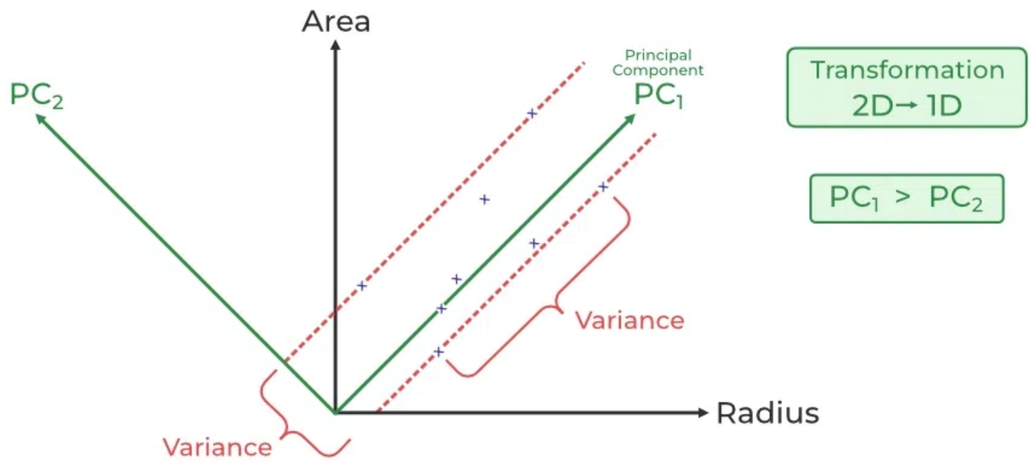
PCA

PCA (Principal Component Analysis) is a powerful dimensionality reduction technique used in data analysis and machine learning. It aims to transform a dataset consisting of high-dimensional data points into a lower-dimensional space while preserving the most important information.

Here's a breakdown of PCA and its different types:

Core Idea:

- Imagine a dataset with many features (dimensions). PCA finds a new set of features (called principal components) that capture the maximum variance in the data.
- These principal components are typically uncorrelated (linearly independent) and ordered by the amount of variance they explain.



Types of PCA:

While the core principle of PCA remains the same (finding directions of maximum variance), there are different approaches for performing PCA:

1. **Eigenvalue Decomposition (Classic PCA):**
    - This is the most widely used and theoretically well-understood method.
    - Steps:
        1. Center the data (subtract the mean value of each feature from each data point).
        2. Calculate the covariance matrix (captures the linear relationships between features).
        3. Find the eigenvalues and eigenvectors of the covariance matrix.
        - Eigenvalues represent the variance explained by each principal component.
        - Eigenvectors represent the directions (axes) of the principal components in the original feature space.
        4. Choose the top k eigenvectors (corresponding to the k largest eigenvalues) and project the data onto them for dimensionality reduction.

2. **Singular Value Decomposition (SVD):**
    - SVD is a more general technique that can be used for matrix factorization and dimensionality reduction. It's often preferred for large datasets due to its computational efficiency:
        1. Decompose the centered data matrix (A) using SVD: $A = U * \Sigma * V^T$.
        2. Σ is a diagonal matrix containing the singular values, representing the importance of each basis vector in U.
        3. Choose the top k columns of U (corresponding to the k largest singular values) and project the data onto them.