****

**دانشگاه تهران**

**دانشکده فنی-مهندسی کامپیوتر**

**دپارتمان الگوریتم ها و محاسبات**

**گزارش تمرین شماره ی سه**

**طراحی الگوریتم انتشار بازگشتی**

**نیلوفر آقایی ابیانه**

**810890001**

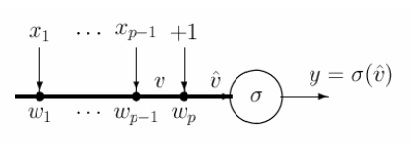
***چکیده***

در این پروژه ، از روش های ***یادگیری***[[1]](#footnote-1) برای طبقه بندی مجموعه ای از داده ها استفاده شده و نتایج حاصل تحلیل می شوند. در اینجا سه الگوریتم گیت OR، گیت XOR، و الگوریتم انتشار بازگشتی شبیه سازی شده اند. برای این کار از روش یادگیری ***پرسپترن چند لایه ای***[[2]](#footnote-2) استفاده شده است. در الگوریتم های گیت OR، گیت XOR از ***تابع فعال سازی***[[3]](#footnote-3) ***پله واحد*** و در الگوریتم ***انتشار بازگشتی***[[4]](#footnote-4) از ***تابع فعال سازی*** sigmoid استفاده شده است.

1. ***مقدمه***

در علم کامپیوتر، الگوریتم های مختلفی برای طبقه بندی وجود دارد، یکی ازاین الگوریتم ها روش های یادگیری است. در این روش با استفاده از مجموعه ای داده که به آن داده ی آموزش[[5]](#footnote-5) گفته می شود، ماشین، آموزش داده می شود؛ سپس با کمک ماشین به دست آمده هر نمونه تست بررسی شده و طبقه بندی می شود.

در روش یادگیری ابتدا بردار ورودی که مجموعه ویژگی ها ی یک نمونه آموزش است، در مجموعه ای از وزن ها ضرب می شود. سپس حاصل ضرب ها با هم جمع شده و به عنوان ورودی تابع فعال سازی داده می شوند (شکل 1). در آخر یک خروجی به دست می آید که با خروجی مطلوب (آخرین ویژگی نمونه آموزش که در طبقه بندی معادل با کلاس نمونه مورد نظر است)ممکن است اختلاف (خطا) داشته باشد. در این صورت این اختلاف برای به روز کردن وزن ها به کار می رود؛ این روند برای نمونه آموزش بعدی مجددا تکرار می شود. این مشی تا زمانیکه ماشین خطا نداشته یا خطا به یک حد ثابتی برسد ،ادامه می یابد.



شکل 1- ضرب نقطه ای برداز ورودی در بردار وزن ها

که بردار ورودی و بردار وزن ها می باشند.

***2.شبیه سازی گیت OR***

در اینجا نشان داده می شود که چگونه می توان از گیت OR ، برای طبقه بندی استفاده کرد.تابع فعال سازی در این شبیه سازی تابع پله واحد است و خطا از فرمول زیر محاصبه می شود؛

که در آن η، نرخ آموزش[[6]](#footnote-6)، w(n) بردار وزن ها ، e(n) خطا، و x(n) بردار ورودی است.

مجموعه آموزش ، نرخ آموزش ، وزن های اولیه به صورت زیر است :

Trainingset = [0 0 0 , 0 1 1, 1 0 1, 1 1 1]

LR = 0.2

W = [0.7 , -0.3 , 0.5]

که این ماشین 15 بار اجرا می شود و پس از این تعداد تکرار ماشین آموزش دیده و خطا ی آن صفر می شود. وزن های بهینه به دست آمده در جدول 1 آمده است.

|  |  |  |  |
| --- | --- | --- | --- |
| 0.1000- | 0.1000 | 0.7000 | وزن های بهینه |

جدول 1- وزن های بهینه برای OR\_GATE با LR=0.2 و وزن های اولیه 0.7، -0.3 و 0.5

در واقع در صورتی که بردار وزن اولیه به صورت W=[0.7 , 0.1, -0.1 ] باشد، با ورودی های trainingset=[ 0 0 , 0 1 , 1 0 , 1 1] خروجی [0 , 1, 1 ,1] حاصل می شود.

***3.شبیه سازی گیت XOR***

در اینجا نشان داده می شود که چگونه نمی توان از گیت XOR ، برای طبقه بندی استفاده کرد.

تابع فعال سازی در این شبیه سازی تابع پله واحد است و خطا از فرمول زیر محاصبه می شود؛

که η، نرخ آموزش، w(n) بردار وزن ها ، e(n) خطا، و x(n) بردار ورودی است.

مجموعه آموزش ، نرخ آموزش وزن های اولیه به صورت زیر است :

Trainingset = [0 0 0 ; 0 1 1; 1 0 1; 1 1 0]

LR = 0.2

W = [0.7 ; -0.3 ; 0.5]

که این ماشین 15 بار اجرا می شود و پس از این تعداد تکرار ماشین آموزش دیده اما خطا ی آن صفر نمی شود. وزن های به دست آمده در جدول 2 آمده است.

|  |  |  |  |
| --- | --- | --- | --- |
| 0.1000 | -0.3000 | -.1000 | وزن های به دست آمده |

جدول 2- وزن های به دست آمده برای XOR\_GATE با LR=0.2 و وزن های اولیه 0.7، -0.3 و 0.5

در واقع در صورتی که بردار وزن اولیه به صورت W=[-0.1 , -0.3, 0.1 ] باشد، با ورودی های trainingset=[ 0 0 , 0 1 , 1 0 , 1 1] خروجی [1 , 1, 1 ,1] حاصل می شود.

در واقع از پرسپترن تنها برای طبقه بندی های خطی می توان استفاده کرد؛ و در شرایط غیر خطی می توان از پرسپترن چند لایه ای استفاده کرد.

***4.پیدا سازی الگوریتم انتشار بازگشتی با استفاده از پرسپترن چند لایه ای***

در اینجا الگوریتم انتشار بازگشتی با استفاده از پرسپترن چند لایه ای و تابع فعال سازی sigmoid پیاده سازی می شود. این الگوریتم طبقه بندی را روی مجموعه داده ای[[7]](#footnote-7) iris و satimage انجام می دهد؛ که iris مجموعه داده ای با ا بوده و satimage مجموعه داده ای با است. پس در iris و satimage بردارهای ورودی به ترتیب به صورت و می باشند.

در اینجا از یک پرسپترن سه لایه ای استفاده می شود که تعداد عناصر لایه نخست، n1، برابر با تعداد عناصر ویژگی ، تعداد عناصر لایه دوم ، n2، مشخص نبوده، و تعداد عناصر لایه سوم، n3، برابر با تعداد کلاس ها می باشد.

در محیط MATLAB با استفاده از تابع n2finding مقدار مناسب n2 ، به دست می آید. این تابع ابتدا مقدار n2 ، را برابر 2، قرار می دهد. در ادامه تا دو برابر تعداد ویژگی ها در هر مرحله مقدار n2 یک واحد افزایش داده و به ازای آن n2 را آموزش می دهد. در آخر از بین n2 های به دست آمده مقداری که بیشترین کارایی را دارد به عنوان بهترین n2 در نظر می گیرد و باقی الگوریتم را با آن n2 اجرا می کند. البته مقدار n2 به نرخ آموزش نیز بستگی دارد.

پس از اینکه مقدار n2 مشخص شد، الگوریتم 25% کل داده را به عنوان تست در نظر گرفته و باقی داده را به چهار قسمت مساوی تقسیم کرده و در چهار مرحله ، هر بر یک قسمت را به عنوان داده ی اعتبار سنجی[[8]](#footnote-8) و سه قسمت دیگر را به عنوان داده های آموزش برای آموزش ماشین استفاده می کند؛ که به این روش Multifold cross-validation گفته می شود

الگوریتم، با استفاده از داده ها ی آموزش ماشین و تابع MLP ماشین را آموزش می دهد؛ سپس از 25% داده ی باقی مانده برای تست ماشین استفاده می کند واین روند تا زمانیکه اختلاف خطا دو نسل[[9]](#footnote-9) متوالی کمتر از 0.0001 نشود ادامه پیدا می کند؛ در نهایت کارایی ماشین با استفاده از تابع evaluation محاصبه می شود.

این الگوریتم برای به روز کردن وزن ها از فرمول زیر استفاده می کند

که مشتق تابع sigmoid ، η، نرخ آموزش، w(n) بردار وزن ها و e خطا است.

نکته ای که باید به آن توجه کرد این است که الگوریتم در ابتدا وزن ها به صورت تصادفی در بازه ی 0-1 انتخاب می کند و داده های هر مجموعه داده ای را نرمال سازی می کند.

برای اجرای این الگوریتم در محیط MATLAB از توابع زیر استفاده شده است:

* MAIN : تابع اصلی؛ نوع مجموعه داده ای و مقدار نرخ آموزش را از کاربر می گیرد
* N2finding : محاصبه ی مقدار n2 مناسب؛از مقدار n2 ، برابر 2،شروع کرده در ادامه تا دو برابر تعداد ویژگی ها، در هر مرحله مقدار n2 یک واحد افزایش داده و به ازای آن n2 را آموزش می دهد. در آخر از بین n2 های به دست آمده مقداری که بیشترین کارایی را دارد به عنوان بهترین n2 در نظر می گیرد .
* Insertionsort\_onlabels : مرتب سازی آرایه ورودی بر اساس کلاس ها
* Getclassesandfeatures : تعیین مقدار n1 و n3
* Initiaweight : مقدار دهی اولیه وزن ها
* MLP : تابع اصلی برای یادگیری ماشین به ازای هر نمونه train؛
* Normalize : نرمال سازی عناصر یک مجموعه داده
* Evaluation : تابع بررسی خروجی ماشین(کارایی)
* Selectingtestandtrain : تعیین مجموعه داده ای های train و تست
* Earlystoppingmethod : داده ها به دو دسته ی آموزش و اعتبار سنجی تقسیم می کند و پس از آموزش ماشین با استفاده از داده های آموزش ، از داده های اعتبار سنجی برای ارزیابی کارایی ماشین استفاده می کند.

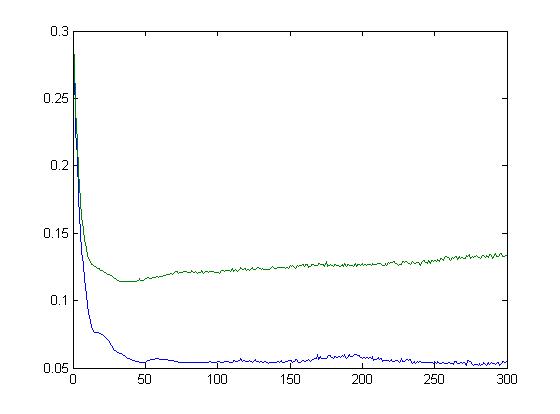
***5.آزمایش ها***

با اجرای تابع MAIN روی مجموعه داده ای satimage با نرخ آموزش 0.2 و n2 مناسب جدول 3 حاصل می شود. . ابتدا پس از جدا کردن داده های تست از آموزش، داده های باقی مانده برای آموزش به چهار قسمت تقسیم می شوند و هربار سه قسمت برای آموزش و قسمت دیگر برای داد ه ی اعتبار سنجی به کار گرفته می شوند.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 25 | 167 | 105 | 157 | نسل |
| 66.6943 | 64.9254 | 70.7539 | 78.6070 | **کارایی** |

جدول 3 – اجرای الگوریتم MLP روی مجموعه داده ای satimage و LR = 0.2 با انتخاب n2 برابر با 5 به عنوان n2 مناسب

با توجه به جدول 3 میانگین و واریانس به ترتیب برابر است با، 70.2451 و 37.0292 .شکل 2 نمودار حاصل از تابع earlystoppingmethod را نشان می دهد. شکل 3 نمودار های حاصل از اجرای چهار بار تابع MLP است.



شکل2- نمودار حاصل از تابع earlysttopingmethod با LR=0.2 و n2=5 5 به عنوان n2 مناسب

با اجرای تابع MAIN روی مجموعه داده ای satimage با نرخ آموزش 0.3 و n2 مناسب جدول 4 حاصل می شود. . ابتدا پس از جدا کردن داده های تست از آموزش، داده های باقی مانده برای آموزش به چهار قسمت تقسیم می شوند و هربار سه قسمت برای آموزش و قسمت دیگر برای داد ه ی اعتبار سنجی به کار گرفته می شوند.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 257 | 441 | 292 | 165 | نسل |
| 77.2162 | 73.5489 | 77.7962 | 74.8756 | **کارایی** |

جدول 4 – اجرای الگوریتم MLP روی مجموعه داده ای satimage و LR = 0.3 با انتخاب n2 برابر با 4 به عنوان n2 مناسب

با توجه به جدول 4 میانگین و واریانس به ترتیب برابر است با، 75.8592 و 3.9661 . شکل 4 نمودار های حاصل از اجرای چهار بار تابع MLP است.

|  |  |
| --- | --- |
|  |  |
|  |  |

شکل 3-خطا های حاصل از چهار بار اجرا ی اجرای الگوریتم MLP روی مجموعه داده ای satimage و LR = 0.2 با انتخاب n2 برابر با 5 به عنوان n2 مناسب

|  |  |
| --- | --- |
|  |  |
|  |  |

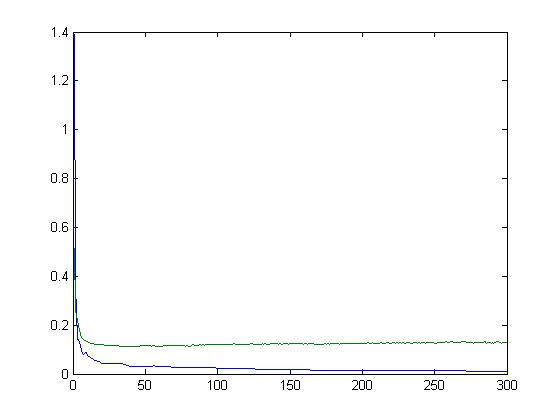
شکل 4-خطا های حاصل از چهار بار اجرا ی اجرای الگوریتم MLP روی مجموعه داده ای satimage و LR = 0.3 با انتخاب n2 برابر با 4 به عنوان n2 مناسب

با اجرای تابع MAIN روی مجموعه داده ای satimage با نرخ آموزش 0.4 و n2 مناسب جدول 5 حاصل می شود. . ابتدا پس از جدا کردن داده های تست از آموزش، داده های باقی مانده برای آموزش به چهار قسمت تقسیم می شوند و هربار سه قسمت برای آموزش و قسمت دیگر برای داد ه ی اعتبار سنجی به کار گرفته می شوند.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 78 | 14 | 28 | 37 | نسل |
| 51.5512 | 45.0249 | 43.5791 | 44.0299 | **کارایی** |

جدول 5 – اجرای الگوریتم MLP روی مجموعه داده ای satimage و LR = 0.4 با انتخاب n2 برابر با 24 به عنوان n2 مناسب

با توجه به جدول 5 میانگین و واریانس به ترتیب برابر است با، 51.5512 و 215.8607 .شکل5 نمودار حاصل از تابع earlystoppingmethod را نشان می دهد. شکل 6 نمودار های حاصل از اجرای چهار بار تابع MLP است.



شکل5- نمودار حاصل از تابع earlysttopingmethod با LR=04 و n2=42 به عنوان n2 مناسب

با اجرای تابع MAIN روی مجموعه داده ای iris با نرخ آموزش 0.2 نتایج در جدول 6 حاصل شد. ابتدا پس از جدا کردن داده های تست از آموزش، داده های باقی مانده برای آموزش به چهار قسمت تقسیم می شوند و هربار سه قسمت برای آموزش و قسمت دیگر برای داد ه ی اعتبار سنجی به کار گرفته می شوند.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 315 | 389 | 544 | 549 | نسل |
| 96.4286 | 92.8571 | 100 | 100 | **کارایی** |

جدول 6 – اجرای الگوریتم MLP روی مجموعه داده ای iris و LR = 0.2 با انتخاب n2 برابر با 3 به عنوان n2 مناسب.

|  |  |
| --- | --- |
|  |  |
|  |  |

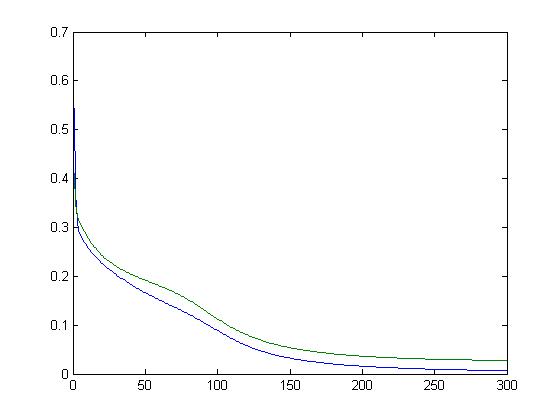
شکل 6-خطا های حاصل از چهار بار اجرا ی اجرای الگوریتم MLP روی مجموعه داده ای satimage و LR = 0.3 با انتخاب n2 برابر با 42 به عنوان n2 مناسب

با توجه به جدول 6 میانگین و واریانس به ترتیب برابر است با، 97.3214 و 11.6922 .شکل 7 نمودار حاصل از تابع earlystoppingmethod را نشان می دهد. شکل 8 نمودار های حاصل از اجرای چهار بار تابع MLP است.

با اجرای تابع MAIN روی مجموعه داده ای iris با نرخ آموزش 0.3 نتایج در جدول 7 حاصل شد. ابتدا پس از جدا کردن داده های تست از آموزش، داده های باقی مانده برای آموزش به چهار قسمت تقسیم می شوند و هربار سه قسمت برای آموزش و قسمت دیگر برای داد ه ی اعتبار سنجی به کار گرفته می شوند.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 428 | 536 | 656 | 655 | نسل |
| 92.8571 | 96.4286 | 85.7143 | 96.4286 | **کارایی** |

جدول 7 – اجرای الگوریتم MLP روی مجموعه داده ای iris و LR = 0.3 با انتخاب n2 برابر با 8 به عنوان n2 مناسب

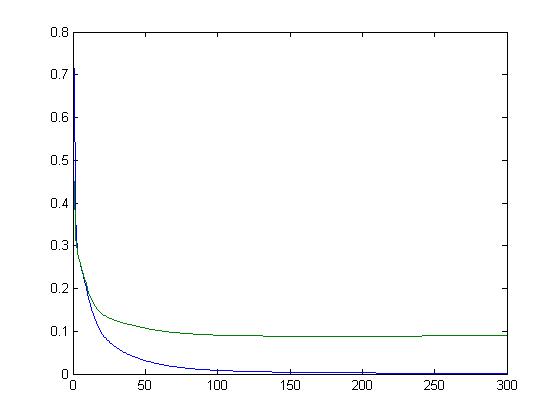


شکل7- نمودار حاصل از تابع earlysttopingmethod با LR=0.2 و n2=3 به عنوان n2 مناسب

|  |  |
| --- | --- |
| C:\Users\pishi\Desktop\1-3.jpg | C:\Users\pishi\Desktop\1-2.jpg |
| C:\Users\pishi\Desktop\1-5.jpg | C:\Users\pishi\Desktop\1-4.jpg |

شکل 8-خطا های حاصل از چهار بار اجرا ی اجرای الگوریتم MLP روی مجموعه داده ای iris و LR = 0.2 با انتخاب n2 برابر با 3 به عنوان n2 مناسب

با توجه به جدول 7 میانگین و واریانس به ترتیب برابر است با، 92.8571و 25.5102 . شکل 10 نمودار حاصل از تابع earlystoppingmethod را نشان می دهد. شکل 11 نمودار های حاصل از اجرای چهار بار تابع MLP است.



شکل9- نمودار حاصل از تابع earlysttopingmethod با LR=0.3 و n2=8 به عنوان n2 مناسب

|  |  |
| --- | --- |
|  |  |
|  |  |

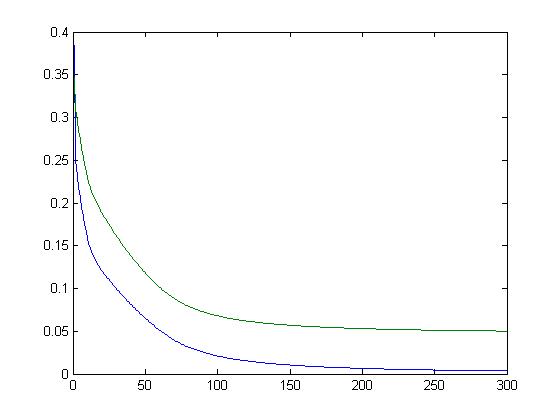
شکل 10-خطا های حاصل از چهار بار اجرا ی اجرای الگوریتم MLP روی مجموعه داده ای iris و LR = 0.3 با انتخاب n2 برابر با 8 به عنوان n2 مناسب

با اجرای تابع MAIN روی مجموعه داده ای iris با نرخ آموزش 0.4 نتایج در جدول 7 حاصل شد. ابتدا پس از جدا کردن داده های تست از آموزش، داده های باقی مانده برای آموزش به چهار قسمت تقسیم می شوند و هربار سه قسمت برای آموزش و قسمت دیگر برای داد ه ی اعتبار سنجی به کار گرفته می شوند.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 213 | 88 | 158 | 194 | نسل |
| 96.4286 | 100 | 89.2857 | 96.4286 | **کارایی** |

جدول 7 – اجرای الگوریتم MLP روی مجموعه داده ای iris و LR = 0.4 با انتخاب n2 برابر با 3 به عنوان n2 مناسب

با توجه به جدول 7 میانگین و واریانس به ترتیب برابر است با، 95.5357و 20.1956 . شکل 11 نمودار حاصل از تابع earlystoppingmethod را نشان می دهد. شکل 12 نمودار های حاصل از اجرای چهار بار تابع MLP است.



شکل11- نمودار حاصل از تابع earlysttopingmethod با LR=0.4 و n2=3 به عنوان n2 مناسب

***5.تحلیل نتایج***

به طور کلی بر اساس نتایج به دست آمده، الگوریتم ***انتشار بازگشتی*** برای طبقه بندی داده ها نسبت به الگوریتم ***های طبقه بند اولین نزدیک ترین همسایه***[[10]](#footnote-10)، ***طبقه*** ***بند بیز***[[11]](#footnote-11)، ***k امین نزدیک ترین همسایه[[12]](#footnote-12) و پنجره ی پارزن*** [[13]](#footnote-13)

بهتر عمل می کند.

در طی انجام آزمایش ها به این نتیجه رسیده می شود که چنانچه مقدار n2 و تعداد نسل ها مناسب انتخاب شوند این الگوریتم از کارایی بسیار بالایی برخوردار است.

|  |  |
| --- | --- |
|  |  |
|  |  |

شکل 12-خطا های حاصل از چهار بار اجرا ی اجرای الگوریتم MLP روی مجموعه داده ای iris و LR = 0.4 با انتخاب n2 برابر با 3 به عنوان n2 مناسب

1. Learning [↑](#footnote-ref-1)
2. Multi Level Perceptron [↑](#footnote-ref-2)
3. Activation Function [↑](#footnote-ref-3)
4. Back- propagation [↑](#footnote-ref-4)
5. Training Data [↑](#footnote-ref-5)
6. Learning Rate [↑](#footnote-ref-6)
7. dataset [↑](#footnote-ref-7)
8. Validation data [↑](#footnote-ref-8)
9. epoch [↑](#footnote-ref-9)
10. One-nearest Neighbor [↑](#footnote-ref-10)
11. Bayes Classifier [↑](#footnote-ref-11)
12. K-nearest Neighbor [↑](#footnote-ref-12)
13. Parzen Window [↑](#footnote-ref-13)