In this work, we present a kernelized dictionary learning framework for carrying out regression to model signals having a complex nonlinear nature. The optimization is carried out where the regression weights are learnt together with the dictionary and coefficients. Relevant formulation and dictionary building steps are provided. To demonstrate the effectiveness of the proposed technique, elaborate experimental results using different real-life datasets are presented. The results show that non-linear dictionary is more accurate for data modeling and provides significant improvement in estimation accuracy over the other popular traditional techniques especially when the data is highly non-linear.

**RATIONALE AND USEFULNESS**

- We are living in the era of data deluge. Even though visual data has dominated/dominating this deluge, the equation is changing with huge data emanating from the Internet of Things (IoT) Machines.
- In order to understand the data and make effective use of them, it is necessary to have appropriate data-driven methods to capture the nature of data. With this understanding, one can carry out different inference tasks like, classification, clustering and regression.
- For any data analysis, it is necessary to identify dependent variables also known as responses or predicands, and independent variables or predictors. The relationship between the predictors and responses is described by a regression function.
- This function approximation approach is useful to model the data, to characterize different states of the data generating source.
- The changes can be further leveraged towards arriving at appropriate predictive and prescriptive analytics results.

**METHOD**

Given a multivariate data of \( N \) samples, let \( X \in \mathbb{R}^{L \times N} \) represents the independent variables of feature vector length \( L \) and \( y \in \mathbb{R}^{N} \) represent dependent variable. We propose to incorporate a ridge regression penalty into the kernel dictionary learning framework for carrying out a joint optimization where the dictionary atoms, coefficients and the regression weights are learnt together. This technique offers significant improvement in estimation accuracy over the other traditional techniques especially when the data is highly non-linear.

**EXPERIMENTAL RESULTS**

Useful experimental results are obtained with different real-life datasets demonstrating the potential of the proposed algorithm in effective modeling of the data. The technique offers significant improvement in estimation accuracy over the other popular traditional techniques.

**CONTACT**

kriti.kumar@tcs.com, m.ghandra@tcs.com, achannaanil.kumar@tcs.com
angshul@iiitd.ac.in