

به نام خدا

گزارش پروژه چهارم کاربرد های هوش مصنوعی

سید مهدی میرفندرسکی 9723093

ستایش ثانوی 9628024

•تاثیر حذف کلمات پرتکرار و کم تکرار در دقت بدست آمده

با حذف کلمات پرتکرار و کم تکرار ممکن است نتوانیم به بعضی از کلمات احتمال درستی بدهیم و به عبارتی احتمال false positive یا false negative را زیاد میکند ولی از طرفی چون کلمات پرتکرار و کم تکرار ارزش زیادی در تشخیص زبان ندارند حذف آنها امن است و تاثیر بد زیادی روی دقت نخواهد داشت.

•تاثیر مقدار λ و ϵ دقت بدست آمده

با توجه به فرمول با تغییر لاندا ها وزن احتمال ها تغییر میکند و اگر بخواهیم تاثیر احتمال جفت کلمات را بیشتر کنیم میتوانیم لاندا مربوط به آن را کاهش دهیم. هرچه مقدار کل لاندا کمتر باشد دقت تشخیص بیشتر میشود.

در مورد اپسیلون هم برای از بین بردن احتمال صفر یک کلمه لازم است تا اگر کلمه ای در داده آموزشی نبود در داده تست آن را صفر نگیرد و این مورد دقت را افزایش میدهد و هرچه مقدار آن کمتر باشد دقت بیشتر می شود.

•بهترین دقت دستیافته و تحلیل تاثیر پارامترها در آن

بهترین دقت دست یافته زمانی است که وزن احتمال جفت کلمات کمتر باشد و مقدار اپسیلون هم برای جلوگیری از احتمال صفر داشته باشیم با مقدار هرچه کمتر.