The University of Chicago

Department of Statistics

Master's Thesis Presentation

JOHN LABIAK
Department of Statistics
The University of Chicago

Distance Metric Learning for K Nearest Neighbor Phonetic Frame Classification

FRIDAY, August 20, 2010, at 1:00 PM
110 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

In this paper, we consider the problem of classification. Given a set of labeled training examples, each one having associated with it some set of measured variables or features, the classification problem involves inferring the class label of a new test example, given only the associated set of features as a predictor of class membership. One simple decision rule for assigning test examples to classes is the K nearest neighbor (K-NN) decision rule. The K-NN decision rule requires that a valid distance metric be defined to measure distances between examples. For this purpose, it is common to use the Euclidean distance metric. The Euclidean distance metric, however, ignores the structure or statistical regularities (i.e. correlations) present within the data that likely informs the decision. In this paper, we consider two algorithms: the large margin nearest neighbor (LMNN) algorithm and the locality preserving projections (LPP) algorithm, for learning a distance metric that is adapted to the specific classification task we wish to solve. We consider the task of inferring the phonetic class labels of 10 ms frames of speech from the TIMIT acoustic-phonetic continuous speech corpus, after observing a set of correctly classified frames. We found that using a distance metric learned from the data using the LMNN or LPP algorithm significantly improved the classification accuracy of a K-NN classifier relative to using the Euclidean distance metric. We also found that classification using the K-NN decision rule, for all distance metrics considered, yielded significantly lower misclassification rates on the phonetic frame classification task than classification using a baseline Gaussian mixture model (GMM) classifier. Future work will seek to apply the methods developed in this paper to replace GMMs in the construction of state-dependent probability distributions within the hidden Markov model (HMM) framework that is commonly used in automatic speech recognition systems. Future work might also consider extensions to the distance learning approaches considered here, such as multiple metric LMNN and kernel LPP.

Information about building access for persons with disabilities may be obtained in advance by calling Sandra Romero at 773.702-0541 or by email (sandra@galton.uchicago.edu).