

به نام خدا

تمرین سوم یادگیری ماشین آماری

محمد مهدی عبدالله پور - ۹۵۳۱۰۵۷

سوال اول)

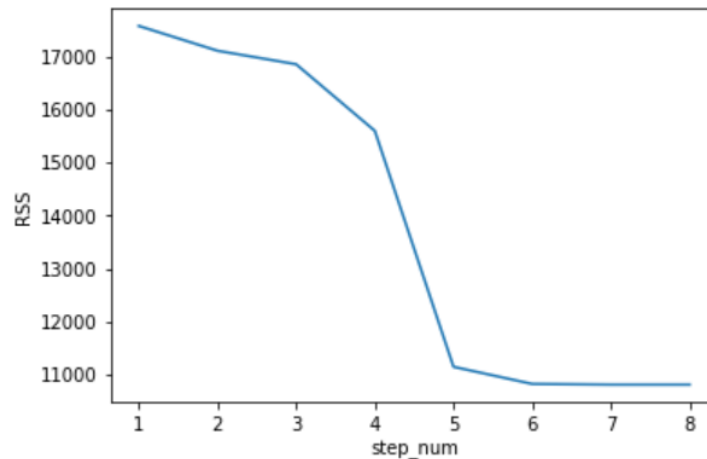
الف. (طبق نمودار های داخل ژوپیتر)

ویژگی اول و هفتم رابطه تقریبا مستقیم دارند. بقیه ویژگی ها نشان می دهند مقادیری از هدف فقط در بازه خاصی از آن ویژگی ممکن است.

ب. جزییات خواسته شده در ژوپیتر آمده.

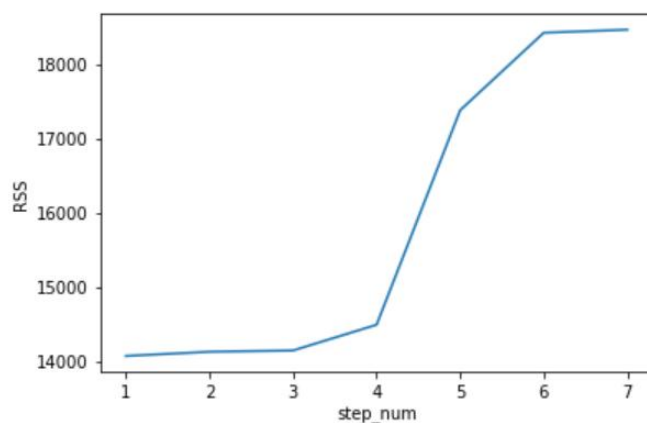
ج. ویژگی ۷، چون معیار  $RSS$  کمتری نسبت به بقیه دارد. در مقدار  $R^2$  هم بهتر از بقیه است.

د. موارد خواسته شده در ژوپیتر انجام شده است.

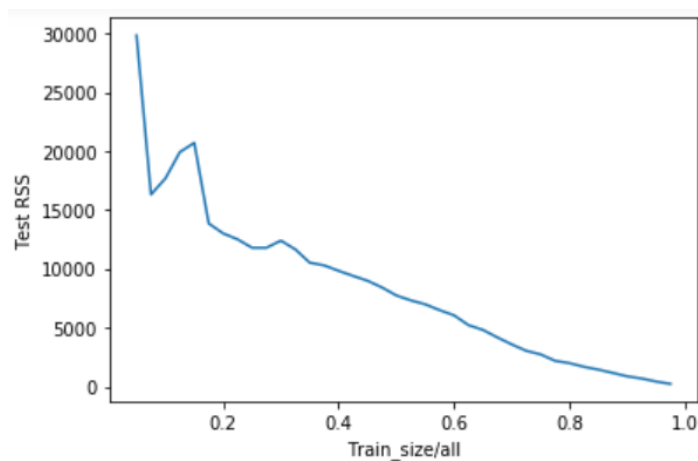


این معیار با افزایش ویژگی ها کاهش می یابد. بعد از قدم پنجم دیگر کاهش چشمگیری نداریم.

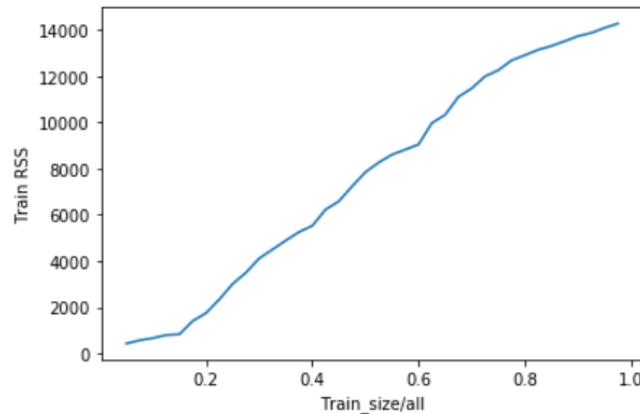
ه. موارد خواسته شده در ژوپیتر ثبت شده است.



با کاهش ویژگی‌ها، این معیار افزایش می‌یابد ولی تا قدم چهارم افزایش خاصی نداریم و تقریباً معامله خوبی داشته ایم.



با افزایش درصد داده آموزشی، معیار خطا روی تست کاهش یافته است چون داده تست کمتر داریم و RSS جمع روی تعداد آن‌هاست و تعداد کمتر مقدار کمتر را نتیجه می‌دهد + اینکه مدلی بهتری را با داده‌های بیشتر آموزش داده ایم که خطای کمتری در کل دارد.



با افزایش درصد داده آموزشی، معیار خطا روی آموزشی افزایش یافته است چون داده آموزشی بیشتر داریم و RSS جمع روی تعداد آن هاست و تعداد بیشتر مقدار بیشتر را نتیجه می دهد.

نتیجه گیری کلی: با استفاده از model selection، توانستیم مجموعه کوچکتري از ویژگی ها با کارایی مناسب برای پیش بینی بدست بیاوریم.

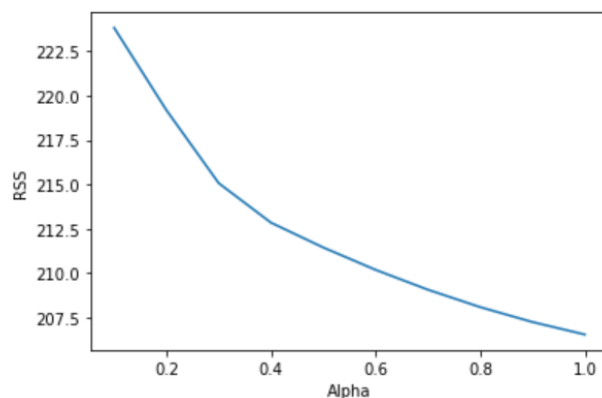
سوال دوم)

الف. (طبق نمودار های داخل ژوپیتتر)

رابطه ویژگی ها با هدف بصورت مستقیم است. طبیعی هم است چون اگر کسی در درسی نمره بالا گرفته احتمالا در درس های دیگر هم بالا می گیرد.

ب. مقادیر نامشخص را با اعداد تصادفی گرفته شده از توزیع گوسی تخمین شده از نمرات دیگر آن شخص پیدا می کنیم. حال وقتی نمودار رسم می کنیم توزیع ها خیلی بهتر است.

د.



نمودار در  $\lambda = 0.3$  زانو میزند، پس مقدار خوبی است.

و. با  $0.001$  ضرایب بزرگتر هستند از زمانی که  $\lambda$  را  $0.3$  در نظر گرفتیم.

نتیجه گیری کلی: استفاده از  $\lambda$  باعث می شود تا  $\lambda$  model selection و آموزش با هم انجام شود.

(سوال سوم)

الف) مقادیر  $\lambda$  را با استفاده از اعداد تصادفی تولید شده از توزیع گوسی تخمین زده شده از آن ویژگی پر میکنیم.

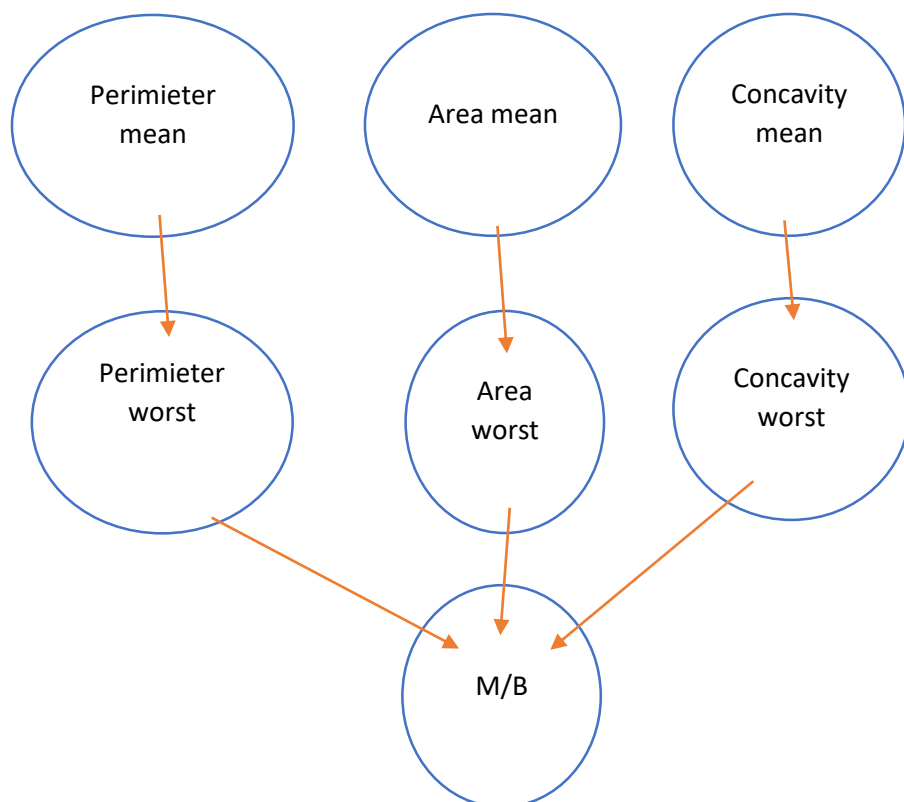
برای گسسته سازی، اگر ویژگی از میانگین بیشتر بود  $1$  و در غیر آن صورت صفر در نظر می گیریم.

ب. با توجه به نمودارها (که در فایل ژوپیتتر وجود دارد)، قدرت جدا کنندگی ویژگی های  $0$  و  $1$  و  $2$  و  $3$  و  $4$  و  $5$  از بقیه بیشتر است.

د. سه زیر مجموعه از ویژگی هایی که جدا کنندگی خوبی دارند را در نظر می گیریم. از  $\lambda$  و  $\lambda$  ها یکی از آن ها را در نظر می گیریم. نتیجه بهتر از حالت قبل می شود.

## Model 1

فرض می‌کنیم perimeter mean به perimeter worst وابسته است. برای area و concavity هم همینطور.



## Model 2

فرض می‌کنیم perimeter worst به perimeter mean وابسته است. برای area و concavity و texture هم همینطور.

و smoothness worst هم به radius mean وابسته است و compactness هم به area و perimeter وابسته است. (طبق فرمولش)

