



دانشگاه تهران

پردیس دانشکده های فنی

دانشکده مهندسی برق و کامپیوتر

یادگیری ماشین

تمرین پنجم

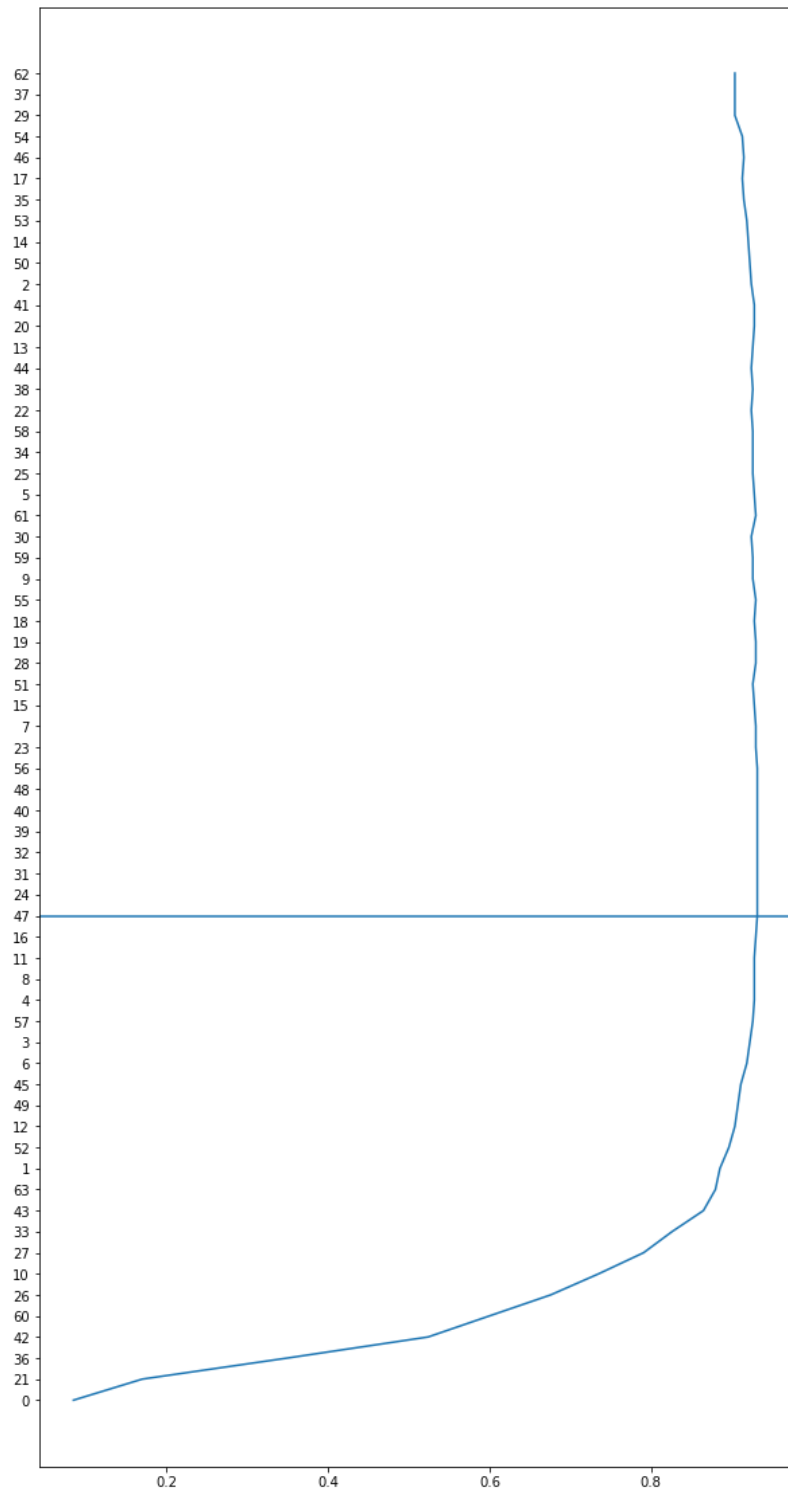
امیرحسین عباسکوهی

۸۱۰۱۹۷۵۳۹

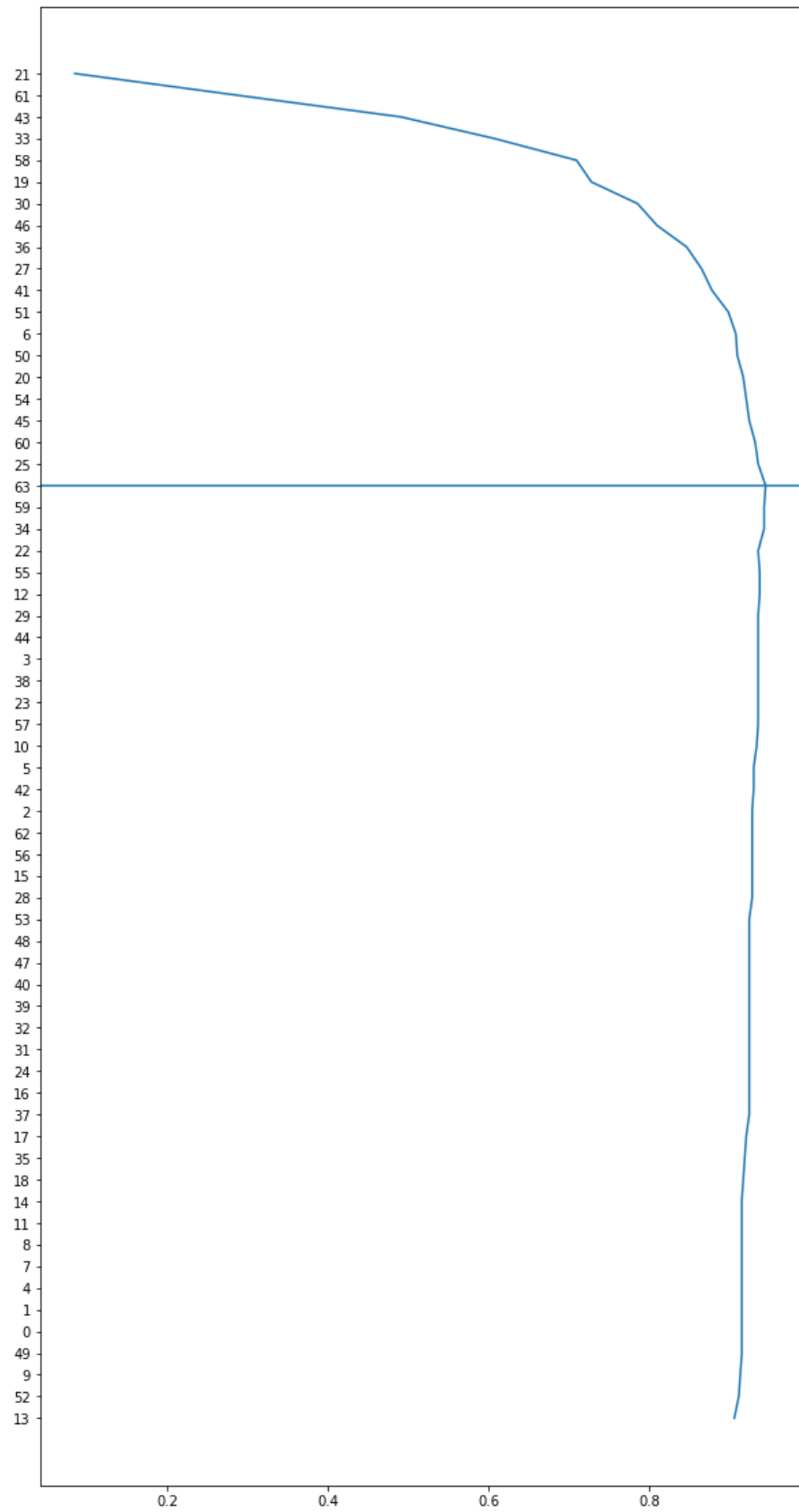
استاد ابولقاسمی

سوال ۶)

نمودار مربوط به forward selection را در زیر میبینیم:



در اینجا هم تصویر backward selection را میبینیم:

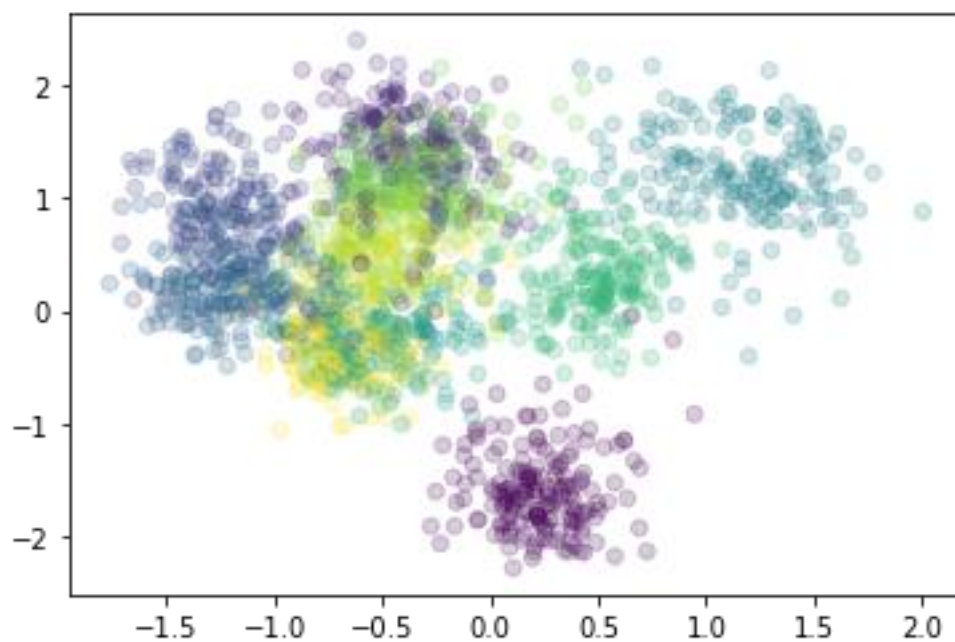


همانطور که قابل مشاهده است می توان گفت تقریباً forward selection نموداری برعکس backward selection دارد که این موضوع به علت ذات برعکس این دو روش است که در forward در هر مرحله یک فیچر اضافه میکنیم اما در backward در هر مرحله کم میکنیم. از آن جایی که دیتاستی که در این سوال با آن کار میکنیم، دیتاست تصویری است پس می توان گفت اکثر فیچرها یک نشان دهنده یک پیکسل هستند اهمیت دارد به همین دلیل میبینیم با افزایش تعداد فیچر دقت بالاتر می رود اما نکته ای که در این جا وجود دارد این است که پس از مدل دقت fluctuate میکند. از طرفی میبینیم که از ۶۴ فیچر موجود در کل بر اساس نمودارها حدود ۲۴ فیچر اهمیت دارد که این موضوع به این دلیل است که اعداد تمام پیکسلها را استفاده نکرده اند و فیچرها مهم تری وجود دارد. همانطور که قابل مشاهده است دو روش به دقت تقریباً مشابهی در بهترین حالت با تعداد متفاوت رسیده اند. در یکی ۲۴ فیچر انتخاب و در دیگری ۲۰ فیچر حذف شده اند که نشان دهنده تاثیر الگوریتم دو روش است.

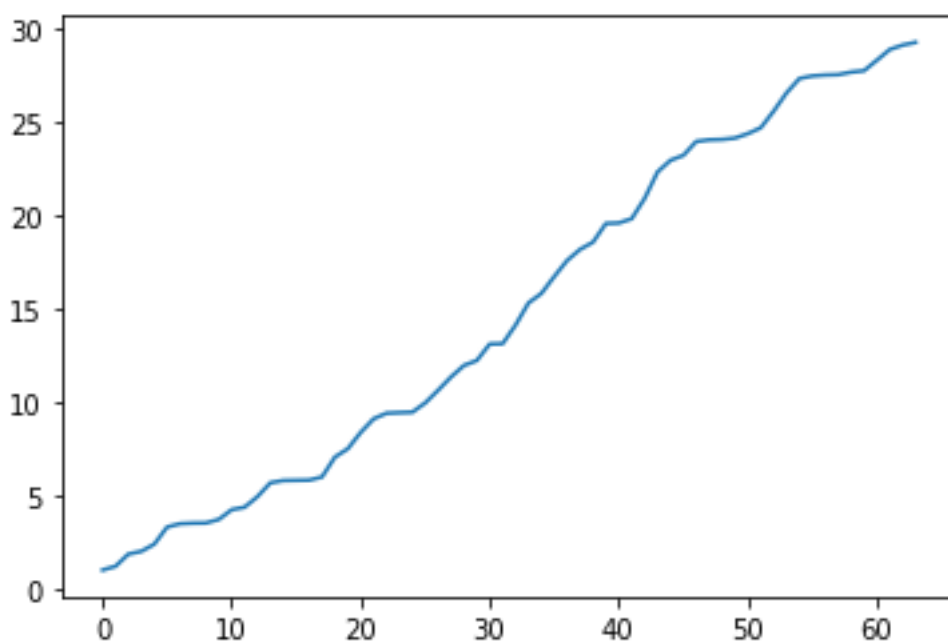
در کل روش backward بیشتر مورد پسند است زیرا روش forward تاثیر so-called suppressor تولید می کند. این suppressorها زمانی اتفاق می افتند که تخمین زننده ها و پیش بینی کننده ها زمانی عالی هستند که دیگری ثابت نگه داشته شوند.

سوال (۷)

با استفاده از LDA میتوانیم داده ها را در فضای دو بعدی رسم کنیم:



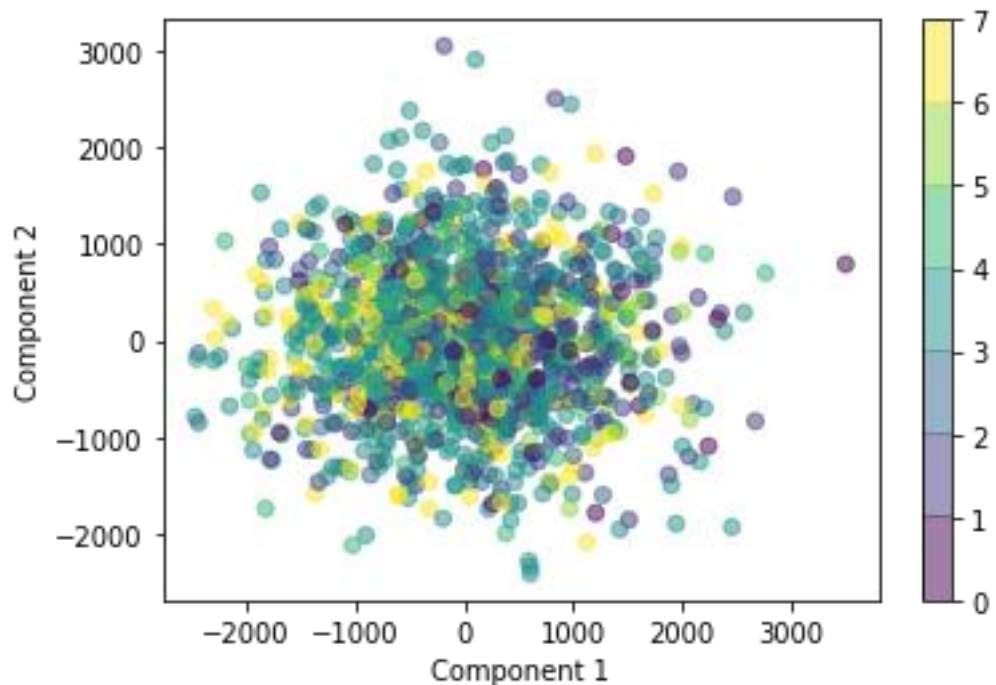
اما وقتی نمودار trace را رسم میکنیم با نمودار زیر مواجه میشویم:



در اینجا میبینیم که با افزایش تعداد فیچر ها به طور کلی trace افزایش پیدا میکند. در واقع این همان multidimensional fisher discriminant است. (با افزایش فیچر ها به طبع داده های بیشتر در کل داریم و به همین دلیل trace بیشتر می شود)

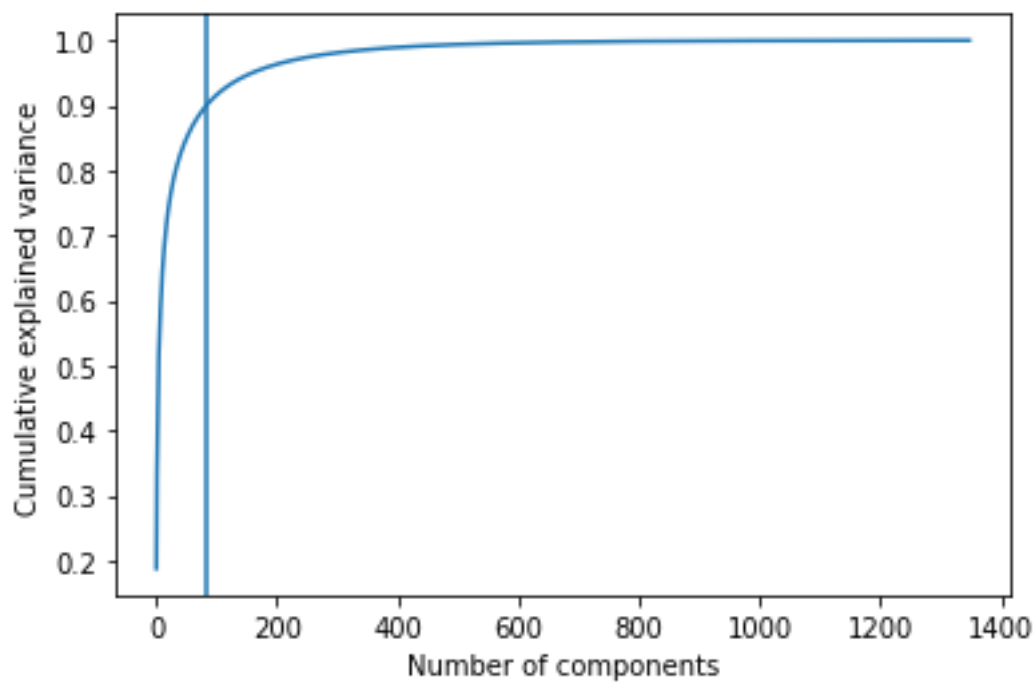
سوال ۹)

تصویر داده ها در فضای دو بعدی را در زیر میبینیم:



Principal component ها، متغیرهای جدیدی هستند که به صورت ترکیبات خطی یا مخلوط‌های اولیه ساخته می‌شوند. این ترکیبات به گونه‌ای انجام می‌شوند که متغیرهای جدید ناهمبسته هستند و بیشتر اطلاعات درون متغیرهای اولیه فشرده یا فشرده می‌شوند. وقتی به دو کامپوننت کاهش می‌دهیم یعنی دو کامپوننت برتر را انتخاب می‌کنیم.

نمودار cumulative explained variance را در زیر داریم:



همچنین تصاویر برای ۳۰ کامپوننت اول:



به نظر می‌رسد که تصاویر اولیه (از چپ بالا) با زاویه نور بر روی صورت ارتباط داشته و سپس بردارهای اصلی به نظر می‌رسد که ویژگی‌های خاصی مانند چشم‌ها، بینی و لب دارند. با توجه به نمودار cumulative explained variance متوجه می‌شویم برای تعداد کامپوننت انتخاب شده ما اطلاعات لازم را بازیابی نکرده ایم و در نتیجه تصاویر در این بخش‌ها اطلاعات کافی را ندارند.



برای تصاویر recover شده اگر تعداد component را تغییر متوجه میشویم که با افزایش این تعداد تصویر واضح تری داریم که باز هم به نمودار explained variance برمیگردد که برای مقدار ۱۲۰ که تقریباً همان نقطه واریانس بالای ۰.۹ به دست آمده، اطلاعات جامع تری ذخیره شده و به نتیجه مطلوب رسیده ایم.