

حمیدرضا همتی

9631079

پروژه 3

نتایج به ازای مقادیر γ و ϵ متفاوت:

Lamda3 = 0.7

Lamda2 = 0.2

Lamda1 = 0.1

Epsilon = variable

```
E:\educational\universities curses and slides\artificial intelligence\project\NLP\venv\Scripts\python.exe  
lamnda3: 0.7 lamnda2: 0.2 lamnda1: 0.1 epsilon: 1e-05  
percentage: 86.11918604651163 %
```

```
"E:\educational\universities curses and slides\artificial intelligence\project\NLP\venv\Scripts\python.exe"  
lamnda3: 0.7 lamnda2: 0.2 lamnda1: 0.1 epsilon: 0.01  
percentage: 82.44912790697676 %
```

```
E:\educational\universities curses and slides\artificial intelligence\project\NLP\venv\Scripts\python.exe  
lamnda3: 0.7 lamnda2: 0.2 lamnda1: 0.1 epsilon: 0.1  
percentage: 76.4171511627907 %
```

```
"E:\educational\universities curses and slides\artificial intelligence\project\NLP\venv\Scripts\python.exe"  
lamnda3: 0.7 lamnda2: 0.2 lamnda1: 0.1 epsilon: 1  
percentage: 70.60319767441861 %
```

Lamda3 = 0.5

Lamda2 = 0.25

Lamda1 = 0.25

Epsilon = variable

```
"E:\educational\universities courses and slides\artificial intelligence\project\NLP\venv\Scripts  
lamnda3: 0.5 lamnda2: 0.25 lamnda1: 0.25 epsilon: 1e-06  
percentage: 85.13808139534885 %
```

```
"E:\educational\universities courses and slides\artificial intelligence\project\NLP\venv\Scripts\python  
lamnda3: 0.5 lamnda2: 0.25 lamnda1: 0.25 epsilon: 1e-05  
percentage: 86.04651162790698 %
```

```
lamnda3: 0.5 lamnda2: 0.25 lamnda1: 0.25 epsilon: 0.0001  
percentage: 86.0828488372093 %
```

```
"E:\educational\universities courses and slides\artificial intelligence\project\NLP\venv\scri  
lamnda3: 0.5 lamnda2: 0.25 lamnda1: 0.25 epsilon: 0.1  
percentage: 72.92877906976744 %
```

تأثیر گذاری اپسیلون:

اگر فقط تاثیرات bigram و unigram را در نظر بگیریم در شرایطی که مثلاً در test کلمه جدیدی ظاهر شود که در مجموعه train نباشد آنگاه احتمال unigram و bigram صفر میشود.

اپسیلون برای این به فرمول زیر اضافه شده است که در این موارد به صفر نرسیم.

اگر این مقدار خیلی بزرگ باشد تاثیر bigram و unigram ها کم رنگ تر میشود و دقت را کم میکند

و اگر هم خیلی کم باشد باز هم خوب نیست. مقدار مناسب برای این ضریب طبق محاسبات مختلف 10 به توان منفی 5 یا 6 است.

$$\hat{P}(c_i|c_{i-1}) = \lambda_3 P(c_i|c_{i-1}) + \lambda_2 P(c_i) + \lambda_1 \epsilon$$

تأثیر تغییرات λ ها :

مقدار احتمالاتی $P(c_i | c_{i-1})$ مقدار احتمالاتی bigram است. این مقدار، احتمال وجود یک ترکیبی از کلمه ها است.

مقدار احتمالاتی $P(c_i)$ مقدار احتمالاتی unigram است که مشخص میکند یک کلمه خاص با چه احتمالی وجود دارد

خب قاعداتاً احتمال حضور ترکیبی از کلمه ها تأثیر گذار تر است تا احتمال حضور یک کلمه خاص در جمله.

با توجه به این توضیحات بهتر است که λ_3 که ضریب $P(c_i | c_{i-1})$ است از λ_2 که ضریب $P(c_i)$ است بیشتر باش تا دقت افزایش یابد

برای بهینه سازی و پیدا کردن بهترین حالت این ضرایب λ در این مدل مسائل از شبکه عصبی استفاده میکنند.