

بسمه تعالی

گزارش تمرین عملی سوم هوش مصنوعی

استاد: دکتر عبدی

محمد حسین جلالی 97101456

بخش اول:

در این بخش با تقلید از شکل های موجود در فایل تمرین! 3 نوع تابع را در قالب 3 فایل پایتون پیاده کردم.

توضیح کلی کد:

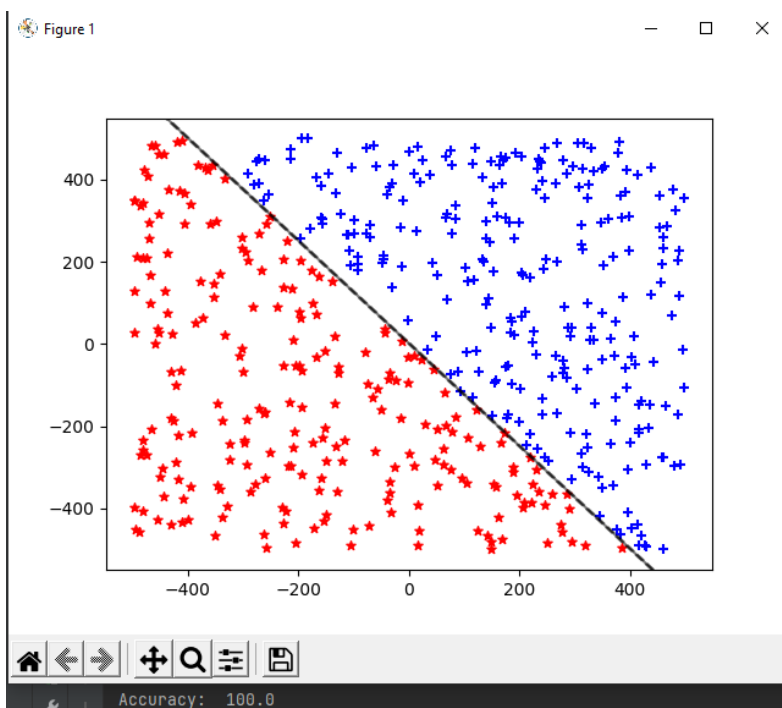
کد 3 فایل تقریبا مشابه است و صرفا در نوع هسته و تعریف تابع تفاوت وجود دارد:

در ابتدا باید داده آزمایشی را تولید کنیم. از تابع eval هم برای این کار استفاده شده است که حقیقتا موجود جذابیت. بعد از این که به صورت رندوم داده ها را تولید کرده و لیست داده ها و لیبل ان ها آماده شد نوبت به لرن کردن می رسد. با هسته خطی داده ها را لرن می کنیم.

بعد از این کار داده های تست را ساخته و رسم میکنیم. بعد از ان خط جدا کننده و همینطور خطوط مارجین را با استفاده از تابع پلاتی که زدیم رسم می کنیم. سپس دقت حاصل در داده های تست را حساب کرده و چاپ میکنیم.

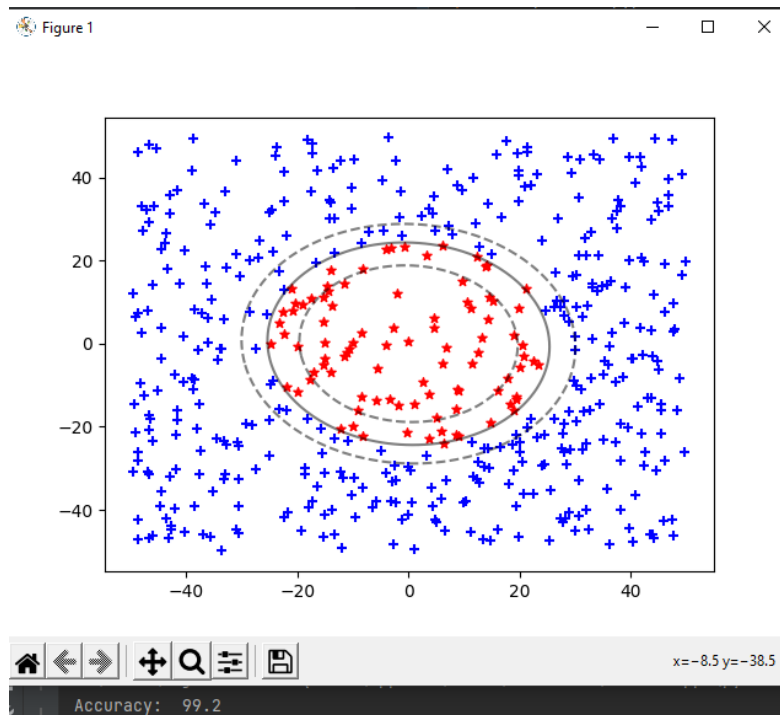
حال به بررسی 3 حالت می پردازیم:

حالتی که تابع خطی است به راحتی با هسته خطی به دقت بسیار بالایی می رسمیم:

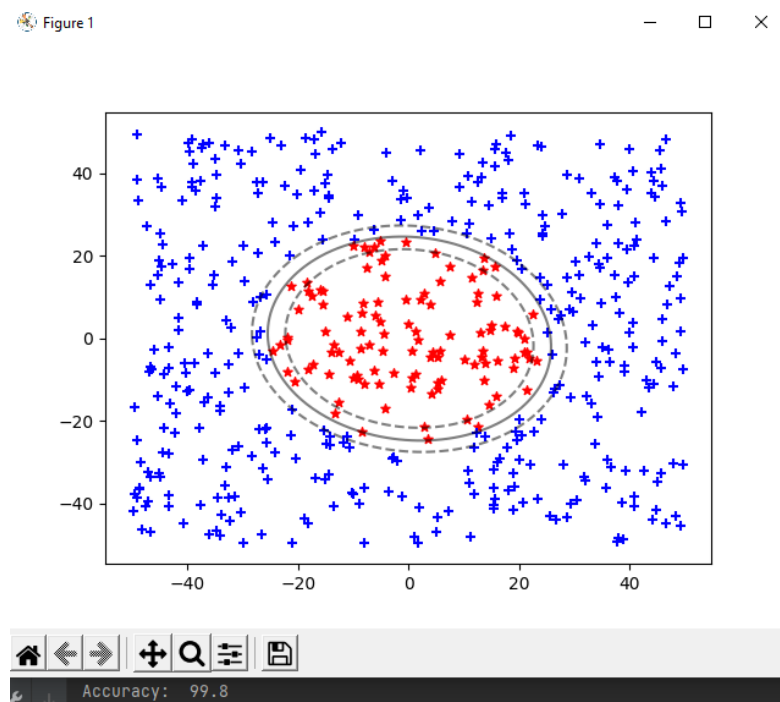


خطوط مارجین و خطوط اصلی تقریبا روی هم افتاده اند. در حالت هسته rbf دقت حدود 99 درصد می شود.

حالت دایره ای:

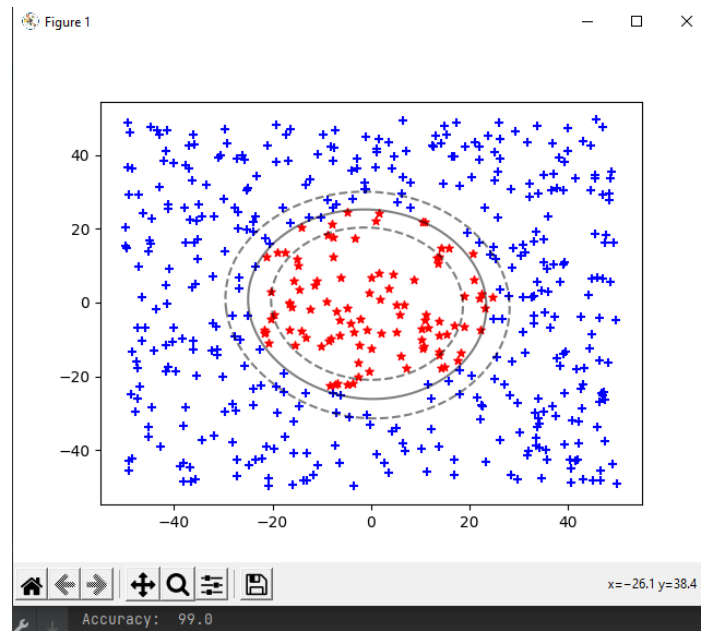


در حالت کرنل ما در حالت poly درجه 2 با ثابت 0 بود. حالت درجه 3 با ثابت 2:



در حالت درجه 5 دقت کمتر شد! حال بقیه هسته ها را نیز تست میکنیم:

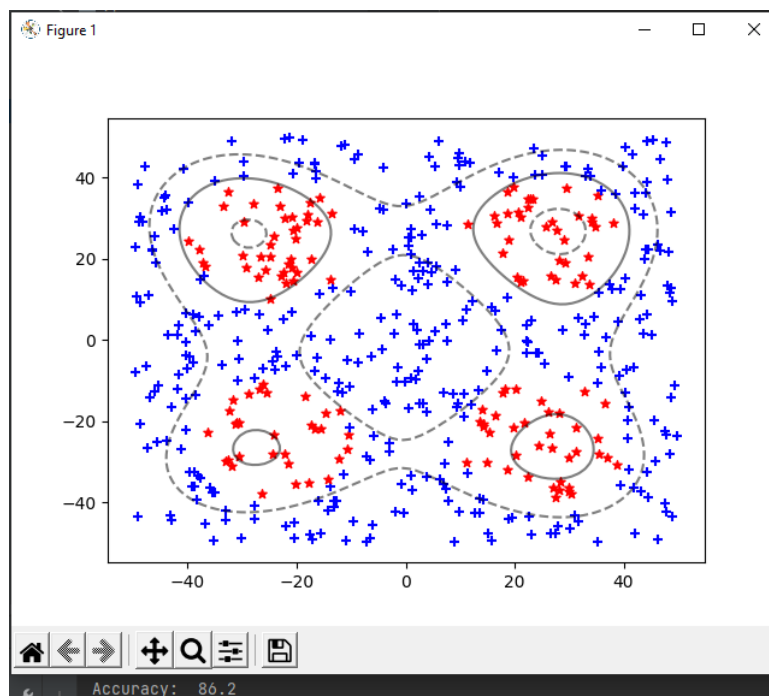
هسته rbf:



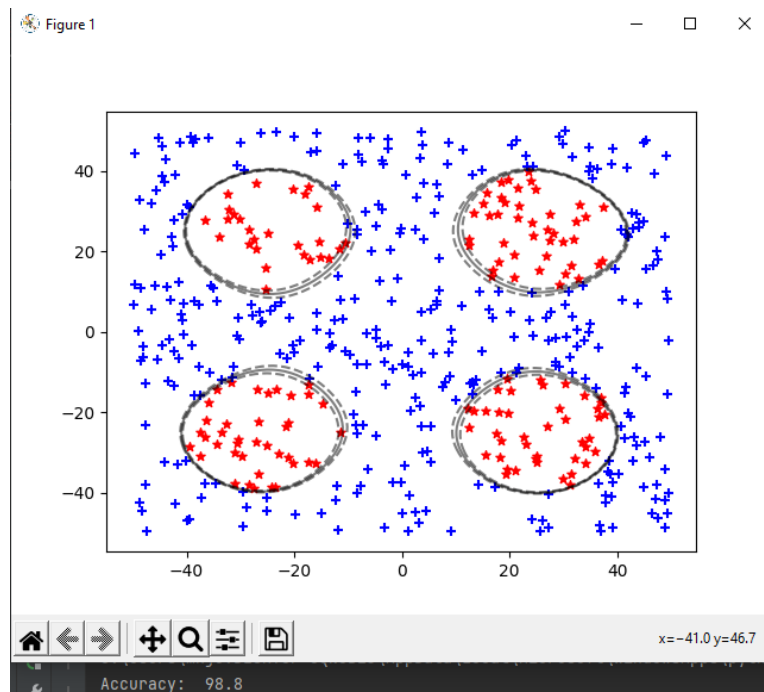
هسته سیگموئید هم عملکردی بسیار بد داشت و دقت آن در حدود 60 درصد بود! در کل ظاهراً هسته poly بهترین هسته ممکن برای این حالت است.

حالت 4 دایره:

هسته rbf:



دقت 86.2 با توجه به عجیب و غریب بودن تابع به نظر منطقی است. ولی هسته poly با درجه بالا عملکردی بسیار بهتر دارد:



با کمی آزمایش و خطا درجه 8 و ثابت 6 عملکرد بسیار خوبی داشت. (98.8%) هسته سیگموئید هم مثل همیشه عملکرد مناسبی نداشت.

در کل با توجه به نوع تابع هسته ها عملکردهای متفاوتی دارند. ولی مزیت هسته rbf عملکرد مناسب در همه حالات است که شایسته احترام است.

بخش دوم:

در این بخش از داده USPS که برای پروژه هم هست استفاده کردیم. ابتدا داده ها را مناسب استفاده کردیم و لیبل گذاری مناسب هم انجام دادیم. برای این کار طبق توصیه پاورقی عمل شده است. (تبدیل به بردار...) هسته ابتدایی که با آن تست را انجام دادیم حالت دیفالت یا همان rbf است که این دقت را به ما می دهد (تغییر دادن پارامترها تغییر خاصی ایجاد نکرد):

94.8181365221724

هسته سیگموئید عملکرد خوبی ندارد و اگر از هسته poly با درجه 10 و ثابت 2 استفاده کنیم:

93.97110114598904

درجه 5 با ثابت 3:

94.22022919780767

حدوداً عملکرد مشابهی هر دو هسته داشتند ولی با وجود آزمایش های مختلف هسته rbf عملکردی بهتر داشته است.

نکته: برای ران کردن کد باید عکس ها در دو پوشه train و test در پوشه ای که کد در آن هست قرار داده شود.

بخش سوم:

در این بخش ابتدا داده های تست و ترین را جدا کردیم. به این صورت که حدود 20 درصد داده ها را برای تست در نظر گرفته و داده های دیگر برای ترین استفاده شده است. در ادامه مثل بخش عکس ها برای استفاده آماده کردیم.

هسته rbf:

96.66666666666667

هسته poly با درجه 5 و ثابت 3:

97.0

حال نسبت داده های ترین و تست را تغییر میدهم:

اگر عکس ها را به دو دسته مساوی تقسیم کنیم دقت به حدود 94 درصد میرسد.

اگر برای کاهش مدت ترین داده های ترین حدود یک 10 ام داده های تست باشند دقت 90 درصد و اگر داده های تست 20 برابر داده آزمایشی باشد دقت به 85 درصد می رسد. جالب اینجاست که اگر تنها از یک صدم داده ها برای ترین استفاده کنیم دقت 67 درصد خواهیم داشت!

نکته: برای ران کردن کد باید پوشه های دسته ها همه در پوشه ای به نام data باشند.