Mathematical Foundation for Machine Learning

Tirtharaj Sapkota [1]
[1]Purbanchal University School of Science and Technology (PUSAT), Purbanchal University, Nepal

***Abstract*—**The concept of Artificial Intelligence (AI) and Machine Learning (ML) have gained popularity in recent years because of their several applications in diverse fields. Even though they are sometimes used interchangeably and are related to each other, they are quite different in terms of their use, data sets and many more. AI can be described as a branch of Computer Science that is basically concerned with building intelligent machines. ML, on the other hand can be considered as a subset of AI and includes techniques to implement AI. It has the ability to learn itself from the data being provided without having to program explicitly. Mathematics and ML have a very close relation. ML has gained so much success because it has mathematics as a backbone. Mathematics is the foundation of almost all ML algorithms. All ML models are created using results and concepts from mathematics. One of the aims of this paper is to illustrate that knowledge of few important mathematical concepts are very essential in understanding the ML algorithms. Also, this paper contributes to a deeper understanding and tries elucidate the need of Mathematics in designing and developing ML algorithms using few simple examples. An attempt has been made to familiarize the readers with the concept of Data Science which is pretty fresh in IT industry. It is the use of mathematics and statistical tools and techniques integrated with Machine Learning models to solve complex problems which aids in complex decision making.

***Keywords*—** *Artificial Intelligence, Machine Learning, Mathematics, Statistics, Data Science, ML Models*

I. INTRODUCTION

Artificial Intelligence (AI) and Machine Learning (ML) are probably the most frequently used concepts in technology world in last few decades. AI is a concept that helps a machine to be intelligent, meaning that a machine can perform various task which if performed by human being would require intelligence [1]. The world is continually transforming into an ecosystem where everything is expected to be performed by intelligent machines. Intelligent machine means the machines havingthe capacity to perform activities in the same intelligent way as we human do. Intelligence can be described as the capacity to work dynamically according to the current situation. Traditional computers are being described as unintelligent or dumb as they are only able to perform whatever has been instructed in advance, unable to react to events according to the current circumstances. The authors in [2] states that Ml focus on dealing with building computers which automatically improves through experience. The authors further states that ML lies at the intersection of statistics and computer science. The goal of AI and ML is to design and develop machines that can emulate the way human being perform actions i.e., to design and develop intelligence machines.

Data and ML models are the two basic concepts that form the foundation for any machine learning techniques. We cannot define or describe any machine learning methods without these two components. Basically, what happens in ML applications is that, we take a ML model and feed lots and lots of data to this model. The model in turn, learns from the data provided and is then is able to make intelligent decisions when new data is being provided at later point of time. This means that we first train the ML model to do some specified tasks and then ask it to do some similar tasks. Machine Learning can be mostly classified as Supervised, Unsupervised and Reinforcement Learning Algorithms. Supervised ML algorithms learns from Labelled data. Labelled data means that the data has some kind of tags attached to it such as name, a type etc. Labelling makes the training process efficient and simple. The goal is to make predictions using the learned data. Classification and Regression are few tasks that fall under supervised learning. Classification is applicable in predicting discrete random variables whereas Regression basically deals with predicting continuous random variables. Decision Tree classifier, Random Forest classifier and k-nearest neighbor are few algorithms categorized under classification task whereas Logistic Regression, Polynomial Regression, Support Vector Machines (SVM) are categorized under Regression tasks. In unsupervised learning algorithm, the data does not have anylabel and the ML model learns from unlabeled data. The goal for the model is to read the unlabeled data and find some interesting patterns from the data. Clustering and Association are important tasks that is classified under unsupervised learning. Clustering means grouping of data having similar pattern. Association task generally finds the important relationship between data points i.e., how one data is associated with other. K-means clustering, hierarchical clustering, principal component analysis are few examples of clustering algorithms whereas Apriori and Eclat are Association algorithms. Agent and the Environment are the most important modules in a Reinforcement Learning Algorithm. Unlike supervised or unsupervised algorithm, Reinforcement Learning algorithms does not need pre-defined data sets instead it learns real time data using environment and reward. The agents try to learn the optimal actions to take by interacting with the environment. Self-driving car, ChatGPT are few recent applications of Reinforcement learning algorithms.

II. LITERATURE REVIEW

The authors in [3] advocates that Machine Learning is a technology in the field of Computer Science which find solutions to many problems like prediction, classification etc.

by building mathematical models. In this regard, concepts of mathematics, statistics, probability and calculus are crucial. Mathematical concepts are necessary for building appropriate models and receiving precise results. Likewise, for prediction and classification task, the concept of probability is important. statistical concept is necessary for sampling in ML while calculus is used for multivariate related results. The authors in [4] emphasizes that mathematics is at the heart of machine learning. The author focuses on describing the concepts of mathematics such as linear algebra, probability theory and statistics, multivariate calculus, partial derivatives, and algorithmic optimizations that are indispensable for creating efficient ML systems. The authors in [5] brings out the importance of mathematics in Machine Learning Model building. Machine Learning along with the combinations of Statistics, Calculus, Linear Algebra, and Probability becomes more powerful technique for solving complex problems. The authors in [6] suggested that all those who wants to work in the field of ML and AI must realize that Mathematics is a fundamental domain to focus. Every idea that we acquire in ML or any small algorithms that we design or build to solve a problem has connection to mathematics either directly or indirectly. While solving problems using machine learning, we obtain knowledge regarding practical applications of ML rather than obtaining theoretical knowledge. The author advocates that the best way to become familiar with notions of mathematics for ML is to consider some real time use cases and solving and understanding the mathematics behind it. The author in [7] provides some meaningful insights about the importance of Mathematics in Machine Learning and the topics and useful resources to gain proficiency in these topics. Some of the most important topics according to the author are Linear Algebra, Probability Theory and Statistics, Multivariate Calculus, Algorithms and Complex Optimizations etc. The author also provides some data related to the amount of mathematics required and the level of mathematics to understand these topics. In [8], a detail description of a number of mathematical constructions like Lagrange multipliers and some selected subtopics from matrix analysis from a machine learning perspective have been stated. These concepts are widely used in applied machine learning. The author in [9] states that ML techniques that are used in creating intelligent machine depends heavily on mathematical optimization, statistics, and algorithm design. A basic framework of machine learning has been shown in the paper and some mathematical problems associated with it are pointed out.

Almost all papers that are considered for literature review advocates that Mathematics is extremely important in machine learning. The more understanding of mathematics one has, the better he is in ML implementations. Knowing mathematics relevant to ML definitely helps with building up the skill sets and appreciating how a model works.

III. RESEARCH METHODOLOGY

Smart web search has been considered as a component of the research methodology for this paper. Researchers often use smart search strategies to efficiently gather information from diverse sources on the internet. A number of search engines such as https://google.com, https://ask.com, https://bing.com, https://search.yahoo.com has been used to construct this paper. Few meta search engines were also considered such as www.surfwax.com, www.HotBot.com. Academic search Engines such as www.scholar.google.co.in and www.academicindex.net were also useful. Social search engines such as blogs, microblogs and search Facebook were also considered for literature survey. Last but not the least, ChatGPT, owned by OpenAI was also useful during the research.

IV. RESULTS AND DISCUSSIONS

Suppose we want to make an intelligent system that will observe an animal image and tell us whether it is a lion or a tiger. To implement this, we will feed lots of different images of lions and tigers to our ML model. Now the ML model are ready for training. During the training, it looks over all the images that were provided and tries to find patterns in the images. Once it learns from data (images of lions and tigers), we provide a new image which was not given to the model during training, the ML system can predict whether the image is a lion or a tiger. This is just a simple example of ML applications that is trying to predict a discrete variable, hence a classification task. There are many more complex applications where ML techniques are significant. So, the important concept to note here is that we feed data to ML models and it learns from those data. 'Google Assistant' is a good example of ML applications. It is first trained with huge amount of speech data and it learns from these data. Whenever a user asks something, it understands and response with appropriate feedback. Having said this, what exactlya ML Model is? Consider the following values of X and Y in the Table 1 shown below.

Table 1. Relationship between independent and dependent variable

X	1	2	3	4	5
Y	5	7	9	11	13

Here, X is an independent variable and Y depends on X, hence it is a dependent variable. We can see that for values of X there are corresponding values in the Y row. The value of Y is related to X. But how is Y related to X?

Our job is to discover the link between these two variables. If we can understand how Y is related to X, it will be easier for us to understand how ML Model works. This is a very basic concept used in ML models. It is important to find ways or methods which can depict the relationship between these variables. How can we find a function or an equation that relates the two variables? This is the most important task of an ML-Model. One way to show the relationship is by plotting a graph as show in Figure 1.

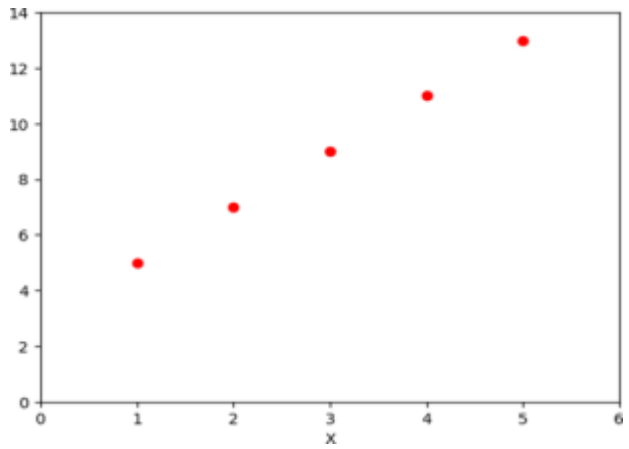

Figure 1. Relationship between X and shown using Points

By looking at the plot, we can derive that if the value of X increases the value of Y also increases. These two values are directly proportional. But what is the exact relationshipbetween X and Y? This question needs a clearer answer. If we join all the points, we can see that they form a straight line as shown in Figure 2. So, if we can find the exact equation of that straight line, it may tell us about the relationship between X and Y.The equation of the line is **Y=mX + C**, where X is X-value, Y is Y-value, m is the slope and C is the intercept. As we already have X and Y values, our intention is to find **m** and **C** so that relationship between X and Y is determined. The slope of the line isrepresented by **m**. It tells us how slanting the line is or at what angle the line is oriented. The slope tells 'For a change in value of X, what is the change in value of Y?' For example, in our case we can see that if the X value change by one unit, the Y value changes by 2 units. For different values of m, we will have different line orientations.

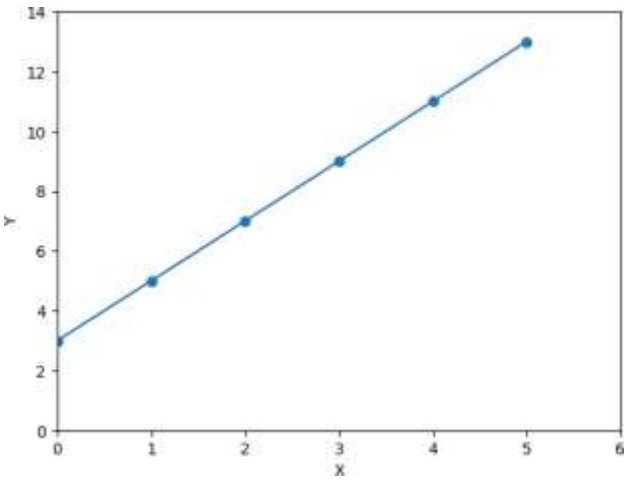

Figure 2. Relationship between X and Y shown as a Line

Intercept C is the distance of the line from the origin (0,0). It is point where the line intersects at y-axis. We can find the value of **m** by considering any two given points. Let us take p1 as (2,7) and p2 as (3,9), the second and the third points respectively givenin the table. Using the formula m=(y2-y1)/(x2-x1) we can find the value of m as 2. To find the value of C, we may consider any given point say (4,11), and using the calculated m value in the equation (i.e. Y=2X+C), we find the value of C as 3. Therefor the equation of the straight line is Y=2X+3. This is the relationship between X and Y. When we have this line, we can consider any value of X and find the corresponding value of Y, irrespective of the points given in the table. Figure 3 clearly shows the equation of a line along with the slope and the intercept.

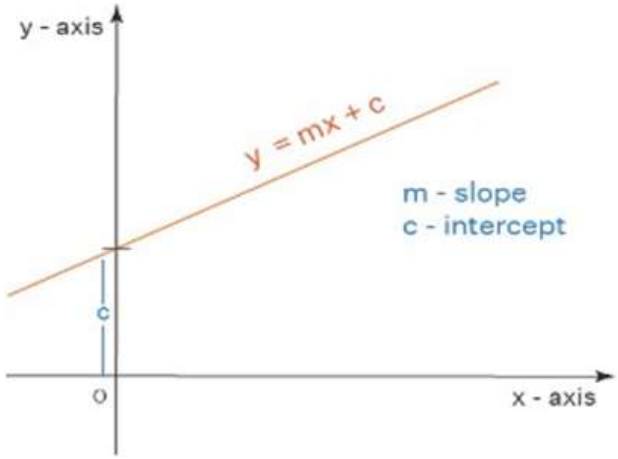

Figure 3. Slope and intercept of a line

Machine Learning is very similar to this and is all about fitting the model to the data points sothat we get a relation or a function something like Y=2X+3 in this case. Once we have this equation (relation between X and Y), we can test any points on this equation. Fitting this machine learning model to data points is called ML Model Training. Therefore, ML Model is a function that depicts the relationship between Features (Independent variable) and Target variable (Dependent variable). It tells how the input variables are converted into output variables. This concept can mathematically be represented as shown below

ML_Model : Input → Output

Anyone who is willing to go deeper into the design and development of machine learning techniques, there are few important mandatory mathematical topics that one has to understand. These topics include Linear Algebra, Calculus (Differential and Integral), Statistics and Probability. While building ML models we need to deal with matrix, vectors etc. which are integral part of Linear Algebra. Differential Calculus is all about splitting the data into smaller pieces and see what is the change going on inside the points or find some interesting insights in the smaller parts. Integral calculus deals with integrating the smaller data into larger or complex data. Gradient descent is a Machine Learning Technique that makes use of integral calculus to find solutions to problem. For analyzing the data sets when building Machine Learning Models, statistical measures such as mean, median, mode,

correlation etc. are quite useful. Before the training data is provided to the ML Model, the data has to be cleaned. In such a situation the missing values can be replaced by the mean or median depending on the nature of data. If there are outliers in the data, then replacing a missing value by a median may be more useful than replacing by the mean value. We can use the concept of correlation to see how closely each variable or features are related to each other or find a variable that is more useful for analysis. Therefore, statistics is an important concept useful in building Machine Learning models. Another important component of Machine Learning is Probability. Machine Learning is used for complex decision making. Probability can be used to find the likelihood of some decisions. Concept of random variables, types of random variables (discrete or continuous), probability distribution like Normal distribution, Poisson distribution are very essential. Using the various probability distribution methods, we can analyze the data from different dimensions and derive important insights from the data.

V. CONCLUSIONS

From the above discussion we can observe that Mathematics and Machine Learning have a very close relation. Without the acquaintance of some basic mathematical concepts, understanding the ML algorithms will be quite difficult. Data Science is pretty fresh concept in IT industry. It is the use of mathematics and statistical tools and techniques integrated with Machine Learning models to solve complex problems and make accurate decisions. Data scientist apply machine learning algorithms to various types of data to create AI systems that performs tasks that would normally require human intelligence. If someone wants to shape their carrier as a datascientist in the tech industry, then the mathematical concept stated above are mandatory.So, anyone who is passionate about Machine Leaning and wants to go deep into it, the familiarity with some related Mathematics may be very crucial.

REFERENCES

[1] A. B. Simmons and S. G. Chappell, "Artificial intelligence-definition and practice," in *IEEE Journal of Oceanic* Engineering, vol. 13, no. 2, pp. 14-42, April 1988.

[2] Jordan, Michael I., and Tom M. Mitchell. "Machine learning: Trends, perspectives, and prospects." Science 349.6245 (2015): 255-260.

[3] Dr. Mithun B Patil, Prof. Santosh Jadhav, Dr. Sopan Talekar, Dr Vipul V Bag, "The Role of Mathematics in Machine Learning", Journal of Data Acquisition and Processing Vol. 38 (1) 2023, ISSN: 1004-9037.

[4] Srinivas Pyda, Srinivas Kareenhalli, " Mathematics for Machine Learning", Mathematics and Computing,2018, Volume 253, ISBN: 978-981-13-2094-1.

[5] Lamba, Shweta and Saini, Preeti and Kukreja, Vinay and Sharma, Bhanu, Role of Mathematics in Machine Learning (April 25, 2021). Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2021, Online Available from SSRN: https://ssrn.com/abstract=3833931 or http://dx.doi.org/10.2139/ssrn.3833931.

[6] Dr. Suresh Dara et al., " ROLE OF MATHEMATICS IN MACHINE LEARNING', International Research Journal of Modernization in