatsheets · 持续更新中 | | | |
| 知识图谱 | 图机器学习 | 深度强化学习 | 自动驾驶 |
| Stanford · CS520 | Stanford · CS224W | UCBerkeley · CS285 | MIT · 6.S094 |



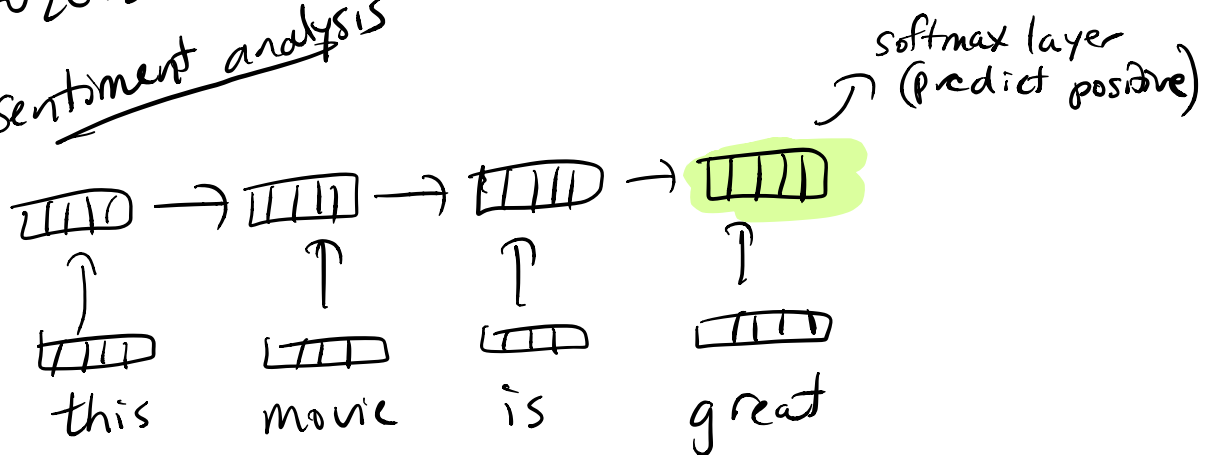
微信公众号

资料下载方式 2: 扫码点击**底部菜单栏**称为 **AI 内容创作者?** 回复 [添砖加瓦]

how we used to solve NLP tasks:

~2013

sentiment analysis



1. randomly initialize the model params

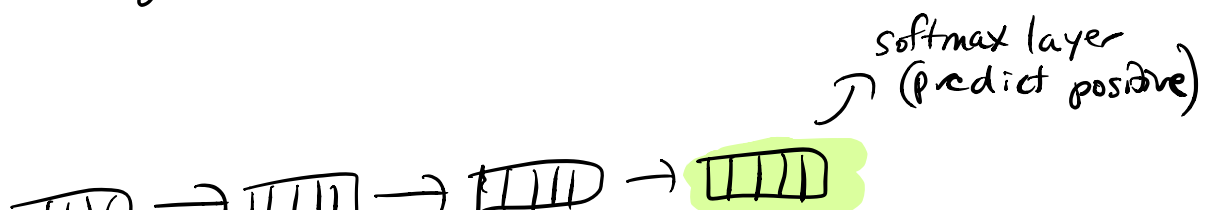
$W_h, W_e, C, \dots, V, W_o$ all trained from scratch

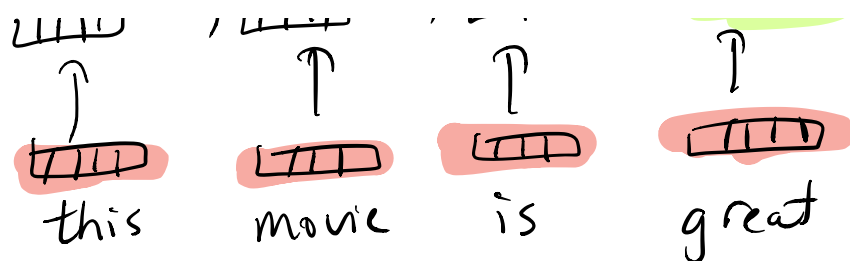
2. Update all parameters by backprop using cross entropy loss from labeled training set

model has to learn how language works from only a small labeled dataset

~2014-2017

why train everything from scratch?
how can we leverage lots of unlabeled data?





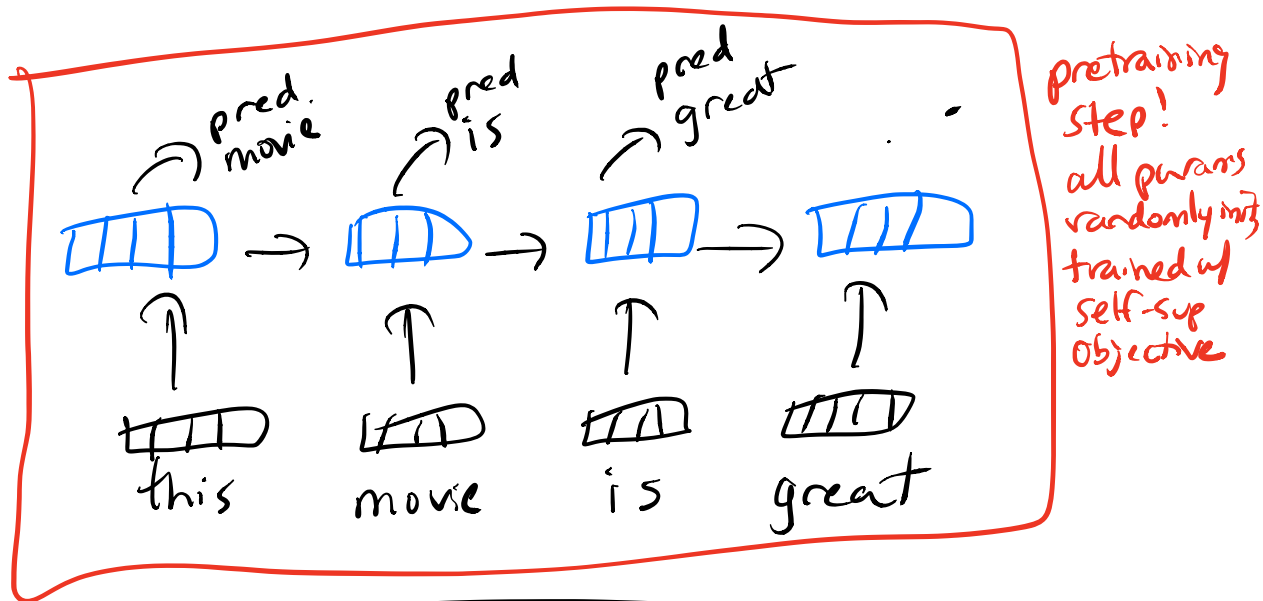
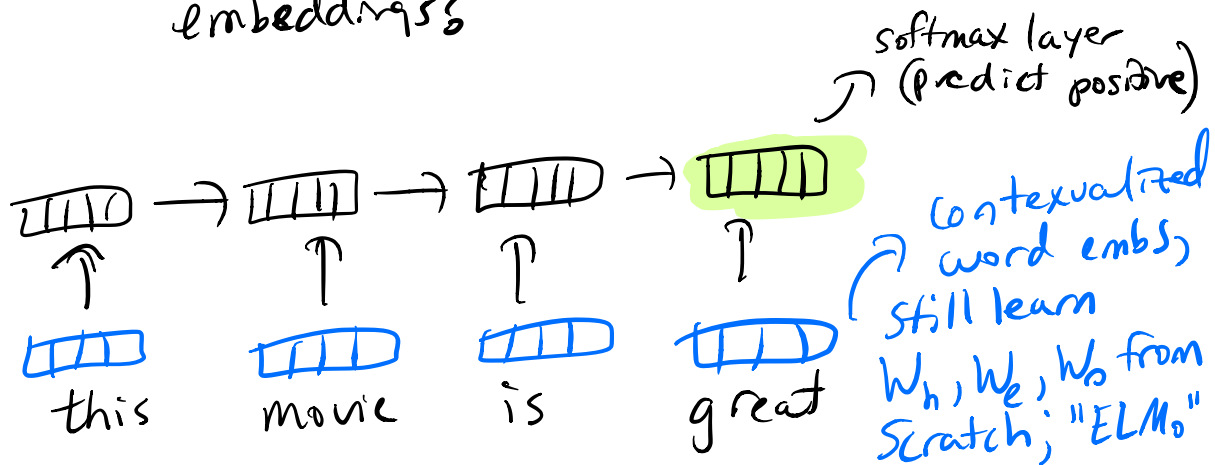
we can pretrain the c_i 's (word embs) using another objective fn that takes adv. of unlabeled data (self-supervised)

- Word2Vec, GloVe
- instead of starting w/ a random word embedding space, we start from a pretrained space in which word embs. capture some linguistic prop.
- train all other params from scratch (W_h, W_c, W_o)

~2018

- issues w/ c_i : word embeddings are static, only one vector per word type regardless of context
- The rest of the model (in our case, the RNN) is responsible for learning composition from scratch given just labeled data

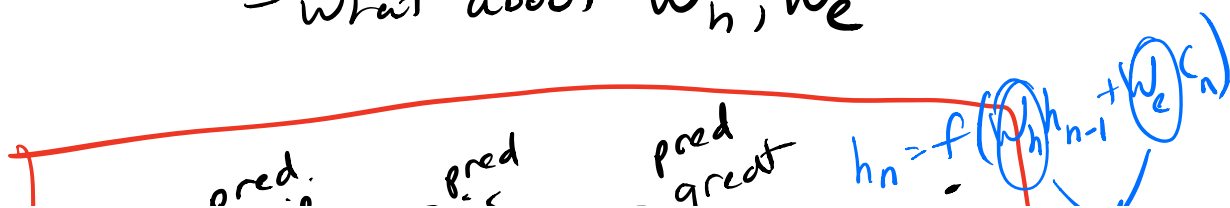
— what if we use the hidden states of a NLM instead of static word embeddings?

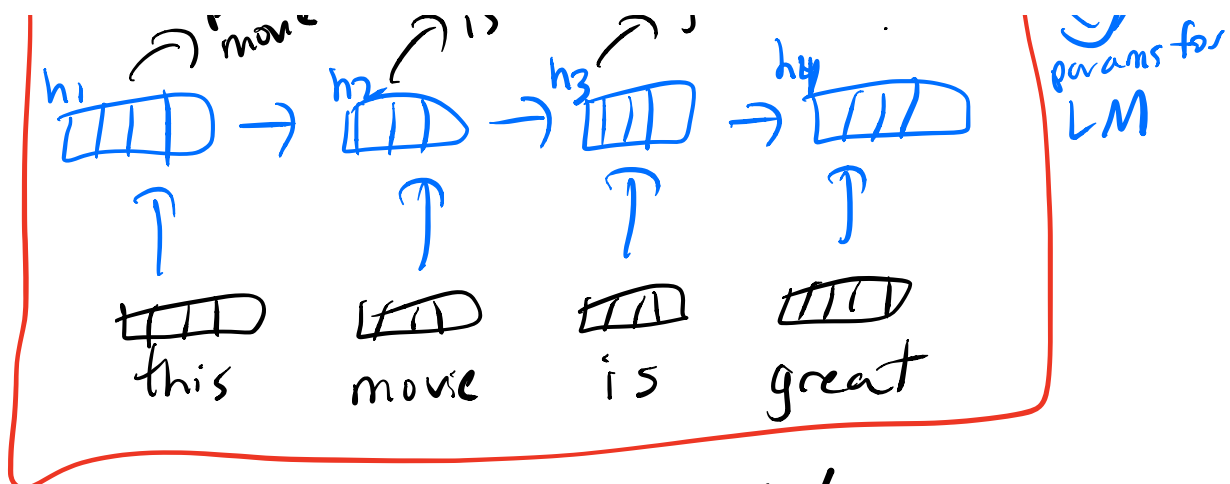


~2019

can we share more params than just the word embeddings?

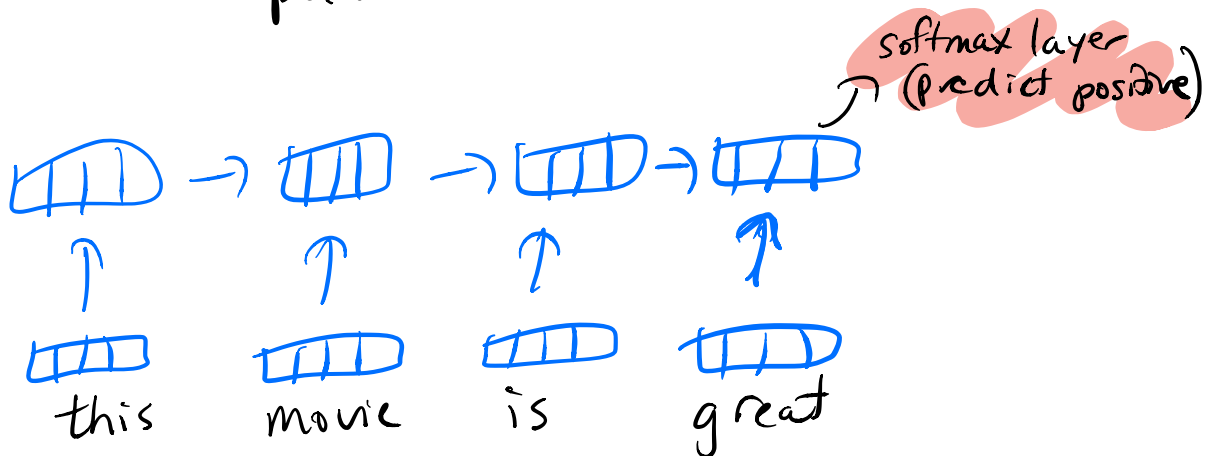
— what about W_h, W_e





initialize my sentiment RNN w/ the
LMs W_h, W_e, C, \dots, W

- all we have to do is learn W_0 from scratch, all the other params are transferred



this is current paradigm in NLP,
popularized by BERT model

- we init our sentiment model w/
a pretrained NLM, and then

backprop the error from a diff.
task (downstream task) into these
parameters. this is called fine-tuning

UMass · CS685 | Advanced Natural Language Processing (2020)

CS685 (2020) · 课程资料包 @ShowMeAI



视频

中英双语字幕



课件

一键打包下载



笔记

官方笔记翻译



代码

作业项目解析



视频 · B 站 [扫码或点击链接]

<https://www.bilibili.com/video/BV1BL411t7RV>



课件 & 代码 · 博客 [扫码或点击链接]

<http://blog.showmeai.tech/umass-cs685>

NLP

语言模型 问答系统 文本生成 BERT

语义解析

模型蒸馏

GPT-3

注意力机制

迁移学习

知识推理

transformer

Awesome AI Courses Notes Cheatsheets 是 [ShowMeAI](#) 资料库的分支系列, 覆盖最具知名度的 **TOP50+** 门 AI 课程, 旨在为读者和学习者提供一整套高品质中文学习笔记和速查表。

点击课程名称, 跳转至课程**资料包**页面, **一键下载**课程全部资料!

| 机器学习 | 深度学习 | 自然语言处理 | 计算机视觉 |
|--|-------------------|--------------------|-------------------|
| Stanford · CS229 | Stanford · CS230 | Stanford · CS224n | Stanford · CS231n |
| # Awesome AI Courses Notes Cheatsheets · 持续更新中 | | | |
| 知识图谱 | 图机器学习 | 深度强化学习 | 自动驾驶 |
| Stanford · CS520 | Stanford · CS224W | UCBerkeley · CS285 | MIT · 6.S094 |



微信公众号

资料下载方式 2: 扫码点击**底部菜单栏**称为 **AI 内容创作者?** 回复 [添砖加瓦]