STATISTICAL ANALYSIS AND DATA MINING

Research Article

Fusion and inference from multiple data sources in a commensurate space

Zhiliang Ma, David J. Marchette, Carey E. Priebe ✉

Cited by: 4

📄
About

🔒
Access

ols

⌖
are

Abstract

Given objects measured under multiple conditions—for example, indoor lighting versus outdoor lighting for face recognition, multiple language translation for document matching, etc.—the challenging task is to perform data fusion and utilize all the available information for inferential purposes. We consider two exploitation tasks: (i) how to determine whether a set of feature vectors represent a single object measured under different conditions; and (ii) how to create a classifier based on training data from one condition in order to classify objects measured under other conditions. The key to both problems is to transform data from multiple conditions into one commensurate space, where the (transformed) feature vectors are comparable and would be treated as if they were collected under the same condition. Toward this end, we studied Procrustes analysis and developed a new approach, which uses the interpoint dissimilarities for each condition. We impute the dissimilarities between measurements of different conditions to create one omnibus dissimilarity matrix, which is then embedded into Euclidean space. We illustrate our methodology on English and French documents collected from Wikipedia, demonstrating superior performance compared to that obtained via standard Procrustes transformation. © 2012 Wiley Periodicals, Inc. Statistical Analysis and Data Mining 5: 187–193, 2012

Citing Literature ⌄

About Wiley Online Library ⌄

Help & Support ⌄

Opportunities ⌄

Connect with Wiley ⌄

WILEY