Dimensionality Reduction Quiz

1. Why do we want to reduce dimensionality?
   A. High dimensional data tends to be sparse, which is not good for many supervised methods
   B. It can make a problem that is not linearly separable into one that is linearly separable
2. Dimensionality reduction is possible when there is a redundancy in the original data.
3. Methods for dealing with data that exhibits the problem of high dimensionality:
   A. Using a set of feature detectors smaller in number than the original dimensions the data
   B. Selecting a subset of features using information gain
   C. Projecting the data onto a smaller set of dimensions
   D. Assuming things about the data, such as that the dimensions are independent, or the data is smooth or symmetric
   E. Ignoring dimensions which help little in the prediction task
4. PCA achieves the following:
   A. Dimensions that are orthogonal
   B. Dimensions of maximum variance
5. The definition of an eigenvector of a matrix is that when the eigenvector is multiplied by the matrix, the result is a multiple of the same Eigenvector.
6. PCA for dimensionality reduction:
   A. Starting with correlated data
   B. Transform the data to zero mean and unit variance
   C. Compute the covariance matrix
   D. Compute the Eigenvectors/values of the covariance matrix
   E. Pick the eigenvectors with the largest eigenvalues
   F. Project the data onto the eigenvectors
   G. This leaves us with uncorrelated data
7. The eigenvalues of a matrix are the roots of the equation set(lambda*I - covariance matrix) = 0, which are the lengths of the vectors that arise when the eigenvectors are multiplied by the matrix.
8. Methods commonly used to pick the principal components:
   A. Form a scree plot of the eigenvalues, pick the ones above the knee and use the associated eigenvectors as principal components
   B. Order the eigenvectors by eigenvalue, use the largest n which account for 90% of the variance as principal components
9. Ways of addressing PCA issues
   A. Sensitivity to large values: centre the data (zero mean and unit variance)
   B. Sensitivity to outliers: not a particular issue for PCA
   C. Can't use class labels: use Linear Discriminant Analysis (LDA)
   D. Can't help in reducing the size of the data: not a particular issue for PCA
   E. Data arises from a subspace which is not linear in the original dimensions: try top find a transform which linearises the data

10. PCA and LDA
    A. PCA
        a. Picks new dimensions by maximising variance along the dimensions across the dataset
        b. Uses eigenvectors/values from the covariance matrix of the data
        c. Can fail in classification if the classes are distributed uniformly across the directions of maximum variance
    B. LDA
        a. Picks new dimensions by maximising discrimination along the dimensions between classes
        b. Uses eigenvectors/values from the between-class and within-class covariance matrices
        c. Can fail in classification if the discriminating data is in the variances of the classes rather than the means