

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]

Today: implementing neural LMs
in PyTorch

→ cross entropy loss

— used in NLMS, as well
as other classification tasks

What is a loss fn?

↳ intuitively, tells us how bad a model
is doing at predicting the training data

↳ in NLMS, how bad is the model
at predicting the next word

assume we have a training ex

"students opened their"

⇒ books

└──────────────────┘

model input

└──────────┘

desired model
prediction

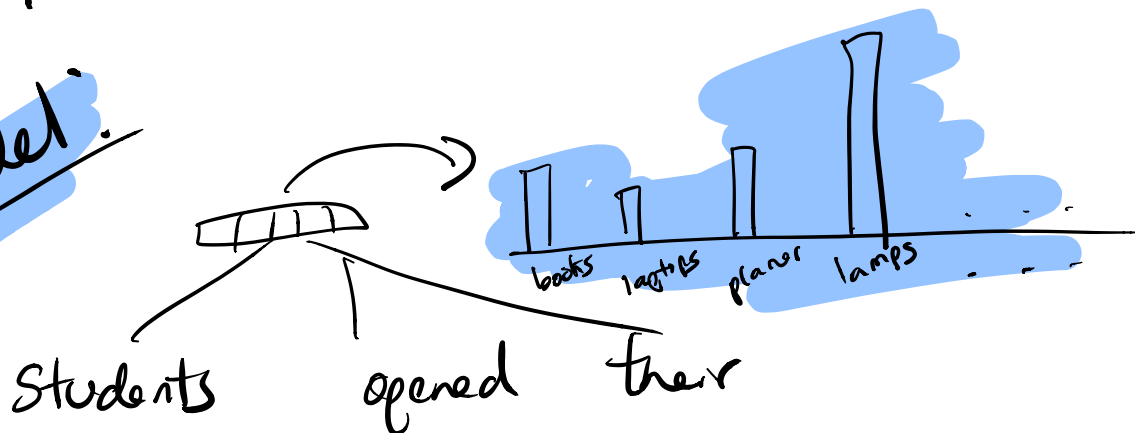
$P(\text{"books"} \mid \text{"students opened their"})$

↳ maximize this probability!

we want to minimize the
negative log probability of "books"
 $L = -\log(p(\text{books} | \text{students opened their}))$

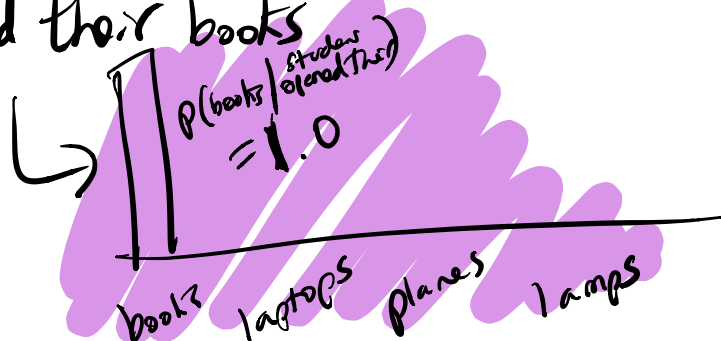
why is this called the cross-entropy loss?

model:



data:

Students opened their books



Cross entropy of two distributions p and q

\Rightarrow quantifies distance between the distributions

$$H(p, q) = - \sum_{w \in V} p(w) \log q(w)$$

↑
1 for "books"
0 every other word
type

↑
predicted
conditional
prob

$$= - \log q(\text{books} | \dots)$$

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]