



دانشگاه یزد

گروه مهندسی کامپیوتر

رشته تحصیلی: هوش مصنوعی و رباتیکز

نام درس: تشخیص الگو

گزارش تمرین عملی ۴

استاد مربوطه: دکتر مهدی یزدیان

تهیه کننده: حمیدرضا نادمی

پاییز ۱۳۹۷

- a) Develop PCA takes $X(D \times N)$ return $Y(d \times N)$ (d is the number of features selected by the PCA algorithm).

تعداد ویژگی‌ها (d) به عنوان ورودی از کاربر دریافت می‌شود و طبق الگوریتم PCA یک ماتریس $d * N$ محاسبه و برگردانده می‌شود.

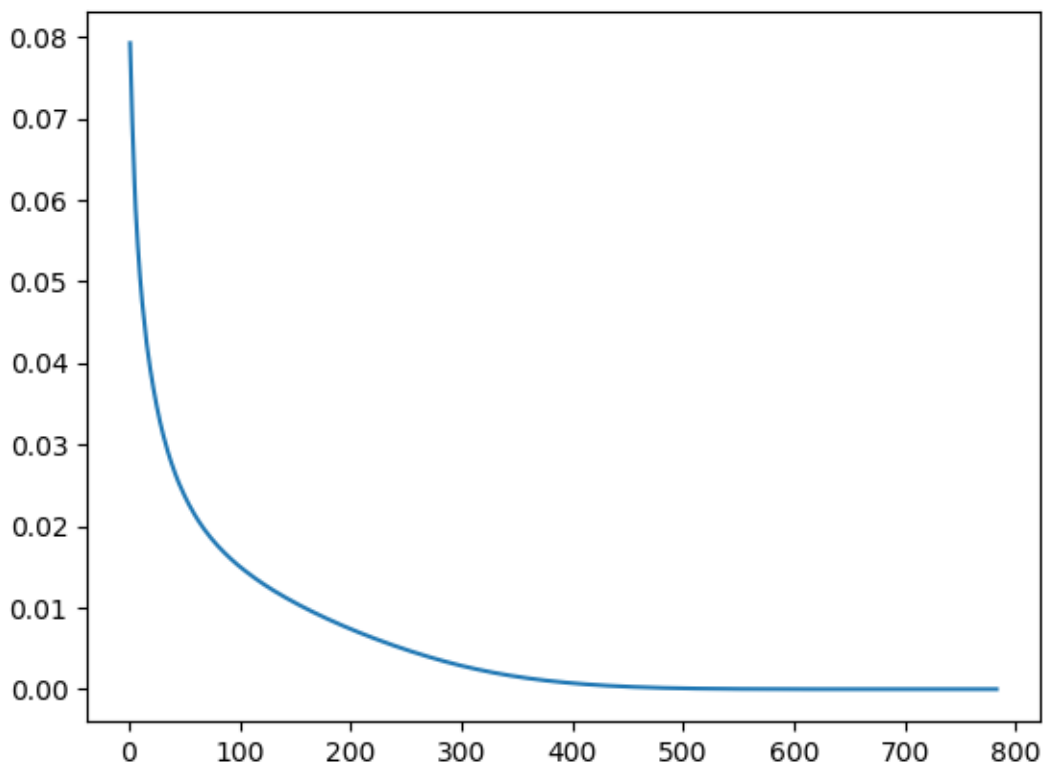
- b) Propose a suitable d using proportion of variance (POV) =95%.

طبق فرمول زیر:

$$POV = \frac{\sum_{i=1}^d \lambda_i}{\sum_{i=1}^D \lambda_i}$$

مقدار d بدست آمده **151** می‌باشد.

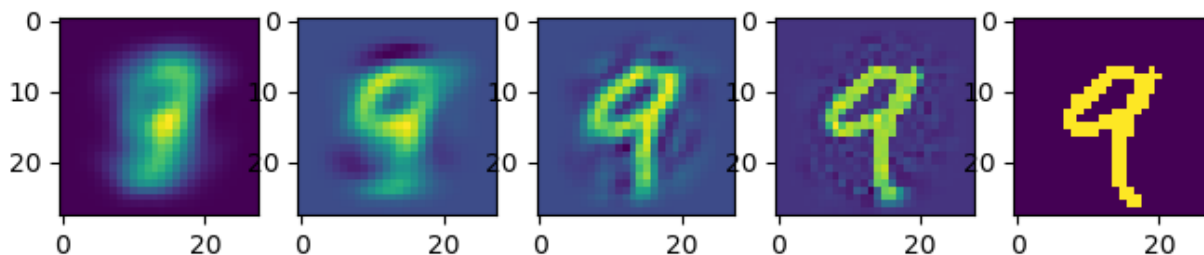
- c) Develop PCA reconstruction takes $Y_{PCA}(d \times N)$ and return $\hat{X} (D \times N)$.
For different values of $d = \{1, 2, 3, 4, \dots, 784\}$ reconstruct all samples and calculate the average mean square error (MSE). Plot d (x-axis) versus MSE (y-axis). Discuss about the results.



Plot d (x-axis) versus MSE (y-axis)

در ابتدا که مقدار d کم است تصویر دوباره ساخته شده (reconstruct) شده نسبت به تصویر واقعی خطای زیادی دارد زیرا از تمام feature های هر عکس فقط تعداد کمی از آنها را برداشتیم. هرچه مقدار d بیشتر می شود مقدار خطا کمتر می شود تصویر های تولید شده به تصویر اصلی نزدیکتر هستند.

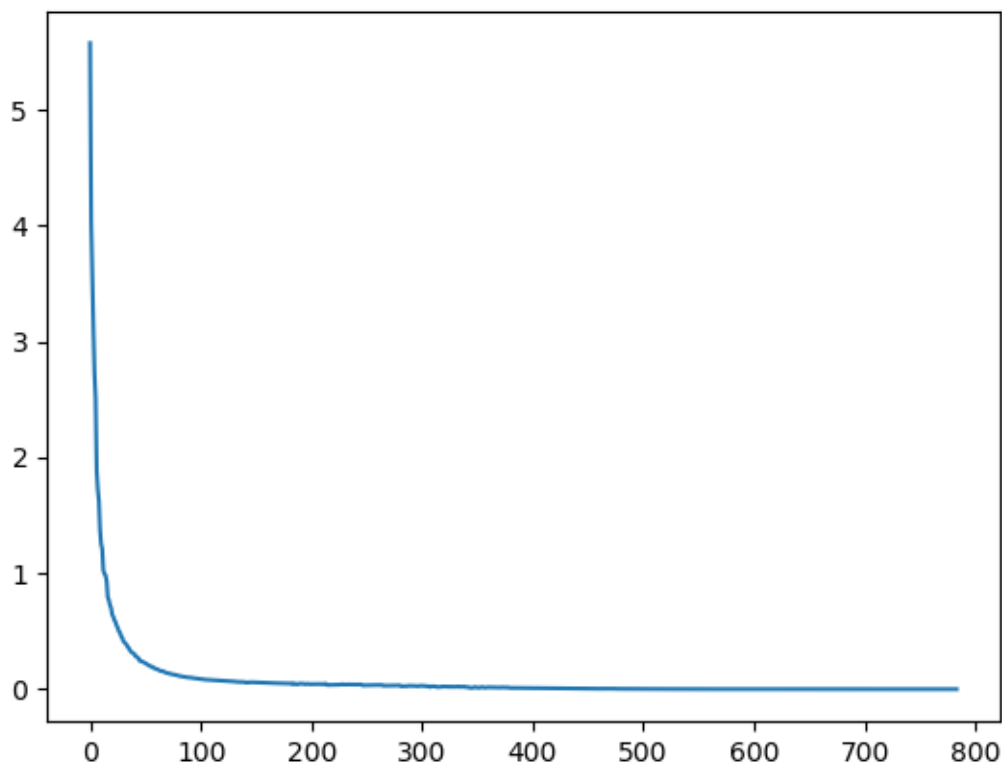
- d) Reconstruct 5th sample and show it as a 'png' image for $d = \{1, 10, 50, 250, 784\}$. Discuss about the results.



تصاویر reconstruct شده برای $d=1, 10, 50, 250, 784$ (به ترتیب از چپ به راست)

هرچه تعداد ویژگی (d) بیشتر می شود تصویر دقیقتر می شود، زیرا تعداد بیشتر از ویژگی ها را جهت ایجاد تصویر برداشته ایم.

- e) For different values of $d = \{1, 2, 3, 4, \dots, 784\}$ plot d (x-axis) versus eigenvalues (y-axis). Discuss about the results.



plot d (x-axis) versus eigenvalues (y-axis)

با توجه به نمودار مشاهده میکنیم که از یک d مشخص به بعد عدد مقدار ویژه کوچک و ناچیز می شود (یعنی از یک d به بعد هرچه تعداد ویژگی ها بیشتر شود روی دقت و کیفیت تصویر تاثیری ندارند و فقط باعث پیچیده تر شدن مدل می شوند)، با پیدا کردن این d و بردارهای ویژه متناظر با آنها و دور ریختن $D - d$ مقدار ویژه و بردار ویژه متناظر با آنها در ساده شدن مدل استفاده کرد.