

"به نام یزدان پاک"

گزارش کار پروژه چهارم هوش

نام استاد:

استاد روشن فکر

اعضای گروه:

کیانا آقا کثیری 9831006

سارا تاجرنیا 9831016

تاریخ آزمایش: 1400/4/27

گزارش

گزارشی شامل موارد زیر تهیه کرده و در فایل تحویلی اضافه کنید:

- تاثیر حذف کلمات پرتکرار و کم تکرار در دقت بدست آمده
- تاثیر مقدار λ و ϵ دقت بدست آمده
- بهترین دقت دست یافته و تحلیل تاثیر پارامترها در آن

1. تاثیر

کلمات پر

تکرار و کم

تکرار در

تعیین نتیجه نهایی بسیار مهم است به طوری که مثلا تکرار کلماتی چون to, the, I, is, am, you, ... که هیچ یک کلماتی با بار منفی یا مثبت نیستند و تقریبا در همه ی جملات پیدا میشوند پس تعداد تکرار بالایی دارند. این تکرار سبب میشود در فرمول ما احتمالات به دست آمده به طور چشم گیری بالا روند در حالی که این کلمات در بار معنایی جمله عملا بی تاثیرند.

2. مقدار لاندا و اپسیلون به این دلیل مورد تاملند که به وجود لاندا در فرمول برای سنگین کردن بار مثبت یا منفی دو کلمه پشت سر هم نسبت به کلمه ای که به صورت تنها آمده است سبب بیشتر شدن احتمال برای رسیدن به جواب درست است.

وجود اپسیلون هم به خاطر این است که ممکن است کلمه در دیکشنری وجود نداشته باشد.

3. رسیدن به جواب با روش بایگرم دقیق تر است که در این روش باید تعادلی میان لاندا ها به وجود آورد به طوری که جمع آنها برابر ۱ شود و به هریک به اندازه ای که باید ارزش داده شود طوری که :

$$Landa\ 3 > Landa\ 2 > Landa\ 1$$

$$Landa\ 3 + Landa\ 2 + Landa\ 1 = 1$$

اپسیلون هم باید عدد نسبتا بسیار کوچکی باشد تا تاثیر بسیار زیادی روی نتیجه نهایی نتواند بدارد (بین 0 و 1).