1. Logistic Regression
   1. x will vote for party A if

x will vote for party B if

x will vote for party A or B randomly when

* 1. The threshold T of the probability p must first be defined to obtain a decision boundary of a more general form

1. Logistic Regression
   1. z will approach argmax where zj approaches 0 for all j except the argmax will approach 1
   2. In contrast, z will approach the same uniform value, i.e., 1/K
   3. Case 1 ():  
      Case 2 ():  
      Combining the two cases, we have
2. Feedforward Neural Network
3. Convolutional Neural Networks
   1. Conv (128, 32, 7, 7, 2)

      2. Conv1 (128, 16, 1, 1, 1)
      3. Conv1 (128, 16, 1, 1, 1)
      4. Conv2 (16, 32, 7, 7, 2)
      5. Conv2 (16, 32, 7, 7, 2)
      6. Conv1 + Conv2
4. Principal Component Analysis
   1. 100%. In other words, the variance is completely preserved. By using two principal components obtained by PCA to project S linearly onto another two-dimensional space spanned by the principal components together, 100% of the total variance can be explained by the two principal components together since S also in a two-dimensional space. Thus, there is no dimensionality reduction and it is equivalent to linear transformation that maps S to the same space.
   2. In PCA, the total variance is the sum of the variance of the principal component which is the eigenvalue as shown in the lecture notes. The maximum percentage of total variance that can be explained is expressed by the sum of eigenvalues of the used principal components divided by the sum of eigenvalues of all principal components. By using one principal components, the maximum percentage of total variance is .
5. Clustering – Partitional Clustering
6. Clustering – Hierarchical Clustering