SPARSE REPRESENTATION METHOD FOR WHOLE GRAPH EMBEDDING

by

Kaveen Gayasara Liyanage

A comprehensive exam submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Electrical Engineering

MONTANA STATE UNIVERSITY
Bozeman, Montana

December 2022

ACKNOWLEDGEMENTS

Part of the proposal is funded by the Department of Homeland Security Science and Technology Directorate under contract number 70RSAT22CB0000005.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ALGORITHMS

ABSTRACT

Sparse representation has gained popularity in the domais of signal and image processing domains.

Graph representation has gained wide popularity as a data representation method in many applications. Graph embedding methods convert graphs to a vector representation and are an important part of a data processing pipeline. In this paper, we utilize sparse dictionary learning techniques as a graph embedding solution. Sparse representation has notable applications in signal image processing. Inspired by the Graph2Vec algorithm, we aim to modify the Doc2Vec model training portion of the Graph2Vec by incorporating unsupervised dictionary learning. We investigate the viability of using the sparse dictionary learning technique KSVD for graph data. We train the dictionary on Weisfeiler-Lehman graph sub-tree kernel features. Furthermore, we use graph-based labeled data sets to compare classification results with several existing graph embedding methods. Findings show that using the learned sparse coefficients as features for a supervised machine learning algorithm provides on-par classification results when compared to other graph embedding methods.

INTRODUCTION

Sparse representation has gained

Graph representation has gained wide popularity as a data representation method in many applications. Graph embedding methods convert graphs to a vector representation and are an important part of a data processing pipeline. In this paper, we utilize sparse dictionary learning techniques as a graph embedding solution. Sparse representation has notable applications in signal image processing. Inspired by the Graph2Vec algorithm, we aim to modify the Doc2Vec model training portion of the Graph2Vec by incorporating unsupervised dictionary learning. We investigate the viability of using the sparse dictionary learning technique KSVD for graph data. We train the dictionary on Weisfeiler-Lehman graph sub-tree kernel features. Furthermore, we use graph-based labeled data sets to compare classification results with several existing graph embedding methods. Findings show that using the learned sparse coefficients as features for a supervised machine learning algorithm provides on-par classification results when compared to other graph embedding methods.

Random references[1].

figure example

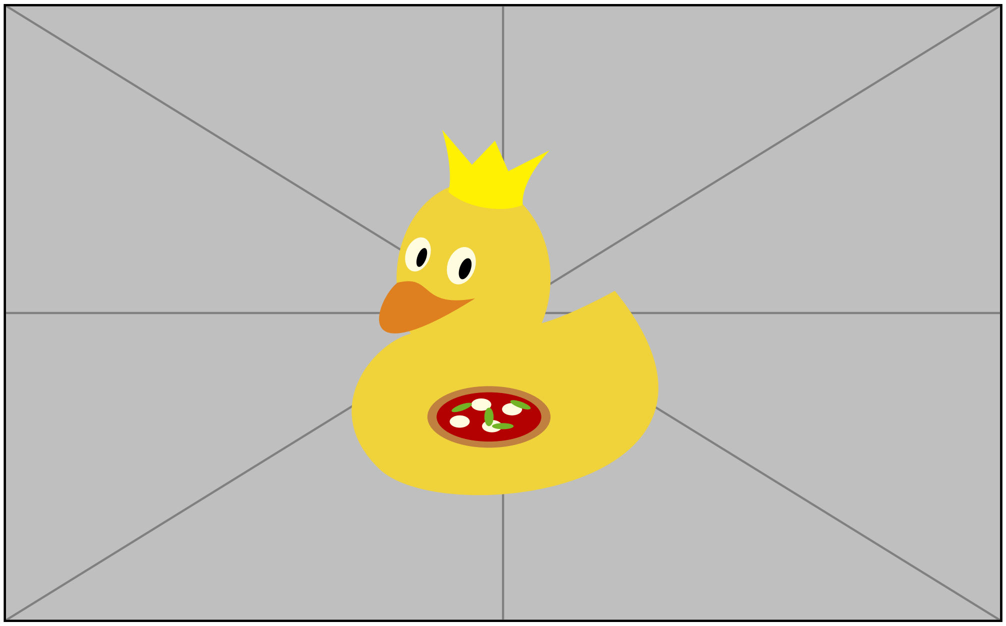

Figure 1.1: Example duck.

BACKGROUND

Sparse Representation and Dictionary Learning

Let $\mathbf{F} = [F_1,\ F_2,\ \ldots F_d]$ be a set of d features F which are collected or curated. The goal is to rank the feature set $\mathbf{F}$ by evaluating a mean-removed training set of n samples: $\bar{\mathbf{Y}} = [y_1,\ y_2,\ \ldots, y_n] \in \mathbb{R}^{d\times n}$. The set of k classes is defined as $\mathbf{C} = [C_1,\ C_2,\ \ldots, C_k]$, where each of the x samples is assigned to a class C. In sparse representation the input sample is represented as a linear combination of dictionary elements D in a over-complete $(d \ll m)$ dictionary $\mathbf{D} = [D_1, D_2, \ldots, D_m] \in \mathbb{R}^{d\times m}$, where the number of dictionary elements used, s, is far less than the number of dictionary atoms: $s \ll m$. The set coefficients $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_n] \in \mathbb{R}^{m\times n}$ of the linear combination is called the sparse coefficients. Each coefficient vector contains s nonzero entries; the remaining $m - s$ entries are exactly zero. The general objective function used for calculating the sparse representation is given by (2.1) with an ℓ_0 constraint on sparsity.

$$\underset{\mathbf{D},\boldsymbol{\alpha}}{\operatorname{argmin}}\, ||\bar{\mathbf{X}} - \mathbf{D}\boldsymbol{\alpha}||_F + ||\boldsymbol{\alpha}||_0 \tag{2.1}$$

The KSVD algorithm is an efficient iterative method that solves the objective function by, first fixing the $\mathbf{D}$ and optimizing the $\boldsymbol{\alpha}$ using orthogonal matching pursuit (OMP)[2]. Second, it fixes $\boldsymbol{\alpha}$ then optimize $\mathbf{D}$ with generalized K-means and singular value decomposition (SVD). Figure**??** represent the input signal, learned dictionary, and the sparse coding as a matrix. Since the learned dictionary is over-complete, the spread of the dictionary atoms can give an insight into which features are more relevant for the representation. Hence, we will be defining simple metrics that will quantify the spread of the dictionary elements in each of the features.

In LC-KSVD[3], the algorithm employs an objective function comprised of reconstructive, discriminative, and classification costs. The LC-KSVD dictionary is initiated by equally dividing among the classes. Then class labels are used to force the dictionary atom learning to be constrained into the predefined dictionary locations. However, this method does not takes into account class imbalances.

Frozen dictionary learning is a modification of traditional sparse coding dictionary learning that attempts to learn a dictionary that can effectively model imbalanced datasets[4]. In this algorithm first, the dictionary learning step is carried out using an algorithm, then the learned dictionary elements are held constant and the dictionary is augmented with additional elements. For the next step, the dictionary is trained again on data containing abnormalities. The frozen elements of the dictionary represent the normal aspects of the data, hence the new elements (non-frozen) learn to represent the anomalous aspects of the data that are not present in the normal data. The frozen dictionary approach could be generally used and applied to the problems including data with or without abnormalities.

Figure**??** illustrates the representation of the input signals with the learned dictionary atoms in a two-dimensional input feature domain with sparsity, $S = 2$. y_1, y_2 are the input signals belonging to two separate classes. With LC-KSVD the algorithm forces dictionary atoms $d_1, \ldots, d_3$ to be used for y_1 in class 1 and $d_4, \ldots, d_6$ to be used for y_2 in class 2. Fig.**??** illustrates the learned classifier boundary in the (sparse) domain of d_1 and d_4. Due to the class constraints in the dictionary learning process, the majority of the input data will only be represented by the dictionary atoms corresponding to their class. Thus in the sparse domain, most signals will lie in a class-specific subspace of the entire dictionary. Hence, a linear classifier can be utilized for the classification with satisfactory results. Further, there is a direct relationship between the classifier domain and the input domain.

Bibliography

[1] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, pp. 4311–4322, nov 2006.

[2] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, pp. 40–44, IEEE, IEEE Comput. Soc. Press, 1993.

[3] Z. Jiang, Z. Lin, and L. S. Davis, "Learning a discriminative dictionary for sparse coding via label consistent k-SVD," in *CVPR 2011*, IEEE, jun 2011.

[4] B. T. Carroll, B. M. Whitaker, W. Dayley, and D. V. Anderson, "Outlier learning via augmented frozen dictionaries," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, pp. 1207–1215, jun 2017.