subject : ۹۹۶۴۲۱۰۴ - محمد یارمقدم date :

1) 1a.

i) momentum باعث می‌شود که گرادیان فقط به گرادیان mini_batch مرتبط با جوش وابسته نباشد دقیقاً مانند کاری که SGD انجام می‌دهد. در بعضی از موارد اگر ضریب یادگیری خیلی بزرگ باشد ما هیچ وقت به راه‌حل بهینه نخواهیم رسید و converge نمی‌شویم. در طرف مقابل اگر خیلی هم کوچک باشد converge خیلی آرام صورت خواهد گرفت. استفاده از momentum باعث می‌شود که ضریب یادگیری stable باشد و به سرعت تغییر نکند و می‌توانیم ضریب یادگیری را با مقدار بزرگ‌تری مقداردهی اولیه کنیم تا جوش خودش را تنظیم کند تا converge زودتر اتفاق بیفتد.

ii) پارامترهای مدل که آپدیت های کوچک‌تری دریافت می‌کنند، آپدیت‌های بزرگ‌تری دریافت خواهند کرد. شرایط بالانس را درنظر می‌گیریم. پارامتر های مدل که آپدیت های بزرگ‌تری دریافت می‌کنند، ضریب یادگیری بهینه خود را خواهند داشت. پس می‌توانیم ضریب های یادگیری را به عنوان نرمال سازی آپدیت پارامتر adaptive در نظر بگیریم.

b)

i) هدف متغیر γ این است مقدار expected متغیر h_{drop} را همچنان h قرار دهد.

پس چون مقدار expected ها به صورت مقابل است:

$d = 1 - P_{drop} \Rightarrow$

$$\gamma = \frac{1}{1 - P_{drop}}$$

ii) این تکنیک regularization مورد مدت است تا از overfitting جلوگیری کند.

چون overfitting فقط در مرحله training اتفاق می افتد.

(۲)

a)

stack	Buffer	New dependency	Transition
[ROOT]	[I, parsed, this, sentence, correctly]		initial config
[ROOT, I]	[parsed, this, sentence, correctly]		shift
[ROOT, I, parsed]	[this, sentence, correctly]		shift
[ROOT, parsed]	[this, sentence, correctly]	I ← parsed	left_arc
[ROOT, parsed, this]	[sentence, correctly]		shift
[ROOT, parsed, this, sentence]	[correctly]		shift
[ROOT, parsed, sentence]	[correctly]	this ← sentence	left_arc
[ROOT, parsed]	[correctly]	parsed → sentence	right_arc
[root, parsed, correctly]	[]		shift
[ROOT, parsed]	[]	parsed → correctly	right_arc
[ROOT]	[]	ROOT → parsed	right_arc

b) هر کلمه می تواند با تمامی کلمات موجود در متن dependency تشکیل دهد پس اگر طول متن n کلمه باشد هر کلمه با n-1 کلمه دیگر بتواند رابطه داشته باشد

$Sum = (n-1)^n$

۹۹۴۴۲۱۰۴ محمد یارمقدم

subject: date:

۲) f)

i) I was heading to a wedding faring my death.

ارتباط wedding → fearing اشتباه است. از آنجا که یک head به verb phrase

اشتباه متصل شده است پس از نوع verb phrase attachment error خواهد بود. صحیح

آن heading → fearing است.

ii) It makes me want to rush out and rescue people from dilemmas of their own making.

رابطه and ← rescue اشتباه است. چون یک head به second conjuct اشتباه متصل

شده است پس از نوع Coordination attachment error خواهد بود. رابطه درست

به صورت rush → and خواهد بود.

iii) It is on loan from a guy named Joe O'Neil in Midland, Texas.

رابطه named → Midland غلط است. چون یک head به prepositional phrase اشتباه

متصل است پس از نوع prepositional phrase attachment error است. رابطه صحیح به شکل

guy → Midland است.

iv) رابطه most ← element غلط است و از نوع Modifier Attachment error است.

رابطه درست most ← crucial است.

MAXnote

Scanned with CamScanner

subject : گزارش اجرای آموزش / آزمون date :

درست داده های کوچکتر که با دستور python run.py -d اجرا شد هر epoch از train شامل ۴۸ جمله بود که evaluation که معیار UAS بود (که در بهترین حالت) از مقدار ۵۹ آغاز شد و در epoch آخر به مقدار ۶۹٫۶۸ رسید اما در مواردی از اجرا در محدوده ۲۰ تا ۳۰ به کار خود پایان می داد. در موارد دیگری UAS از epoch ۴ یا ۵ شروع به کاهش می کرد که نشان دهنده overfitting در train داده ها بود.

درست داده های اصلی در بهترین حالت UAS از مقدار ۸۲ آغاز و با مقدار ۸۸٫۹ پایان یافت اما در مواردی مقادیر کم هم داشت. میزان loss یا خطا در انتهای کار طبق پیش بینی کمتر از ۰٫۰۸ شد. مقدار دقیق آن ۰٫۰۶۵ شد.

عکس هر دو مورد در ادامه گزارش پیوست می شود.