

Machine Learning and Data mining

Serie 5. KNN

Prof. J. Savoy

November 1, 2021

1. Take the titanic dataset and use all attributes to predict the class 'Survived' with a k-nearest neighbours classifier, which one do you think is the best distance measure? and why?
 - (a) Manhattan distance
 - (b) Euclidian distance
 - (c) Cosine distance
2. Build a KNN model with your selected stock / market index using your best distance measure, determine the number attributes that is capable of giving the best prediction. (Select attributes in ascending order(ie 3, 5, 7, ...) and determine the accuracy for the selected attributes, compare the accuracies to find which is the best)