

به نام خدا

گزارشکار پروژه ۴ مبانی هوش

مهدی مرادی - محمد متقی

به ازای λ و ε های مختلف نتایج متفاوتی حاصل می شود که در ادامه چند مورد از آن های آورده شده است :

حالت اول:

$$\lambda_3 = 0.9$$

$$\lambda_2 = 0.09$$

$$\lambda_1 = 0.01$$

$$\varepsilon = 0.0001$$

Accuracy: 0.7954971857410882

Accuracy: 0.7636022514071295

حالت دوم:

$$\lambda_3 = 0.1$$

$$\lambda_2 = 0.2$$

$$\lambda_1 = 0.7$$

$$\varepsilon = 0.0001$$

Accuracy: 0.7504690431519699

Accuracy: 0.7392120075046904

حالت سوم:

$$\lambda_3 = 0.15$$

$$\lambda_2 = 0.7$$

$$\lambda_1 = 0.15$$

$$\varepsilon = 0.0001$$

Accuracy: 0.7448405253283302

Accuracy: 0.7936210131332082

حالت چهارم:

$$\lambda_3 = 0.6$$

$$\lambda_2 = 0.2$$

$$\lambda_1 = 0.2$$

$$\varepsilon = 0.05$$

Accuracy: 0.8180112570356473

Accuracy: 0.3789868667917448

حالت پنجم:

$$\lambda_3 = 0.1$$

$$\lambda_2 = 0.15$$

$$\lambda_1 = 0.75$$

$$\varepsilon = 0.05$$

Accuracy: 0.8611632270168855

Accuracy: 0.28330206378986866

حالت ششم:

$$\lambda_3 = 0.15$$

$$\lambda_2 = 0.7$$

$$\lambda_1 = 0.15$$

$$\varepsilon = 0.05$$

Accuracy: 0.8198874296435272

Accuracy: 0.39587242026266417

نتیجه:

با بررسی آزمون انجام شده در فایل (resualts.txt) ملاحظه می شود که هرچی e و l_1 بیش تر باشه به یک سمت بایاس میشود یا مثبت ها رو خیلی خوب و بادقت میابد یا منفی ها رو خیلی بد یا برعکس .

همچنین با بررسی آزمون انجام شده در فایل (results_relaxed.txt) در می یابیم که 12 باید بسیار بزرگتر از 13 باشد.

با حذف کلمات پرتکرار مشخصاً دقت پایین می آید (یا ممکن است حروف ربط باشند و تغییری چندانی نکنند) چون داریم داده حذف می کنیم و در نتیجه تعداد داده ورودی برای یادگیری کم می شود.

با حذف کلمات کم تکرار مشخصاً دقت بسیار پایین می آید ، زیرا هم داده های ورودی برای یادگیری را کم کردیم و هم برخی از این کلمات کم تکرار تنها در یک قسمت مثبت یا منفی قرار دارند .