

علت تفاوت دو نمودار ، تفاوت‌های دو روش PCA و tSNE است:

- 1) PCA یک projection خطی است و لذا وابستگی‌های غیرخطی داده‌ها را نمی‌تواند اعمال کند، در حالیکه tSNE غیر خطیست.
- 2) PCA به شدت توسط داده‌های پرت تحت تاثیر قرار می‌گیرد، برخلاف tSNE.
- 3) PCA تنها حالت کلی داده‌ها را حفظ می‌کند و به همسایه‌های هر نقطه آنچنان دقت نمی‌کند اما tSNE سعی در حفظ locality نیز دارد. در واقع برای کلاس‌بندی داده‌ها، روش tSNE بهتر عمل کرده و داده‌های با label یکسان را حتی اگر به صورت توده‌های دارای اشتراک باشند، بهتر دسته‌بندی می‌کند.
- 4) PCA یک روش قطعی بوده در حالیکه tSNE یک روش احتمالاتی است. همچنین روش tSNE دارای hyperparameter است.
- 5) PCA با چرخاندن بردار ویژه‌ها برای یافتن راستای با بیشینه واریانس به دست می‌آید ولی tSNE بر اساس کمینه کردن فاصله‌ی هر نقطه در داده‌ی اصلی و داده‌ی کاهش بعد یافته به دست می‌آید. در واقع در PCA ما می‌توانیم میزان فاصله مد نظر را تعیین کنیم اما در tSNE ما با استفاده از هایپرپارامترها میزان فاصله را تعیین می‌کنیم.
- 6) در مجموع PCA به خوبی tSNE عمل نمی‌کند.