

Principal Component Analysis (PCA)

- يعتبر أحد التقنيات (unsupervised techniques) التي تُستخدم لتقليل أبعاد البيانات (Dimensionality Reduction)
- يتم تطبيقها على Features وليس Target
- تقوم بتحويل Features إلى مجموعة متغيرات جديدة تسمى (Principal Components) PC.
- يتم ترتيب PC بناءً على مقدار variability بحيث يمثل العنصر الأول (most variance)، أما العنصر الأخير فهو يمثل (Least variance).
- على سبيل المثال: لو كان لدينا 10 Features سوف نقوم بترتيبهم بناءً على مقدار variability ثم اختيار جزء من هذه Features (أهم 3 Features).
- استخدامات PCA:
 - في Data Visualization
 - في Dimentionality Reduction لزيادة سرعة Model.
 - الافتراضات التي تطرحها PCA:
 - وجود علاقة من نوع Linear Relationship بين المتغيرات Features.
 - استخدام variance لتحديد أهمية العناصر بحيث كلما زاد variance رادت الأهمية
 - جميع Features لها نفس Unit، في بعض الأحيان نحتاج لعمل Scale للبيانات.

Mathematical PCA:

- <https://builtin.com/data-science/step-step-explanation-principal-component-analysis>