

An Efficient Approach from Iterative Learning Algorithm

B. Himaja

Dept of CSE

Besant Theosophical College

Abstract:- A framework which takes into account machine learning for the analysis of massive datasets is proposed. The framework maps the algorithms to the respective platform so as to extract maximum resource efficiency. In addition, the framework takes into account a data projection technique called as Elastic Dictionary to form sparse representation of the underlying data. By this way, the resource efficiency is optimized leading to reduction in the cost associated with the performance. The framework represents a model and shows the performance metrics in accordance with their respective runtime and storage. An additional application program interface takes into account the applicability of the framework to the underlying platform or datasets. The framework is based on the union of both the content and platform aware methodologies so as to make the machine learning algorithms to utilize the resources efficiently.

Keywords: Cholesky, Factorization, and Elastic Dictionary.

INTRODUCTION:

Learning algorithms tend to use the dependencies and patterns which are present across the datasets. In these cases the objective function is solved by iteratively updating the parameters of interests which may require matrix multiplication that involves gram matrix but in cases when data cannot fit on a single node and therefore can require a spread architecture, repetition based updation exhibit large computation and communication costs. A sustainable computing environment requires efficient utilization of resources. The most challenging environments are those which take into account iterative updates on non-sparse(dense) datasets. This conditions involve algorithms put into place for to solve regression, deep learning or SVM used in learning applications. Example of dense datasets may include video and image datasets which stands for the substantial amount of the generated data of the modern. In this paper a framework is introduced which is a machine learning based framework for performance optimization which takes into account both platform and content aware techniques. The new viewing involving projection of data can play a vital role in the mapping onto modern multi-core and distributed architectures drives the efficiency. We observe that degree of freedom to factorize datasets to reduced dimensions involving the same error (approximated error) can be used to map computations with respect to the underlying platform. The respective implementations of this paper can be given as:

(i) Implementation of a framework which supports a range of applications in the machine learning domain are dependent upon the repetitive optimization over condensed datasets to achieve convergence.

(ii) A performance evaluation metric for the framework which is used for the optimization in terms of runtime,

energy and memory performance for the execution of a machine learning algorithms on a given system.

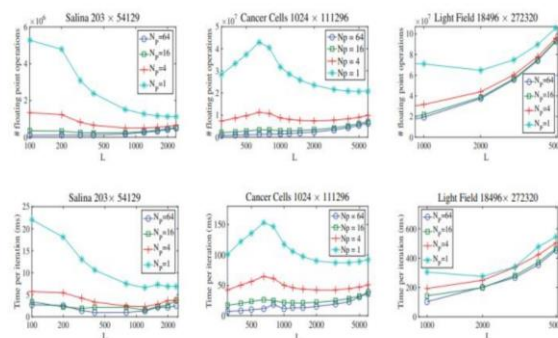
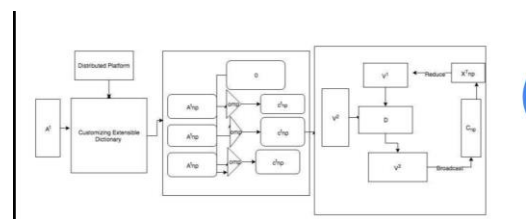
(iii) Implementation of extensible dictionary which is a parametric data projection technique which can produce many possible lower dimensional subspaces for a given approximation error. To minimize the cost associated with the performance, the parameters in the extensible dictionary are tailored with respect to the corresponding platform.

(iv) A mapping methodology to implement optimized execution of the machine learning techniques over the underlying architecture which takes into account load balancing and minimization of communication concurrently.

EXISTING SYSTEM:

In the existing system Therefore, the framework is based on the union of both the content and platform aware methodologies so as to make the machine learning algorithms to utilize the resources efficiently. Prior works listed in this section are associated to the framework

Techniques to Lessen Dimensions: Matrices involving a $M \times N$ dimensionality matrix computation methods using Singular value decomposition and principal component analysis becomes redundant on larger datasets as they involve $O(M^2N)$ complexities which also involves large constant parameters. Column Subset Selection (CSS) are randomized algorithms which have been developed to address these challenges [2].



In this methodology data is projected upon lower dimensional subspace. A subset of data columns is taken into account for learning the projection basis which is also known as dictionary. The randomized CSS techniques are applicable to large scale data as they are scalable [10]. Methods such as Farahat [4] are adaptive and lower the dictionary size to achieve an approximation error.

Techniques that are Content Aware: For accelerating some machine learning algorithm previous work have exhibited the use of importance data factorization methods. But these techniques are however better at some targeted specific machine learning algorithms but are not generic in nature unlike the proposed framework. The proposed framework works upon generic learning algorithms. An earlier method called Rankmap highlights the effectiveness of the data transformation techniques and also works upon generic learning algorithms but the system is not platform aware unlike our framework. In this paper we use Rankmap[13] as a comparison basis. There exists one another method called Stochastic Gradient Descent [17] which does not take into account efficient memory usage and the results are not assured.

Techniques that are Platform Aware: There exists an extensive amount of previous research concerned with the depiction of the ML algorithmic techniques on to dynamic platforms which efficiently run the algorithms by mapping them on to distributed architectures. It does not involve changes to the underlying data configuration [1,15]. Data provided by the respective application present is worked upon instead which is not dependent upon the corresponding architecture.

PROPOSED SYSTEM:

The framework provided here works upon large scale datasets or Gram matrices and facilitates iterative learning algorithms to run efficiently upon this massive and dense data. There exists many machine learning algorithms which operate upon the relations existing upon the present data samples such as Power method to evaluate PCA [6] or methods for solving SVM [5]. All these methods have an operational cost associated with them during an iterative update which arises majorly from the multiplication of the Gram matrix. A gram matrix is given as the multiplication of the transpose of a matrix to the matrix itself and can be given as $B^T B$ where B can be termed as the given data matrix. While operating on larger and denser datasets an amend tends to have a lot of operational cost due to the huge number message passing mechanism amongst the participating nodes. An overall flow regarding the framework is given in Fig. 1. The framework relies on the fact that machine learning applications are open to dynamic solution outputs which offers a path to trade accurate results with optimized utilisation of the resources [2,10]. In this paper we also introduce a data transformation method called Elastic dictionary

RESULT ANALYSIS:

Factor analysis is a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called

factors. For example, it is possible that variations in six observed variables mainly reflect the variations in two unobserved (underlying) variables. Factor analysis searches for such joint variations in response to unobserved latent variables. The observed variables are modelled as linear combinations of the potential factors, plus "error" terms.

It is a theory used in machine learning and related to data mining. The theory behind factor analytic methods is that the information gained about the interdependencies between observed variables can be used later to reduce the set of variables in a dataset. Factor analysis is commonly used in biology, psychometrics, personality theories, marketing, product management, operations research, and finance.

Factor analysis is related to principal component analysis (PCA), but the two are not identical.[1] There has been significant controversy in the field over differences between the two techniques (see section on exploratory factor analysis versus principal components analysis below). PCA can be considered as a more basic version of exploratory factor analysis (EFA) that was developed in the early days prior to the advent of high-speed computers. Both PCA and factor analysis aim to reduce the dimensionality of a set of data, but the approaches taken to do so are different for the two techniques.

EVALUATION:

The framework is implemented by using eigen libraries for carrying out computations involving linear algebra. In C++ a message passing system has been used. The inputs for the application program interface includes approximation error a data matrix A and the function which executes upon Gram matrix. In each continuous execution stage the procedure involves to select subsets, denoted B_i , which is selected in a random manner from rows of B. A single execution step involves a subset which is taken from rows of the matrix of B at random. It involves computation using in place of B_i . The model is operated by comparing the assumed running time trend with the corresponding precise measurement one.

CONCLUSION:

A framework which takes into account machine learning for the analysis of massive datasets has been proposed. The framework takes into account a data projection technique which is Elastic Dictionary which helps to form sparse representation of the underlying data so as to optimize the resource efficiency and so as to reduce the cost associated with the performance. An application programme interface has been provided which confirms the applicability of learning algorithms to respective platforms and datasets. We show that by tuning of parameters associated with Elastic Dictionary in terms of the underlying platform we can achieve ideal performance.

REFERENCES:

- [1] Demmel J, Eliahu D, Fox A, Lipshitz B, and Spillinger O., "Communication optimal parallel recursive rectangular matrix multiplication". IPDPS'13.

- [2] Drineas P ,and Mahoney M. “On the Nystrom method for approximating a gram matrix for improved kernel-based learning”. JMLR’05.
- [3] Dyer E, Goldstein T, Patel R, Kording K, and Baraniuk R. “Self expressive decompositions for matrix approximation and clustering”. arXiv:1505.00824 (2015).
- [4] Farahat A, Elgohary A, Ghodsi A, and Kamel M, “Greedy column subset selection for large-scale data sets” Knowledge and Information Systems (2014).
- [5] support vector machines” SIAM J. on Optimization (2002).
- [6] Figueiredo M, Nowak R, and Wright S, “Gradient projections for sparse reconstruction: Application to compressed sensing and other inverse problems” IEEE J. Select. Top. Signal Processing 1, 4 (2007).
- [7] Fine S, and Scheinberg K, “Efficient SVM training using low-rank kernel representations” JMLR’02.
- [8] Fowlkes C, Belongie S, Chung F, and Malik J. “Spectral grouping using the Nystrom method.” TPAMI’04.
- [9] Gilbert A., Strauss M., Tropp J., and Vershynin R. “One sketch for all: Fast algorithms for compressed sensing”, STOC’07.
- [10] Gittens A., and Mahoney M. “Revisiting the Nystrom method for improved large-scale machine learning”, ICML’13.
- [11] Low Y, Gonzalez J, Kyrola A, Bickson, D. Guestrin, C. andHellerstein, J. M. “GraphLab: A new parallel framework for machine learning”, UAI’10.
- [12] Malewicz G., Austern M., Bik, A., Dehnert J., Horn I., Leiser N., and Czajkowski G. “Pregel: a system for large-scale graph procesing”,