**PCA**

PCA, or Principal Component Analysis, is a technique used in data analysis to reduce the dimensionality of a large dataset while retaining the maximum amount of information. It accomplishes this by identifying the principal components, or the dimensions that explain the most variation in the data. PCA can be used for several purposes, such as data visualization, noise reduction, and feature extraction. It is often used as a preprocessing step before performing other data analysis techniques such as clustering or regression. In summary, PCA is a powerful tool for reducing the dimensionality of large datasets while retaining as much information as possible. It is a widely used technique in data analysis and can help uncover patterns and relationships in the data that would be difficult to detect otherwise.

**Kernel PCA**

Kernel PCA (KPCA) is a variant of the standard PCA that is used when the underlying data is **nonlinearly** related. Traditional PCA is based on linear projections, which may not be effective in capturing the underlying nonlinear structure of the data. In contrast, KPCA applies a nonlinear mapping to the data using a kernel function, which transforms the data into a **higher-dimensional space** where it may become linearly separable. KPCA then applies standard PCA to this transformed space to extract the principal components.

The kernel function used in KPCA is typically chosen based on the specific problem and the underlying data. Popular kernel functions include the radial basis function (RBF) kernel, polynomial kernel, and sigmoid kernel. We have used **RBF** kernel for all dataset as it performs better in most of cases. The RBF kernel is widely used and is defined as a function of the distance between pairs of data points. The kernel function allows KPCA to capture complex patterns and relationships that may not be apparent in the original data. It has several advantages over traditional PCA, but selecting the appropriate kernel function and parameters can be challenging.

The **gamma** parameter plays a crucial role in Kernel PCA, particularly when using the RBF kernel. Its value determines the degree of influence that each training example has on the classification of new points, and selecting the optimal value is often a challenging task that requires careful experimentation. We have performed a grid search over a range of gamma values and choose the value that results in the best performance. Optimal value of gamma was different on the each dataset.