Survey Paper on Linear Programming in Machine Learning

Vivekananda Gogi
Department of Industrial Engineering and Management
R V College of Engineering,
Affiliated to VTU, Bengaluru, India

Kruthi P.R.
Department of Computer Science and Engineering,
R V College of Engineering,
Affiliated to VTU, Bengaluru, India

Shivani Deshpande
Department of Computer Science and Engineering,
R V College of Engineering,
Affiliated to VTU, Bengaluru, India

Abstract:- Linear programming (LP) is one of the well-researched and simple way to obtain optimized solutions to problems. A variety of Computer Science problems use linear programming to find the optimal solutions if they exist and Machine Learning problems are one of them.

Machine Learning a branch of data Science in Computer Science. This paper present two case studies namely Linear Programming in Reinforcement Learning and Wasserstein barycenter to review the synergy between Linear programming and complex problems of Machine Learning.

Keywords:- *Linear Programming, Optimisation, Machine Leaning.*

I. INTRODUCTION

Optimization is the way of living since we have limited time and resources. We want to make most of both of them, from using your time efficiently to solving supply chain problems to allocating resources in companies – every problem essentially needs optimizing. It's also a relevant to the problems faced in Machine Learning. Linear programming (LP) is one of the well-researched and simple way to obtain optimized solutions to problems. Using LP we can solve some very complex optimization problems by making a few assumptions that simplifies the problem. As a data science analyst, we to come across many applications and problems to be solved which can be done using LP efficiently.

Linear programming, is a well-known field of mathematics that deals with finding efficient solutions to systems which has multiple linear equalities and inequalities. The efficient solution meaning to maximise or minimise the objective function. LP is widely used in solving management problems like allocating resources, staff, maximising profits and minimising losses.

The main components of LP are as follows:

- Decision variables - These are the quantities to be determined.
- Objective function - This represents how each decision variable would affect the cost, or it is the value that needs to be optimized.
- Constraints - These represent how each decision variable would use limited amounts of resources.
- Data - These quantify the relationships between the objective function and the constraints

Linear programming is currently the popular optimization approach used in most of the businesses, as it's model of objective and constraint provides an efficient framework to separate soft constraints from hard constraints. This separation of constraints is very important in modelling real life problems and situations.

Machine learning is a branch of data science which focuses on the using data and algorithms to imitate the way that humans learn, and help computers and machines to learn on its own, gradually improving the accuracy. Usually, given more examples and data, a learning machine will ideally learn better. That is, its performance will improve over time, after looking at more examples or data.. LP can be applied here to maximise the learning and minimise the error rate.

There are three types or approaches s of machine learning: Supervised, Unsupervised and Reinforcement Learning

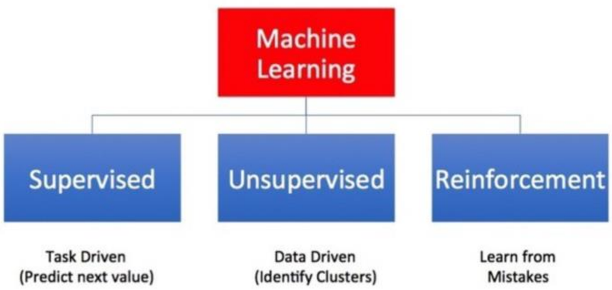

Fig 1:- Basic approaches of Machine Learning

Unlike Machine Learning, there is no learning in Linear programming. The accuracy and the success of LP strongly depends on the availability of domain experts to accurately express the objective function and constraints. If we are able to translate a Machine Learning problem into an objective function accurately, we would be able to solve the problem using LP easily and effectively.

In this paper we would be discussing various ways of how LP techniques helps in solving machine learning problems.

II. RELATED WORK

This section, includes the discussion about academic works that have already taken place. The focus is on analyzing the methods of LP used in Machine Learning discipline.

It is noticeable that use of LP models has grown extensively irrespective of the working domain. This paper [2] brought out the importance of linear programming techniques for the industry. It has become important and essential to examine the linear programming-based systems as the LP applications are gaining acceptance extensively.

This paper provides insights about the research taken up by the writers about the iterative swapping algorithm (ISA) for the computation of the Wasserstein barycenter. The paper also indicated that algorithm can be designed on the equivalence of barycenter problem to the n – coupling problem. The principle of Wasserstein barycenter is achieved through high dimensional linear program technique. The paper also suggests the approximation of such linear program using a multi-index assignment problem which is further used in a recursive manner to calculate the optimal solution. [3]

III. MOTIVATION

Linear programming is generally an important tool in the construction and optimization of algorithms, as well as, in the construction of proofs. Many results that arise in machine learning that can be cast in the framework of linear programming essentially to find an optimal solution. Hence some basics of LP help a lot in designing approximation algorithms for a class of problems in computer science called as NP-hard problems.

➢ *Case Study I*

Reinforcement Learning (RL) is a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences. As compared to unsupervised learning, reinforcement learning is different in terms of goals. While the goal in unsupervised learning is to find similarities and differences between data points, in the case of reinforcement learning the goal is to find a suitable action model that would maximize the total cumulative reward of the agent.

Markovian Decision Models are mathematical frameworks to describe an environment in RL and almost all RL problems can be formulated using them. An MDP consists of a set of finite environment states S, a set of possible actions A(s) in each state, a real valued reward function R(s) and a transition model P(s', s | a). When trying to solve this MDP, we often run into the so-called three curses of dimensionality, referring to a state space, outcome space and action space all exploding when the size of the problem grows.

For any given state, we have the assumption that the state's true value is reflected by:

$$\boldsymbol{V*(s) = r + \gamma\, max_{a\in A} \sum_{s'\in S} P(s'|s,a).V*(s')}$$

That is, the true value of the state is the reward we get for being in it, plus the expected future rewards of acting optimally from now until infinitely far into the future, discounted by the factor γ, which captures the idea that reward in the future is less good than reward now.

In Linear Programming, we find the minimum or maximum value of some function, subject to a set of constraints.

We can represent the problem of finding the value of a given state as:
minimize $V(s)$
subject to the constraints:

$$\boldsymbol{V(s) \geq r + \gamma \textstyle\sum_{s'\epsilon S} P(s'|s,a) * V(s'), \forall a \in A, s \in S}$$

If we are able to formulate the linear program by writing a program like the one above for every state and then minimize $\sum V(s)$, subject to the union of all the constraints from all these sub-problems then we have reduced the problem of learning a value function to solving the LP.

➢ *Case Study II*

- *Wasserstein Barycenter Problem*

Unsupervised learning problems have input variables and no corresponding output data format. Algorithms help in finding out the interesting patterns hidden in the data. Unsupervised learning technique tries to obtain a distributed model of the data in order to extract more insights about it.

The Wasserstein barycenter is an important notion in the analysis of high dimensional data with a broad range of applications in applied probability, economics, statistics, and in particular to clustering and image processing.

The problem of Wasserstein Barycentre focuses on gaining a solution by using the technique of weighted mean of summation of probability distribution in such a way that the weighted distance is found to minimal between men and the probability distribution in the summation obtained. Using the Sinkhorn Algorithm, the regularised Wasserstein distance is considered. The regularised Wasserstein distance is found to be optimal and much faster in computation in comparison to the original Wasserstein barycentre.

In Data Science field, Wasserstein distance is defined as the minimal transportation cost between a supply distribution and a demand distribution. The aim of the Wasserstein Barycenter Problem is to compute the points in such a way that the value is found to be the minimal value of the sum of its Wasserstein distances to each of a set of distributions:

$$WD\,(s, d^k) = \min \sum_{i=1}^{N} \sum_{j=1}^{N} c_{ij}\, x_{ij}$$

$$\text{s.t.} \quad \sum_{j=1}^{N} x_{ij} = s_i \;\; \forall i = 1, \dots, N$$

$$\sum_{i=1}^{N} x_{ij} = d_j \;\; \forall j = 1, \dots, N$$

$$x_{ij} \geq 0, \qquad \forall\, i, j$$

IV. CONCLUSION

Linear programming is a method of optimizing operations with some constraints. The main objective of linear programming is to maximize or minimize the numerical value.

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

Machine learning involves a strong connection with mathematics. The algorithms in machine learning discipline are designed based on the concepts of Mathematics. The best algorithm is adopted on the basis of training complexity, number of features, accuracies, etc. An essential field of mathematics is Linear programming. Linear programming mainly includes the study of various fields like vectors, matrices, planes, mapping and linear transformations.

The foundation of machine learning is based on Linear programming which has vital role in development of various machine learning algorithms. Use of Linear programming in Machine Learning domain has enabled the algorithms to run on huge data sets. Neural Networks models are also designed using Linear programming techniques. Data science field is found to use Linear algebra techniques for optimized results.

Introduction and familiarity to linear programming and optimization theory enables a person to understand the Machine Learning and Data Science field.Linear programming also helps to create better supervised as well as unsupervised Machine Learning algorithms.

REFERENCES

[1]. Jean-Michel Reveillac “Linear Programming", 2015 IEEE Network Operations and Management Symposium - 4 December 2015 , Pages 209-237, Publisher – IEEE.
[2]. J. S. Aronofsky, "Growing Applications of Linear PRogramming”, ACM , Communications of the ACM, Volume 7, Issue 6, June 1964, pp 325–332 Publisher: ACM
[3]. G. Puccetti, L. Rüschendorf and S. Vanduffel, “On the computation of Wasserstein barycenters” , Journal of Multivariate Analysis, Volume 176, March 2020,Elsevier.
[4]. A. Charnes,C. E. Lemke, “Computational problems of Linear Programming”, ACM '52: Proceedings of the 1952 ACM national meeting (Pittsburgh)May 1952 Pages 97–98.
[5]. Kouichi Genda, “On Demand Network bandwidth reservation combining machine learning and linear programming” , 17th Internation Conference on Network and Service Management,2021.
[6]. Patrick Nadeem Ward, ”Linear Programming in Reinforcement Learning”, Computer Science McGill University, Montreal,April,2021.
[7]. Rajat Mittal , “Applications of Linear programming in Machine Learning”,IIT Kanpur.