



آکادمی هوش مصنوعی
howsam.org | هوسم

دوره دیپ کاتالیست

پروژه اول

مدل سازی زبان

(Language Modeling)



فهرست مطالب

فهرست مطالب



» آنچه گذشت...

» پیش‌نیاز

» اجرای تئوری و عملی چک‌لیست پروژه

» جمع‌بندی



آنچه گذشت...

آنچه گذشت...



«« عملکرد ما:

«« از چکلیست انجام پروژه گفته‌یم.

«« پروژه تخمین سن

«« تشریح تئوری

«« تشریح عملی پروژه

«« طراحی چالش

آنچه گذشت...



«« عملکرد شما:

«« عدهای اصلا مشاهده نکردن! 😕

«« عدهای مشاهده کردند، اما مشارکت نکردن! 😏

«« عدهای مشاهده نکردن، اما مشارکت کردند! 😊

«« عدهای هم مشاهده کردند و هم مشارکت کردند! 😃

آنچه گذشت...

نقطه ضعف‌ها

- «« تحلیل‌گر نبودید!
- «« تحلیل داده و مدل
- «« دیباگر نبودید!
- «« کم خلاق بودید!
- «« کم پایه بودید! 😊
- «« باهم همکاری نکردید!

نقطه قوت‌ها

- «« کدنویسی‌تان خوب بود. ⚡
- «« منظم بودید.
- «« کمی خلاق بودید!
- «« کمی پایه بودید!

Pinned message #16

گروه دیپکاتالیست هوسم
59 members

...استاد من خودم از یک فیس دیتکشن که تو گیت هاب بودا
...استاد من خودم از یک فیس دیتکشن که تو

02:40, 3.1 MB 3:17 PM

گروه دیپکاتالیست هوسم
Voice message

پیشنهاد خوبیه
خودم بهش فکر کرده بودم
حتماً امتحان میکنم 3:23 PM

گروه دیپکاتالیست هوسم
Voice message

استاد دیتاست ها برای face زیاده، حالا اگه بیایم اینا هارو
ترکیب کنیم و با یک توزیع خوبی از سن های مختلف یک
دیتاست جدید تولید کنیم (با فیس دیتکشن یکسان) که
زیرمجموعه‌ی همه‌ی اینها باشد مدل ترین شده روی
این دیتاست می‌توانه نتیجه‌ی بهتری تو دنیا واقعی
بگیری؟ 3:29 PM

گروه دیپکاتالیست هوسم

سلام استاد من خطای تستم به 4.55 رسید.
بعد اومدم همین مدل رو رو دیتاست CACD2000 تست
کردم (البته حجم دیتاست زیاد بود فقط روی ۱۰۰۰ تا داش
تست کردم) خطای ۲۳ شد.
شما گفتید در این حالت ممکنه خطای یک پرش بزرگ داشته
باشه ولی ۲۳ خیلی زیاد نیست؟ ← 11 12:46 PM

گروه دیپکاتالیست هوسم
Voice message

سلام استاد من خطای تستم به 4.55 رسید...
04:30, 5.3 MB 1:01 PM

گروه دیپکاتالیست هوسم
Voice message

ممنون! نکات که گفتید رو بررسی میکنم نتیجه‌شوند.

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

آکادمی هوش مصنوعی
howsam.org | هوسم



آکادمی هوسم از هرگونه نشر و کپی برداری از مطالب این مجموعه آموزشی راضی نیست.

۴



آکادمی هوشمصنوعی
howsam.org | هوسم

۳

AI Project Management ▾ + Add Page

Search (Ctrl+E) 🔍

گام هشتم-مستندسازی و ارائه گام هفتم-کاربردسازی گام نهم-بهبود روش گام پنجم-آزمایش گام سوم-انتخاب روش مناسب گام چهارم-باده‌سازی روش انتخاب گام دوم-بررسی روش‌های موجود گام اول-سازمانی کلیات پروژه

راهمنا باده‌سازی روش ۱

دیناست مناسب و استاندارد موجود هست؟

- دیناست مربوطه است، ولی استاندارد نیست. نیاز به اصلاح دارد!
- استفاده از دیناست‌های موجود برای ساختن دیناست مناسب با اهدافمان خودمان باید دیناست جمع‌آوری کنیم.
- تحلیل اکتشافی داده (EDA)
- مشاهده و بررسی نمونهها
- رسم نمودارهای تحلیلی مختلف با هدف برقراری ارتباط با داده (دیناست چه می‌گوید?)
- آماده سازی و پیش پردازش داده
- رفع ایرادهای ابتدایی در دیناست (مثلاً حذف اصلاح داده‌های بدون برچسب، خواص,...)
- تقسیم داده به بخش‌های Train , Test , Validation
- داده افزایی یا Data Augmentation
- ترمال‌سازی و استانداردسازی داده
- پیاده‌سازی ماژول فراخوانی داده
- دیناست در پایتون آماده است؟
- دستور مناسب برای فراخوانی در پایتون وجود دارد؟ (ImageFolder)
- کاستوم دیناست در اینترنت موجود هست؟
- ChatGPT می‌تواند کمک کند.
- خودمان باید کاستوم دیناست بنویسیم؟

مرحله ۲: مدل

○ پیاده‌سازی مدل / روش اریحات انتخابی

○ حالتهای رایج در پیاده‌سازی مدل:

○ مذا و مث منظقل ماد. ماته. ح. ممجد است

Show desktop

آکادمی هوسم از هرگونه نشر و کپی‌برداری از مطالب این مجموعه آموزشی راضی نیست.



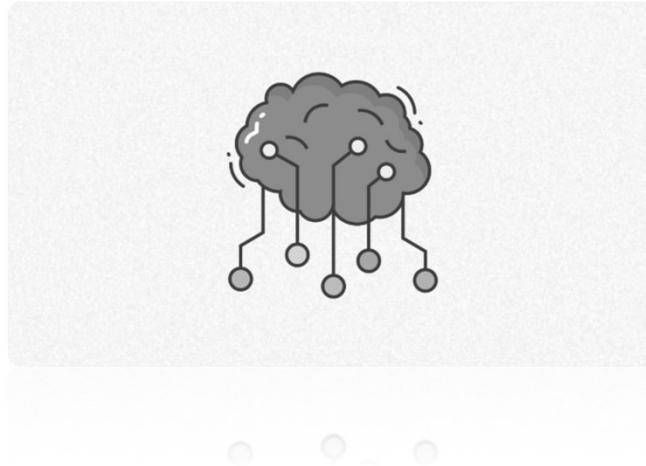
هر پروژه یک چالش!



بزن بريم!

پیش‌نیاز

پیش‌نیازهای پروژه اول

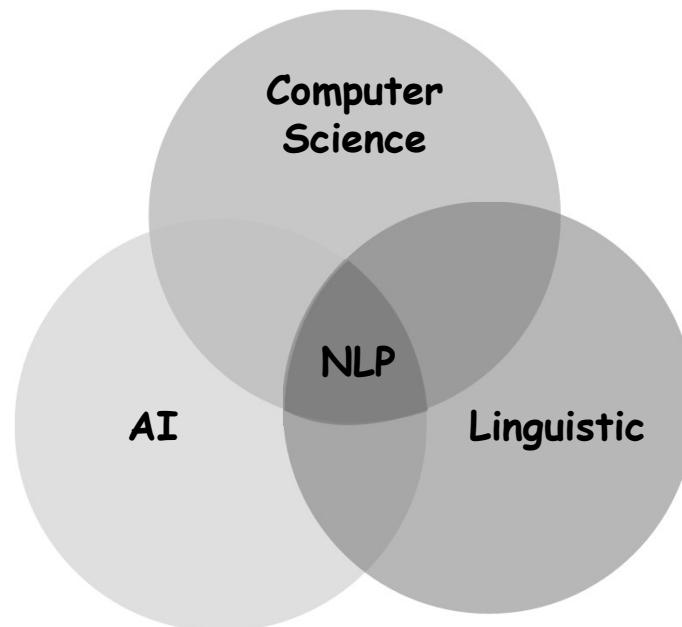


- ﴿﴿ تئوری و کدنویسی شبکه بازگشتی و ترنسفورمر
- ﴿﴿ آشنایی با پردازش زبان طبیعی و مدل‌سازی زبان
- ﴿﴿ مطالب مرتبط با دوره یادگیری عمیق ۲۰۲۲:
- ﴿﴿ فصل شش: شبکه عصبی بازگشتی
- ﴿﴿ فصل هفت: شبکه ترنسفورمر
- ﴿﴿ فصل یازده: پردازش زبان طبیعی

پردازش زبان طبیعی

Natural Language Processing (NLP)

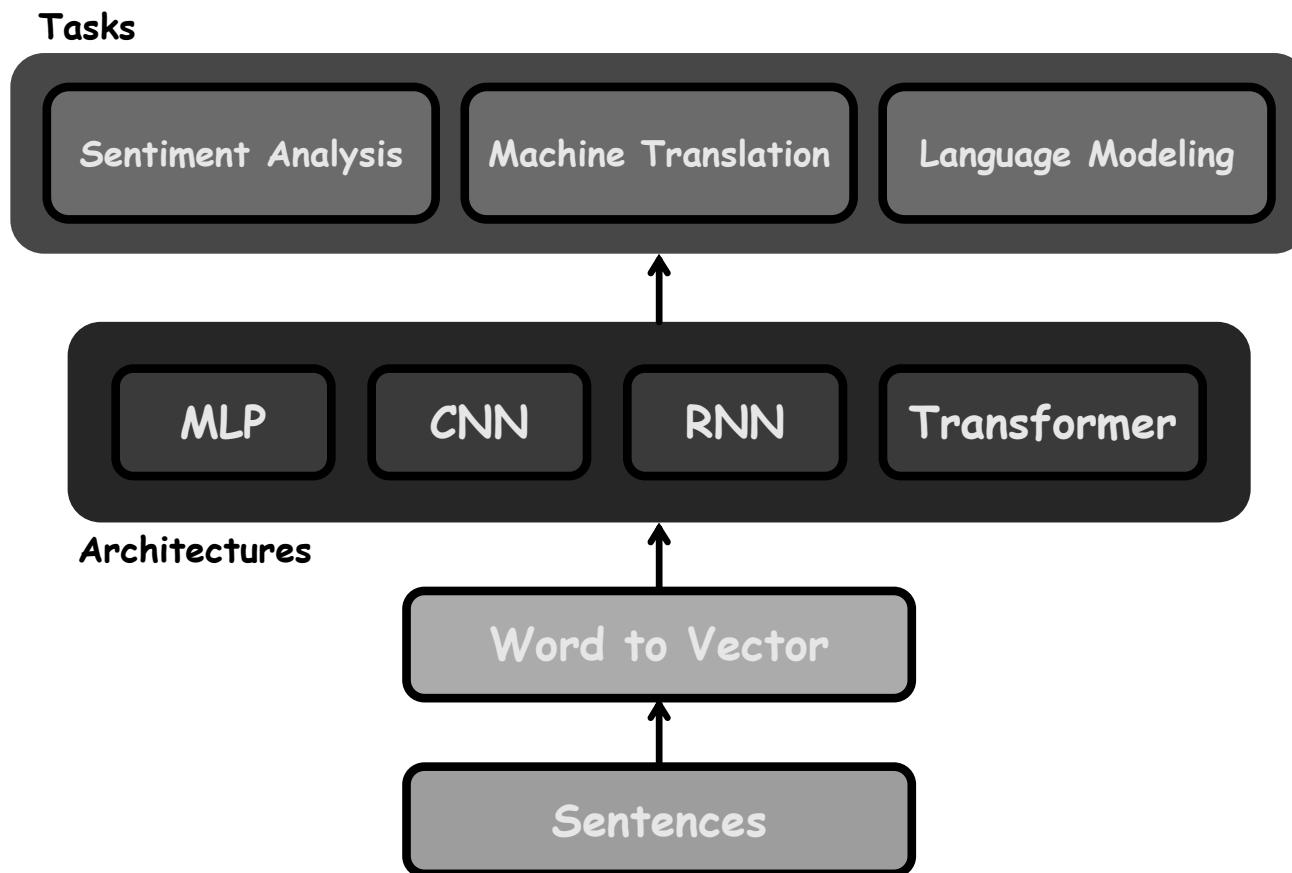
«درک زبان انسان توسط رایانه»



آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

https://en.wikipedia.org/wiki/Natural_language_processing

NLP



آکادمی هوسام از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

متن

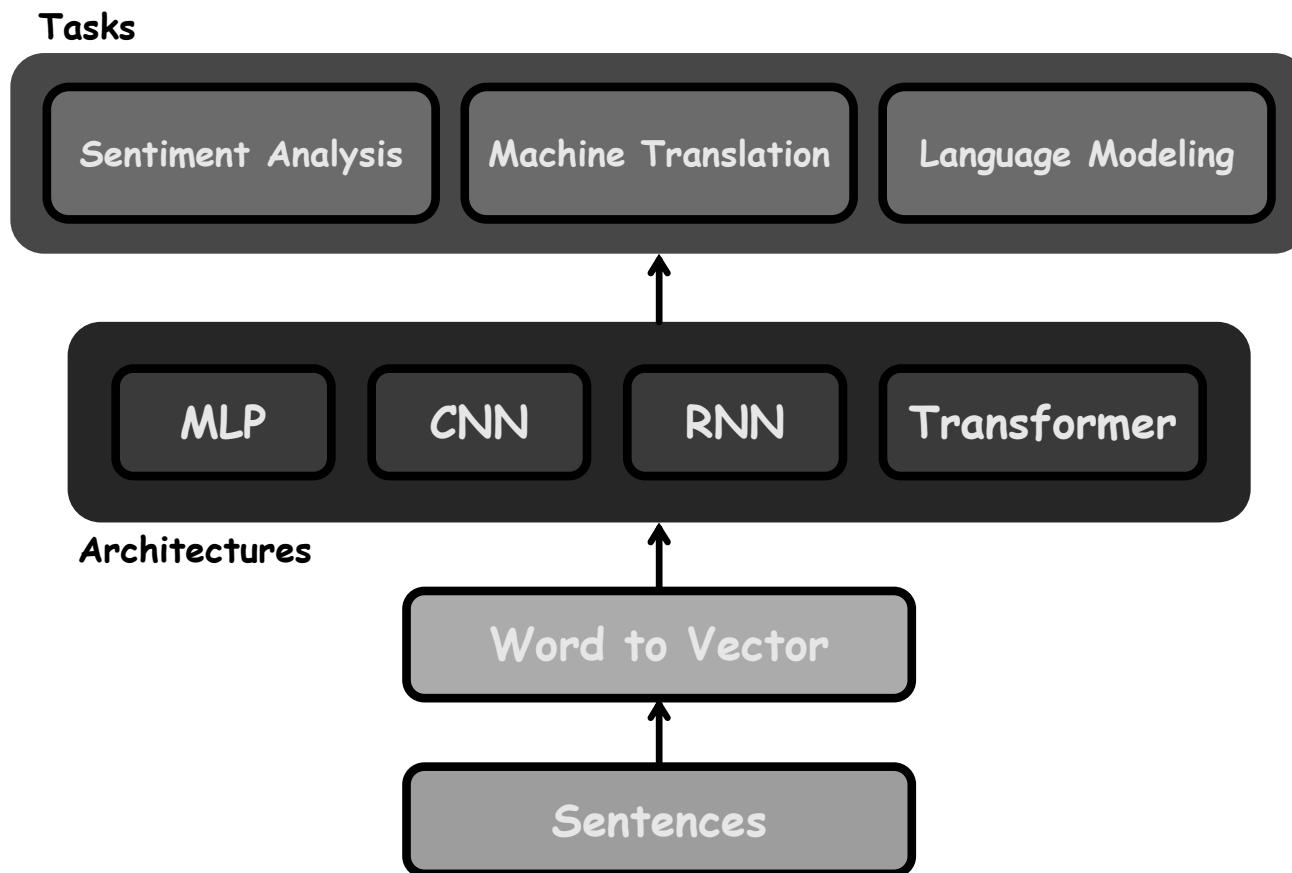
متن

خیلی عالی بود.

ناامید کننده هست، فکر نمی‌کردم اینقدر بد باشد.

خوب بود، اما من در کارهای آینده انتظار بیشتری از شما دارم.

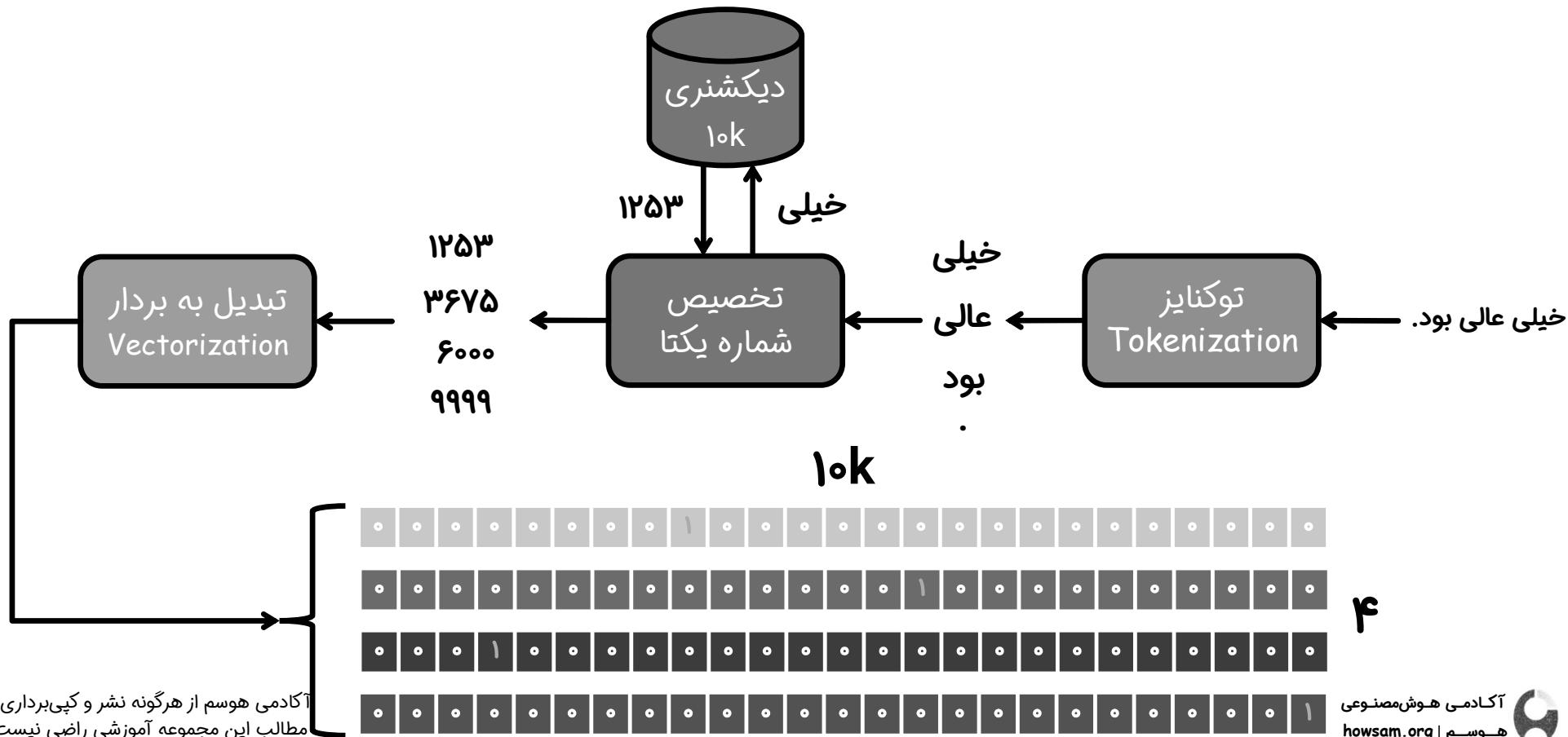
NLP



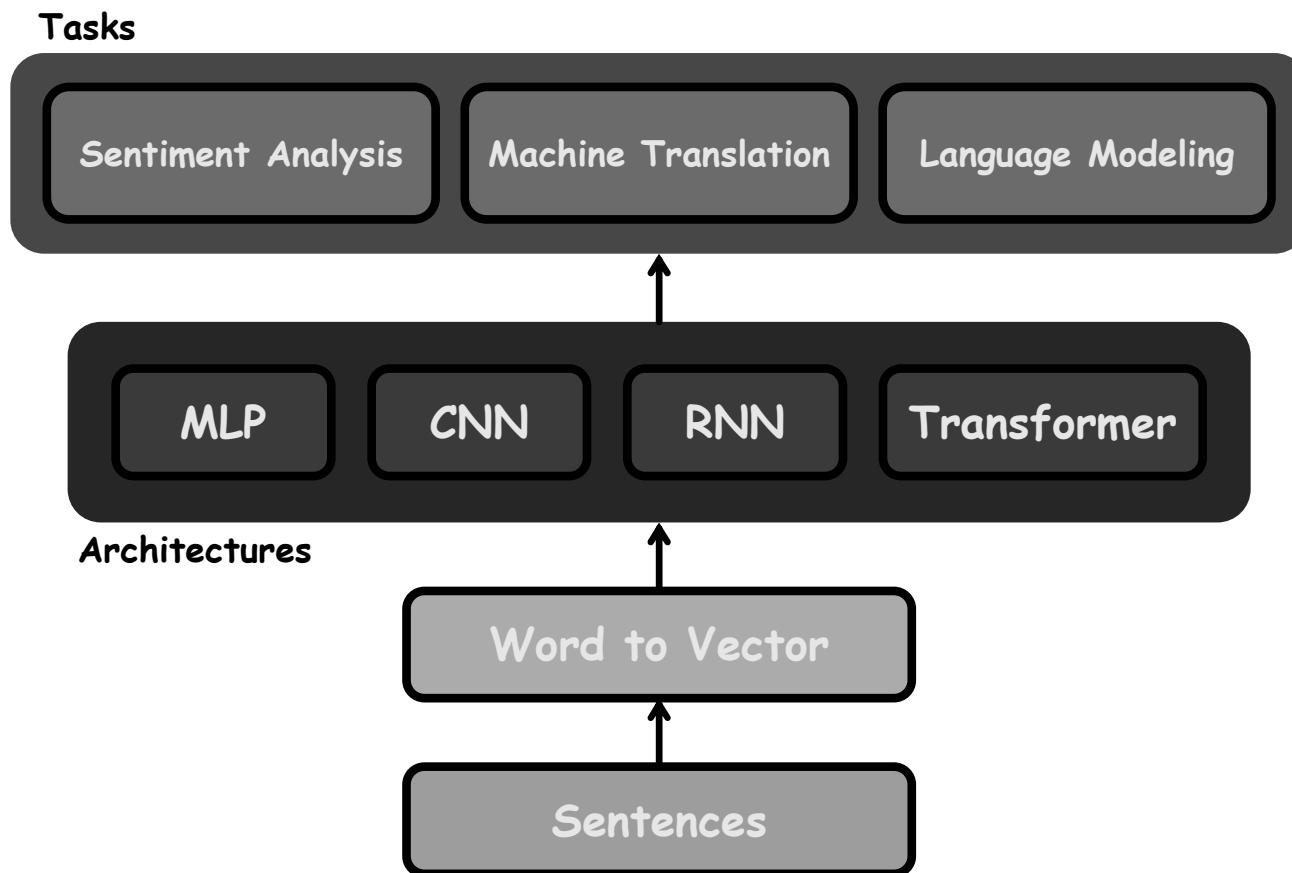
آکادمی هوسام از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

Word to Vector

Word to Vector



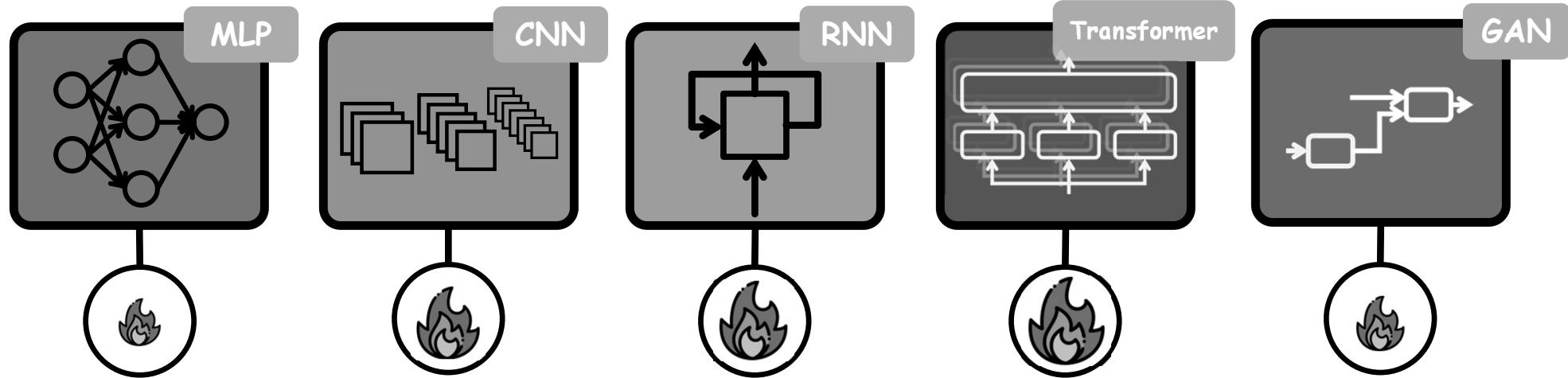
NLP



آکادمی هوسام از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

شبکه‌های رایج در NLP

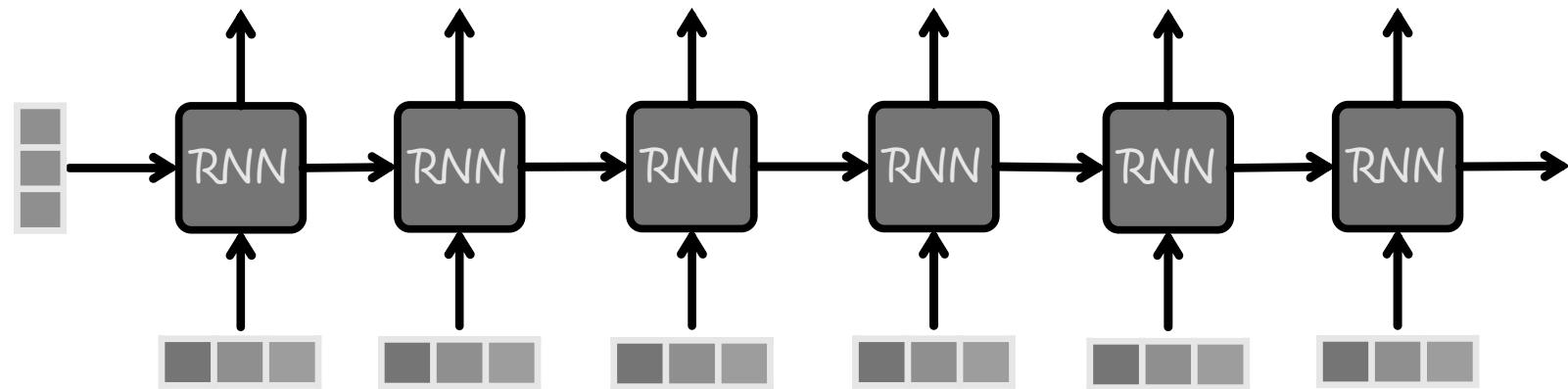
شبکه‌های رایج در NLP



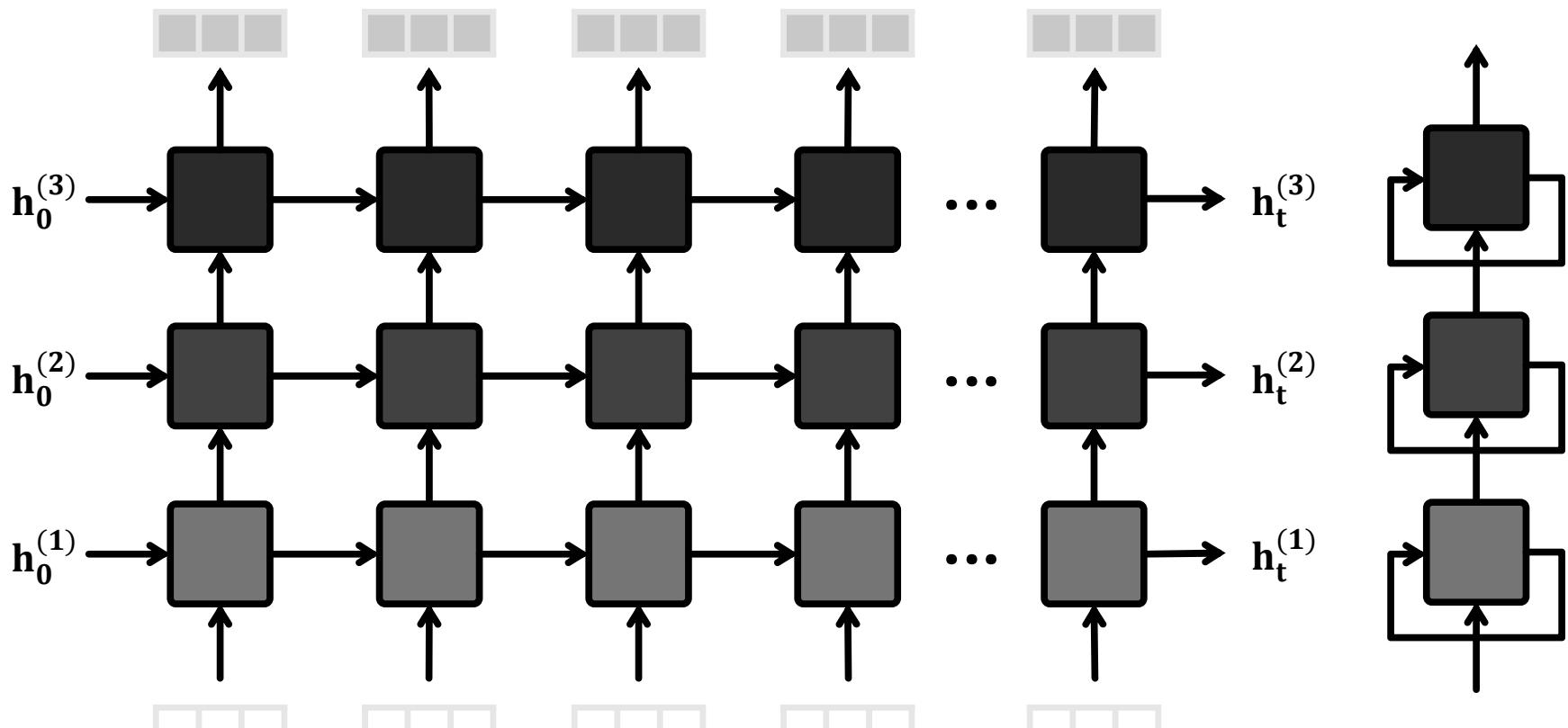
آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

شبکه بازگشتی

شبکه بازگشتی



شبکه بازگشی چند لایه (Multi-layer) (چند لایه)

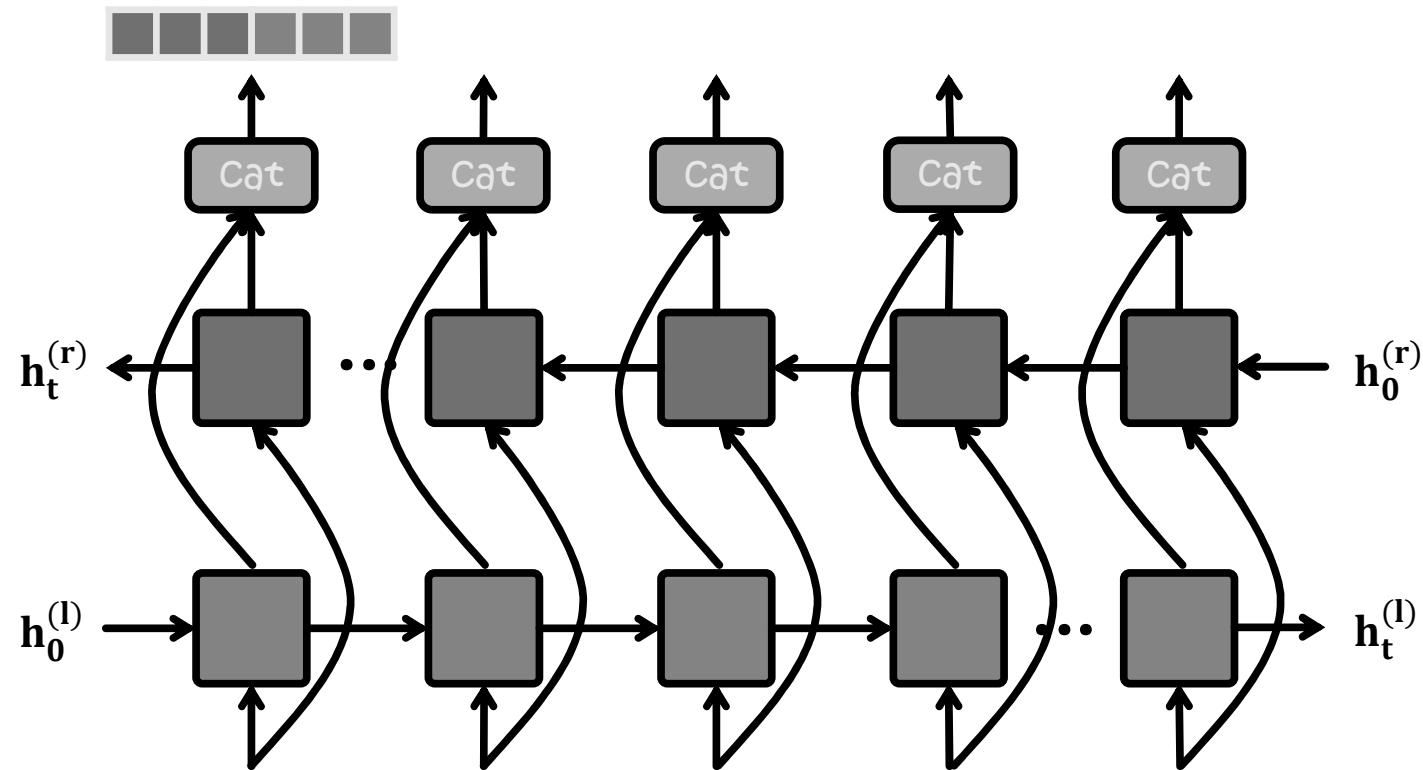


اکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

شبکه

```
# batch_size: 1  
  
for x, y in train_loader:  
  
    h1i, h2i, h3i = zeros(...)  
  
    for xi in x: # shape(x): LxF  
  
        o1i, h1i = RNN1(xi, h1i)  
  
        o2i, h2i = RNN2(o1i, h2i)  
  
        o3i, h3i = RNN3(o2i, h3i)  
  
    ...
```

شبکه بازگشتی دووجهی (Bidirectional)

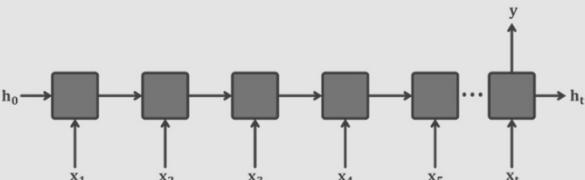
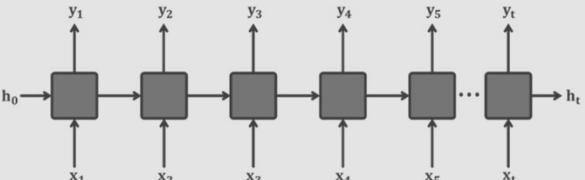
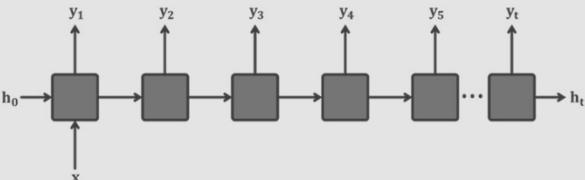


آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

شبکه

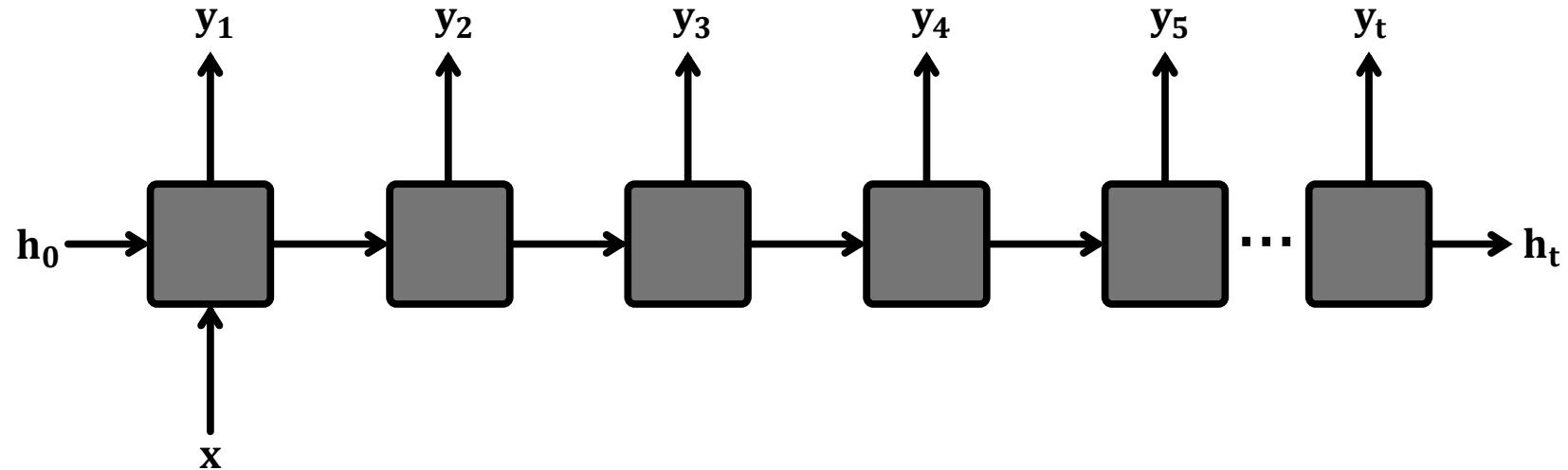
```
# batch_size: 1  
for x, y in train_loader:  
    for xrli, xlri in zip(x, x.reverse()):  
        orli, hrli = RNNrl(xrli, hrli)  
        olri, hlri = RNNlr(xlri, hlri)  
        oi = torch.cat(orli, olri)
```

ورودی و خروجی شبکه بازگشتی

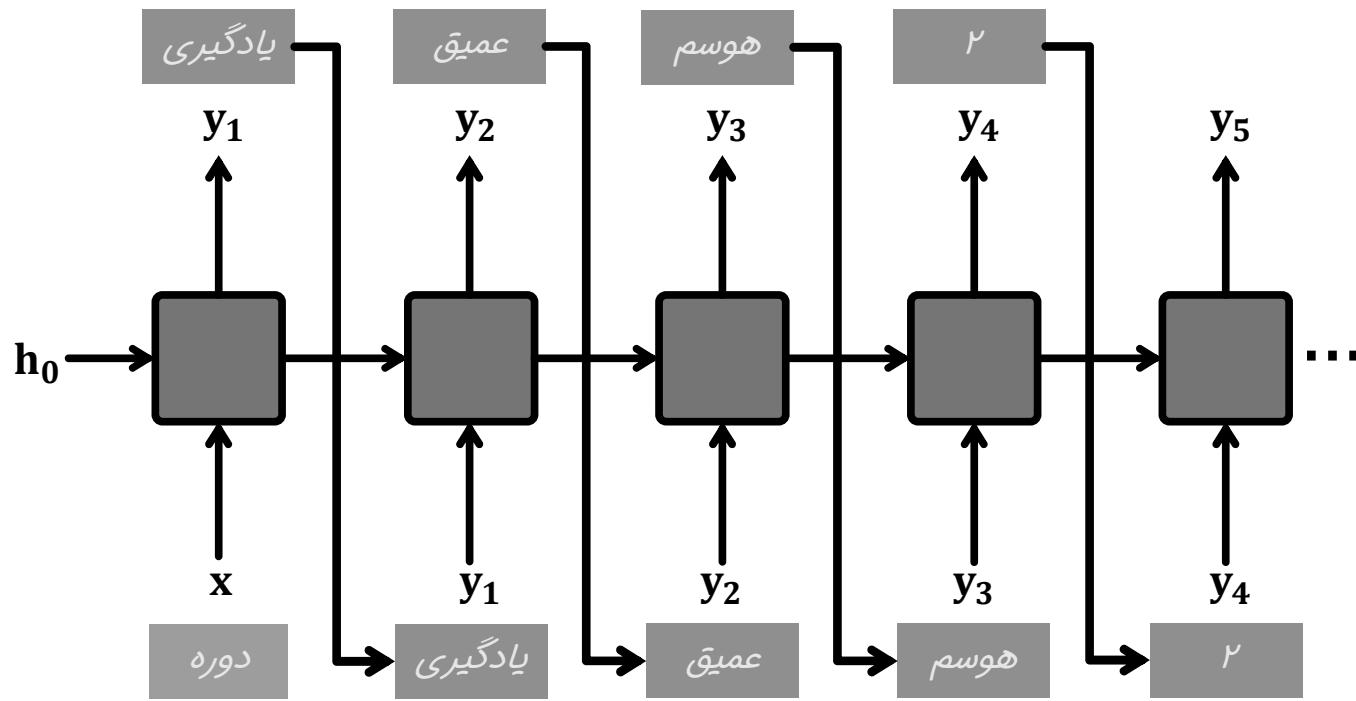
معماری	خروجی	ورودی	حالت
	One	Many	Many to One
	Many	Many	Many to Many
	Many	One	One to Many

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

حالت One to Many



حالت One to Many



آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

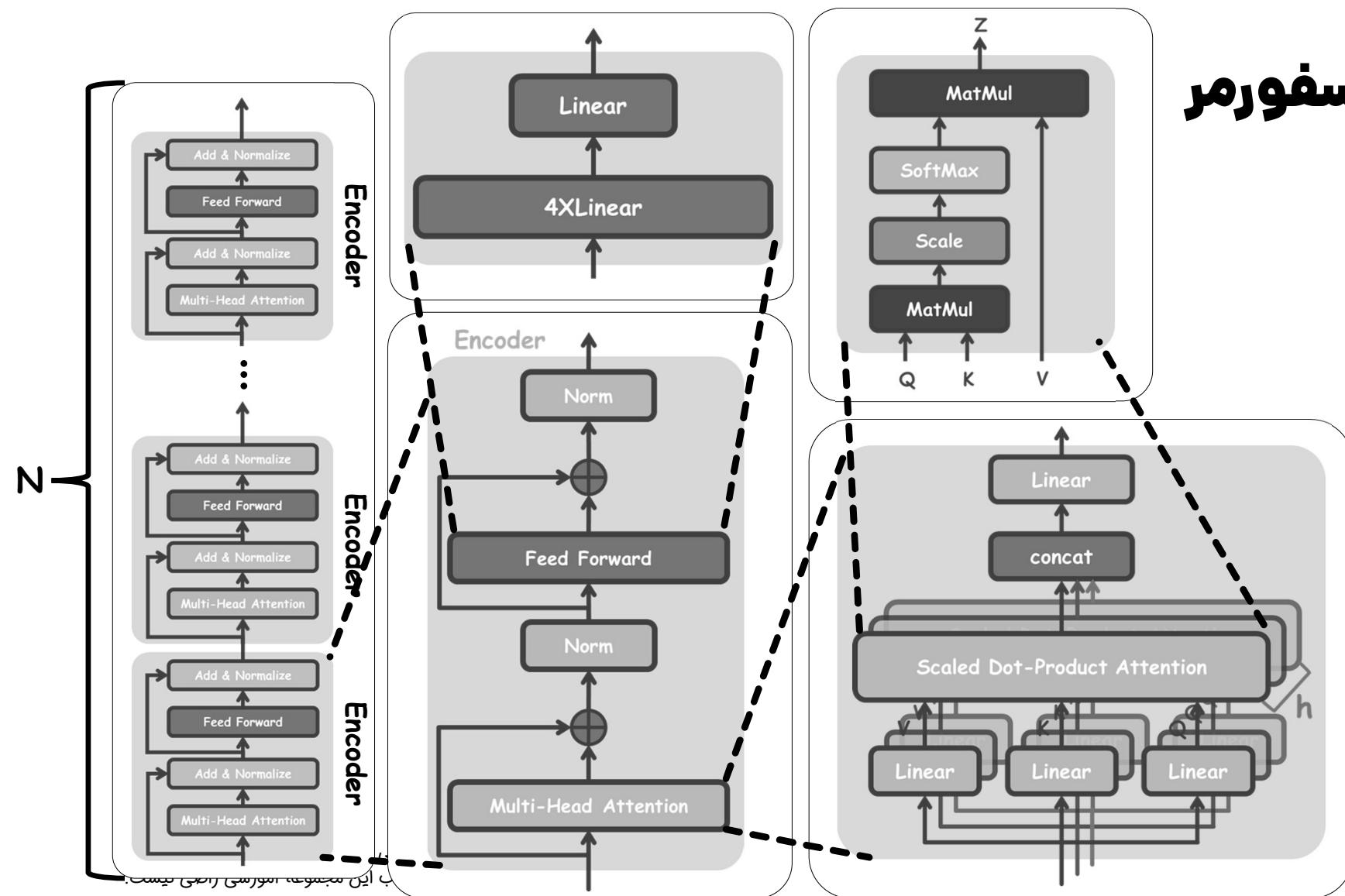
دوره یادگیری عمیق هوسم



دوره

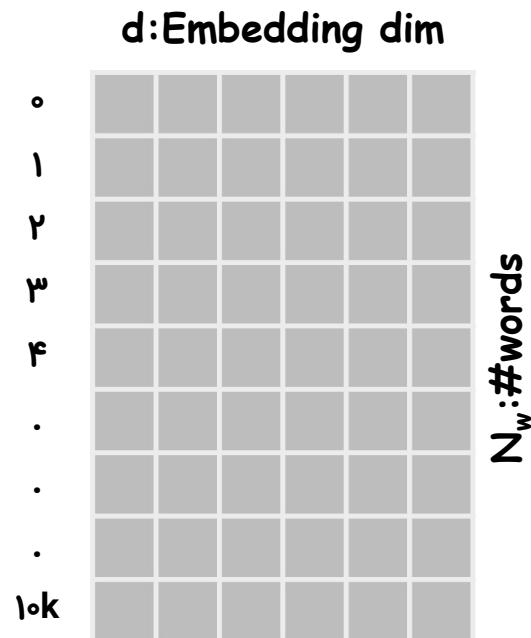
شبکه ترنسفورمر

شبکه ترسیفورمر



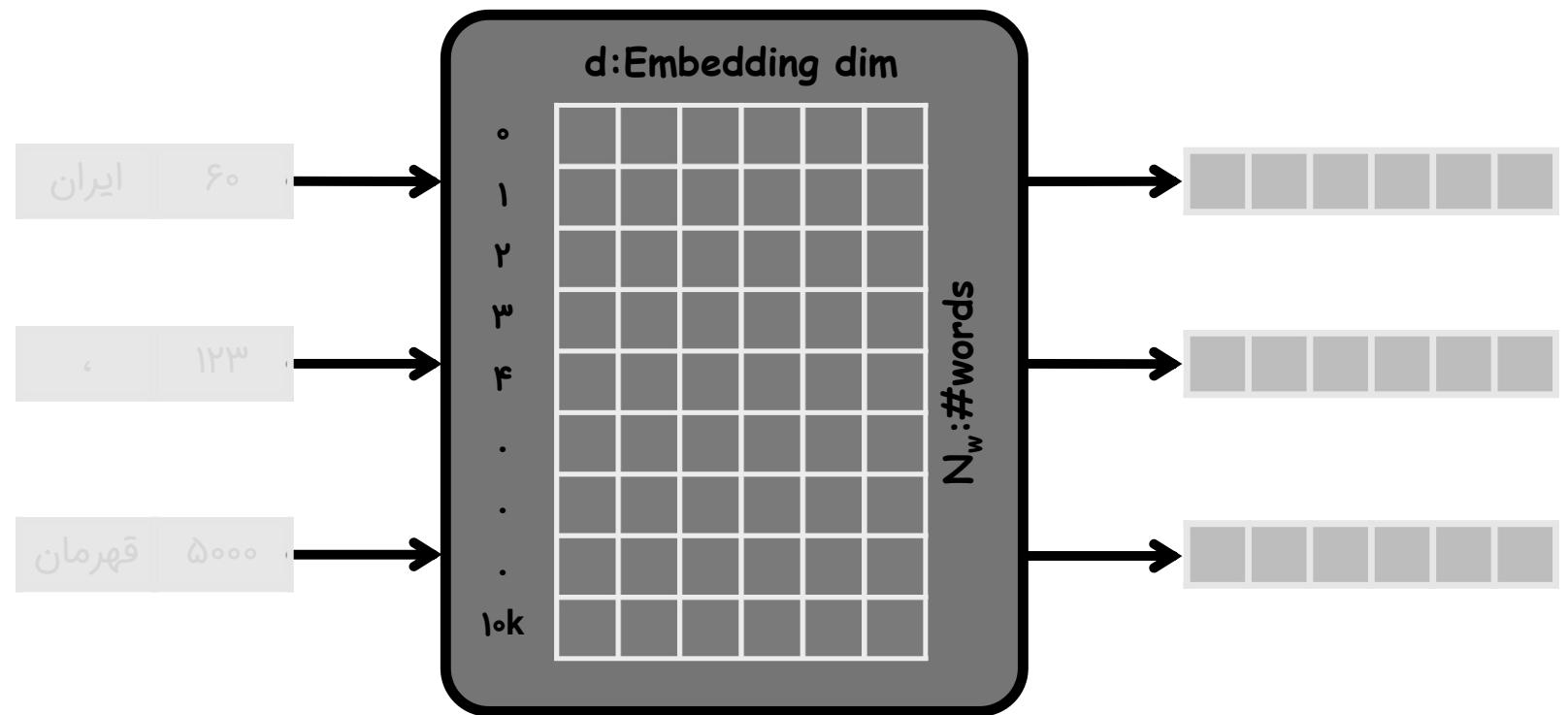
لابه اميد ينگ

لا^يه امبدینگ (Embedding)



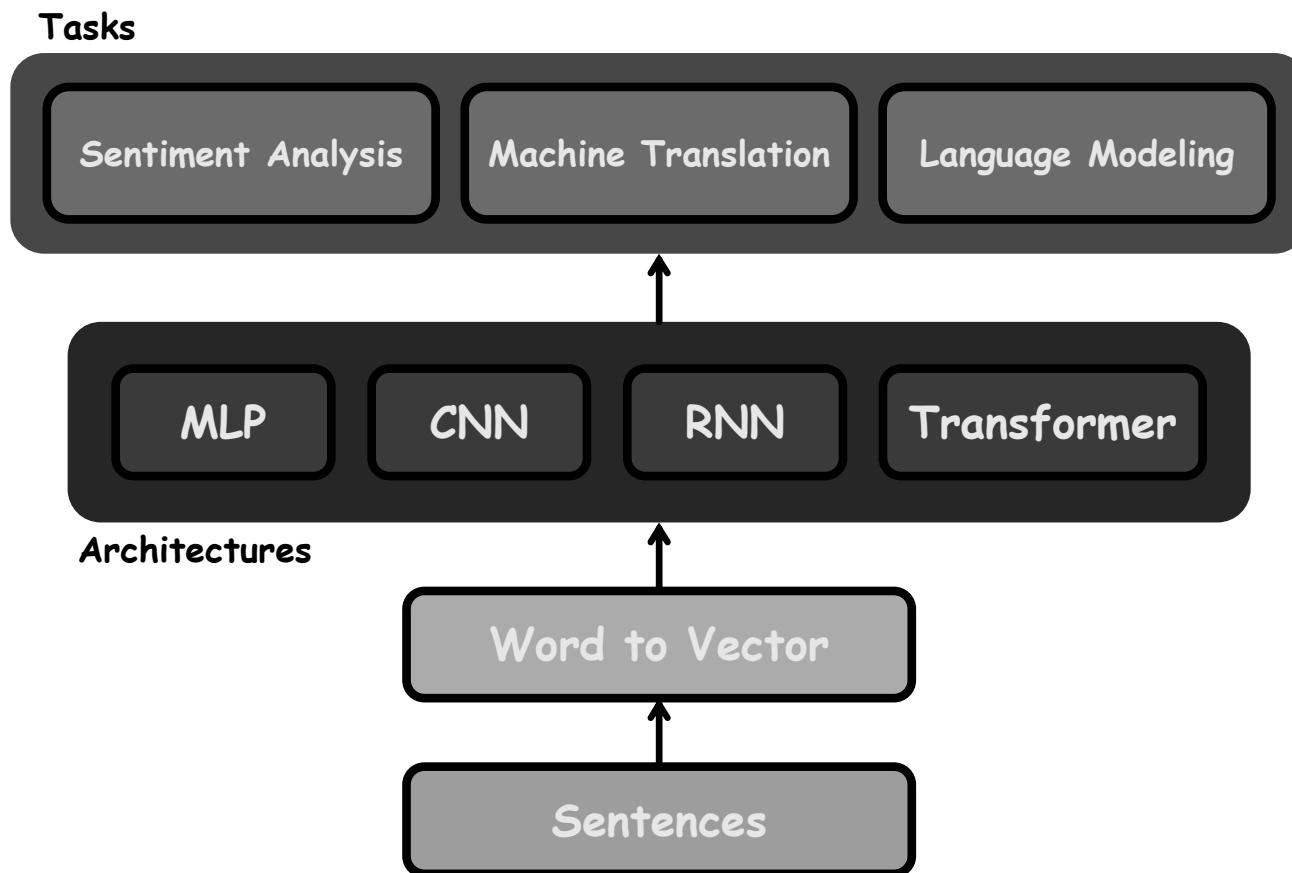
آکادمی هوسم از هرگونه نشر و کپی برداری از
مطلوب این مجموعه آموزشی راضی نیست.

لا^جه امبدینگ (Embedding)



آکادمی هوسم از هرگونه نشر و کپی برداری از
مطلوب این مجموعه آموزشی راضی نیست.

NLP



آکادمی هوسام از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

تَسْكِنَة

تسک‌ها

Movie Reviews

Cats & Dogs

.....

I absolutely loved this cat and that dog!

?

☆☆☆☆☆

...

☆☆☆☆☆

☆☆☆☆☆

new! 3+

I saw a cat |

I saw a cat on the chair
I saw a cat running after a dog
I saw a cat in my dream
I saw a cat book

Persian

سلام، حالت چطوره؟

←→

English

Hi how are you?

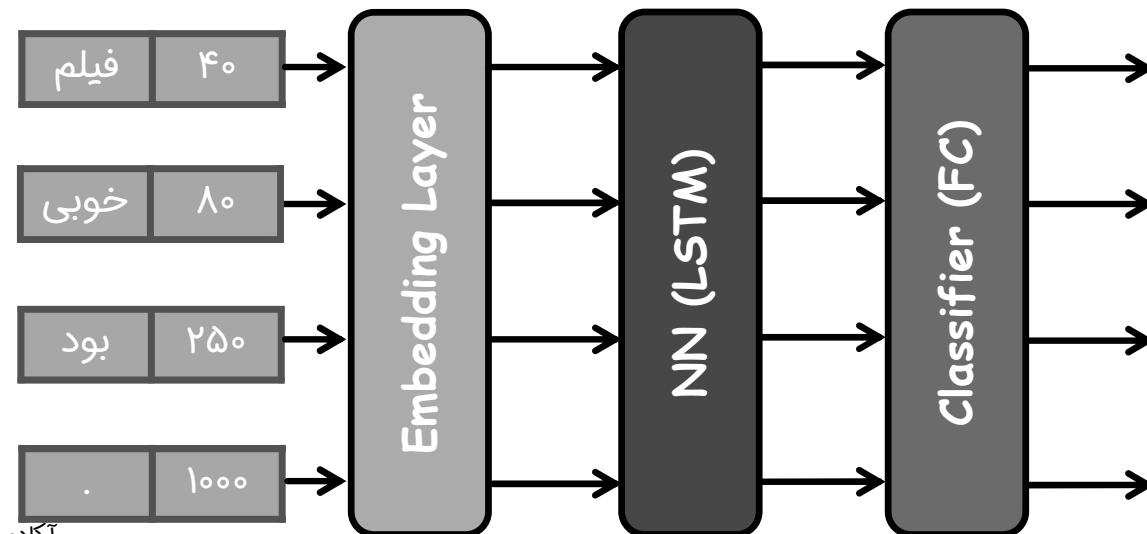
Transfer Learning

Transfer Learning

«دو مولفه با امکان ترانسفر لرنینگ:

«لایه امبدینگ

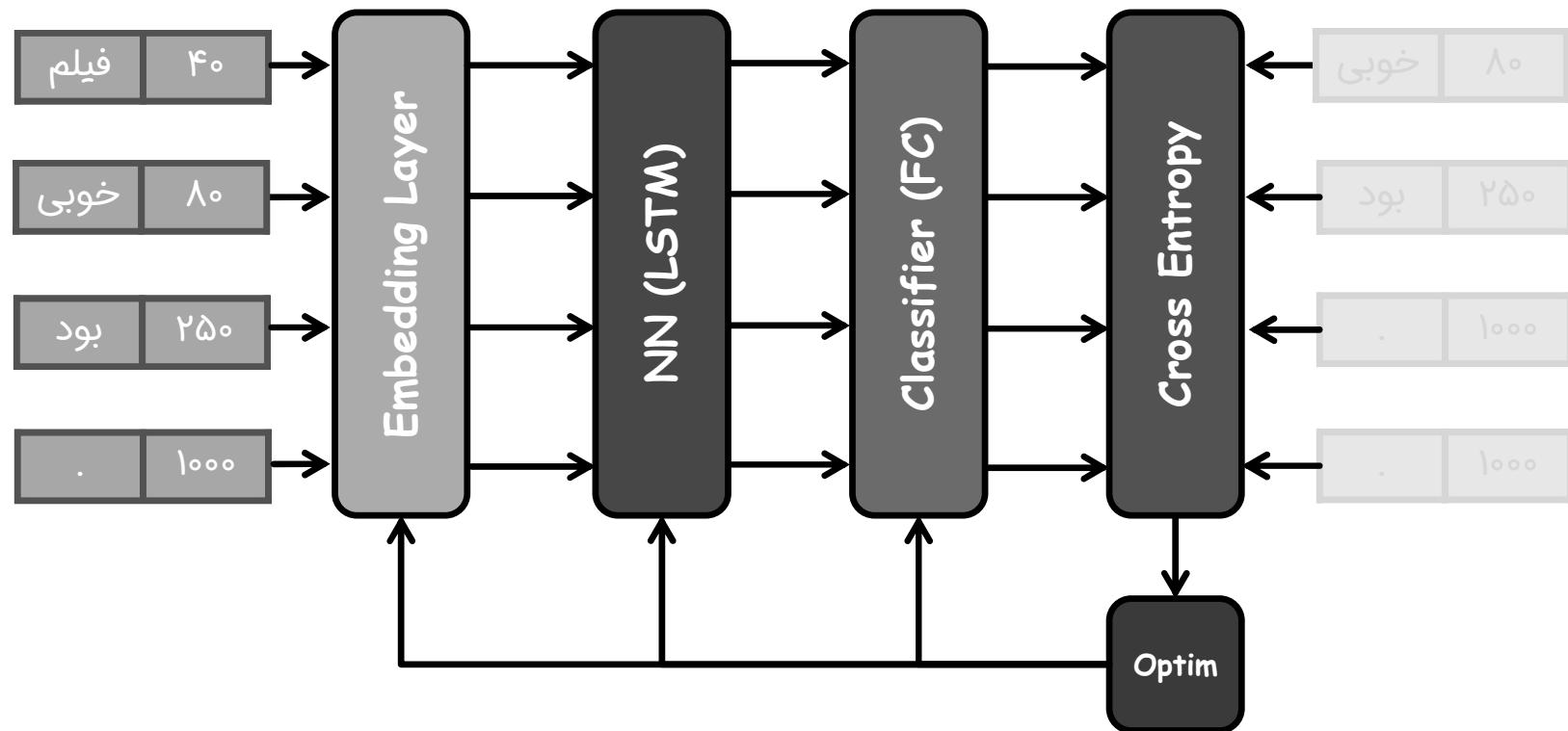
(Backbone) «بدنه



آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

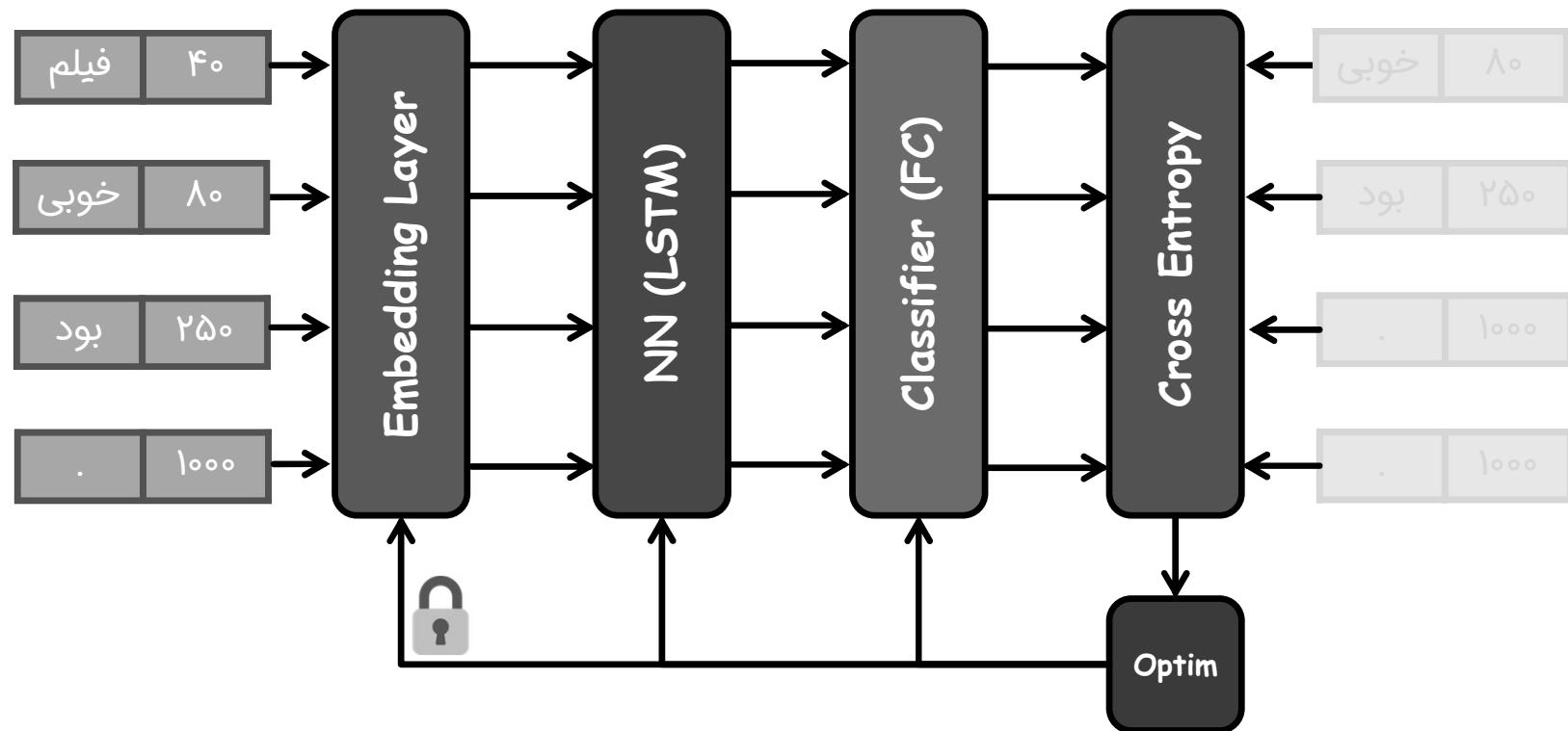
امبینگ (Embedding)

From scratch



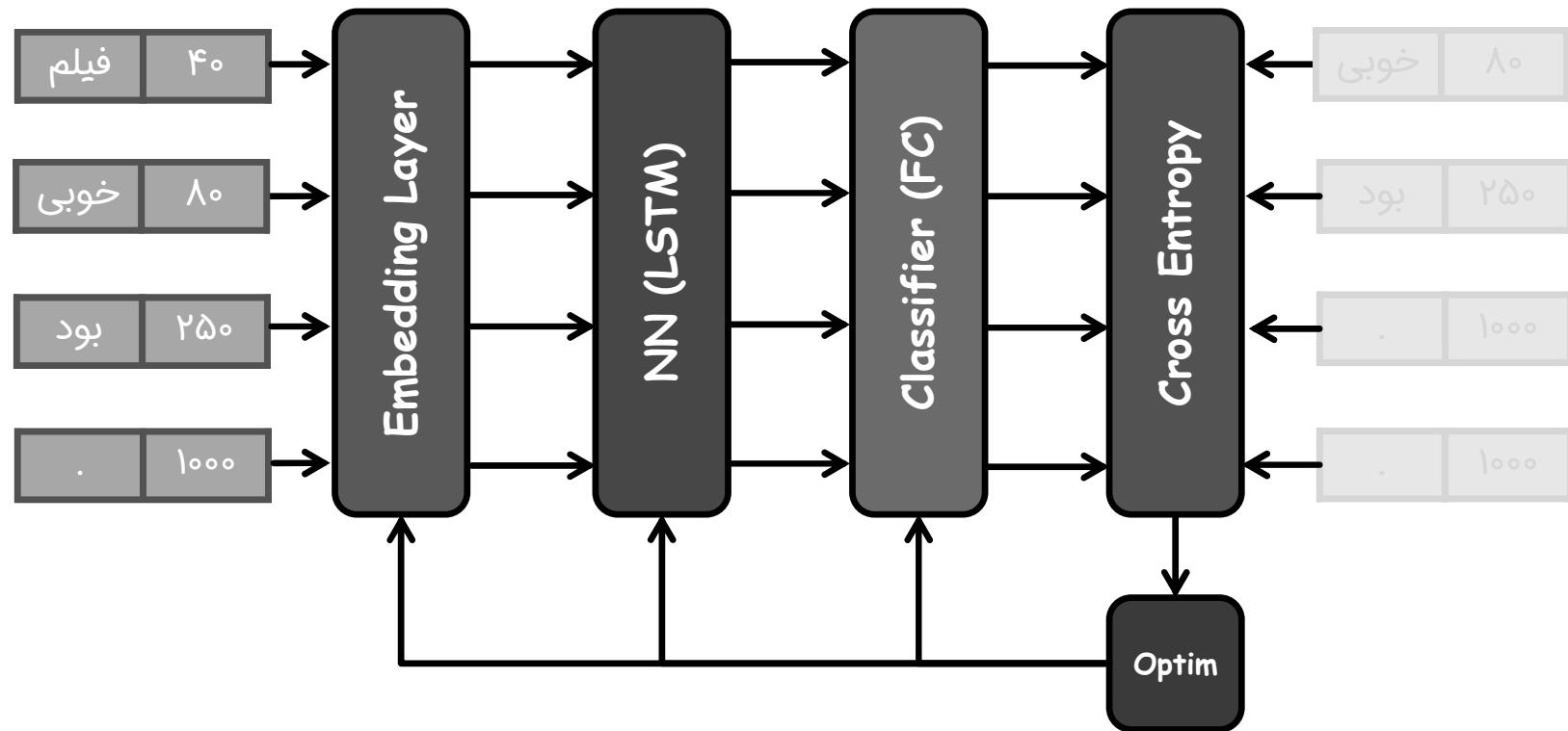
آکادمی هوسام از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

pre-trained fixed



آکادمی هوسام از هرگونه نشر و کپی برداری از
مطلوب این مجموعه آموزشی راضی نیست.

pre-trained + fine-tune

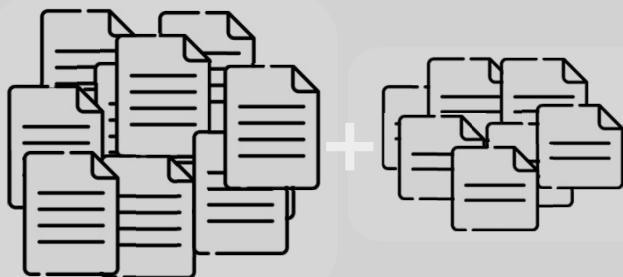


آکادمی هوسام از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

لایه امبدینگ

pre-trained + fine-tune

- » روابط بین کلمات خوب و
- » قابلیت تطبیق با تسک



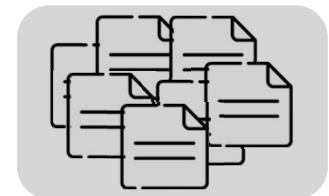
pre-trained fixed

- » روابط بین کلمات خوب، اما
- » عدم تطبیق با تسک

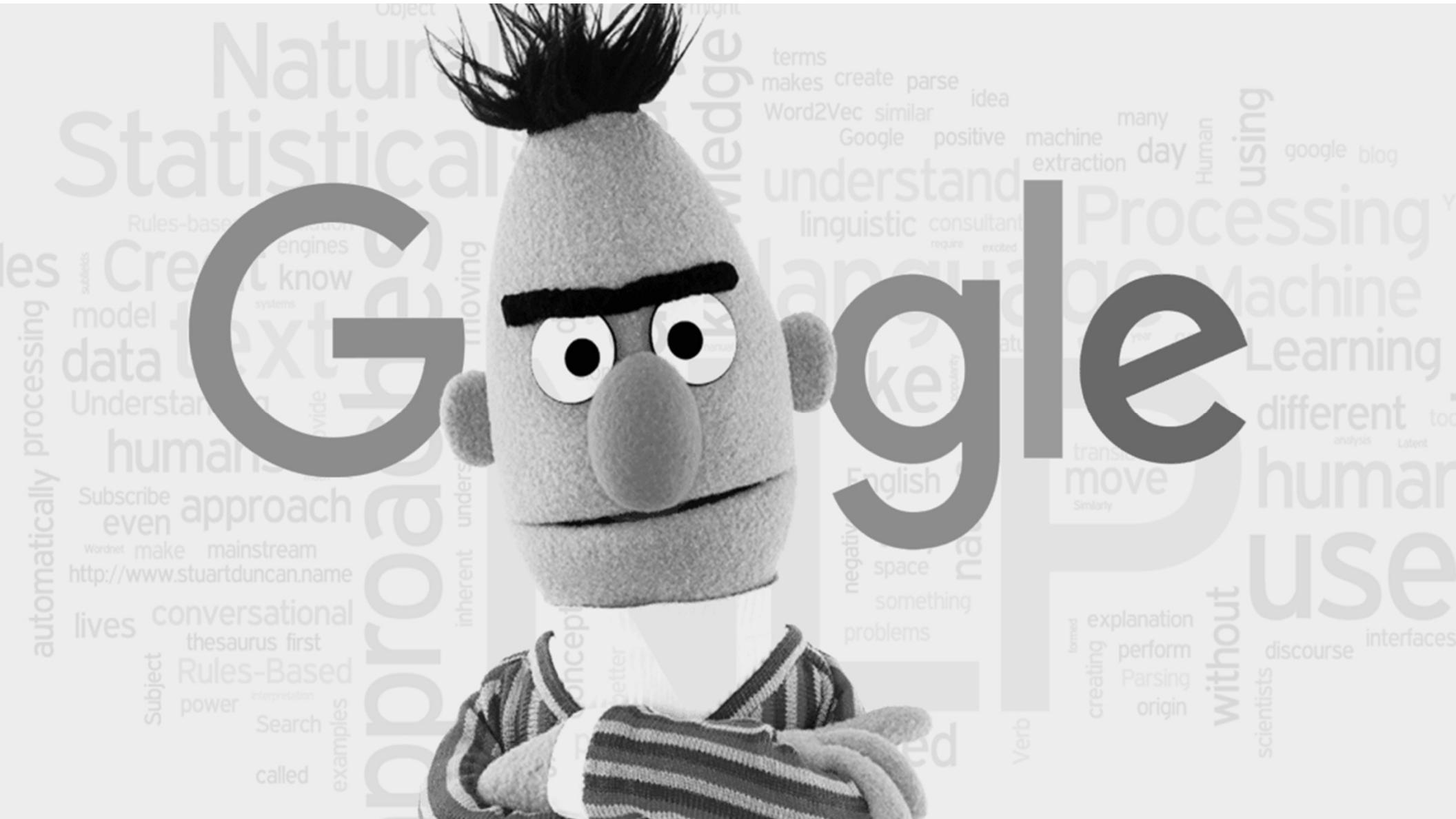


From scratch

- » امکان عدم یادگیری روابط بین کلمات
- » داده ناکافی



بُنه (Backbone)



BERT شبکه

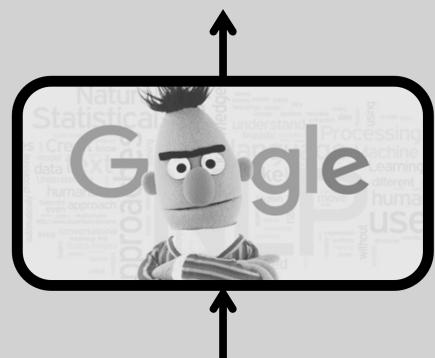
BERT Base

۱۱۰ میلیون پارامتر

$L = 12 \ll$

$H = 768 \ll$

$A = 12 \ll$



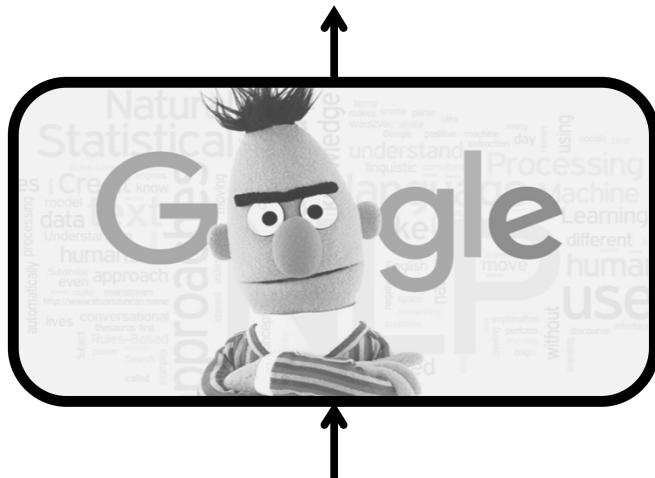
BERT Large

۳۴۰ میلیون پارامتر

$L = 24 \ll$

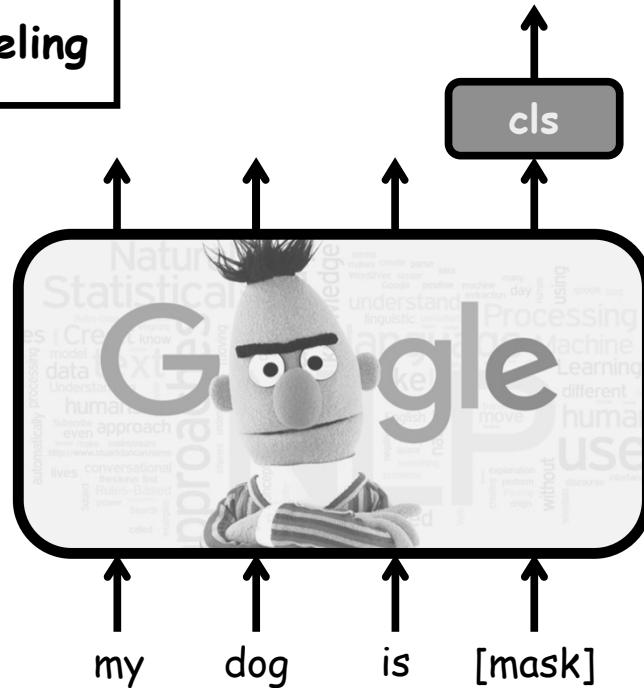
$H = 1024 \ll$

$A = 16 \ll$



آموزش شبکه BERT

Masked Language Modeling



آکادمی هوسم از هرگونه نشر و کپی برداری از
مطلوب این مجموعه آموزشی راضی نیست.

««« خودناظر (Self-supervised)

««« کتاب و ویکی پدیا

««« کتاب با ۸۰۰ میلیون کلمه

««« ویکی پدیا با ۲۵۰۰ میلیون کلمه

my dog is hairy

X

my dog is [mask]

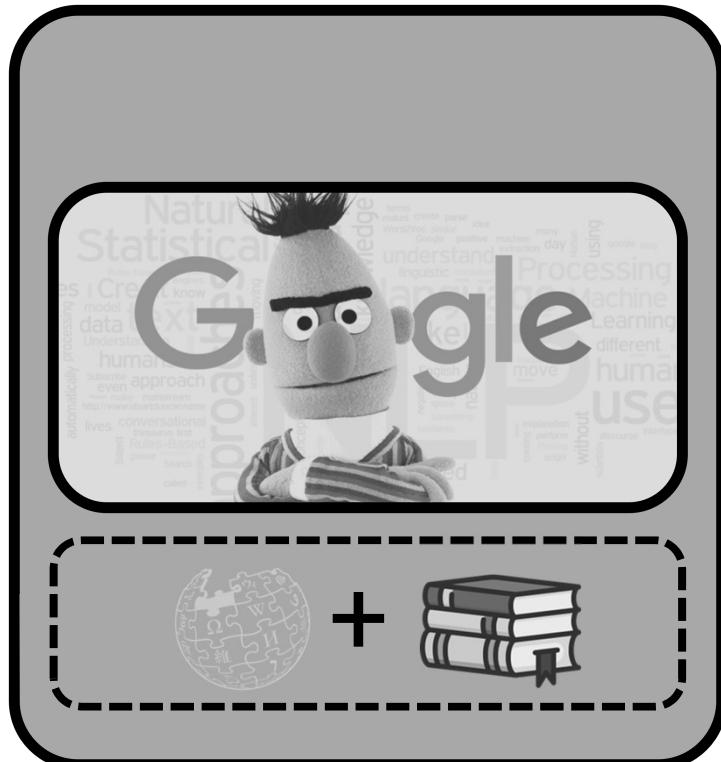
Y

hairy

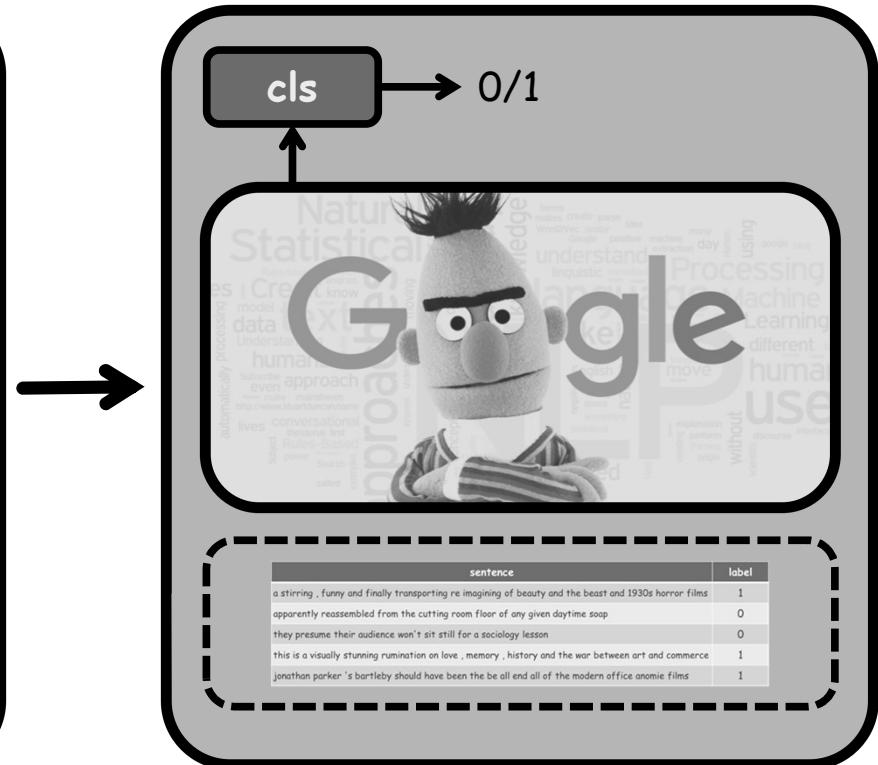


حالا ترنسفر لرنینگ

Self-supervised Learning



Supervised Learning



آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

مدل سازی زبانی

Language Modeling

مثال

دوره |

دوره رنسانس

دوره قاجار

دوره زبان

دوره MBA

دوره پایتون

...

آکادمی هوسام از هرگونه نشر و کپی برداری از
مطلوب این مجموعه آموزشی راضی نیست.

آکادمی هوش مصنوعی
howsam.org | هوسام

مثال

دوره ویدیویی |

دوره ویدیویی آموزش زبان انگلیسی

دوره ویدیویی ایزی چارت

دوره ویدیویی آموزش فارکس

دوره ویدیویی آموزش زبان آلمانی

دوره ویدیویی آموزش زبان

...

مثال

دوره ویدیویی یادگیری ا

دوره ویدیویی یادگیری زبان

دوره ویدیویی یادگیری عمیق

دوره ویدیویی یادگیری ماشین

...

مثال

دوره ویدیویی یادگیری عمیق ا

دوره ویدیویی یادگیری عمیق هوسم

دوره ویدیویی یادگیری عمیق فارسی

دوره ویدیویی یادگیری عمیق اندروانگ

دوره ویدیویی یادگیری عمیق با پایتون

دوره ویدیویی یادگیری عمیق با متلب

...

احتمال شرطی

دوره ویدئویی یادگیری عمیق هوسم

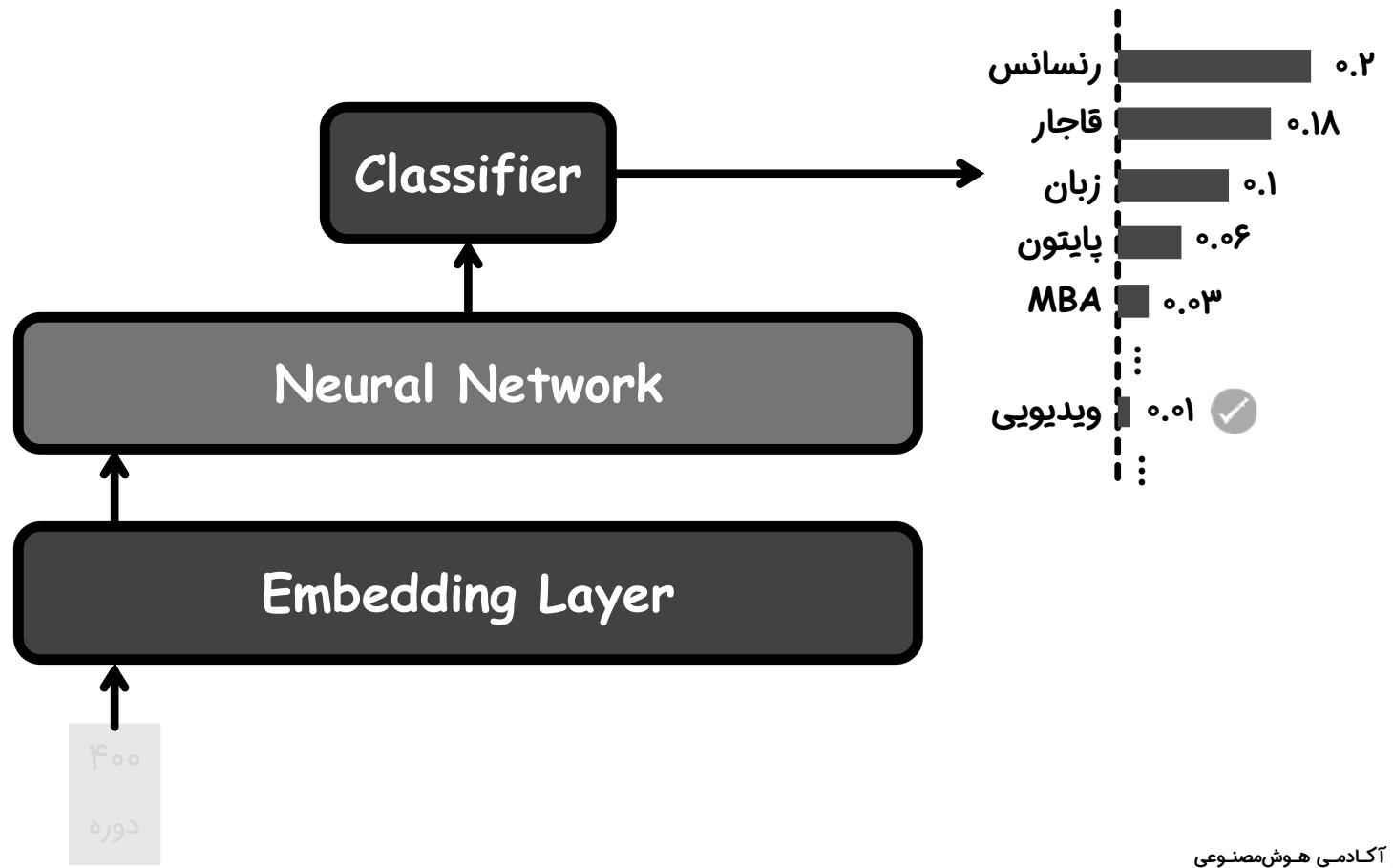
$P(\text{دوره ویدئویی یادگیری عمیق هوسم}) =$

$(\text{دوره ویدئویی یادگیری عمیق} | \text{هوسم}) P \times (\text{دوره ویدئویی یادگیری} | \text{عمیق}) P \times (\text{دوره ویدئویی} | \text{یادگیری}) P \times (\text{دوره} | \text{ویدئویی}) P$

احتمال شرطی

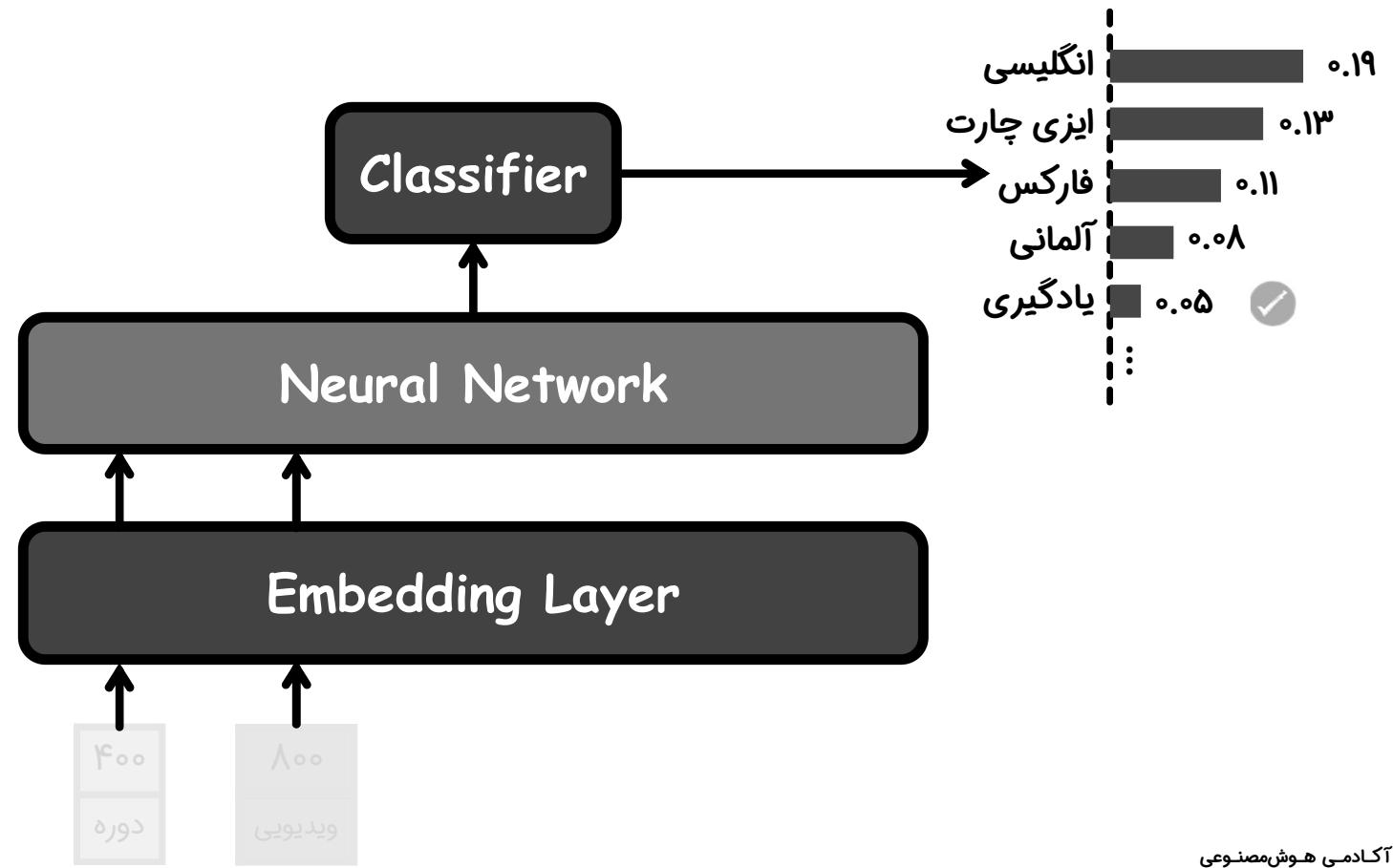
$$P(y_1, y_2, \dots, y_n) = P(y_1) \cdot P(y_2|y_1) \cdot P(y_3|y_1, y_2) \cdot \dots \cdot P(y_n|y_1, \dots, y_{n-1}) = \prod_{i=1}^n P(y_t|y_{<t})$$

Language Modeling



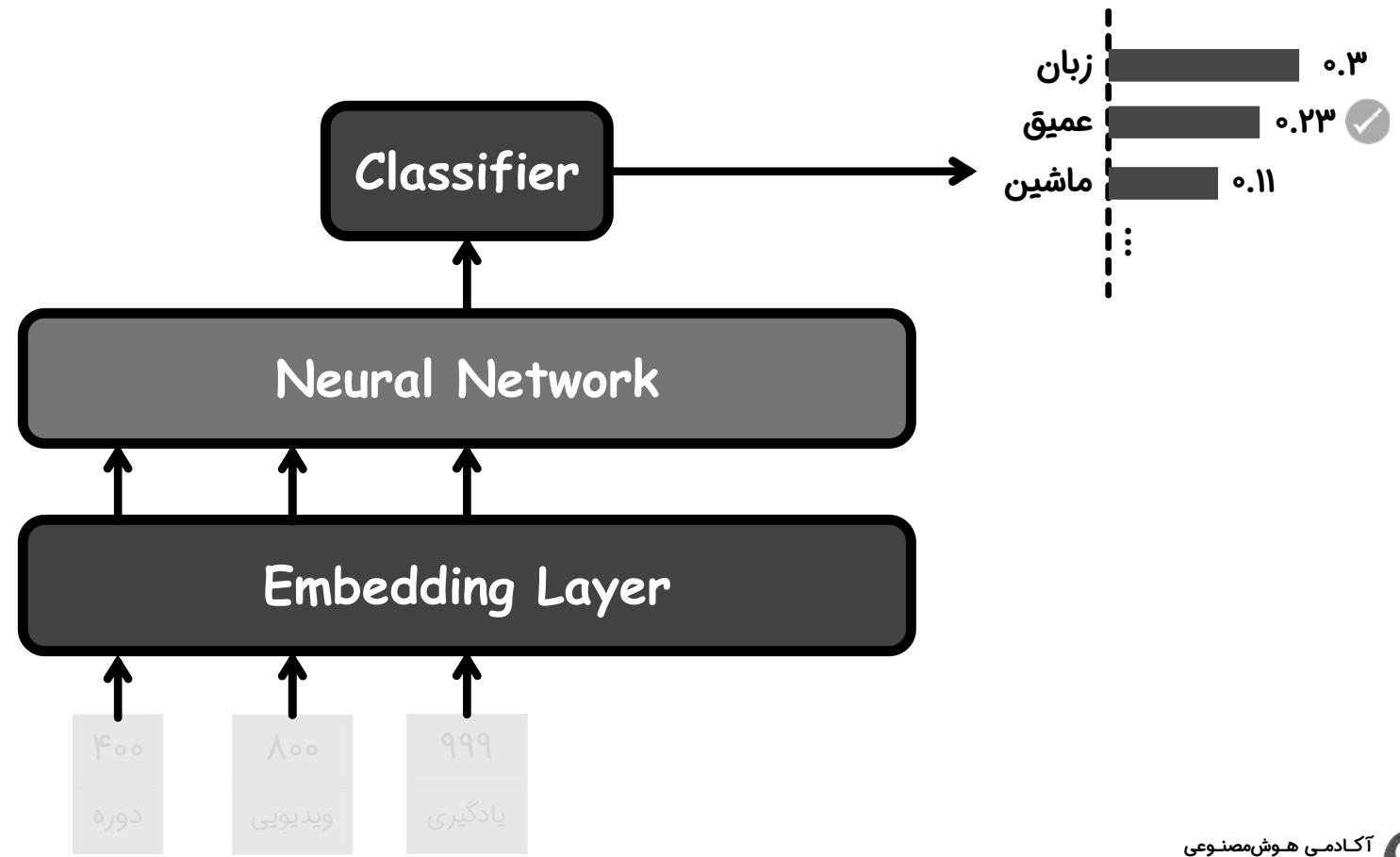
آکادمی هوسام از هرگونه نشر و کپی برداری از
مطلوب این مجموعه آموزشی راضی نیست.

Language Modeling



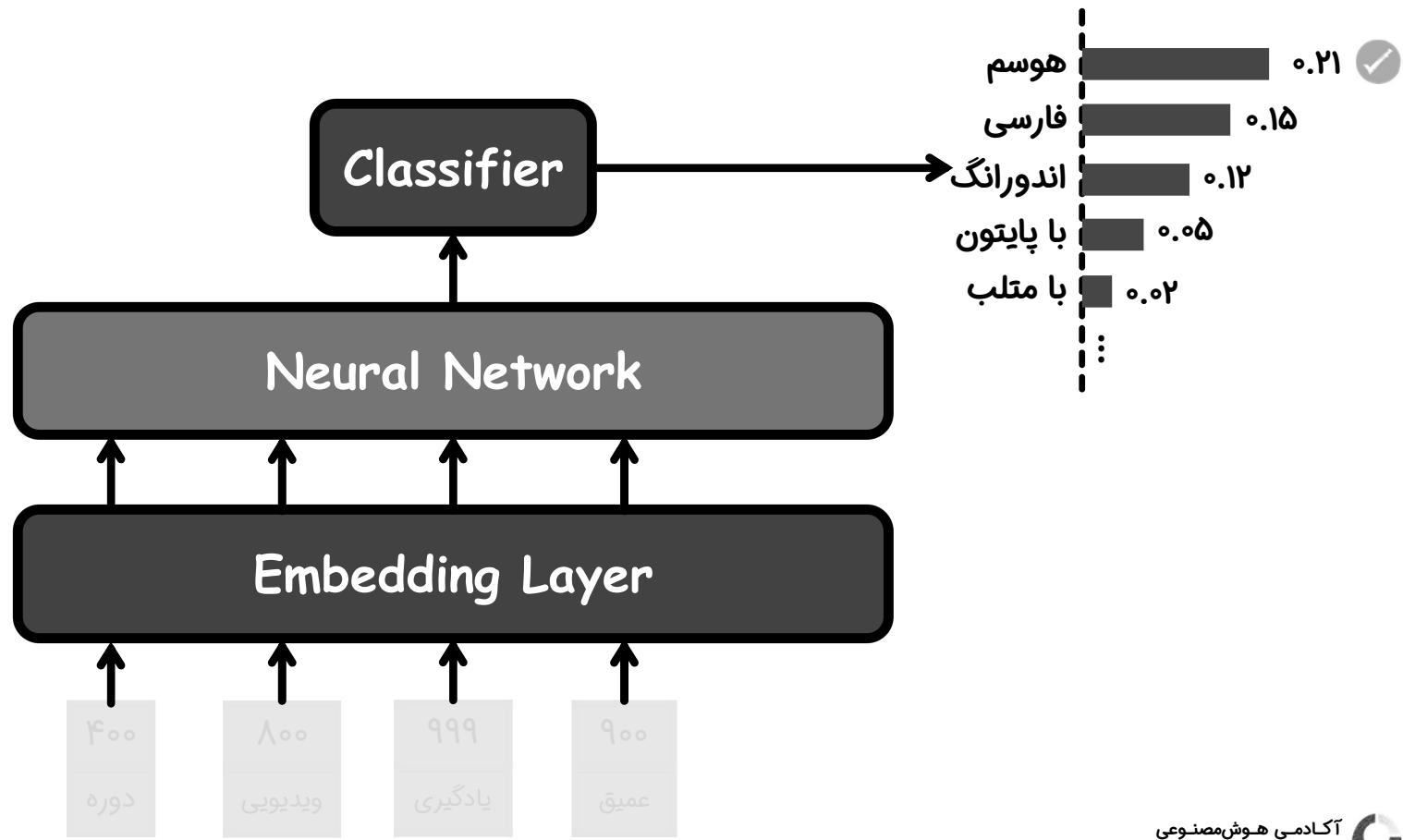
آکادمی هوسام از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

Language Modeling



آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

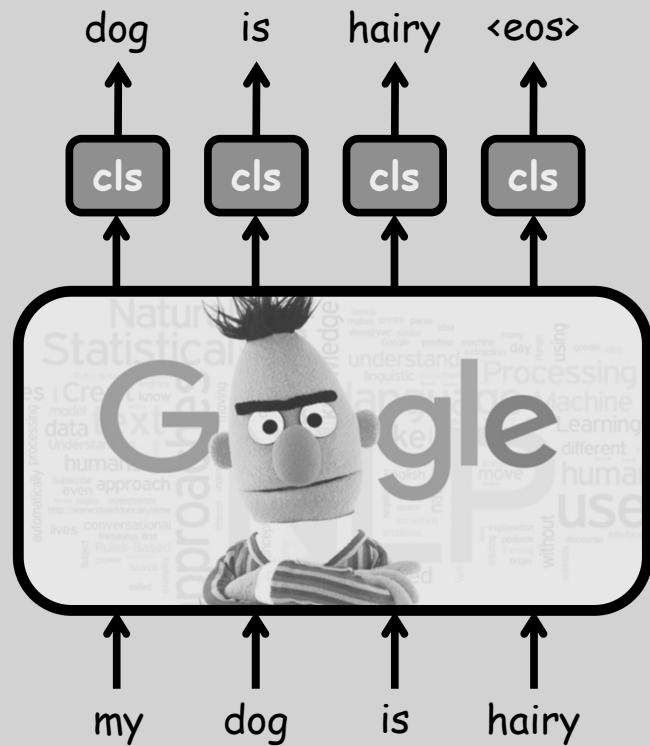
Language Modeling



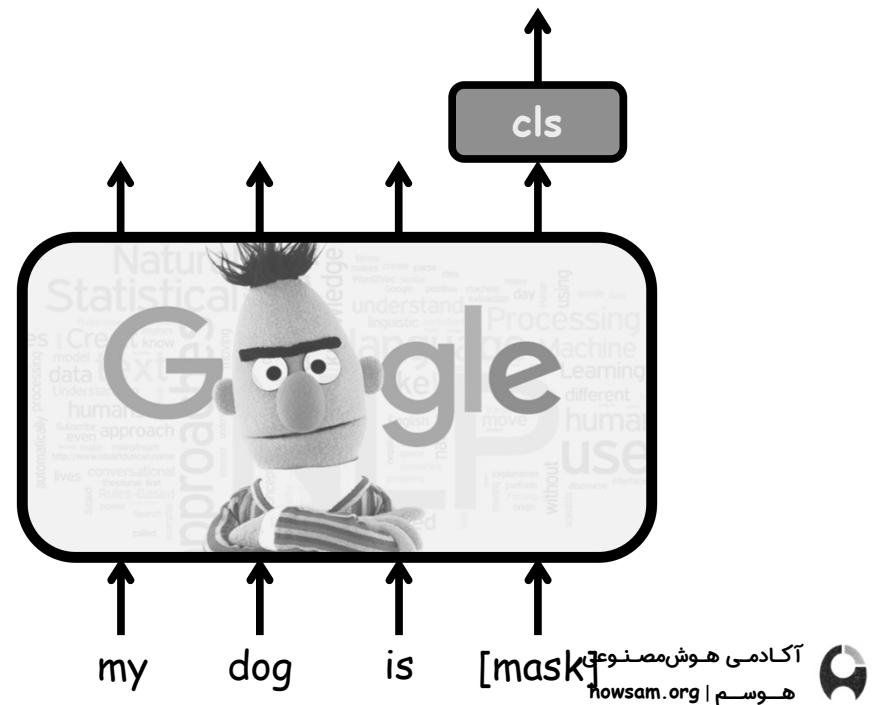
آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

آموزش LM

آموزش LM علی Causal Language Modeling



آموزش LM با ماسک Masked Language Modeling



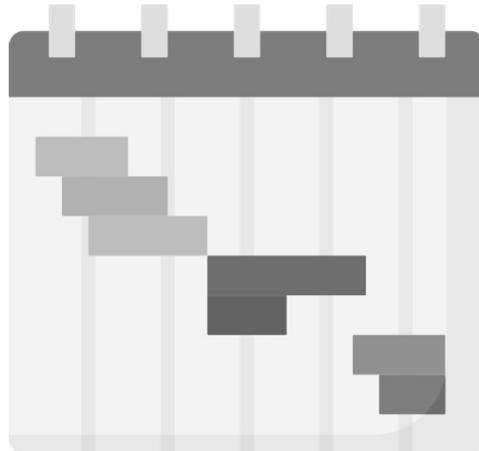
معیار ارزیابی Perplexity

$$PP(Y) = \sqrt[n]{\frac{1}{P(y_1, y_2, \dots, y_n)}}$$

تَمَامٌ

چک لیست انجام پروژہ

چک لیست انجام پروژه



﴿﴿﴿ گام ۱: بیان مساله

﴿﴿﴿ گام ۲: بررسی روش‌های موجود

﴿﴿﴿ گام ۳: انتخاب روش مناسب

﴿﴿﴿ گام ۴: پیاده‌سازی روش انتخابی

﴿﴿﴿ گام ۵: آزمایش

﴿﴿﴿ گام ۶: بهبود روش

﴿﴿﴿ گام ۷: کاربرد سازی

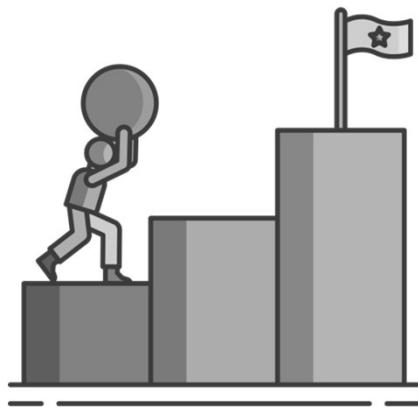
﴿﴿﴿ گام ۸: مستندسازی و ارائه

مدل‌سازی زبانی

Language Modeling

I think the movie is not to be the first episode

چالش!



﴿﴿ جزئیات چالش به شرح زیر هست:

- ﴿﴿ در شروع، ویدئوهای تئوری چکلیست پروژه در اختیار شما قرار می‌گیرد!
- ﴿﴿ یک نوتبوک تمپلیت به شما داده می‌شود.
- ﴿﴿ حالا شما باید ویدئوها را ببینید و کد هر بخش را بنویسید.
- ﴿﴿ برای این چالش یک هفته فرصت دارید!
- ﴿﴿ پایان هفته اول، ویدئوهای کدنویسی منتشر می‌شود.
- ﴿﴿ تا پایان هفته دوم، فرصت بررسی ویدئوهای کدنویسی را دارید.
- ﴿﴿ راستی یک کامپیوتونیتی خوب داریم که می‌توانید درمورد پروژه صحبت کنید.
- ﴿﴿ با حضور مدرس و کمک مدرس‌ها!

ڭام ا: بیان مسالە

گام ا: بیان مساله

»»» تعریف دقیق و مختصری از مساله

»»» مدل‌سازی زبان (Language Modeling) یک تسک مهم در پردازش زبان طبیعی (NLP) است که هدف آن تولید یک متن منسجم و معنادار براساس متن ورودی (prompt) است.

»»» از جمله کاربردهای مدل‌سازی زبان:

»»» تولید متن (Text Generation)

»»» ترجمه ماشینی (Machine Translation)

»»» بازشناسی گفتار (Speech Recognition)

»»» تحلیل معنایی متن (Sentiment Analysis)

»»» سوال و جواب (Question Answering)

»»» ...

I think the movie is not to be the first episode

آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطالب این مجموعه آموزشی راضی نیست.



گام ۳: بررسی روش‌های موجود

گام ۲: بررسی روش‌های موجود

»»» روش ۱: استفاده از شبکه بازگشتی

»»» یکی از رایج‌ترین رهیافت‌های مدل‌سازی زبان، استفاده از شبکه بازگشتی، لایه Embedding و لایه کلاسیفایر (fc) هست.

»»» یک نمونه لینک مرتبط

آکادمی هوسم از هرگونه نشر و کپی‌برداری از مطالب این مجموعه آموزشی راضی نیست.



گام ۲: بررسی روش‌های موجود

﴿﴿﴿ روش ۱: استفاده از شبکه بازگشتی

﴿﴿﴿ روش ۲: توسعه روش ۱

﴿﴿﴿ استفاده از شبکه‌های بازگشتی توسعه یافته مانند EGRU، QRNN، Mogrifier LSTM وغیره

[Mogrifier LSTM](#) ﴿﴿﴿ لینک

﴿﴿﴿ تنظیم و بهینه‌سازی مدل‌های زبانی LSTM

[paperswithcode](#) ﴿﴿﴿ سایت

آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطالب این مجموعه آموزشی راضی نیست.



گام ۲: بررسی روش‌های موجود

﴿﴿﴿ روش ۱: استفاده از شبکه بازگشته

﴿﴿﴿ روش ۲: توسعه روش ۱ با AWD-LSTM

﴿﴿﴿ روش ۳: استفاده از شبکه ترنسفورمر

﴿﴿﴿ این روش همان ساختار روش ۱ را دارد، با این تفاوت که شبکه ترنسفورمر جایگزین شبکه بازگشته شده است.

﴿﴿﴿ یک نمونه لینک مرتبط

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



گام ۲: بررسی روش‌های موجود

﴿﴿﴿ روش ۱: استفاده از شبکه بازگشته

﴿﴿﴿ روش ۲: توسعه روش ۱ با AWD-LSTM

﴿﴿﴿ روش ۳: استفاده از شبکه ترنسفورمر

﴿﴿﴿ روش ۴: توسعه روش ۳

﴿﴿﴿ استفاده از شبکه‌های ترنسفورمر توسعه یافته مانند GPT-2، BERT و غیره

﴿﴿﴿ به کارگیری ترنسفر لرنینگ

﴿﴿﴿ لینک مرتبط

آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطالب این مجموعه آموزشی راضی نیست.



گام ۲: بررسی روش‌های موجود

﴿﴿﴿ روش ۱: استفاده از شبکه بازگشته

﴿﴿﴿ روش ۲: توسعه روش ۱ با AWD-LSTM

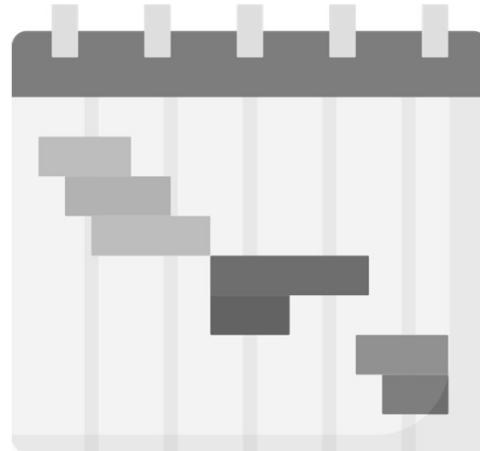
﴿﴿﴿ روش ۳: استفاده از شبکه ترنسفورمر

﴿﴿﴿ روش ۴: توسعه روش ۳ با GPT-2 Pretrained

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



چک لیست انجام پروژه



﴿﴿﴿ گام ۱: بیان مساله

﴿﴿﴿ گام ۲: بررسی روش‌های موجود

﴿﴿﴿ گام ۳: انتخاب روش مناسب

﴿﴿﴿ گام ۴: پیاده‌سازی روش انتخابی

﴿﴿﴿ گام ۵: آزمایش

﴿﴿﴿ گام ۶: بهبود روش

﴿﴿﴿ گام ۷: کاربرد سازی

﴿﴿﴿ گام ۸: مستندسازی و ارائه

گام ۳: انتخاب روش مناسب

یادآوری: روش‌های موجود

﴿﴿﴿ روش ۱: استفاده از شبکه LSTM

﴿﴿﴿ روش ۲: توسعه روش ۱ (AWD-LSTM)

﴿﴿﴿ روش ۳: استفاده از شبکه ترنسفورمر

﴿﴿﴿ روش ۴: توسعه روش ۳ (GPT-2 Pretrained)

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



گام ۳: انتخاب روش مناسب

﴿﴿﴿ روش ۱: استفاده از شبکه بازگشتی

﴿﴿﴿ شبکه انتخابی: LSTM. چرا؟

﴿﴿﴿ انتظار داریم به نتایج اولیه امیدوارکننده برسیم!

﴿﴿﴿ دردامه:

﴿﴿﴿ به کارگیری تکنیک‌های توسعه روشن ۱ (مانند AWD-LSTM)

﴿﴿﴿ حرکت به سوی شبکه ترنسفورمر

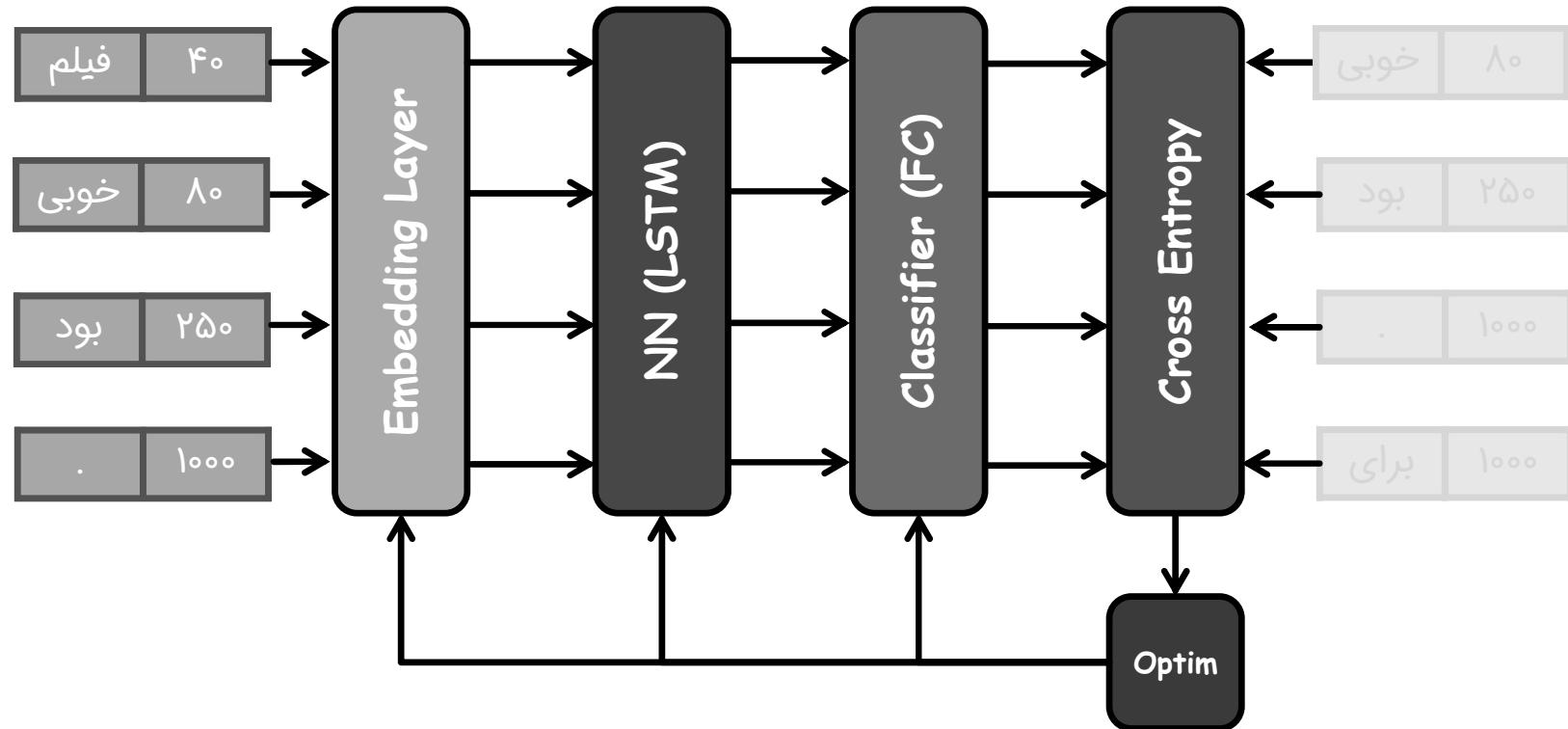
﴿﴿﴿ به کارگیری توسعه‌های روشن ۲ (GPT-2 یا BERT)

﴿﴿﴿ در صورت امکان!

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

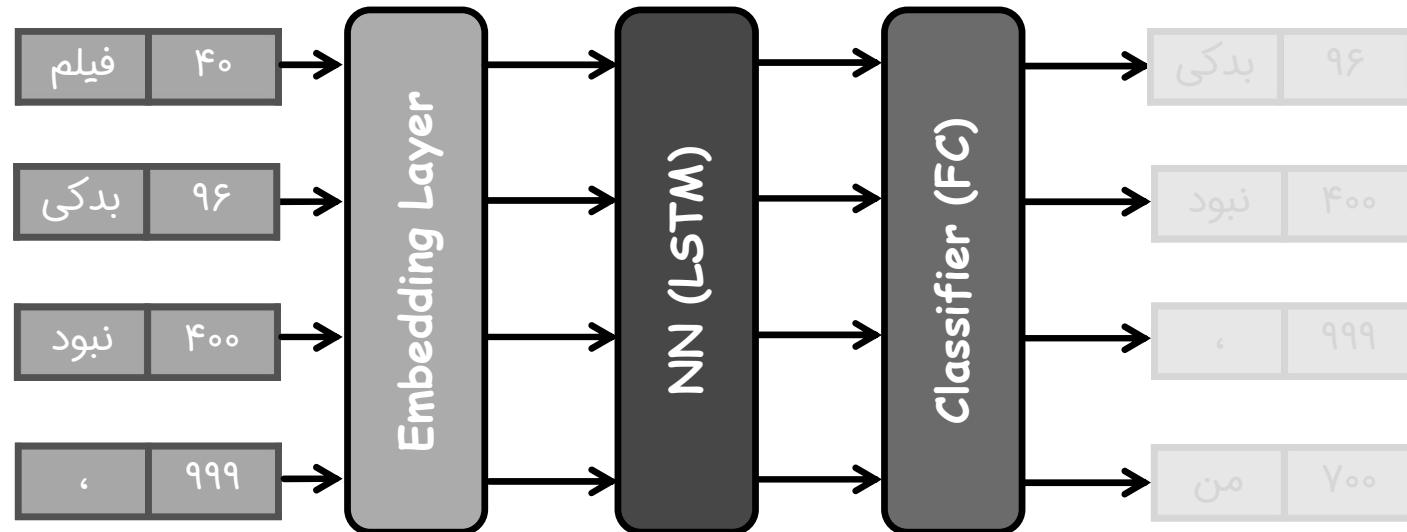


کام ۳: انتخاب روش مناسب {آموزش}



آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

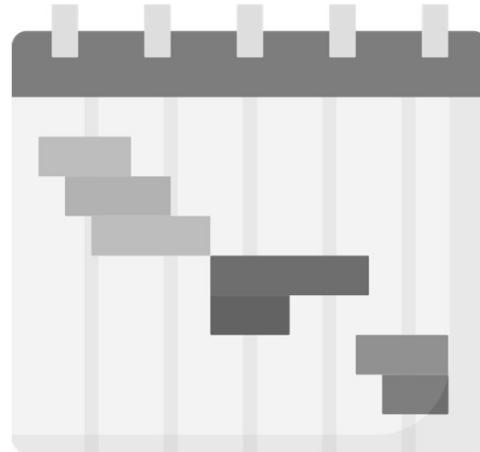
گام ۳: انتخاب روش مناسب {تولید}



آکادمی هوسم از هرگونه نشر و کپی برداری از مطالب این مجموعه آموزشی راضی نیست.

آماده‌ای؟!

چک لیست انجام پروژه



﴿﴿﴿ گام ۱: بیان مساله

﴿﴿﴿ گام ۲: بررسی روش‌های موجود

﴿﴿﴿ گام ۳: انتخاب روش مناسب

﴿﴿﴿ گام ۴: پیاده‌سازی روش انتخابی

﴿﴿﴿ گام ۵: آزمایش

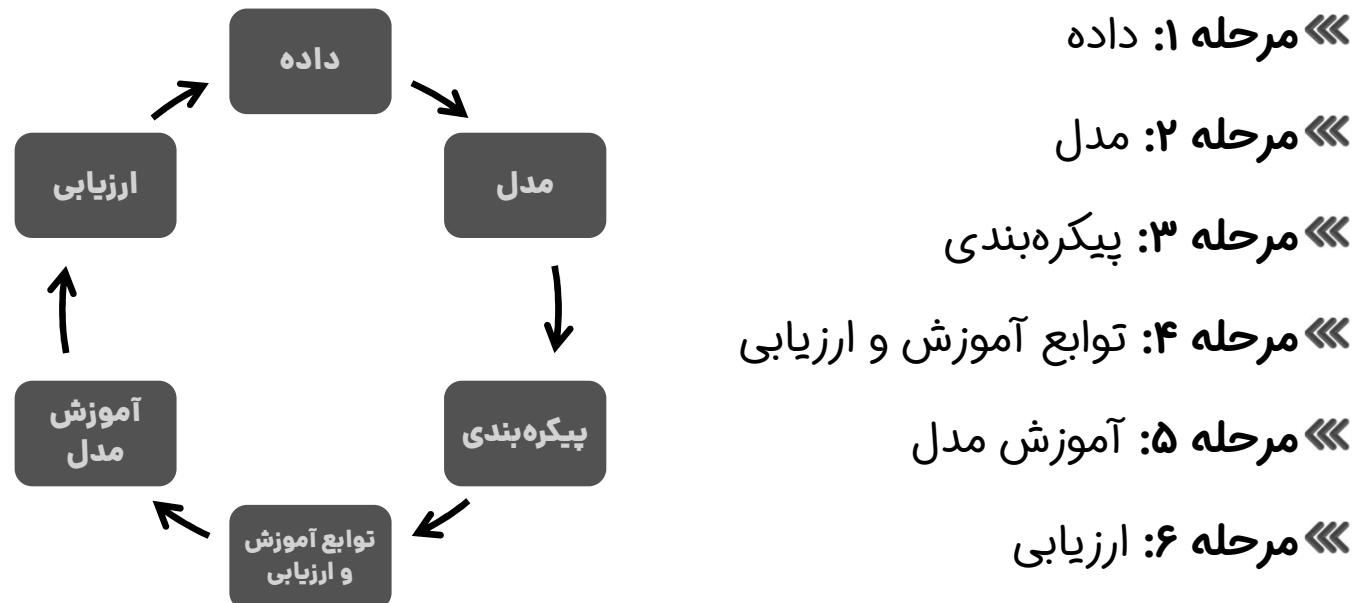
﴿﴿﴿ گام ۶: بهبود روش

﴿﴿﴿ گام ۷: کاربرد سازی

﴿﴿﴿ گام ۸: مستندسازی و ارائه

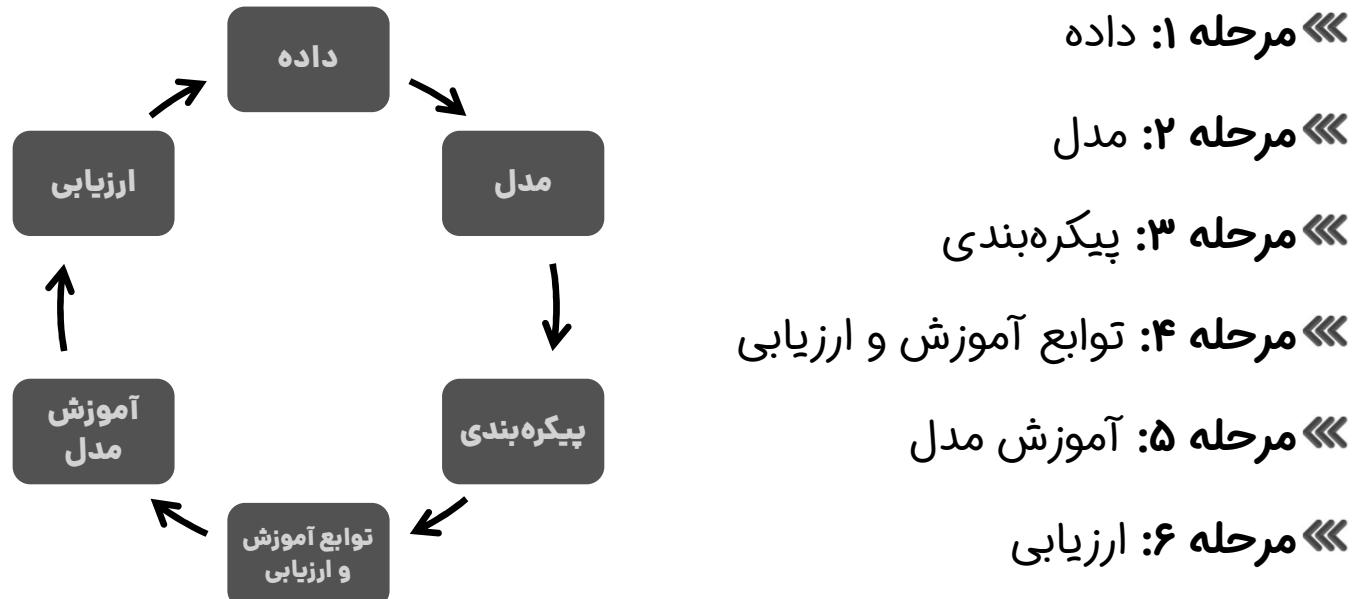
گام ۴: پیاده‌سازی روش انتخابی

مراحل پیاده‌سازی پروژه دیپ‌لرنینگ



آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطلوب این مجموعه آموزشی راضی نیست.

مراحل پیاده‌سازی پروژه دیپ‌لرنینگ



آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطلوب این مجموعه آموزشی راضی نیست.

مرحله ا: داده

مرحله ا: داده



««مجموعه کارهای لازم برای پروژه مدلسازی زبانی:

««تهیه داده

از دیتاست **WikiText2** استفاده کنید.

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



مرحله ا: داده



«« مجموعه کارهای لازم برای پروژه مدلسازی زبانی:

«« تهیه داده

«« ساخت دیکشنری

«« دیکشنری (Vocab) از روی داده آموزش بسازید.

«« کمترکارها را وارد Vocab نکنید.

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



مرحله ا: داده



««مجموعه کارهای لازم برای پروژه مدلسازی زبانی:

««تهیه داده

««ساخت دیکشنری

««تحلیل اکتشافی داده (همان EDA)

««حتما تلاش کنید تا به درگ خوبی از دیتاست برسید!

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



مرحله ا: داده



«« مجموعه کارهای لازم برای پروژه مدلسازی زبانی:

«« تهیه داده

«« ساخت دیکشنری

«« تحلیل اکتشافی داده (همان EDA)

«« آمادهسازی و پیشپردازش داده

«« تغییر شکل داده به شکل $N \times B \times L$ یا $M \times L$

«« N : تعداد کل بچه ها

«« B : سایز بچ

«« L : طول هر نمونه در بچ

«« $N \times B \times M$

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



ساختار بج داده در LM

خیلی عالی بود.

نامید کننده هست، فکر نمی‌کردم اینقدر بد باشد.

خوب بود، اما من در کارهای آینده انتظار بیشتری از شما دارم.

ساختار بچ داده در LM

خیلی	عالی	بود	.
نامید	کننده	هست	.
خوب	بود	،	اما
بد	فکر	من	در
اینقدر	نمی‌کرم	کارهای	آینده
باشد	.	انتظار	بیشتری
از	.	دaram	شما
.	.	.	.

آکادمی هوسام از هرگونه نشر و کپی برداری از مطالب این مجموعه آموزشی راضی نیست.

ساختار بچ داده در ~~CLS~~

.	بود	عالی	خیلی
.	هست	کننده	ناامید
.	اما	،	بود

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطلوب این مجموعه آموزشی راضی نیست.

ساختار بج داده در LM

بود خوب . باشد بد اینقدر نمی‌کرم فکر ، هست کننده نامید . بود عالی خیلی

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطلوب این مجموعه آموزشی راضی نیست.

آکادمی هوشمصنوعی
howsam.org | هوسم

ساختار بچ داده در LM

.	باشد	بد	اینقدر	نمی‌کرم	بیشتری	فکر	,	هست	کننده	ناامید	.	بود	عالی	خیلی
.	دارم	شما	از			انتظار	آینده	کارهای	در	من	اما	,	بود	خوب
...
...
...
...
...
...
...
...
...
...
...
...
...
...

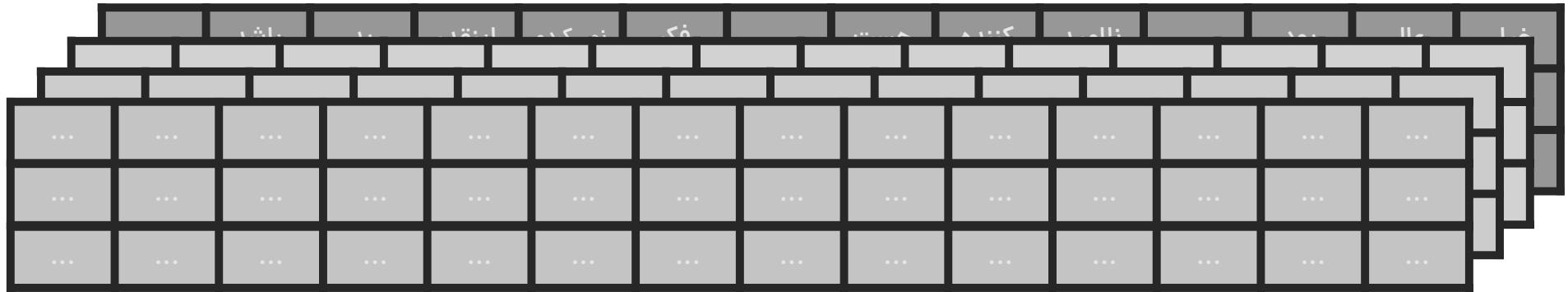
کمی برداری از
ضی نیست.

ساختار بچ داده در LM

.	باشد	بد	اینقدر	نمی‌کرم	فکر	,	هست	کننده	ناامید	.	بود	عالی	خیلی
.	دارم	شما	از	بیشتری	انتظار	آینده	کارهای	در	من	اما	,	بود	خوب
...
...
...
...
...
...
...
...
...
...
...
...
...

کبی برداری از
ضی نیست.

ساختار بچ داده در LM



آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

مرحله ا: داده



﴿﴿﴿ مجموعه کارهای لازم برای پروژه مدلسازی زبانی:

﴿﴿﴿ تهیه داده

﴿﴿﴿ ساخت دیکشنری

﴿﴿﴿ تحلیل اکتشافی داده (همان EDA)

﴿﴿﴿ آماده‌سازی و پیش‌پردازش داده

﴿﴿﴿ پیاده‌سازی ماثول فراخوانی داده

﴿﴿﴿ کاستوم دیتاست برای فراخوانی یک نمونه/بچ داده

آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطالب این مجموعه آموزشی راضی نیست.



مرحله ا: داده



﴿﴿﴿ مجموعه کارهای لازم برای پروژه مدلسازی زبانی:

﴿﴿﴿ تهیه داده

﴿﴿﴿ ساخت دیکشنری

﴿﴿﴿ تحلیل اکتشافی داده (همان EDA)

﴿﴿﴿ آماده‌سازی و پیش‌پردازش داده

﴿﴿﴿ پیاده‌سازی ماثول فراخوانی داده

آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطالب این مجموعه آموزشی راضی نیست.



تمرین



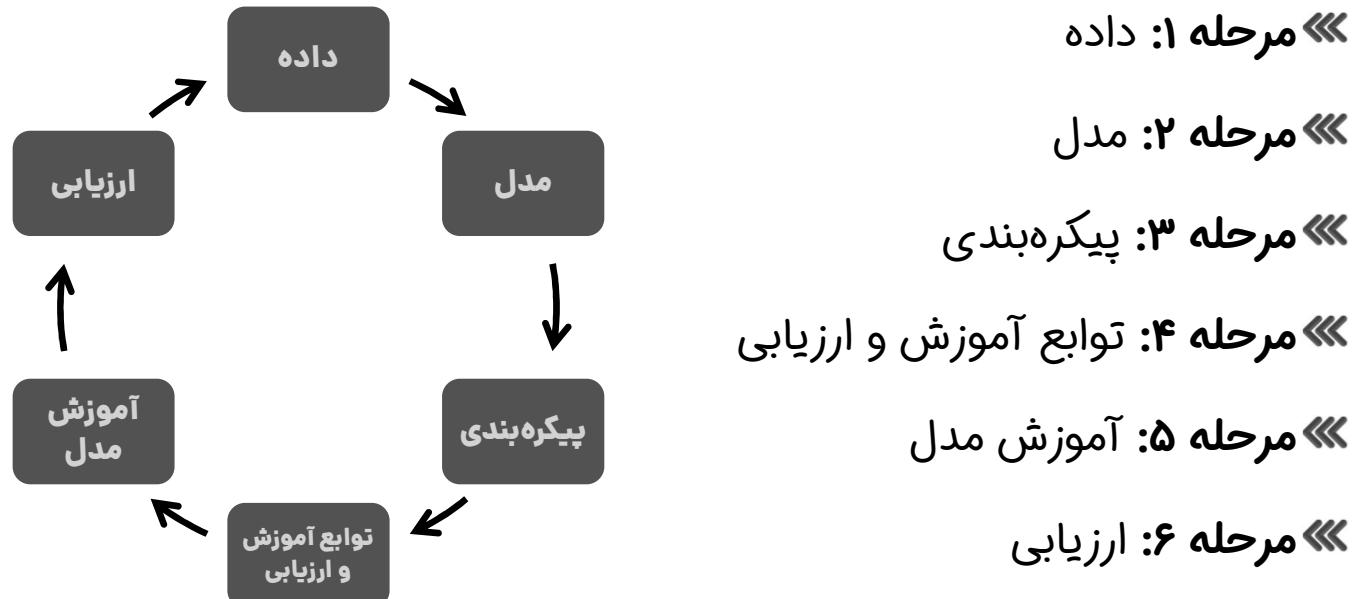
﴿﴿ بخش داده دیتاست WikiText2 را پیادهسازی کنید.

آکادمی هوسام از هرگونه نشر و کپی برداری از مطالب این مجموعه آموزشی راضی نیست.



<کدنویسی/>

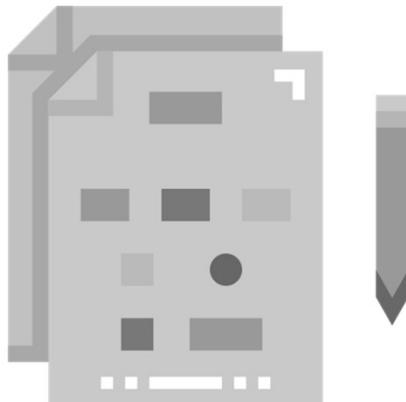
مراحل پیاده‌سازی پروژه دیپ‌لرنینگ



آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطلوب این مجموعه آموزشی راضی نیست.

مرحله ۲: مدل

مرحله ۲: مدل



«« بخش `__init__`

«« لایه امبدینگ

«« سایز امبدینگ؟

«« می‌تواند همراه با وزن‌های آموزش دیده باشد!

«« شبکه LSTM

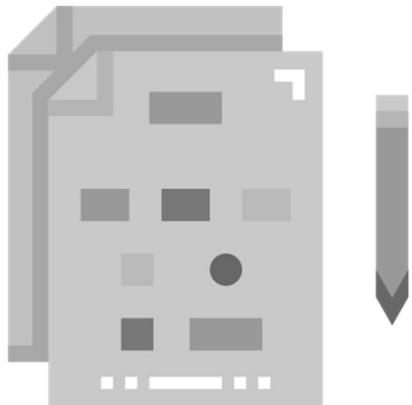
«« تعداد لایه‌ها و سایز `fc` داخلی؟!

«« لایه دسته‌بندی

«« تعداد کلاس‌ها؟!

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

مرحله ۲: مدل



«بخش forward»

«ابتدا لایه امیدینگ»

«سپس، استفاده از شبکه بازگشتی»

«نهایتاً، دسته‌بندی»

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



تمرین



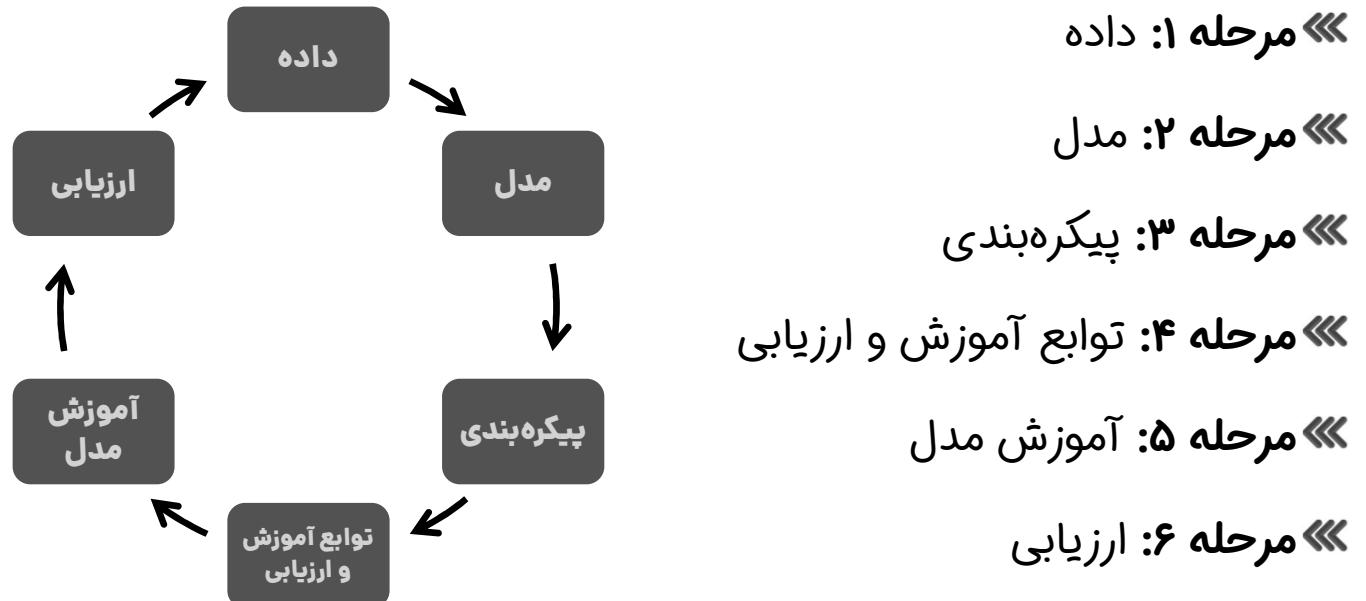
- ﴿ بخش داده دیتاست WikiText2 را پیادهسازی کنید.
- ﴿ مدل روش انتخابی را پیادهسازی کنید.

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



<کدنویسی/>

مراحل پیاده‌سازی پروژه دیپ‌لرنینگ



آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطلوب این مجموعه آموزشی راضی نیست.

مرحله ۳: پیگرد بندی

مرحله ۳: پیکره‌بندی (Configurations)



﴿﴿﴿ مجموعه تنظیمات لازم برای راهاندازی/آموزش مدل

﴿﴿﴿ تعریف مدل با آرگومان‌های مناسب

﴿﴿﴿ تعریف تابع اتلاف مناسب

﴿﴿﴿ تعریف بهینه‌ساز

﴿﴿﴿ تنظیم هایپرپارامترها

﴿﴿﴿ تعریف معیار ارزیابی

Perplexity ﴿﴾

آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطالب این مجموعه آموزشی راضی نیست.



مرحله ۴: توابع آموزش و ارزیابی

همان قبليها!

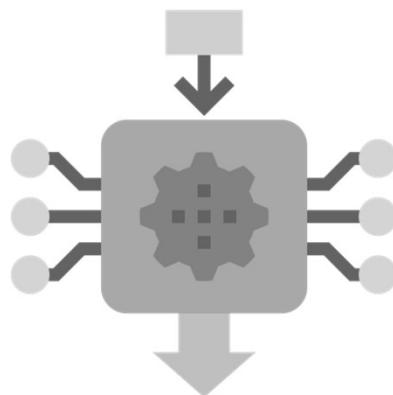
مرحله ۵: آموزش مدل

دستور پخت آموزش بهینه مدل

مرحله ۶: ارزیابی

همان تابع Evaluate!

به تولید هم نیاز داریم!



» متن تولیدی مدل برای یک متن ورودی دلخواه از کاربر
» این کار در قالب تابعی با عنوان **generate** انجام می‌شود.

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

تابع generate

«« تابعی بنویسید که:

«« متن ورودی کاربر را دریافت کند.

«« توکنایز و تبدیل به اندیس را انجام دهد.

«« بعنوان ورودی به مدل آموزش دیده داده شود و خروجی کلاسیفایر دریافت شود.
← ↑
«« خروجی، برداری به طول تعداد توکن‌های دیکشنری هست.

«« خروجی از یک سافت‌مکس با Temperature مشخص رد شود.

«« حال با فرض توزیع مولتی‌نومیال یک نمونه انتخاب کنید.

«« اگر «unk» بود، یک بار دیگر انتخاب کنید.

«« اگر به «eos» رسیدیم، تابع generate پایان یابد.

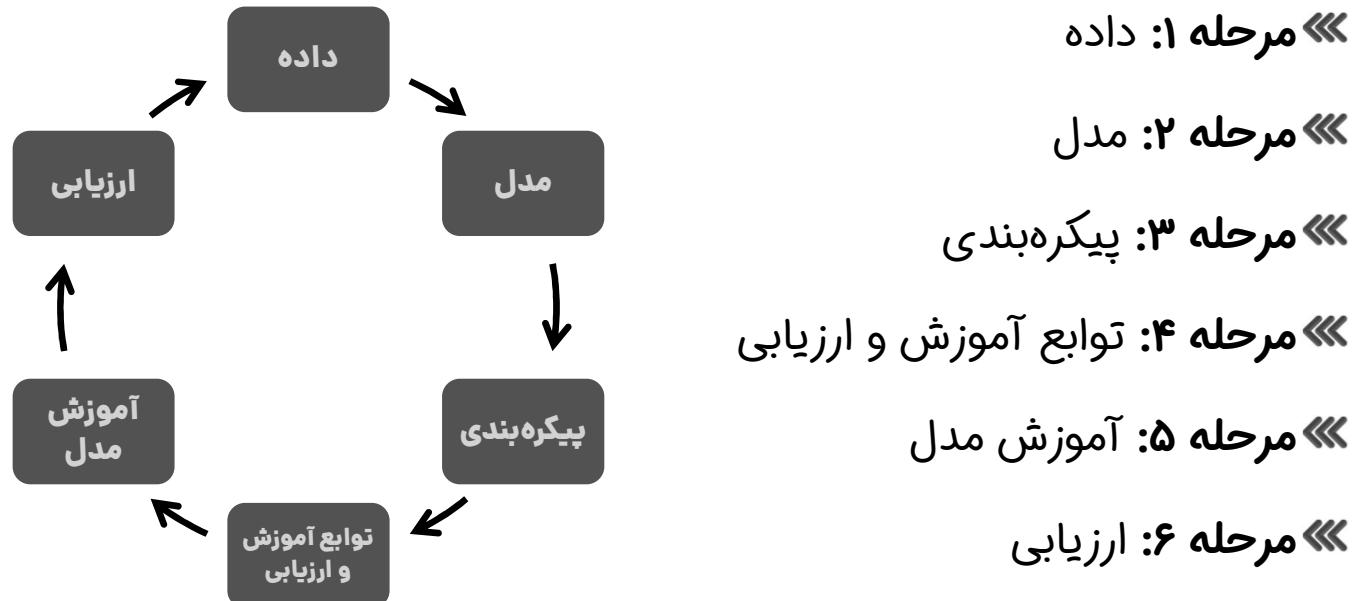
«« توکن جدید به ورودی اضافه شود.

→ «« برگردیم به مرحله ۳

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.

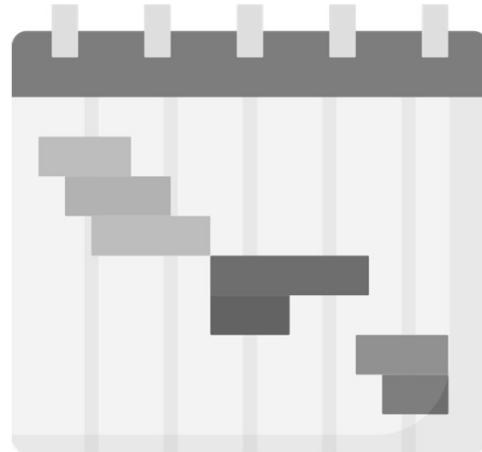


مراحل پیاده‌سازی پروژه دیپ‌لرنینگ



آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطلوب این مجموعه آموزشی راضی نیست.

چک لیست انجام پروژه



﴿﴿﴿ گام ۱: بیان مساله

﴿﴿﴿ گام ۲: بررسی روش‌های موجود

﴿﴿﴿ گام ۳: انتخاب روش مناسب

﴿﴿﴿ گام ۴: پیاده‌سازی روش انتخابی

﴿﴿﴿ گام ۵: آزمایش

﴿﴿﴿ گام ۶: بهبود روش

﴿﴿﴿ گام ۷: کاربرد سازی

﴿﴿﴿ گام ۸: مستندسازی و ارائه

تمرین



﴿﴿ بخش داده را روی دیتاست UTKFace پیاده‌سازی کنید.

﴿﴿ مدل روش انتخابی را پیاده‌سازی کنید.

﴿﴿ بخش پیکره‌بندی، توابع آموزش، ارزیابی و غیره را آماده کنید.

﴿﴿ مدل را آموزش دهید.

﴿﴿ تابع generate را بنویسید.

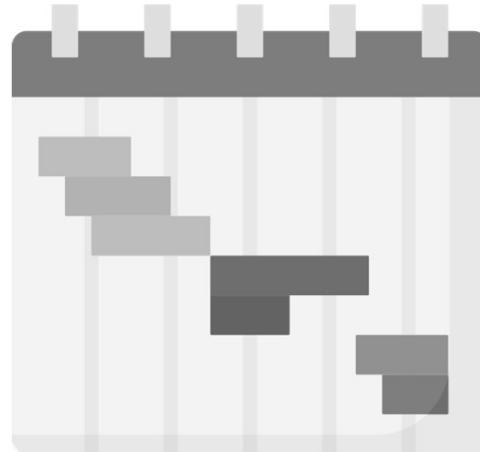
آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



<کدنویسی/>

توقف!

چک لیست انجام پروژه



﴿﴿﴿ گام ۱: بیان مساله

﴿﴿﴿ گام ۲: بررسی روش‌های موجود

﴿﴿﴿ گام ۳: انتخاب روش مناسب

﴿﴿﴿ گام ۴: پیاده‌سازی روش انتخابی

﴿﴿﴿ گام ۵: آزمایش -

﴿﴿﴿ گام ۶: بهبود روش

﴿﴿﴿ گام ۷: کاربرد سازی

﴿﴿﴿ گام ۸: مستندسازی و ارائه

توقف!

«پروژه را تا انتهای پیاده‌سازی کردیم!

«نتیجه تولید جمله جالب هست، اما

«با کیفیت نیست!

«باید به فکر چاره باشیم...

«عملکرد مدل را بهبود دهیم.

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطلوب این مجموعه آموزشی راضی نیست.



گام ۲: بررسی روش‌های موجود

﴿﴿﴿ روش ۱: استفاده از شبکه بازگشته

﴿﴿﴿ روش ۲: توسعه روش ۱ با AWD-LSTM

﴿﴿﴿ روش ۳: استفاده از شبکه ترنسفورمر

﴿﴿﴿ روش ۴: توسعه روش ۳ با GPT-2 Pretrained

آکادمی هوسم از هرگونه نشر و کپی‌برداری از
مطالب این مجموعه آموزشی راضی نیست.



گام ۳: انتخاب روش مناسب

﴿﴿﴿ روش ۱: استفاده از شبکه بازگشتی

﴿﴿﴿ شبکه انتخابی: LSTM. چرا؟

﴿﴿﴿ انتظار داریم به نتایج اولیه امیدوارکننده برسیم!

﴿﴿﴿ دردامه:

﴿﴿﴿ به کارگیری تکنیک‌های توسعه روشن ۱ (مانند AWD-LSTM)

﴿﴿﴿ حرکت به سوی شبکه ترنسفورمر

﴿﴿﴿ به کارگیری توسعه‌های روشن ۲ (GPT-2 یا BERT)

﴿﴿﴿ در صورت امکان!

آکادمی هوسم از هرگونه نشر و کپی برداری از
مطالب این مجموعه آموزشی راضی نیست.



چالش جدید!

Yashu Seth · Sep 12, 2018
@yashuseth · Follow
In this blog post, I give a walkthru of the AWD-LSTM paper by @Smerity Nitish Keskar @RichardSocher and try to explain a set of awesome techniques that have taken language modeling to new heights.
What makes the AWD-LSTM great? [yashuseth.blog/2018/09/12/awd...](http://yashuseth.blog/2018/09/12/awd-lstm-walkthrough/) via @yashuseth

Smerity · Sep 12, 2018
@Smerity · Follow
Oh that looks amazing! Better than what I would have and should have written up as a blog post when publishing the AWD-LSTM work! Nitish has his Twitter handle at @StrongDuality :)

﴿﴿﴿ می خواهیم کارمان را بهبود دهیم:

﴿﴿﴿ کار AWD-LSTM را مطالعه کنید.

﴿﴿﴿ مقاله | گیتهاب | مدیوم | نوتبک | بلاگ

﴿﴿﴿ چکلیستی از ایده‌های مقاله آماده کنید.

﴿﴿﴿ از ایده‌های ساده تا ایده‌های پیچیده

﴿﴿﴿ یکی یکی ایده‌ها را روی کدتان اعمال کنید.

﴿﴿﴿ میزان اثرباری ایده‌ها را در جدولی یادداشت کنید.

﴿﴿﴿ این فرآیند را تا رسیدن به نتایج مطلوب و نزدیک به مقاله ادامه دهید.

﴿﴿﴿ هشدار: هدف، صرفا افزایش دقیقت نیست! می خواهیم کار اصولی را بیاموزیم!

آکادمی هوسم از هرگونه نشر و کپی برداری از مطالب این مجموعه آموزشی راضی نیست.



موفق باشید!