



درس مبانی و کاربردهای هوش مصنوعی

گزارش پروژه سوم



بررسی تاثیر تغییر پارامترهای مدل آموزش داده شده:

برای بررسی تاثیر تغییر پارامترهای مدل، ابتدا نتایج مدل آموزش داده شده و دقت بدست آمده توسط این مدل به ازای پارامترهای مختلف به صورت زیر می باشد:

شماره مدل	دقت مدل Backoff	دقت مدل Bigram	دقت مدل Unigram	پارامترهای مدل
1	0.5625	0.8194	0.8125	$L1 = 0.001, L2 = 0.233, L3 = 0.766,$ $Eps = 0.001$
2	0.6845	0.8194	0.8125	$L1 = 0.00001, L2 = 0.23333, L3 =$ $0.76666,$ $Eps = 0.001$
3	0.8194	0.8194	0.8125	$L1 = 0, L2 = 0.5, L3 = 0.5,$ $Eps = 0.001$
4	0.8386	0.8194	0.8125	$L1 = 0, L2 = 0.001, L3 = 0.999,$ $Eps = 0.001$
5	0.7252	0.8194	0.8125	$L1 = 0.001, L2 = 0.233, L3 = 0.766,$ $Eps = 0.000001$
6	0.7896	0.8194	0.8125	$L1 = 0.00001, L2 = 0.23333, L3 =$ $0.76666,$ $Eps = 0.000001$
7	0.8194	0.8194	0.8125	$L1 = 0, L2 = 0.5, L3 = 0.5,$ $Eps = 0.000001$
8	0.8386	0.8194	0.8125	$L1 = 0, L2 = 0.001, L3 = 0.999,$ $Eps = 0.000001$

مدل 4 با 8 و مدل 3 با 7 در اصل یکی هستند، به این دلیل که ضریب $L1$ که در Eps ضرب می شود برابر صفر می باشد، بنابراین مقدار Eps تاثیری در خروجی مدل backoff ندارد.

با توجه به مدل های 1 و 2 و 5 می توان نتیجه گرفت که هر چه حاصل ضرب Eps با $L1$ کوچک تر شود، دقت مدل backoff نیز بیشتر می شود؛ چرا که این مقدار همانند یک بایاس عمل می کند و می تواند مرزهای تصمیم گیری بین داده

ها را جابجا کند؛ بنابراین بیشترین مقدار دقت بدست آمده از طریق پارامترهای متفاوت برابر با حدوداً 0.83 می باشد؛ که این دقت زمانی بدست می آید که مقدار $L1 = 0$ ، $L2 = 0.001$ و $L3 = 0.999$ باشد.

با توجه به اینکه مدل بیگرام جزئیات بیشتری از جمله را در نظر می گیرد، به همین دلیل دقیق تر از مدل یونیگرام می باشد؛ بنابراین هر چه ضریب $L3$ افزایش یابد نتایج بدست آمده از طریق مدل backoff نیز بهتر می شود؛ همچنین بیشتر بودن دقت مدل 4 نسبت به مدل 3 نیز می تواند درستی این ادعا را نشان دهد.

در نهایت با توجه به دقت های بدست آمده از جدول فوق نتیجه گرفته می شود که هر چه ضریب $L3$ بزرگ تر از $L2$ باشد و ضریب $L1$ و Eps (به عبارت حاصل ضرب این دو مقدار کمتر باشد) نیز کوچک تر شود؛ دقت مدل بیشتر می شود؛ یعنی مدلی که رابطه زیر بین پارامترهایش برقرار باشد می تواند به نتایج مطلوبی دست یابد:

$$Eps \times L1 \ll L2 < L3$$

که رابطه فوق نشان می دهد که هر چه ضریب $L1$ بزرگ تر از $L2$ باشد، و همچنین حاصل ضرب $Eps \times L1$ نیز عدد کوچک تری شود و بسیار کوچک تر از $L2$ باشد؛ دقت مدل بیشتر می شود.