iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** III **Month of publication:** March 2021

DOI: https://doi.org/10.22214/ijraset.2021.33316

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue III Mar 2021- Available at www.ijraset.com

A Review of Grey Scale Normalization in Machine Learning and Artificial Intelligence for Bioinformatics using Convolution Neural Networks

Divya Kothari[1], Ajay Kumar Sharma[2], Mayank Patel[3]
[1]*Research. Scholar,* [2, 3]*Associate Professor, Department of Computer Science & Engineering, GITS, Udaipur*

Abstract: ***These days, machine learning is a modern trending area and is an artificial intelligence technology. To make computers function in a certain way without being specifically programmed, machine learning uses certain statistical algorithms. The algorithms obtain an input value and by using some statistical techniques, forecast an output for this. Machine learning's primary goal is to create intelligent machines that can think and function like human beings. Machine learning systems can perform sophisticated processes by gathering and analyzing data, rather than adopting pre-programmed rules, by allowing computers to perform complex jobs smartly. Exciting developments in machine learning have been seen in recent years, which have expanded its abilities across a variety of applications. Growing data availability has helped machine learning systems to be developed on a broad pool of examples, while the computational capabilities of these systems have been supported by increasing computer processing power.***

Keywords: Machine learning, Deep learning, Bioinformatics, Convolution neural network

I. INTRODUCTION

Machine learning helps machines to replicate and manipulate human-like behavior. Using machine learning, each interaction, each action carried out, becomes knowledge that the computer can learn and use as training for the next occasion [1]. And there have been computational developments within the discipline itself, which have provided significant strength to machine learning. As a consequence of these developments,devices that operated at considerably below-human levels just a few years ago now can perform better than humans in certain particular tasks.A report [2] says Machine learning may indeed be a key enabler for a variety of scientific fields, advancing forward the parameters of science by processing the enormous amounts of data now being produced in fields such as life sciences, particle physics, astronomy, social sciences, and more. For researchers to examine these huge datasets, recognizing previously unforeseen trends or extracting unusual observations, machine learning may become a key tool.

In this paper, we address the analysis state of machine learning, discuss what it is, why it has succeeded in improving deep rooted approaches of conventional neural networks, and, most significantly, how you can attempt to incorporate deep learning into research activities to tackle both new and existing problems and develop improved, wiser user devices and applications. This research is an analysis of this type of data processing that allows computers to learn and do what is intuitive to humans, i.e. to learn from the past. It includes the basics of machine learning that answers what, how and why, the meaning, applications and implementations. To recognise and check its potential in bioinformatics, the tech framework of machine learning is addressed. The purpose of this study is to provide perspective into why machine learning seems to be the potential.Since it's an emerging new technology and this technology is not known to most people. We bring forward some important postulates of this conception with our research work.

II. MACHINE LEARNING

Machine learning is a branch of artificial intelligence that gives systems the ability to learn automatically without being specifically programmed or without human interference, and to develop themselves from experience. The main purpose is to allow computers learn from experience automatically.

The system will learn to differentiate between the pattern and make a relatively good prediction due to the large amount of data produced by computers, sensors and social media users [3]. The learning factor also produces systems that can be flexible and, once they have been implemented, continue to enhance the precision of their outcomes [4]. In 2015, for instance, researchers developed a machine learning system that, in a limited range of vision-related tasks, exceeded human capacities on identifying individual handwritten digits [5].

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue III Mar 2021- Available at www.ijraset.com

A. Requirements for Building good Systems for Machine Learning

1) *Data:* To predict the performance, input data is required.
2) *Algorithms:* Machine Learning to determine data patterns relies on some statistical algorithms.
3) *Automation:* It is the ability to automatically make devices run.
4) *Iteration:* The entire process is iterative, i.e. process repetition.
5) *Scalability:* In size and scale, the power of the computer can be increased or decreased.
6) *Modeling:* The models are generated by the modelling process according to demand.

B. Machine Learning Techniques are Grouped into Several Groups. [1]

1) *Supervised Learning:* The machine is provided with input and output along with feedback during the training in this process. The computer's accuracy of predictions during training is also evaluated. The primary objective of this training is to teach computers how to map input to output.
2) *Unsupervised Learning:* In this case, no such training is provided, leaving computers to find the output on their own. Transactional data is often applied to unsupervised learning. It is used in more difficult assignments. To arrive at any conclusions, it utilises another iteration technique known as deep learning.
3) *Reinforcement Learning:* Three components are used in this type of learning: agent, environment, behaviour. An agent is the one who perceives the world, the one with whom an agent communicates and behaves in that environment. In reinforcement learning, the primary purpose is to find the best policy possible.

C. Machine Learning Functioning

Machine learning makes use of methods equivalent to data mining processes. In terms of target function (f), mapping the input variable (x) to an output variable, machine learning algorithms are defined (y).

It is possible to represent this as:
$Y=f(x)$

There is also an e error that is distinct from the x input variable. Thus the equation's more generalised form is:
$Y=f(x) + e$, $y=f(x)$

The mapping from x to y for predictions is performed on the computer. To make the most reliable predictions, this methodology is known as predictive modeling. This function has different assumptions.

D. It Operates on the 3 Criteria that Follow

Finding flaws in algorithms for machine learning.
Develop methods for testing these possible vulnerabilities.
Implementing these preventive measures with a view to improving algorithm security.

E. Machine Learning Advantages

1) *Prediction:* Decision-making is quicker. By prioritising repetitive decision-making processes, machine learning offers the best possible results. These methods construct a database of training samples and evaluate the results with other cases in the database compared to fresh data is provided as input using a similarity measure to find the closest match and make the judgement [6].
2) *Adaptability:* Machine learning offers the ability to quickly adapt to modern, evolving environment conditions
3) *Innovation:* Machine learning uses sophisticated algorithms that enhance the overall capacity for decision making. This helps to build new services and models for companies.
4) *Insight:* Machine learning helps to understand unique data patterns and what particular steps can be taken based on them.
5) *Business Development:* The overall business process and workflow of machine learning would be quicker and therefore lead to the overall growth and acceleration of the business.
6) *Accurate Results:* With machine learning, the accuracy of the result will be increased with less probability of error.

F. Software for Machine Learning

IBM's 'Watson' uses machine learning in various ways. One of these is natural language processing – the form of machine learning which allows computers to process written or verbal information – which Watson uses to extract information from the vast collection of published research papers and case reports, and use this information to recommend treatment options [7].

1) Tensor Flow: It is a library of open source applications for machine learning by Google. [2] TensorFlow, along with documentation, tutorials and other support tools, offers a library of numerical computations.
2) Amazon Web Services: Amazon has released developer toolkits along with applications ranging from interpretation of images to facial recognition. AlsoAmazon is developing the use of delivery drones, with the first successful delivery taking place in December 2016 [9].
3) *Caffe:* It is a platform for deep learning and is used in the field of voice, vision and expression in different industrial applications.
4) *Veles:* This is another deep learning framework written in the language of C++ and uses python to communicate between the nodes.

III. DEEP LEARNING

Deep Learning is part of the wider field of machine learning and is focused on learning about data representation. It is based on artificial neural network interpretation. The algorithm for Deep Learning uses multiple computing layers. The output of the previous layer is used by each layer as an input to itself. An algorithm or an unsupervised algorithm may be supervised by the algorithm used. Deep Learning is specifically created to manage complex input and output mappings. It is basically designed to work like human brain. It is another hot subject along with machine learning for the M.Tech thesis and project. One of said most popular deep learning algorithms are: [8]

Convolutional Neural Network (CNN)

In the last few years, we have witnessed an exponential growth in research activity into the advanced training of convolutional neural networks (CNNs) [10].

A. Deep Learning Advantages

Eliminates unnecessary costs by identifying flaws and failures in the system.

Deep learning helps to recognise defects that have been left untraceable in the system. It detects defects that are otherwise difficult to detect.

Deep learning can inspect unusual shapes and patterns that are difficult for machine learning to identify. Deep learning can inspect irregular shapes and patterns.

It is good at pattern recognition problems and is data-driven [11].

IV. BIOINFORMATICS MACHINE LEARNING

The word bioinformatics is a mixture of two bio-informatics concepts. Bio means biology-related, and informatics means records. Thus, bioinformatics is a field that uses a quantitative and statistical approach to deal with the analysis and understanding of biological data. As biological data grows exponentially, it is important to pay attention to the efficient storage and information management, as well as to extract useful information from that data. In order to turn this diverse data into meaningful information, suitable analytical techniques must also be used. These analytical tools and methods, or you can say machine learning tools, allow more detailed data to be grasped and provide knowledge in terms of testable models by which we can obtain system predictions. There are several biological areas in which machine learning tools can be used to retrieve information, followed by neural network applications in bioinformatics;

1) In the recognition of gene area coding
2) Problems in the recognition of genes
3) Recognition and analysis of signals generated from regulatory sites
4) Detecting sequence, grouping, and characteristics
5) Genetic and Genomic Data Expression
6) Processing an image and signal

Machine Learning find its application in the following subfields of bioinformatics: Genomics, Proteomics, Microarrays, System Biology and Text mining.

V. LITERATURE REVIEW

We observe that supervised learning offers broad opportunities in machine learning in the perspective of training and testing machines and is quite intriguing topic for further research. Supervised learning is further categorized into numerous topics in the viewpoint of research. Some of them are listed in table 1.

A. Regression and Classification are Broad Categories of Supervised Learning

In regression a single output value is generated in regression using training data. This value is a predictive interpretation that, after evaluating the strength of correlation between the input variables, is calculated. Regression, for instance, can help to predict the cost of a property based on its location, size, etc.

Classification allows the data to be divided into groups. You may use classification to decide whether or not a person is a loan defaulter if you are thinking about extending credit to an individual. When the supervised learning algorithm labels input data into two separate groups it is called binary classification. Multiple classifications suggest the classification of information into more than two groups.

Table 1 : Potential fields of supervised learning that can be improvised in training and efficiency.

Supervised learning technique	Area of research	Potential fields that can have improvised training and increased efficiency
Classification	Bioinformatics	Use of deep learning to train models to give better results.[12]
Regression	Speech recognisation	Security systems that use voice recognition to grant customers access to their accounts.[13]
Classification	Banking	Robot bank-tellers that use machine learning to respond to customer queries.[14]
Classification	Image recognisation	Human diseases can be detected through various medical imaging techniques like MRI[15]
Classification	Language processing	Natural language processing techniques to analyze text and detect inappropriate statements which are indicative of phishing attacks[16]
Classification	Medical	To predict how well patients will respond to different drugs used in treating depression[17]
Regression	Finance	Machine learning algorithms can help address monetary policy-making [18]
Regression	Weather prediction	To optimise the temperature requirements of places by predicting temperatures like Google DeepMind [19]

VI. CONCLUSION

This paper describes a description of the mechanism of machine learning. The different machine learning algorithms based on techniques of machine learning are also described. Examples of applications for machine learning and the requirements are presented. Machine learning is currently automating repetitive technological tasks in many areas, but machine learning applications in these aspects are diversifying, from machines offering legal advice to health tracking medical apps using machine learning.This research presented the findings of study to investigate the most well-known machine learning and deep learning methods used for classification. For organising the data, classification is very necessary so it can be easily accessed. In several different application fields, from banking to medicine, from business to bioinformatics, these methods have gained a lot of significance.

VII. FUTURE WORK

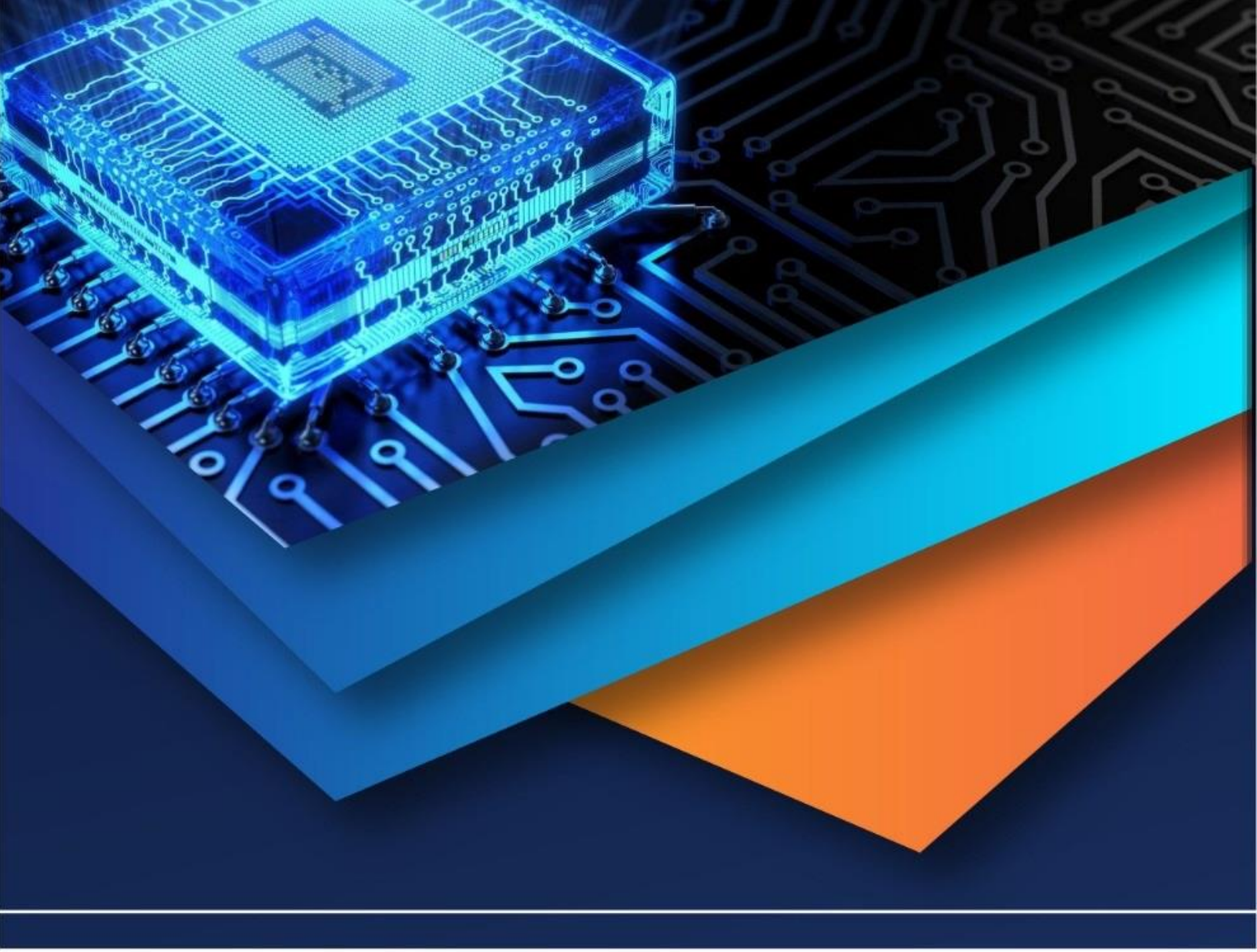
We propose two potential categories by this review in which further research work is attainable.

A. Enhancing performance of Convolution neural network can be achieved by perform a grayscale normalization to reduce the effect of illumination's differences. Thus, making CNN work faster. Keras model needs an extra dimension in the end which correspond to channels. Our images are gray scaled so it use only one channel.

B. Implementation can be done for Gene Expression analysis in which acute myeloid leukemia patients can be classified into classes to help diagnose the disease with two datasets containing the initial (training samples) and independent (test samples) datasets.

REFERENCES

[1] JafarAlzubi, AnandNayyar and Akshi Kumar, Machine Learning from Theory to Algorithms: An OverviewJournal of Physics: Conference Series*: Volume* 1142 , Dec 2018.
[2] A report by Royal Society, April 2017, “Machine learning: the power and promise of computers that learn by example “, Page 41 ISBN: 978-1-78252-259-1.and Information Technologies, Vol. 7 (3) , 2016, 1174- 1179.
[3] PariwatOngsulee 15th International Conference on ICT and Knowledge Engineering (ICT&KE), nov 2017,**DOI:** 10.1109/ICTKE.2017.8259629.
[4] Shalev-Shwartz S, Ben-David S. 2014 Understanding machine learning: from theory to algorithms. Cambridge, UK: Cambridge University Press.
[5] Markoff J. 2015 A learning advance in artificial intelligence rivals human abilities. New York Times. 10 December 2015. See https://www.nytimes.com/2015/12/11/science/an-advance-in-artificial-intelligence-rivals-human-vision-abilities.html.
[6] Sandhya N. dhage, Charanjeet Kaur Raina, “A review on Machine Learning Techniques”, March 16 Volume 4 Issue 3 , International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), ISSN: 2321-8169, PP: 395 – 399.
[7] IBM. IBM Watson Health. See http://www.ibm.com/watson/health/oncology.
[8] A report by Royal Society, April 2017, “Machine learning: the power and promise of computers that learn by example “, Page 41 ISBN: 978-1-78252-259-1.and Information Technologies, Vol. 7 (3) , 2016, 1174- 1179.
[9] Condlifee, J. 2016 An Amazon drone has delivered its first products to a paying customer. MIT Technology Review. See https://www.technologyreview.com/s/603141/an-amazon-drone-has-delivered-its-first-products-to-a-payingcustomer
[10] Joe Lemley; ShababBazrafkan; Peter Corcoran, Deep Learning for Consumer Devices and Services: Pushing the limits for machine learning, artificial intelligence, and computer vision.IEEE Consumer Electronics Magazine (Volume: 6, Issue: 2, April 2017),**Page(s):** 48 – 56.
[11] Shekhawat V.S., Tiwari M., Patel M. (2021) A Secured Steganography Algorithm for Hiding an Image and Data in an Image Using LSB Technique. In: Singh V., Asari V.K., Kumar S., Patel R.B. (eds) Computational Methods and Data Engineering. Advances in Intelligent Systems and Computing, vol 1257. Springer, Singapore. https://doi.org/10.1007/978-981-15-7907-3_35.
[12] H. Gupta and M. Patel, "Study of Extractive Text Summarizer Using The Elmo Embedding," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 829-834, doi: 10.1109/I-SMAC49090.2020.9243610.
[13] Shailja Joshi and Mayank Patel. “Natural Language Processing for Classifying Text Using Naïve Bayes Model.”, PAIDEUMA JOURNAL, Volume 13, Issue 10, ISSN NO : 0090-5674, Page No: 72-77, DOI : 11.3991/Pjr.V13I10.85307
[14] NeelamBadi, Mayank Patel and Amit Sinhal (2019). The Role of Fuzzy Logic in Improving Accuracy of Phishing Detection System. International Journal of Innovative Technology and Exploring Engineering, Volume-8 Issue-8, ISSN: 2278-3075, pp.3162-3164.
[15] Menaria H.K., Nagar P., Patel M. (2020) Tweet Sentiment Classification by Semantic and Frequency Base Features Using Hybrid Classifier. In: Luhach A., Kosa J., Poonia R., Gao XZ., Singh D. (eds) First International Conference on Sustainable Technologies for Computational Intelligence. Advances in Intelligent Systems and Computing, vol 1045. Springer, Singapore. https://doi.org/10.1007/978-981-15-0029-9_9
[16] [16] K. C. Giri, M. Patel, A. Sinhal and D. Gautam, “A Novel Paradigm of Melanoma Diagnosis Using Machine Learning and Information Theory," 2019 International Conference on Advances in Computing and Communication Engineering (ICACCE), Sathyamangalam, Tamil Nadu, India, 2019, pp. 1-7, doi: 10.1109/ICACCE46606.2019.9079975.
[17] PReDicT (Predicting Response to Depression Treatment). See http://predictproject.p1vitalproducts. com/
[18] Condon, C. 2016 Quest for robo-Yellen advances as computers gain on rate setters. Bloomberg. See https://www. bloomberg.com/news/articles/2016-05-24/quest-for-robo-yellen-advances-as-computers-gain-on-rate-setters
[19] DeepMind. 2016 Press release: DeepMind AI reduces Google data centre cooling bill by 40%. See https://deepmind. com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-40/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY