JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

MATHEMATICIANS TOWARDS DATA SCIENCE AND MACHINE LEARNING

Dr. Vidyullata Shekar Jadhav

Assistant Professor
V. P. Institute of Management Studies & Research, Sangli
vsjadhav@vpimsr.edu.in

ABSTRACT

Data Science has become one of the most important aspects in most fields as data volumes grow by the day. Data is critical in every industry, from healthcare to business. Mathematics and data are fundamental to modern science, and many industries are looking for graduates who can work in multiple fields. When one gets graduated, they can have a wide range of career options because of gained skills and knowledge in areas. One will have gained the necessary foundation for further postgraduate study and research in related fields. This paper highlights the relation of mathematics and data science with respect to the career opportunities for the mathematicians in the field of data science which is considered as a sexiest job in 21st century followed with different applications of mathematics in data science.

KEYWORDS: Mathematics, Mathematicians, Data Science, Machine Learning, career, applications

INTRODUCTION

Our world is becoming increasingly digital, and data can be found almost anywhere. It can be found in one's social network timeline, fitness tracker, MRI scans, or bank transactions. All of these massive amounts of data contain valuable information; the question is how to extract that information in such a way that it can be put to good use. Mathematics in Data Science will deal with mathematical models and machine learning algorithms, which are necessary for analyzing data and providing a foundation for evidence-based decisions in a wide range of fields. To optimally employ data science and further develop the field into exciting new directions, one must use the underlying theory from, for example, statistics, numerical analysis, or graph theory. Regardless of the future career path in industry, mathematics is a central training pillar for data scientists. It guarantees to help a company solve problems, gain a competitive

advantage, innovate faster, optimize model performance, and apply complex data effectively to business challenges. Mathematics has gained prominence in cutting edge technologies such as Machine Learning, Artificial Intelligence, Data Science, Deep Learning and many others. Every new technology in today's world is either directly or indirectly related to mathematics with the aim of developing intelligent and simple solutions to problems. Mathematics currently rules the world like a monarch due to its applications in a wide variety of fields. Mathematics with its various mathematical functions, new techniques, theorems, algorithms etc supports many industries like software, medical, automotive, design, robotics and many more (Harsha, 2021).

MATHEMATICS IN MACHINE LEARNING (ML)

Mathematics in Data Science entails working with mathematical models and algorithms that can be used to analyze data, draw conclusions, and make decisions in order to become an expert data scientist or researcher who advances the field. Machine learning is fundamentally data-driven; data is at the heart of machine learning. The ultimate goal is to uncover useful hidden patterns in data. Although data is not always numerical, it is more useful when treated as such. Data can be thought of as vectors, which are objects that follow arithmetic rules. This enable s us to comprehend how linear algebra rules operate on data arrays. A model is a mathematical representation of certain assumptions and beliefs. It is said to learn the process (linear, polynomial, etc.) of how the data it is given was generated in the first place, and then make predictions based on that learned process. We define functions of many variables that represent certain assumptions from a probabilistic standpoint using general ideas from applied math. The model in a regression example approximates a function that maps inputs to real-valued outputs. To account for the term "automatically" in machine learning, each model is distinguished by a cost function that we create to assess how well the assumptions we made while developing our model correspond to reality (yet unseen data or test data). To minimize this cost function, we use numerical optimization methods (partial differentiation, SVD). The new field of mathematical analysis of deep learning is also in the trend (Julius, Philipp, Gitta, & Phil, 2021).

OPPORTUNITIES FOR MATHEMATICIANS IN DATA SCIENCE AND ML

As a mathematician, one will occasionally push the boundaries of current mathematics by inventing new expansions where necessary. With a mathematician's flexible thinking allowing one to come up with practical answers to questions in many fields, one is likely to find oneself working in multidisciplinary teams. Because mathematics is the common language of engineers and an essential link in optimization processes in virtually every field. Using advanced mathematics and statistics, develop data science and machine learning skills (Tirthajyoti, 2020). Prospective students must be willing to study mathematical theory and apply their knowledge in the practical management of big data applications. Modern mathematics is considered an independent, genuine technology, rather than merely an auxiliary science for other sciences to solve technical and

economic problems. Its main distinguishing features are its algorithmic and technology-driven approach, as well as industrial relevance and innovation. The goal of Mathematician in Data Science is to persuade young people to participate in ongoing developments in the most competent and comprehensive manner possible, as well as to provide them with a promising future in the field of mathematics and its applications to digital media and data science. Data science necessitates the use of mathematics. A strong education in specific mathematical subjects is required of any practicing data scientist or anyone interested in pursuing a career in data science. Depending on the job path chosen, most firms require at least a B.A., M.A., or Ph.D. degree as a Data Scientist. Your capacity to apply your data science skills to real-world problems is dependent on your success and mathematical understanding. Machine learning techniques, as well as performing analysis and extracting insights from data, necessitate a math degree for data science professions. While mathematics isn't the only prerequisite for a data science education and job, it is frequently the most crucial. Identifying and comprehending business difficulties and turning them into mathematical challenges is one of the most critical elements in a data scientist's workflow. You may hire a data scientist, machine learning engineer, business intelligence developer, data architect, or another industry expert. Consider the many sorts of mathematical requirements and how they are used in data research. So that you can more effectively pursue your chosen mathematical education, you should better understand your skills and interests.

Mathematics is the foundation of data science, and it is one of the most important aspects of this field. The process of making sense of data, determining the implications of the results, and drawing appropriate conclusions is heavily reliant on mathematics, with some of the key concepts being Linear Algebra, Differential Calculus, Discrete Math, and, most importantly, Statistics. The mathematical concepts at the heart of the four branches, which are regression, classification, dimensionality reduction, and density estimation. If anyone wants to pursue Data Science as a career, you should be familiar with topics such as Linear Algebra, which covers vectors and matrices, Probability distribution to quantify uncertainty, Scalars, and vector calculus, which explain gradient descent (Ben, 2020). *Linear Algebra:* Developing machine learning algorithms necessitates an understanding of how to form linear equations. To inspect and observe datasets, these must be used. Loss functions, regularization, covariance matrices, and support vector machine classification are all examples of how linear algebra is utilized in machine learning. Many machine learning concepts rely on linear algebra. PCA, for example, necessitates eigen values, whereas regression necessitates matrix multiplication (Marc, A., & Cheng, 2020). Furthermore, the bulk of machine learning applications work with data that is multidimensional (data with many variables). The ideal approach to express this type of data is via matrices (Ronald, 2021). *Calculus:* Algorithm training and gradient descent both require multivariate calculi. It's necessary to investigate derivatives, curvature, divergence, and quadratic approximations. Calculus is required for a number of critical machine learning applications. To optimize, you'll need to know how to calculate derivatives and slopes, for example. Gradient

descent is, in fact, one of the most extensively used optimization methods (Tivadar, 2021). It deals with the rate at which quantities change, multivariate calculus plays a very important role in machine learning algorithms, to quantify better, find prices more accurately, and understand relationships better, this concept is widely used, multivariate calculus is nothing but calculus that involves multiple variables, Calculus is also used to find the gradient descent (Ronald, 2021). Example: This concept is used a lot by the data science team in companies now a days, to understand at what the money spent on advertising is helping them with their sales. *Statistics:* Statistics is a completely different subject in and of itself, but it is very important in data science because it is basically the collection and analysis of large amounts of numerical data. It is the fundamental building block for all machine learning algorithms. It is used to comprehend user engagement, retention, conversion, and leads, as well as to translate big data into big picture. *Probability:* This is critical for hypothesis testing and distributions like the Gaussian and probability density functions (Ronald, 2021). *Discrete Mathematics:* It is the investigation of mathematical data that are fundamentally discrete rather than continuous. Similarly, data scientists do not always receive continuous data; they may receive noncontinuous or discrete data, which can be difficult to make sense of; however, with an understanding of this segment of math, algorithms can be simplified and made very useful to someone. It primarily consists of concepts such as sets, subsets, power sets, counting functions, recurrence equations and equations, graph theory, and so on. Discrete math can be used in algorithms for social media analysis, and Google Maps uses it to determine the fastest driving route. Machine learning is driven by four key concepts: statistics, linear algebra, probability, and calculus. While all models require statistical ideas, calculus allows studying and optimizing them. When working with a huge dataset, linear algebra comes in helpful, and probability helps forecast the result of upcoming events (Chakraborty, 2021). Professional career as a Mathematician and data scientist after receiving your master's, one can pursue a PhD programme. This entails spending four years at research institutes studying a specific research area in depth. Mathematics is and will continue to be a subject with a high job guarantee. Because of their analytical and methodological skills, mathematicians are increasingly in demand in high-tech and IT research labs.

OUTCOME OF MATHEMATICIANS IN DATA SCIENCE

To study a variety of taught units in applied math techniques, mathematical modeling, data science, machine learning, and scientific communication. Real-world experience will be gained through group projects with other students as well as the opportunity to work on a project with one of our industrial partners. Our research specialists contribute to projects and modules in areas such as industrial app lied mathematics, probability, scientific computing, machine learning, dynamical systems, asymptotic, imaging, wave and fluid modeling, and mathematical biology. It is now an essential component of many fields, resulting in a high demand for data scientists. From assisting brands in understanding their customers to solving complex IT problems, to its applicability in almost every other field, it is critical to

the operation and growth of any organization or company. A data scientist's average annual pay is Rs.698,412. An entry-level data scientist can earn around $500,000 per year with less than a year of experience. Data scientists with 1 to 4 years of experience may expect to make £610,811 per year (InterviewBit, 2021).

APPLICATIONS OF MATHEMATICS IN DATA SCIENCE AND ML

Businesses of all sizes use data scientists to help them function and prosper on a daily basis. Understanding how to use arithmetic in real-world situations will help you grasp why firms need data scientists and what role math plays in that. Some examples of practical math applications in popular data science and machine learning applications and technologies employed by today's leading companies. *Natural Language Processing (NLP):* For word embedding, NLP employs linear algebra, as well as unsupervised learning approaches like topic modeling and predictive analytics. NLP applications include chatbots, language translation, speech recognition, and sentiment analysis. *Computer Vision:* Linear algebra is also employed in picture rendering and processing in computer vision applications. People think of firms like Tesla and its self-driving cars when they think of computer vision. In addition to agriculture, computer vision is increasingly utilized in healthcare to classify disorders and enhance diagnostics. *Marketing and Sales:* Hypothesis testing, for example, can be used to test the success of marketing efforts. It's also utilized in approaches like causal effects analysis or survey design to better understand consumer behavior, such as why individuals buy from a particular brand, as well as in personalized recommendations using predictive modeling or clustering (Ronald, 2021). *Fraud detection:* Machine learning algorithms and techniques for loss reserve models, claim fraud detection, risk evaluation and prediction, and other applications. *Finance:* Data-driven models in mathematical finance and investment.

CONCLUSION

In a nutshell, mathematics in data science and machine learning is not about crunching numbers, but about understanding what is happening, why it is happening, and how we can experiment with different variables to achieve the desired results. Data science is used to find/identify patterns, and patterns can be represented in a fashion that allows them to be studied by knowing various mathematical terminologies, which is essential for developing statistical models, algorithms, and procedures that make appropriate conclusions. The cutting edge of mathematical application in industry combines modeling, analysis, and interpretation with methods from applied mathematics and the rapidly growing field of data science. This is a career for those who want to use mathematics, statistics, and computer science to solve complex problems in industry.

REFERENCES

[1] Ben, D. (2020, September 23). What's the best way to prepare for machine learning math? Retrieved December 23, 2021, from TechTalks: https://bdtechtalks.com/2020/09/23/machine- learning-mathematics/

[2] Chakraborty, A. (2021, June 1). How to Learn Mathematics For Machine Learning? What Concepts do You Need to Master in Data Science? Retrieved December 23, 2021, from Data Science Blogathon: https://www.analyticsvidhya.com/blog/2021/06/how-to-learn- mathematics- for- machine- learning-what-concepts-do-you-need-to-master-in-data-science/

[3] Harsha, V. G. (2021, April 26). Mathematics as a Monarch. Retrieved December 22, 2021, from Data Science Blogathon: https://www.analyticsvidhya.com/blog/2021/04/mathematics- in-data-science/

[4] InterviewBit. (2021, December 2). Data Scientist Salary in India – For Freshers & Experienced. Retrieved December 23, 2021, from InterviewBit: https://www.interviewbit.com/blog/data-scientist-salary/

[5] Julius, B., Philipp, G., Gitta, K., & Phil. (2021). The Modern Mathematics of Deep Learning. In Theory of Deep Learning. Cambridge : Cambridge University Press

[6] Marc, P. D., A., A. F., & Cheng, S. O. (2020). Mathematics for Machine Learning. Cambridge: Cambridge University Press.

[7] Ronald, V. L. (2021, November 11). Math and Data Science: What Do You Need to Know? Retrieved December 22, 2021, from https://www.simplilearn.com/math-and-data-science- article: https://www.simplilearn.com/math-and-data-science-article

[8] Tirthajyoti, S. (2020). Hands-On Mathematics for Data Scientists: Explore advanced mathematical concepts to improve your abilities to become a successful data scientist. In D.T. Sarkar, Hands-On Mathematics for Data Scientists: Explore advanced mathematical concepts to improve your abilities to become a successful data scientist. Packt Publishing Limited.

[9] Tivadar, D. (2021). The Roadmap of Mathematics for Machine Learning. Retrieved from Mathematics of Machine Learning: https://www.tivadardanka.com/blog/roadmap-of- mathematics-for-machine-learning