به نام خدا

تمرین سری دوم درس یادگیری عمیق

دکتر محمدی

محمد یارمقدم

96462104

گزارش سوال سوم:

الف) معرفی مجموعه داده iris

این مجموعه داده مختص گل های زنبق است. این مجموعه داده شامل 150 عضو است که این 150 عضو در 3 لیبل به طور مساوی تقسیم شده اند، به این معنی که هر لیبل 50 عضو دارد. لیبل های این مجموعه داده عبارت اند از:

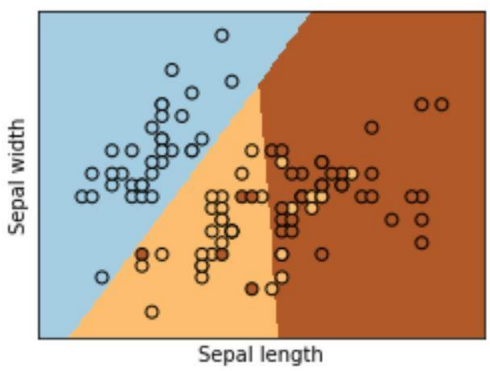
- Setosa
- Virginica
- Versicolor

در هر نمونه چهار feature اندازه گیری شده است. طول و عرض کاسبرگ (sepal) و طول و عرض گلبرگ (petal) چهار ویژگی موجود برای هر نمونه است.

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
count	150.000000	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.057333	3.758000	1.199333	1.000000
std	0.828066	0.435866	1.765298	0.762238	0.819232
min	4.300000	2.000000	1.000000	0.100000	0.000000
25%	5.100000	2.800000	1.600000	0.300000	0.000000
50%	5.800000	3.000000	4.350000	1.300000	1.000000
75%	6.400000	3.300000	5.100000	1.800000	2.000000
max	7.900000	4.400000	6.900000	2.500000	2.000000

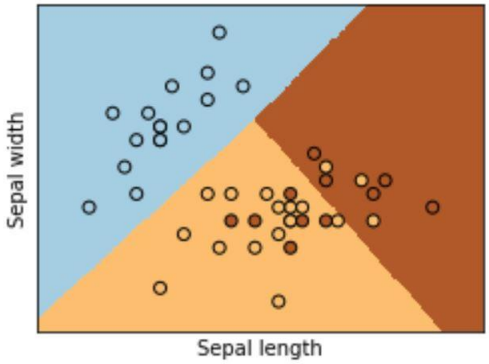
ب)

برای این قسمت از کتابخانه matplot استفاده می کنیم. در ابتدا از ماژول "LogisticRegression" در کتابخانه sklearn یک نمونه می سازیم. سپس داده های آزمون را که در قسمت های بالایی جدا کردیم تا به تابع fit پاس می دهیم. سپس برای تعیین محدوده نمودار بزرگ ترین مقادیر x و y را می یابیم و مش را طبق آن تشکیل می دهیم. سپس نقاط مورد نظر را با تابع scatter به نمودار اضافه می کنیم. در آخر محور ها را نامگذاری می کنیم و با استفاده از تابع show نمودار را نمایش می دهیم.



ج)

در این قسمت نیز همانند قسمت قبل عمل کردیم. با این تفاوت که به جای داده های آموزش (X_train) از داده های تست (X_test) برای ورودی تابع fit استفاده کردیم.



د)

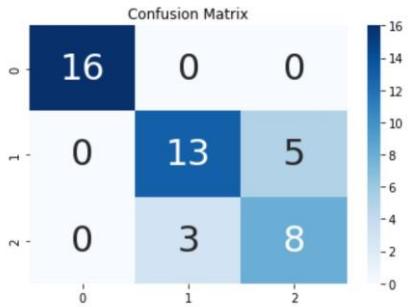
در این قسمت تابع predict به طور جداگانه بر روی داده های تست و آموزش صدا زده شد و از درصد شباهت داده های پیش بینی شده با داده های موردانتظار دقت شبکه به دست آمد.

```
test accuracy = 80.0
train accuracy = 81.9047619047619
Confusion Matrix:
[[16  0  0]
 [ 0 15  3]
 [ 0  6  5]]
```

از آنجا که درصد داده های آموزش برای train شبکه 30 درصد بوده و مابقی داده ها تست هستند، درصد دقت داده ها در قسمت تست به علت فراوانی در داده های تست نزدیک به داده های آموزش است. البته با تغییر random_state دقت تغییر خواهد کرد. اما طبق پیش بینی دقت تست از آموزش کمتر است چون تست اطلاعات خود را از آموزش بدست آورده سات.

ه)

برای بدست آوردن Confusion matrix از ماژول metrics که در کتابخانه sklearn موجود است استفاده شد. ورودی های این ماتریس y_test و y_pred بود که در واقع داده های مورد انتظار و پیش بینی شده شبکه هستند. سپس برای رسم آن از seaborn استفاده شد.



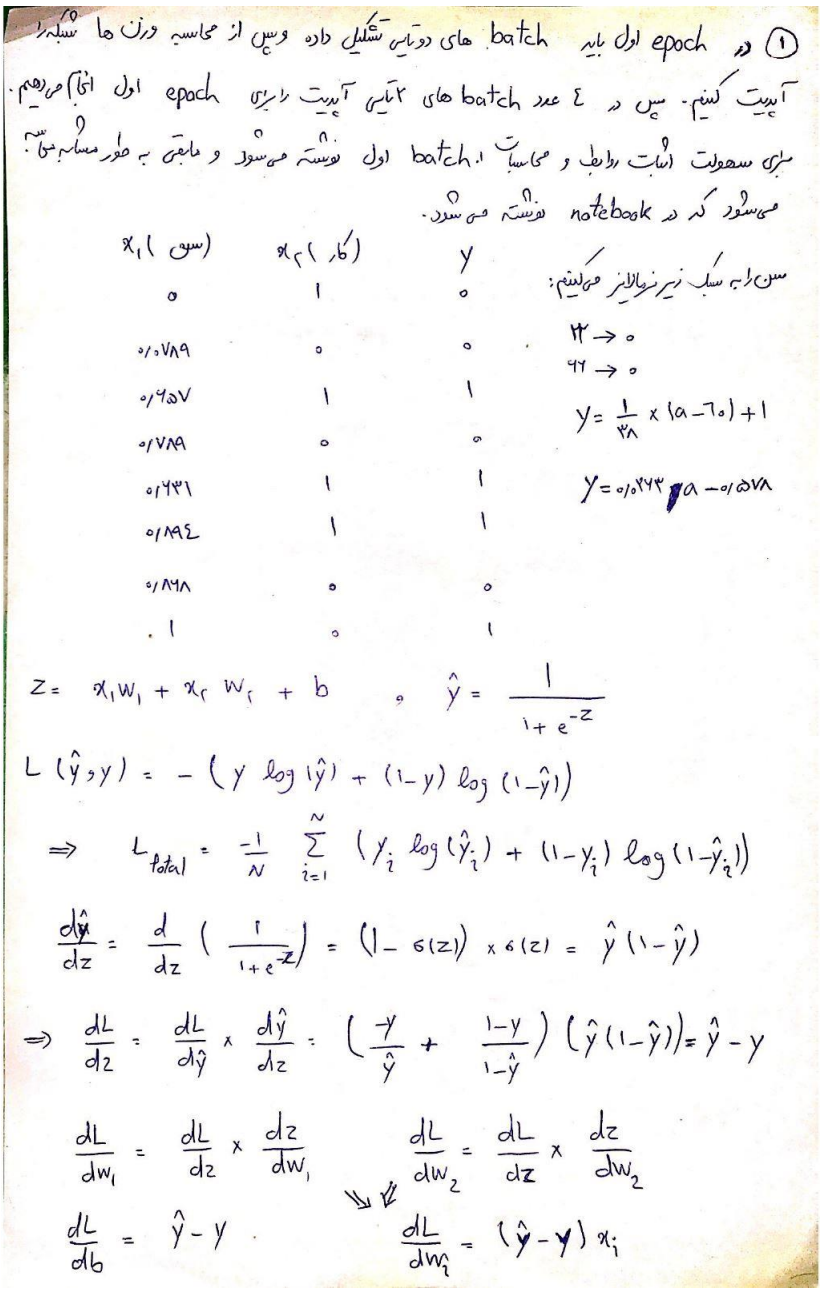
این ماتریس برای خلاصه سازی نتایج تست به کار میرود. ستون های این ماتریس "predicted class" هستند و سطر های آن "actual class" هستند. اعدادی رو روی قطر اصلی قرار دارند به معنی پیش بینی درست است، یعنی کلاس واقعی با پیش بینی درست است پس:

Correct predictions = 16 + 13 + 8 = 37

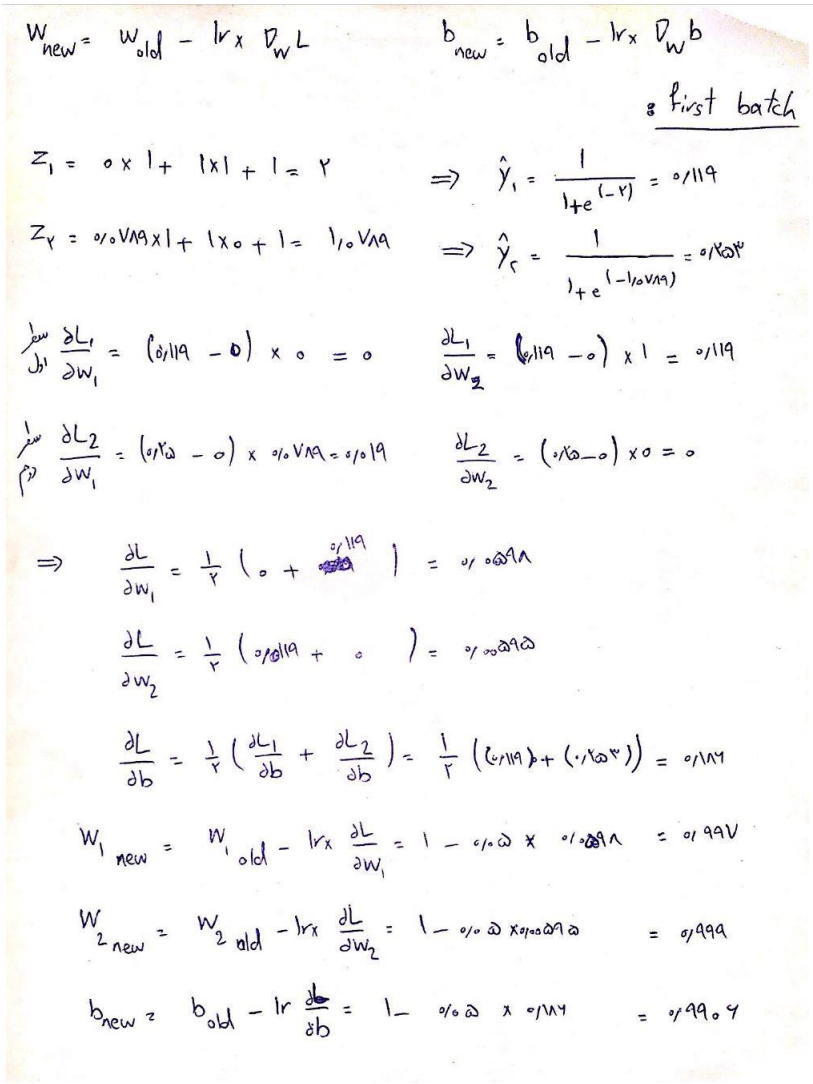
Wrong predictions = 3 + 5 = 8

سوال اول)

در این بخش اثبات ها و مراحل را برای batch اول در اولین epoch و مابقی مراحل را در notebook پیش می بریم.



ادامه سوال اول



می توان دریافت که cost حالت نوسانی دارد و ممکن است پس از هر آپدیت به سمت بهینه شدن حرکت نکند و بالا و پایین بشود.

اما روند به صورت کلی نزولی است و cost در نهایت به روند نزولی داشته و شبکه به سوی بهینه شدن حرکت می کند.

مورد ب)

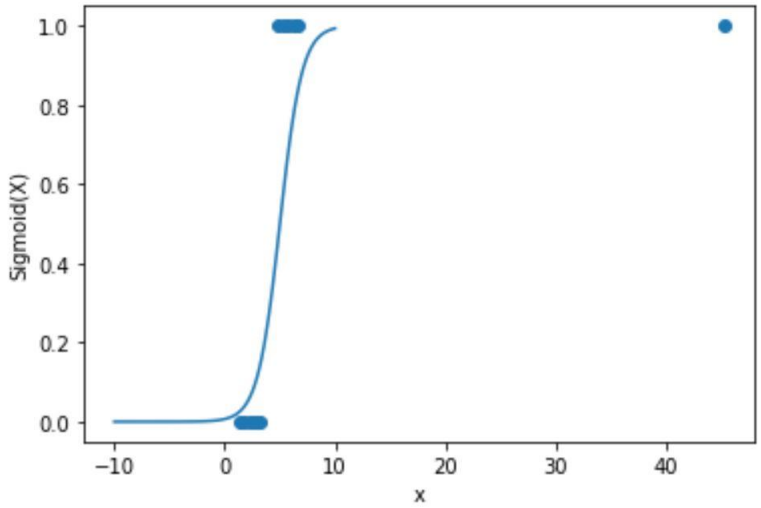
مشکلات الگوریتم گرادیان کاهشی عبارت است از : 1- طولانی بودن روند converge 2- امکان ماندن در مینیمم های محلی 3- انجام پراسس و محاسبات طولانی

اما مزیت این الگوریتم این است که فرآیند موازی سازی محاسبات به علت وجود بردارسازی در آن وجود دارد. اما در الگوریتم گرادیان کاهشی تصادفی این vectorizing وجود ندارد. از آنجا که توضیحات هم گفته شد هزینه در حالت تصادفی نوسانی بوده پس در مجموعه داده ها با

تعداد بالاتر بهینه تر است. از آنجا که در کل داده مشاهده نمی شود پس محاسبات سبک تر است و سریع تر همگرایی صورت میگیرد. در حالت تصادفی نیز ماندن در مینیمم های محلی کمتر صورت میگیرد چون در مشاهده قسمت بعدی از داده امکان رهایی بالاتر است.

سوال دوم)

شکل های مناطق جداسازی در نوت بوک به صورت کامل رسم شده است.

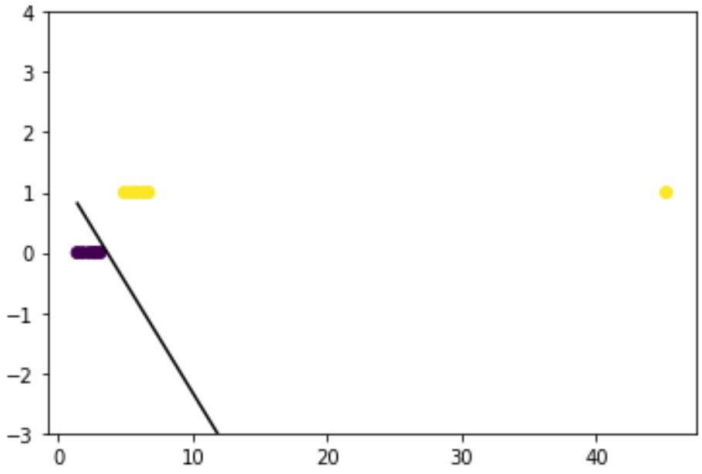


اما در اینجا برای بدست آوردن مرز جدایی بین دو منطقه از روش زیر استفاده می کنیم.

در حالت logistic طبق معادله sigmoid داریم:

$Z = 1 / (1+np.exp(-x)) = 0.5 \rightarrow z = 0 \rightarrow z = x.W + b \rightarrow x = -b/W \rightarrow x = 5.53$

در حالت linear نیز داریم :



$Z = 0.5 \rightarrow x = (0.5 - x) / W$

که از روی شکل مقدار $x = 3.4$ بدست آمد.