Research Journal of Recent Sciences ISSN 2277-2502
Vol. **8(1),** 14-29, January **(2019)** *Res. J. Recent Sci.*

Review Paper

A critical review on machine learning algorithms and their applications in pure sciences

Sovesh Mohapatra[1*] and Shreyans Babel[2]
[1]Seedling Modern High School, Jaipur, Rajasthan, India
[2]Seedling Modern Public School, Udaipur, Rajasthan, India
soveshm@seedlingschools.com

Available online at: www.isca.in, www.isca.me
Received 28th September 2018, revised 10th December 2018, accepted 26th December 2018

Abstract

Today, it is difficult to think about solving any set of problems without the use of Artificial Intelligence. This has grown tremendously across several fields starting from the Management to the Life Sciences. The use of AI has made life simpler and better. Today, its use in the process of high – throughput screening has provided us with several types of advantages such as saving resources, expenditures and many more. The method of Machine learning has led to minimizing the errors involved with the co-relation of different kinds of attributes. Most importantly, it has transformed the Edisonian approach of hit and trial method into a way with full of logic and simulations. Today, using different simulation we can predict several required properties and the after effects of many materials, which led us to save a lot of resources. Here in this review article, we have explicitly presented the machine learning types, different algorithms and along with their uses in several different fields.

Keywords: Artificial intelligence, machine learning, algorithms, supervised learning, unsupervised learning.

Introduction

With the increment of the use of data around the world, Machine learning has become more of a necessity than a helpful source. It empowers the computer systems to get into a method of self-learning without being explicitly programmed. At the point when presented with new information, these computer programs are empowered to learn and train themselves according to the data; the more is the data, the better the computer understands. Machine learning has advanced from the investigation of procedure learning hypothesis and pattern recognization. Utilized in the era of innovation, it is one of the best procedures till date for the world to exercise as the ultimate objective is to predict or analyzeimportant data by forming a couple of trained models through the assistance of algorithms[1–5]. Typically, people are susceptible to creating mistakes throughout the analyses, or while attempting to determine relationships between multiple datasets.

This makes them troublesome to discover answers for specific problems[6,7]. Machine learning is an intense device to enhance the precision and productivity ofthese works. For example, your email inbox appears like a far-fetched put for Machine learning, yet the innovation is to a great extent working on unexpectable places. According to Tom M. Mitchell, A computer program is expected to perform specific tasks in a time-bound manner, learn from its experience from the ever-evolving dataset and improve its performance with time, concerningcertainpre-defined key performance parameters[2,4,8].

Machine learning types, namely categorised as Supervised Learning, Unsupervised Learning, and reinforcement learning, are precisely talked about in this paper. In earlier stages of machine learning, companies which benefitted the most were the information firms and online companies who took a big step and grabbed the opportunity of big data before others. As the use of Machine Learning progressed, several large datasets were analyzed in a short span of time, and return gave out valuable outputs[5,8,9]. At, present, nearly any industry can benefit and profit from it. Having great benefits, it also has some of the challenges which people face or will face with the growth of it. The most significant drawback of machine learning is that it sometimes overfit with the small dataset and produce results that seem correct but are not and also there is a huge problem in getting such large datasets to train the models effectively[3].

Here in this paper, we have reviewed the several numbers of Machine Learning approaches, their algorithms and also the real-life applications solved using Machine Learning.

Types of machine learning algorithms

The Machine Learning is broadly divided under three main types, i.e., Supervised Learning, Unsupervised Learning and Reinforcement Learning.

Supervised Learning: In this type of learning, Algorithms tries to map the function (y=f(x)) by the help with certain provided parameters, namely input variable (x) and output variable (y).

Our main goal is to map a function so accurately so that we can predict a new value of the dependent variable(y) by the given value of the independent variable(x)[10,11]. To get a clear understanding, I'll explain to you with the help of a real-life example. Suppose there are a student and a teacher. The student tries to answer every question, and when he or she is wrong the teacher correction. Learning stops when the student achieves an acceptable level of performance. Here, the teacher is the data set, the questions are the data of the data set, and the student is the algorithm regarding Machine Learning[12].

Supervised learning models do have a acceptable amount of benefits over un supervised learning, Nevertheless, They are not without the restrictions of their own. These frameworks make the decisions which are in the assumptions of the alternatives provided to it by the people[6,13,14] Supervised Learning additionally discovers trouble in overseeing new information. Anyways, in any case its benefits outweigh the costs of its downfalls[15,16].

Supervised learning provides a diverse range of algorithms to choose from: such as- Support Vector Machines, Linear regression, K-Nearest Neighbour Algorithm, Logistic regression, Decision trees, Naive Bayes.

This Supervised learning is again divided into two sub-categories: Regression - In this method, the output is detected by estimation through the model generated on the connection within the two parameters, x, and y, titled as feature and model mutually in the terminology of machine learning[17]. The main motive for regression is to generate an equation that outputs the value of y when given the value of x. As, a straight-line graph, the equation is in the form y= a + bx, where ‘a’ is the y-intercept, and ‘b’ is the slope. There are many kinds of algorithms involved in it such as Linear Regression, Random Forest Trees, Support Vector Regression[5,18].

Classification: It is an approach to find discrete output variables ’y’ by mapping the function in ‘x’ and ‘y.’ Classification is considered an example of supervised learning. It learns from the set of correctly practical data sets. It maps the function and predicts the category or information of the asked perception[7,19].

For example, an email of text can be seenas: “spam and not spam” an email may be given the probabilities of 20% as being “spam” and 80% as being “not spam”. We select the “not spam” label to get the highest predicted outcome.

In short, Classification either predicts clear-cut class names or characterizes information (build a model) in values of the preparation set and the data (class names) in ordering characteristics and utilizations it in arranging new information. There are some classification models. Classification models include logistic regression, decision tree, random forest, gradient-boosted tree, Naive Bayes and many more[6].

Unsupervised Learning: In this type of learning, algorithmstry to map the function (y=f(x)) by the help of a certain parameter, namely input variable (x), where output variable is not provided[16].

Our main goal is to model the function, without the help of the output variable, to learn more about the data. To get a clear understanding, I'll explain to you with the help of a real-life example. Suppose there is a student but no teacher this time. The student tries to answer every question by himself. Since there are no incorrect or correct answers; the student doesn’t depend upon the teacher for correction. Here, the student is the algorithm regarding Machine Learning. Unsupervised machine learning implies to reveal beforehand obscure examples in information, yet more often these examples are poor approximations of what directed machine learning can accomplish[13,19]. Furthermore, since you don't recognize what the results ought to be, there is no real way to decide how precise they are, making directed machine adapting more appropriate to certifiable issues. The best time to utilize unsupervised machine learning is the point at which you don't have information on wanted results, such as deciding an objective market for an altogether new item that your business has never sold. In any case, Supervised learning is the perfect strategy if in case you are attempting to make your present customer base more inclusive[19–21].

Because of the absence of output variable, Unsupervised Learning can be used with regression and classification leaving it impossible to train the algorithm in a way that we use to. Instead, it is classified into two sub-categories:

Clustering- It is a type of unsupervised learning in which its way of working is to divide the data points into the groups similar to it from the data points which are different to it[22].

Association-An Association reveal connections between apparently unconnected information from a large database[23]. Reinforcement Learning - It is a segment of Artificial Intelligence, which empowers machines and computing authorities to subsequently adopt a perfect technique inside a specific setting, remembering the ultimate objective is to get the greatest performance. Reinforcement Learning is characterized by a particular sort of issue, and every one of its answers is classed as Reinforcement Learning calculations[24,25]. In the issue, an agent has to choose the best activity to decide in light of his present state. Reinforcement learning alludes to objective arranged algorithms, which figure out how to achieve a mind-boggling (objective) or expand along a specific measurement over numerous means; for instance, augment the focuses won in a diversion over numerous moves[26]. They can begin from a clean slate, and under the correct conditions, they accomplish exceptional execution[27,28]. These calculations are punished when they settled on the wrong choices and compensated when they make the correct ones – this is reinforcement[27]. Due to the distinctiveness of the issue determination, the feasibility of utilization of Reinforcement Learning is deep[24,29].

Terms in machine learning

Models: A model is the y-variable in the simple linear regression graph, which we predict. It can be the future price of rice, meaning of a video clip, kind of bird shown in the picture, or just anything. When we try to dig deeper into the topic, we find that Machine Learning models are nothing, but the model brand that is developed during the training. The process of learning is through provided ML algorithm and several datasets. ML models are used to get predictions for which we do not know to the outcome of the new data.

Machine learning works by finding a connection between the model and feature. We do this by demonstrating a protest (our model) a cluster of cases from our dataset. Every illustration characterizes how each component influences the label. We allude to this procedure as preparing our model.

Features: A feature is the input variable—the x variable simple linear regression. An easy machine learning project would possibly use one feature, whereas a lot of subtle machine learning project might use numerous options. Features are a column of knowledge given as the input[1]. They're conjointly referred to as attributes or would possibly typically be referred as dimensions. A specific data set can contain a lot of features in it. Picking useful, separating and autonomousfeatures is an urgent advance for compelling algorithms in design recognition, grouping and regression[12,30]. Features are generally numeric, yet auxiliary features, for example, strings and charts are utilized in syntactic example acknowledgment. The idea of "feature" is identified with that of illustrative variable utilized in factual procedures, for example, linear regression[31,32].

Algorithms in machine learning

An algorithm is some definitive computing method that seizes several values, or collection of values, called input and generates some results, called output. An algorithm is thereby a chain of computational measures that alter the input into the output. We could further regard an algorithm as a device for a solution of a well-stated computational problem. In general terms, the declaration of the issue pinpoints the wanted relationship between input and output[33]. The algorithm explains a definite computational method for accomplishing that relationship between input and output[34].

Algorithms which are planned to resolve the exact issue may differ drastically in their effectiveness. For example, we have two algorithms for sorting- merge sort and insertion sort[35]. Due to the difference in the rate of working, both take different time intervals to sort the same data even if they perform the same function. For smaller data-sets insertion sort is relatively very faster than merge sort, but with a much larger data, merge sort is exceptionally fast as compared to insertion sort[8,36].

If we talk about the total system performance, selecting an efficient hardware is as crucial as selectingan effective algorithm. Similarly, as we observe brisk advancement in technology algorithms also share a similar kind of advancement. Unbelievably, algorithms are also crucial for the contemporary computers even after so much advancement in the technology[32]. Although some applications do not heavily depend upon algorithms, there is a slight extent of dependency upon it. Algorithms are at the core of most technologies used in contemporary computers as well as advanced computers. And as the technology advances, the uses of algorithms are observed even more with larger problems to solve[37].

There are many kinds of algorithms used in the Machine Learning. Here are some examples of the algorithms:

Linear Regression: It is a direct way to deal with displaying the connection between a dependent variable and at least one illustrative factor. Its analysis is totally based on predictions. The relations in linear regression are made through linear functions whose unknown parameters are found by using the data[38,39]. The least complex type of the regression condition with one needy and one free factor is characterized by the equation y= c + b*x, where y= dependent variable, c= constant, b= regression coefficient, and x= independent variable. The center thought is to acquire a line that best fits the information, which is called Line of Best Fit[40]. The best fit line is the one for which add up to prediction error (data points) are as little as could reasonably be expected.Error is the separation between the points to the regression line[41]. The variable we are constructing our forecasts concerning is known as the predictor variable and is alluded to as X. At the point when there is just a single indicator variable, the forecasting technique is called simple linear regression. In simple linear, the theme of this segment, the forecasts of Y when plotted as a component of X frame a straight line[41–43].

From its name, we can conclude that it uses the approach of regression and not classification. This is becauseclassification is for more discrete data sets by learning through data provided by those sets whereas regression works by a general trend which is followed throughout the process[18,42].

Linear regression is divided into two sub-categories on the basis on number so independent variables present: i. Simple Linear Regression- It contains only one independent variable to predict the dependent one[42]. [Equation = a + bx]. ii. Multiple Linear Regression- It contains at least 2 independent variables to predict the dependent one[18]. [Equation = $b_0 + b_1x_1 + b_2x_2$].

Logistic regression: It is the proper regression investigation to lead when the reliant variable is dichotomous (paired). The strategic(logistic) regression is an insightful exploration alike entire other regression examinations[44]. Logistic regression is utilized to portray information and to clarify the connection between one dependent paired variable and at least one ostensible, ordinal, interim or proportion level autonomous variables. Logistic Regression predicts the likelihood of a result that can just have two qualities (i.e., a dichotomy)[45,46]. The forecast depends on the utilization of one or a few indicators (numerical and clear-cut). Rather than foreseeing precisely 0 or 1, logistic regression produces a probability—an incentive somewhere in the range of 0 and 1[47].

Consider a situation where we have to group whether an email is spam or not. In the event, if we utilize linear regression for this issue, there is a requirement for setting up an edge in light of which grouping should be possible. Say if the genuine class is dangerous, anticipated constant value 0.4 and the limit value is 0.5, the information point will be delegated not threatening which can prompt genuine outcome progressively[48–50].

From this case, it tends to be gathered that linear regression isn't appropriate for arrangement issues. It is unbounded, and this brings logistic regression into the picture. Their value entirely runs from 0 to 1. Logistic regression is decidedly not a classification calculation all alone. It is just a grouping calculation in the mix with a choice decide that makes dichotomous the anticipated probabilities of the result. Logistic regression is a regression model since it appraises the likelihood of class participation as a (change of a) multi-linear capacity of the features[10,47,51–53].

Likewise, the linear regression, the logistic regression is also sub-categorized into three types[54]:

Binary Logistic Regression: The response with two possible outcomes.

Multinomial Logistic Regression: The response with 3 or more categories without ordering.

Ordinal Logistic Regression: The response with 3 or more categories with ordering.

Naive Bayesian: It is type of classifier which depends on Bayes' hypothesis with the autonomy presumptions among predictors[55]. A Naive Bayesian model is anything but difficult to work, with no entangled iterative parameter estimation which makes it especially helpful for vast data sets. Notwithstanding its effortlessness, the Naive Bayesian classifier frequently does shockingly well and is generally utilized because it regularly beats more advanced order techniques[56,57]. Naive Bayes classifiers are an accumulation of grouping calculations given Bayes' Theorem. Bayes theorem, proposed by Reverend Thomas Bayes serves as the foundation of Naive-Bayes. It is a non-singleton algorithm but a lineage of algorithms where each one of them has the similar objective and function. The elementary Naive Bayes hypothesis is that every feature develops self-supporting and equivalent participation to its every result[10,55,57].

There are several advantages involved in using this algorithm such as its computation speed and reliability. This algorithm can be useful for both the binary classification as well as multi-classification. But one of the main disadvantages which makes it less popular is while determining the several relations between the different attributes[10].

K-Nearest Neighbours: It is a non-parametric technique, which is used for regression and classification in cases of pattern recognition. No learning is necessary because the training dataset itself represent the model[58–60]. Complex information structures like kd trees are used to store heavy information. The input is composed of the k-nearest training examples present in the feature space in both cases-regression and classification[60–62]. The result of the K-NN classification is a class membership, whereas the result of the K-NN regression is the property value of the object.

The appropriate method of choosing the value of K is very important step in this uses of algorithm[63]. To prevent biases, there are several methods evolved to find the value of K such as: i. One way is the hit and trial way in which the user can try various values of k and select out the best one, but this method is very time-consuming. ii. The second method is the leave-one-out cross-validation method, where we remember the best value of k at training time is dynamic KNN.

Being lazy learning algorithm, it comes with many drawbacks too, and two of the most majordrawbacks are its very less effectiveness and its need on the selection for a good value of k, which can turn bad for the program because the value of k is a must as discussed above. It also requires large memory as well as high time complexity[61,63].

Random Forest: This algorithm was first ever created by the Tim Kam Ho[64,65]. It is counted in the category of Supervised Learning. It is an easy and adaptable language which yields vast outcomes almost every time. The decision tree is the basic building block of Random forest[66–68]. In an understandable way of saying, Decision tree works on the principal of the series of questions. A flowchart is produced to minimize our range of answers from which we work our way towards the prediction we want to make[69].

The models of the decision tree are used by random forest in the process for interpretation of accurate results. With no extra prior knowledge, Random forest learns about the framework of the required object through the help of provided dataset bit by bit and creates a flowchart, by the help of questions, where it tries to minimize the error percentage and give the best possible outcome[11,67]. The model adapts any connections between the information (features) and the qualities we need to predict (target).

There are several advantages involved in this algorithm such the use of it in both the cases of Classification and Regression, then the way it handles the missing values in the datasets and finally the most important thing in this is the greater number of trees you put in, the better predictions you get. But there is a big disadvantage, i.e., over fit arises very easily,and very difficult to determine too[67,70,71].

Ordinary Least Squares (OLS): It is type of regression algorithm and is a factual technique for examination, which gauges the connection between at least one free factors and a reliant variable; The approach evaluates the connection by constraining the total of the squares in the quality within the observer and anticipated estimations of the dependent variable arranged as a linear graph[72,73]. Ordinary Least Squares (OLS) is the most well-known estimation strategy for linear models. For whatever length of time that your model fulfils the OLS presumptions for linear regression, you can breathe a sigh of relief realizing that we are getting the ideal assessments[29,72]. Regression is a ground-breaking examination that can investigate different factors at the same time to answer complex research questions. In any case, in case we don't satisfy the OLS assumptions, we most likely won't have the aptitude to trust in the results. The co-efficient of the regression function evaluations should tend to be perfect. They should not be successfully too elevated or too less[74–77]. As such, they ought to be unprejudiced or rectify on average. Knowing that getting a 100% accurate estimate is not possible, our main goal remains to be minimizing the inconsistency between the estimated and actual value.

In the same way as other measurable examinations, ordinary least squares (OLS) regression has hidden presumptions. At the point when these established suppositions for linear regression are valid, ordinary least squares deliver the best approximations. Notwithstanding, if a portion of these suspicions is not valid, you may need to utilize healing measures or utilize other estimation strategies to enhance the outcomes[29,74,76].

Merge sort: It is a type of algorithm which is also known as a divide-and-conquer algorithm, works on the divide-and-conquer principal and is given separating a record into a few sub-records until the point that every sub-record comprises of a solitary component and blending those sub-records in a way that outcomes into an arranged record[78,79]. The merge () function is utilized for combining two parts. The merge (arr, l, m, r) is the key process that supposes that arr[l..m] and arr[m+1..r] are arranged and consolidates the two arranged sub-clusters into one[78].

The conception behind this type of sorting is that –Firstly, we split the uncategorized record into X sub-records. Each sub-record should contain one element. Secondly, taken eighbouring sets of two singleton records and consolidate them to frame a list of 2 components. X will now change over into X/2 arrangements of size. The final step is to repeat the above steps till a lone grouped record is acquired. While looking at two sublists for merging, the primary component of the two recordsis thoughtabout. While arranging in rising order, the component that is of a lesser value turns into another component of the arrangedrecord. This technique is rehashed until both the littler sublists are vacant and the recently consolidated sublist includes every one of the components of both the sublists[78,80,81].

Support Vector Machines–It depends on the idea of decision planes that characterize choice boundaries. A choice plane is one that isolates an arrangement of articles having distinctive class enrollments. Support vector machines break down information utilized for order and regression examination[82,83]. They have supervised learning models which supports both classification and regression. Notwithstanding, it is more typical in classification issues. We plot each data component as a dot in n-dimensional space beside the estimation of every component in this algorithm. By then, we carry out grouping by determination of the hyperplane that isolates the two classes in a remarkable way[84–86].

Support vectors are the information coordinates closest towards the hyperplane. The purpose of the support vector points in the dataset is that whenever evacuated would modify the situation of the splitting hyperplane. Along these lines, they can be viewed as the basic components of an informational collection[87].

A hyperplane is a plane which isolates and groups an arrangement of information points. To be most accurate and get the best possible hyperplane, we want the side of the hyperplane to be correct in keeping in mind that the proportionality between the distance of points from line and assurance of correctly classified data is direct[88,89].

There are several advantages of this algorithm, such as it can be very useful when dealing with small and clean datasets. This algorithm uses a subset of training co-ordinates which makes it more efficient[85]. The disadvantage which carries along is the computational power and time it takes while dealing with large datasets but this has been resolve using the Sequential Minimal Optimization (SMO) algorithm which breaks the dataset into several parts and in every step, it attempts to solve the smallest possible optimization problem atfirst and finally after the entire process is done it rejoins effectively using Osuna's theorem which ensures that it is effectively converged[90,91].

Random Tree: In an easy language, A tree is a linked open-chain graph. A node is a single root of the tree, and the count of nodes determines the size of the tree. The Random tree algorithm can work under both classification and regression problems[66,92]. It is an ensemble of tree predictors that is known as the forest. The classification procedures as per follows: the random trees classifier obtains the input, categorizes it with each tree in the forest, and outputs result that got the most of "votes."In an instance of regression, the classifier's response is seized as standard replies[93]. It is an ensemble learning algorithm that generates lots of individual learners. A random tree is a tree is framed by a random procedure, in mathematics and software computing. Each of the random tree models holds a consisting collection of algebraic methodologies through which parameter requirements are plotted for generation of function equations, which are afterward resolved[93,94]. The portrayal of the information as Tree has the profit contrasted with several different approaches to being significant and effortless to understand. The objective is to from a classification model that estimates the rate of the label based on many input attributes of the Example set. Every internal node of the tree represents one of the input attributes. The count of borders of an interior node is equivalent to the count of possible values of the equal input attribute. Every leaf node portrays a value of the label. Random Trees models are strong when we are dealing with large data sets and counts of fields. Because of the utilization of bagging and field sampling, they are very less apt to over-fitting, and consequently, the outcomes that are observed in testing are probably be repeated when you use fresh facts[10,11,66].

For the training of random tree models, we might want several Input areas and one target area. Target and input fields can remain constant (numeric range) or ungrouped.

There are several more algorithms, but the above mentioned are frequently used in the world of AI. Now the application of these algorithms.

Application of Machine Learning

Completion of repetitive forms: In a database background, a form is a window or screen that comprises several fields or intervals to register data. Each field maintains a field tag so that any user who gazes the form acquires a conception of its contents. In this modernized world, everyfield has its form, which helps to maintain a record of the user who wants to use it, e.g-hotel check-in forms, taxes forms, registration forms. Lately, with the introduction of the internet and the database systems, most of the registration work is done online automatically where the user does not have to carry the form here and there in his hand. However, with the Rapid advancement in technology, almost every site on the internet now owns its form which is the information about the user. Sometimes it gets frustrating for the user because of devoting of a lot of time and effort to have a repeated cycle of filling of the forms everywhere. Also, if there is a large data to be filled by a single individual or a small group, the form fillers are bound to feel tried and make some mistake when doing manually, whereas machineis always right and boost its efficiency and productive[95]. This is where machine learning helps us. Nowadays, due to the high cost of programming community people restricts many from getting an effective form automation system[96].

Hermens and Schlimmer have created an automated form-filling system that studies to foretell its consumers' choices by observation. They employed a progressive version of the decision-tree algorithmtodiscovers rules for predicting the existing entry for each field regarding other fields beforehand fixed. The user could always override the predicted worth, altering the default entry and giving new data for future knowledge. Trials demonstrated that the frame topping understudy spared off to 87% in keystroke exertion and effectively anticipated about 90% of the sections on the shape[97].

Air Traffic Control (ATC): It is the assistance supplied through the people on land, which have the designation to control aircraft from land[98,99].The main role of the designated people internationally is to avert crashes, sort out and assist the stream of air activity, and give data and other help to pilots. Likewise, protected and secure ATC framework is compulsory as its disappointment or false execution may cause severe outcomes, like deaths, financial and environmental damages. As a result, it demands trustworthy methods to be utilized at any price to assure protection, trait, and certainty. Guaranteeing the wellbeing of ATC framework has turned into an essential issue because of increment of air movement and developing advancement[100,101]. Given expanding air activity thickness and moderately set number of airway routes, crash, and pause of flights have expanded. On account of a substantial increment in the limit, new-era ATC frameworks are proposed to be intended to enhance effectiveness and accomplish a required level of security[102]. Even thoughcomputerized support to ATC the framework is accessible yet at the same time, it is intensely reliant upon human collaboration causing mishaps because of the disappointment of communication and basic leadership.

Professionals confront a huge amount of issues in communicating between themselves about the sensory Engine abilities required to fly or land even a minor plane. In this type of situation, a knowledge-based-system would be beneficial[97,102,103].

Intrigued by the same thoughts, four researchers Sammut, Hurst, Keizer, and Michie collected information and pieces of evidence one pilot's movements and actions to create a flight simulator. After dividing the data into different chores- such as taking off, landing- they used decision tree algorithm, resulting in an application where the model is given the purpose of actions for a given task and sensor readings, and it creates a perfect replica of a professional[97]. The introduction of machine learning in the field of flight control has brought us with a lot of unconditional benefits like raise incomes, trim expenses and give travelers a more customized travel adventure. There is a handful of opportunities for machine learning and artificial intelligence to grow in the field of flight control which hasn't been explored that can bring a great boost to our traveling lives[104].

Signal Processing: Signal implies data and processing implies task[105]. Signal Processing tells how information in the shape of the signal is worked or altered to get the needed signal and how the framework processes this signal[106]. Signal processing is an extremely broad field. Any data created on this earth is a signal[107]. It could be a picture, a simple voice or advanced signal of any processor and so on. The extraction of data from intricate signals within the existence of noise by transformation of the signals or using different processing algorithms to improve the signal quality or encrypting the data for safety reasons and significant other reasons, each one of traits goes underneath the Digital signal processing (DSP)[108]. An age back, the focal query of signal processing was the manner by which we can break down signals - that is, the way we can extricate data from a finesignal. What the era asked for was a decent contradiction, and in the late 1980s, they got it, through individuals like Albert Benveniste[109,110]. They kind of went the other way, saying that:

We should demonstrate the signals, as opposed to examining them, and got the data in the form of opening branches with tree-like structure called as hierarchical modeling problem. Now, Because of the opening branches statement, we know that application of machine learning would be used here. The brief response is that both areas have delivered proposals to critical issues with the help of each other[111].

One of the signs of Machine learning in the course of the most recent decade is an enormous expansion of models that are to a great degree high-dimensional (data). Also, Signal processing systems have been progressed quickly as of late and have discovered numerous applications in relatively every field of innovation. During its advancement, it began from adaptive signal processing (1975) to machine learning. From machine-learning, it transitioned its way towards neural network, and presently, it is named with artificial intelligence, which becomes a rapidly growing field today[112].

Self-driving Cars: They are also called an autonomous, or a driverless car is a car which includes the capability of perceiving its surroundings and guiding itself less or any human interference[113–115]. At the present day, the machine learning algorithms are widely utilized to discover the answers to diverse challenges emerging in the manufacturing of self-driving cars. With the inclusion of sensor data processing in a car, it is vital to upgrade the uses of machine learning to execute new tasks. Along with its self-driving features, these cars are also self-sustained with features like medical aided functions or driver speech and gesture reorganization. They also have radar, GPS systems, and computer technologies along with sensory invention to navigate the path to the destination[116].

The most significant positive outcome we get from self-driving Cars is the fact that it reduces road accidents[117]. It provides increased mobility to the person inside the car as well as lowest fuel consumption and saves travellers from a long distance drive. All these benefits provided a diverse opening for long-term benefits like reducing expenses, etc. One of the most significant jobs by machine learning algorithm is to continuously read the environmentand predicting the changes which may result to enact there[118].

Still, if we consider the long distance goals, the long-term benefits outweigh the cost it is paying or has a chance of paying.Theselong-term benefits also give rise to several things which may not allow self-automated cars on roads[35].Unresolved problem like problems with safety, technical issues, the risk of loss of privacy and security concerns create a lack of employment in the transport industry. Still, now, there is a lot of vast scope in this field to improve the condition of cars with reducing all its limitations[119].

Natural Language Processing

Natural language processing (NLP) is a region of computer science worried about the co-operations among computers and human (normal) dialects. NLP is a part of artificial intelligence (AI)[120,121]. There is a colossal measure of data put away in free content documents, similar to patients' restorative records, for instance. Before deep learning-based NLP models, this data was blocked off to computers and couldn't be concluded in any sort of efficient way. In any case, NLP enables examiners to filter through enormous troves of free content to discover significant data in the records. The improvement of NLP applications is testing since computers customarily expect people to "talk" to them in a programming dialect that is exact, unambiguous and exceptionally organized, or through a set number of unmistakably spoken voice directions. Human talk isn't constantlyexact. If we observe the correlation between NLP and machine learning, we observe that there is a common overlap between the two, where machine learning helps NLP to understand the dialect[122,123].

Up to the 1980s, most normal dialect handling frameworks depended on complex arrangements written by hand[124]. Beginning in the late 1980s, in any case, there was a turn in regular dialect preparing with the presentation of machine learning calculations for dialect handling. Machines being prepared on voice pursuit will order the up and coming age of human interface in our everyday lives, telling your washing machines to begin, or teaching your music i-pod to start your most loved track. This was because of the constant increment in computational power. Decision tree, one of the oldest machine algorithms, produced schemes of strict if-then regulations same as existing hand-written rules. It enhanced natural language processing in many ways such as increasing the speed of Text Classification, Language Modelling, Speech Recognition in the speech[121,124,125]. The future of NLP stands ahead to which we are reducing the distance from future technologies only possible due to NLP. These include advance chat-bots, invisible user-interfaces, and smarter browsing systems[126].

Stock Forecasting: Stock exchange forecasting is a technique that helps the people by making a future estimation of a stock organization or other money-related tasks. If by any luck prediction gets true, it could unimaginably benefit investors concerned with the estimation. Exploration has been done on Stock market forecasting by analysts of various fields including the business and software engineering[23,62]. They also have attempted diverse methodologies to forecast including distinctive strategies and algorithms[127].

As we know, something is better than nothing, and in this case that something exceptionally boosts the chances of earning more and losing less. Machine learning, even if does not affect the stock values, has the ability to persuade or change the mind of any stock investor by predicting its future. Artificial Neural systems (ANNs) are the most usually utilized procedure for stock forecasting[128]. However, most of the time, ANNs experience the ill effects of an over-fitting issue because of the extensive number of parameters to settle, and the little former input about the relevance of the contributions to the analyzed issue[83]. Because of the ability to estimate universally ideal outputs, support vector machines come into action, which could escape the restrictions encountered by the artificial neural networks[127].

Machine Learning provides someimportant advantages as compared to other programs. They can be very useful to manipulate the human brain, which contains importance of money[129]. Its increases the mobility of the person by giving him loads of market's history and adapting trends to automate his monitoring and consulting, which provides a significant head-start from the people not using them, and according to the balance of probabilities, the greater number of entries to any field of stock market would provide you with more opportunities, and finally extra profits[84].

Weather Forecasting: Forecasting is the method of making logical conclusions of the future generally based on elapsed and current evidence or by investigation of patterns throughout the past[130–132]. We, in this part, will specifically talk about weather forecasting. In easy language, weather forecasting – an application of computer science-predicts the climate of the earth at a specific time and area[130,131,133]. Trying since the 19th century, human beings attempt to forecast the weather by collecting the present condition of the environment at a given site and predict how the surroundings will alter concerningcollect data. From observing cloud patterns to use of electric telegraphs for the prediction of weather, the most approved approach for the weather forecast today is through machine learning and artificial intelligence[134,135]. To classify, there are many types of algorithms, but the most important for weather forecasting is the linear regression. Based on the previously known data, the models of the algorithms are trained[8,136].

They can get diverse entries of input like the minimum and maximum temperatures, humidity levels, basic atmospheric pressure, etc. We will take ¾ of the dataset for teaching the model, and the left ¼ will be used as test data. Neural Network is broadly utilized because of its capacity to catch to capture non-linear dependencies of past climate patterns and future climate conditions among all models. To finish up, Machine Learning and Artificial Intelligence have enormously changed the worldview of Weather estimating with high exactness and productivity. Furthermore, inside the following couple of years, greater headway will be made utilizing these advancements to precisely foresee the climate to stop catastrophes like tropical storms, Tornados, and Thunderstorms[137–139].

Drug Discovery: In the field of biology, chemistry, and pharmacology, Drug discovery is a procedure which uncovers modern drugs, which are supportive in the therapies for curing, treating or preventing a disease[140–142]. Formerly, substances were screened for organic movement without prior understanding of the organictarget[140,143]. Following it, small atoms were combined to explicitly focus on a known medical pathway, maintaining a distance from the mass screening of banks of preserved compounds. Being the most used approach today; individual protein clones made it possible to screen giant chemical libraries[144]. Drug discovery methods are expensive processes which involve large donations from pharmaceutical organizations as well as national authorities.

However, with the introduction of machine learning in drug research, there is a splendid possibility for enhancement of both approaches and consequences. Offering huge openings, machine learning is more efficient and easier way to access and comprehend huge quantities of organic data. These technicalities, drastically upgrade the rate and sum of data that can be grouped about the effects of organic compounds[145]. Moreover, machine learning is broadly utilized in target screening, such as by rendering machine vision data to discover and find target sites on cells; e.g., to predict the side-effects of new drugs[146]. In chemical science, machine learning plays a significant part in forecasting diverse compound characteristics, e.g. physical traits along with solubility, virtual screening for various targets, large-scale models for assessing taste or promiscuity of compounds, among others[65,82,147,148].

The drug discovery procedure with machine learning is made on progress in fundamental scientific information, with unspoken curiosity in maximizing both modernism and efficiency[149]. One of the key regions for future is making models that cover various phases of revelations. There is a machine perusing side of the story where algorithms read all the writing and concentrate on the data that is important[150]. Over that, algorithms that are profoundly coordinated with one another will additionally advance into drug revelation to produce more thoughts, making novel extracts that can be utilized to treat sicknesses. In future, these frameworks will be at a meta level, working together and cooperating to tackle any issue in the field of drug development[151].

Diagnosis of Different types of Cancer: Cancer is a collective group of several distinct diseases which can originate in nearly anyplace of the body[152,153]. Cells are the basic foundation of any life on earth[154]. They grow, multiply, and -if get old or damaged- die, resulting in the replacement of new cells at the death site[155,156]. Interference of genetic changes during this process may cause cancerous tumor that result in the growth of cancer, which then grows and spread throughout the body. There are over 100 types of cancers that can affect humans[157].

There has been a huge focus in the area of healthcare for the detection of cancer at an initial stage. A constant development identified with cancer growth has been executed. Researchers have connected distinctive techniques, for example, screening in the beginning period, with a specific end goal to discover kinds of growth before the side effects[158,159]. Also, they have grown new techniques for the early prediction of cancer treatment result.

Nonetheless, From IBM Watson to other real players, a great deal of cash has been spent attempting to make progress in this field; however, with little achievement[159]. With the approach of new advancements in the field of medication, a lot of cancer information has been gathered and are accessible to the medicinal research network. However, the exact prediction of an illness result is the most fascinating and challenging assignments for doctors. Therefore, Machine learning strategies have turned into a well-known apparatus for medicinal specialists[160]. They can adequately now foresee future results of cancer as of now; machine learning systems can find and distinguish examples and connections between them by the help of complicated datasets. It is obvious that the joining of multidimensional heterogeneous information, joined with the utilization of various systems for feature selection and grouping will give promising instruments for interference in the cancer area for the prediction[161,162].

Prediction of Protein Structure: Proteins are huge bio-molecules comprising of single or more-long chains of amino acid deposits. Its framework is a three-dimensional grouping of atoms in an amino acid-chain[163–165]. They share a very significant place of nutritious diets in our daily routine. Due to the dependence of protein on its three-dimensional states, the function of its tertiary structure creates an important foundation of the research field in protein structure prediction[163,166–168]. Present investigations have exhibited that the tertiary structure of the protein can be uncovered by artificially evaluating different protein properties, for example, secondary structure. Regardless, many pieces of evidence show that the technique mentioned above is very costly and demands expensive labor[169–172]. It is a highly challenging task to predict the protein structure just from its sequence. Understanding the complicated sequence structure sequences is one of the biggest hurdles in the area of computational biology. Having the enormous quantities of protein and DNA sequencing data, it is past the extent of our insight and abilities to break down the conduct of these characteristic frameworks. While we can break down a few examples through our experience and facts, a few examples which are multidimensional are difficult to conclude[173].

Artificial Intelligence comprises numerous regions like Computational Intelligence, Machine Learning and Evolutionary Computing whose essential work is to conclude essential data from the information while additionally permitting us foreseeing the data for feasible future situations. Artificial Intelligence has answers for the issues which are past our degree to examine. Likewise, given the intricacy of natural procedures it is difficult to program the computers for all the conceivable results, and consequently, we depend on Artificial Intelligence methods in perceiving traits investigating through numerous measurements of information.

As of now, the best indicators depend on machine learning approaches, specifically neural system models with a settled, and generally short, input window of amino acids, focused on the prediction spot.

However, where machine learning is dealing with tremendous quantities of indicators — in some cases, strictly, a larger number of indicators than perceptions — and consolidating them in nonlinear and exceedingly intuitive ways[174]. This limit enables us to utilize new sorts of information, whose sheer volume or intricacy would already have made breaking down them unimaginable[175]. In protein prediction, there is an immense number of opportunities when taken consideration with Artificial Intelligence and machine learning which we are not even close to thinking. Exploring those potentials will result in a lot of advancement in the field of medicine and biotechnology[176].

Conclusion

There has been immense work done in the last few decades in the field of Artificial Intelligence. Recently, people started using those computational power in the different areas of researches beginning from the conductivity of a metal to the checking of activity of the drugs for a particular disease, the use of Machine Learning has proven to be very successful. Today, there are an enormous amount of data available in this world, but very few are there for the analysis of it, because of which nowadays many new fields start to emerge starting from Data Science to Bioinformatics and Cheminformatics. We can be assured that this world of AI is going to benefit a lot to humanity, converting the toughest jobs to the simplest ones. They are slowly evolving in the field of medicine and lifesciences where they might do wonders and can help the doctors and clinicians to quicken the method of diagnosis. Lastly, it is true that the Machine Learning has shown a lot of benefits, but there remains an immense amount of opportunity to discover and tackle new problems.

References

1. Stefik M.J. (1985). Machine learning: An artificial intelligence approach. R.S. Michalski, J.G. Carbonell and T.M. Mitchell, (Tioga, Palo Alto, CA); *Artificial Intelligence,* 572 pages, $39.50, 25(2), 236-238. doi:10.1016/0004-3702(85)90005-0

2. Angra S. and Ahuja S. (2017). Machine learning and its applications: a review. In Big Data Analytics and Computational Intelligence (ICBDAC), 2017 International Conference on 57-60. IEEE.

doi:10.1109/ICBDACI.2017.8070809

3. Robert C. (2014). Machine Learning, a Probabilistic Perspective. *Chance,* 27, 62-63.
4. Mitchell T.M. (1997). Machine Learning. *McGraw-Hill*. doi:10.1145/242224.242229
5. Alpaydın E. (2014). Introduction to machine learning. Methods in Molecular Biology. doi:10.1007/978-1-62703-748-8-7
6. Ng A. (2012). Supervised learning. *Mach. Learn.* doi:10.1111/j.1466-8238.2009.00506.x
7. Kotsiantis S.B. (2007). Supervised Machine Learning: A Review of Classification Techniques. *Informatica,* 31, 249-268.
8. Nasrabadi N.M. (2007). Pattern recognition and machine learning. *Journal of electronic imaging*, 16(4). doi:10.1117/1.2819119
9. Kohavi R. and Provost F. (1998). Glossary of Terms. *Mach. Learn.* doi:10.1023/A:1017181826899
10. Caruana R. and Niculescu-Mizil A. (2006). An empirical comparison of supervised learning algorithms. Proceedings of the 23rd international conference on Machine learning, 161-168. ACM. doi:10.1145/1143844.1143865
11. Caruana R., Karampatziakis N. and Yessenalina A. (2008). An empirical evaluation of supervised learning in high dimensions. in *Proceedings of the 25th international conference on Machine learning - ICML '08*. doi:10.1145/1390156.1390169
12. Jordan M.I. and Rumelhart D.E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive science*, 16(3), 307-354. doi:10.1016/0364-0213(92)90036-T
13. Hastie T., Tibshirani R. and Friedman J. (2009). The Elements of Statistical Learning. *Bayesian Forecasting and Dynamic Models*. doi:10.1007/b94608
14. Zhang C., Patras P. and Haddadi H. (2018). Deep Learning in Mobile and Wireless Networking: A Survey. *IEEE Commun. Surv. TUTORIALS*.
15. Math Works (2016). Applying Supervised Learning. *Mach. Learn. with MATLAB*. doi:10.1111/j.2041-210X.2010.00056.x
16. Dougherty J., Kohavi R. and Sahami M. (1995). Supervised and unsupervised discretization of continuous features. *In Machine Learning Proceedings*, 194-202. doi:10.1016/B978-1-55860-377-6.50032-3
17. Rasmussen C.E. and Williams C.K. (2006). Gaussian process for machine learning. MIT press. doi:10.1093/bioinformatics/btq657
18. Aiken L.S. and West S.G. (1994). Multiple regression: Testing and interpreting interactions. *Journal-Operational Research Society*, 45, 119. doi:10.2307/2583960
19. Liu X. (2003). Supervised Classification and Unsupervised Classification. *Cfa.Harvard.Edu*.
20. Chelly S.M. and Denis C. (2001). Applying Unsupervised Learning. *Med. Sci. Sports Exerc.* doi:10.1111/j.2041-210X.2010.00056.x
21. Love B.C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic bulletin & review*, 9(4), 829-835. doi:10.3758/BF03196342
22. Jain A.K., Murty M.N. and Flynn P.J. (1999). Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3), 264-323. doi:10.1145/331499.331504
23. Adil S.H. and Qamar S. (2009). Implementation of association rule mining using CUDA. *In Emerging Technologies*, ICET 2009. International Conference, IEEE, 332-336. doi:10.1109/ICET.2009.5353149
24. Barto A.G. and Dietterich T.G. (2004). Reinforcement learning and its relationship to supervised learning. *Handbook of learning and approximate dynamic programming. John Wiley and Sons, Inc*. doi:10.1002/9780470544785.ch2
25. Das S., Dey A., Pal A. and Roy N. (2015). Applications of artificial intelligence in machine learning: review and prospect. *International Journal of Computer Applications*, 115(9), 31-41.
26. Mnih V., Kavukcuoglu K., Silver D., Rusu A.A., Veness J., Bellemare M.G. and Petersen S. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529. doi:10.1038/nature14236
27. Giryes R. and Elad M. (2011). Reinforcement Learning: A Survey. *In Eur. Signal Process*. Conf, 1475-1479. doi:10.1613/jair.301
28. Sutton R.S. and Barto A.G. (2013). Reinforcement learning : an introduction. *Neural Networks IEEE Transactions*. doi:10.1109/TNN.1998.712192
29. Tran H.A. (2011). Statistical Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 4.
30. Abadi M., Barham P., Chen J., Chen Z., Davis A., Dean J. and Kudlur M. (2016). TensorFlow: A System for Large-Scale Machine Learning TensorFlow: A system for large-scale machine learning. in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*. doi:10.1038/nn.3331
31. Langley P. (1994). Institute for the S. of L. and E. Selection of Relevant Features in Machine Learning. *Proc. AAAI Fall Symp. Relev.* doi:10.1.1.43.4648
32. Spring M.L. (2015). Machine Learning in Action. *Engineering*. doi:10.1007/978-0-387-77242-4
33. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O. and Duchesnay É. (2012). Scikit-learn: Machine learning in Python. *Machine Learning*, 12,

2825-2830. doi:10.1007/s13398-014-0173-7.2

34. Muller K.R., Mika S., Ratsch G., Tsuda K. and Scholkopf B. (2001). An introduction to kernel-based learning algorithms. *IEEE transactions on neural networks*, 12(2), 181-201. doi:10.1109/72.914517

35. Bibuli M., Caccia M. and Lapierre L. (2007). Path-following algorithms and experiments for an autonomous surface vehicle. *IFAC Proceedings*, 40(17), 81-86. doi:10.3182/20070919-3-HR-3904.00015

36. Freund Y. and Schapire R.E. (1996). Experiments with a new boosting algorithm. *In Icml*, 96, 148-156. doi:10.1.1.133.1040

37. Kotsiantis S.B., Zaharakis I.D. and Pintelas P.E. (2006). Machine learning: a review of classification and combining techniques. *Artificial Intelligence Review*, 26(3), 159-190. doi:10.1007/s10462-007-9052-3

38. Weisberg S. (2005). Applied Linear Regression. *New York*. doi:10.1080/00049538408255324

39. Article R. (2010). Linear Regression Analysis. *Deutsches Ärzteblatt International*, 107**,** 776-782.

40. Almeida A.M.D., Castel-Branco M.M. and Falcao A.C. (2002). Linear regression for calibration lines revisited: weighting schemes for bioanalytical methods. *Journal of Chromatography B*, 774(2), 215-222. doi:10.1016/S1570-0232(02)00244-1

41. Kutner M.H., Nachtsheim C.J., Neter J. and Li W. (1996). Applied Linear Statistical Models. *Journal Of The Royal Statistical Society Series A General.* doi:10.2307/2984653

42. Zou K.H., Tuncali K. and Silverman S.G. (2003). Correlation and simple linear regression. *Radiology*, 227(3), 617-628. doi:10.1148/radiol.2273011499

43. Meng X., Bradley J., Yavuz B., Sparks E., Venkataraman S., Liu D. and Xin D. (2016). Mllib: Machine learning in apache spark. *The Journal of Machine Learning Research*, 17(1), 1-7.

44. Kleinbaum D.G. and Klein M. (2010). Logistic regression: a self-learning text. *Springer Science & Business Media*. doi:10.1007/978-1-4419-1742-3

45. Lemeshow S. and Hosmer D.W. (1982). A review of goodness of fit statistics for use in the development of logistic regression models. *Am. J. Epidemiol.,* 115(1), 92-106. doi:10.1093/oxfordjournals.aje.a113284

46. Gortmaker S.L., Hosmer D.W. and Lemeshow S. (1994). Applied Logistic Regression. *Contemp. Sociol.* doi:10.2307/2074954

47. Tripepi G., Jager K.J., Dekker F.W. and Zoccali C. (2008). Linear and logistic regression analysis. *Kidney international*, 73(7), 806-810. doi:10.1038/sj.ki.5002787

48. Landwehr N., Hall M. and Frank E. (2005). Logistic model trees. *Machine learning*, 59(1-2), 161-205. doi:10.1007/s10994-005-0466-3

49. Menard S. (2002). Applied logistic regression analysis. Sage, 106. doi:10.4135/9781412983433

50. Peduzzi P., Concato J., Kemper E., Holford T.R. and Feinstein A.R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*, 49(12), 1373-1379. doi:10.1016/S0895-4356(96)00236-3

51. LaValley M.P. (2008). Logistic regression. *Circulation*, 117(18), 2395-2399. doi:10.1161/CIRCULATIONAHA.106.682658

52. Pampel F.C. (2000). Logistic Regression: A Primer. *Sage Quant. Appl. Soc. Sci.*, 132. doi:10.1177/007542429602400208

53. Peng C.Y.J., Lee K.L. and Ingersoll G.M. (2002). An introduction to logistic regression analysis and reporting. *The journal of educational research*, 96(1), 3-14. doi:10.1080/00220670209598786

54. Chernick M.R. and Friis R.H. (2003). Introductory biostatistics for the health sciences: modern applications including bootstrap. *John Wiley & Sons*.

55. Murphy K.P. (2006). Naive Bayes classifiers Generative classifiers. *Bernoulli*. doi:10.1007/978-3-540-74958-5_35

56. Rish I. (2001). An empirical study of the naive Bayes classifier. In IJCAI 2001 workshop on empirical methods in artificial intelligence, New York, 3(22), 41-46. IBM.doi:10.1039/b104835j

57. Kohavi R. (1996). Scaling Up the Accuracy of Naive-Bayes Classifiers: A Decision-Tree Hybrid. *Proc. Second Int. Conf. Knowl. Discov. Data Min.* doi:citeulike-article-id:3157868

58. Taneja S., Gupta C., Goyal K. and Gureja D. (2014). An Enhanced K-Nearest Neighbor Algorithm Using Information Gain and Clustering. *2014 Fourth Int. Conf. Adv. Comput. Commun. Technol.,* 325-329. doi:10.1109/ACCT.2014.22

59. Keller J.M., Gray M.R. and Givens J.A. (1985). A fuzzy k-nearest neighbor algorithm. *IEEE transactions on systems, man, and cybernetics*, 15(4), 580-585. doi:10.1109/TSMC.1985.6313426

60. Dudani S.A. (1976). The distance-weighted k-nearest-neighbor rule. *IEEE Transactions on Systems, Man, and Cybernetics*, 4, 325-327. doi:10.1109/TSMC.1976.5408784

61. Yu X., Pu K.Q. and Koudas N. (2005). Monitoring k-nearest neighbor queries over moving objects. in *Proceedings - International Conference on Data Engineering,* 631-642. doi:10.1109/ICDE.2005.92

62. Sharma R., Kumar S. and Maheshwari R. (2015). Comparative Analysis of Classification Techniques in Data

Mining Using Different Datasets. *Int. J. Comput. Sci. Mob. Comput.,* 4(12), 125-134.

63. Guo G., Wang H., Bell D., Bi Y. and Greer K. (2003). KNN model-based approach in classification. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". Springer, Berlin, Heidelberg, 986-996. doi:10.1007/b94348

64. Lafferty J., McCallum A. and Pereira F.C.N. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *ICML '01 Proc. Eighteenth Int. Conf. Mach. Learn*. doi:10.1038/nprot.2006.61

65. Palmer D.S., O'Boyle N.M., Glen R.C. and Mitchell J.B. (2007). Random forest models to predict aqueous solubility. *Journal of chemical information and modeling*, 47(1), 150-158. doi:10.1021/ci060164k

66. LaValle S.M. and Kuffner J.J. (2000). Rapidly-exploring random trees: Progress and prospects. *4th Work. Algorithmic Comput. Robot. New Dir.* doi:10.1017/CBO9781107415324.004

67. Svetnik V., Liaw A., Tong C., Culberson J.C., Sheridan R. P. and Feuston B.P. (2003). Random forest: a classification and regression tool for compound classification and QSAR modeling. *Journal of chemical information and computer sciences*, 43(6), 1947-1958. doi:10.1021/ci034160g

68. Breiman L. (2001). Random forests. *Machine learning*, 45(1), 5-32.

69. Ho T.K. (1995). Random decision forests. In Document analysis and recognition, proceedings of the third international conference, IEEE, 1, 278-282.

70. Breiman L. (1999). Random Forests. *Mach. Learn*. doi:10.1023/A:1010933404324

71. Pal M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing,* 26(1), 217-222. doi:10.1080/01431160412331269698

72. Horrace W.C. and Oaxaca R.L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, 90(3), 321-327. doi:10.1016/j.econlet.2005.08.024

73. de Souza S.V. and Junqueira R.G. (2005). A procedure to assess linearity by ordinary least squares method. *Analytica Chimica Acta*, 552(1-2), 25-35. doi:10.1016/j.aca.2005.07.043

74. Studenmund A.H. (2005). Ordinary Least Squares. *Using Econometrics. A Practical Guide*.

75. Leng L., Zhang T., Kleinman L. and Zhu W. (2007). Ordinary least square regression, orthogonal regression, geometric mean regression and their applications in aerosol science. *Journal of Physics: Conference Series*, 78(1), 1-5. 012084. IOP Publishing. doi:10.1088/1742-6596/78/1/012084

76. Kilmer J.T. and Rodriguez R.L. (2017). Ordinary least squares regression is indicated for studies of allometry. *Journal of evolutionary biology*, 30(1), 4-12. doi:10.1111/jeb.12986

77. Dhillon P.S., Foster D.P., Kakade S.M. and Ungar L.H. (2013). A risk comparison of ordinary least squares vs ridge regression. *The Journal of Machine Learning Research*, 14(1), 1505-1511.

78. McCool M., Robison A. and Reinders J. (2012). Structured parallel programming: patterns for efficient computation. Elsevier. doi:10.1016/B978-0-12-415993-8.00013-X

79. Bron C. (1972). Merge sort algorithm [M1]. *Communications of the ACM*, 15(5), 357-358. doi:10.1145/355602.361317

80. Cormen T.H., Leiserson C.E., Rivest R.L. and Stein C. (2001). Introduction to algorithms second edition. *Computer*. doi:10.2307/2583667

81. Cole R. (1986). Parallel Merge Sort. *Annual Symposium on Foundations of Computer Science (Proceedings*). doi:10.1109/SFCS.1986.41

82. Byvatov E., Fechner U., Sadowski J. and Schneider G. (2003). Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. *Journal of chemical information and computer sciences*, 43(6), 1882-1889. doi:10.1021/ci0341161

83. Tao X., Renmu H., Peng W. and Dongjie X. (2004). Input dimension reduction for load forecasting based on support vector machines. *Proceedings of the 2004 IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies (DRPT2004),* 2, 510-514.

84. Yeh C.Y., Huang C.W. and Lee S.J. (2011). A multiple-kernel support vector regression approach for stock market price forecasting. *Expert Systems with Applications*, 38(3), 2177-2186. doi:10.1016/j.eswa.2010.08.004

85. Brereton R.G. and Lloyd G.R. (2010). Support vector machines for classification and regression. *Analyst*, 135(2), 230-267. doi:10.1039/B918972F

86. Berwick R. (2003). An Idiot's guide to Support vector machines (SVMs). *Retrieved on October*, 21, 2011.

87. Ekambaram R., Fefilatyev S., Shreve M., Kramer K., Hall L.O., Goldgof D.B. and Kasturi R. (2016). Active cleaning of label noise. *Pattern Recognition*, 51, 463-480. doi:10.1016/j.patcog.2015.09.020

88. Rai P. (2011). Hyperplane based Classification : Perceptron and (Intro to) Support Vector Machines. Piyush Rai Hyperplane Separates a D -dimensional space into two half-spaces.

89. Suciu A.I. (2014). Mathématiques. *In Annales de la Faculté des Sciences de Toulouse*, 23(2), 417-439. doi:10.5802/afst.1412

90. Platt J.C. (1998). Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. *Advances in kernel methods*. doi:10.1.1.43.4376

91. Rasmussen C. (2006). Gaussian processes for machine learning. *Int. J. Neural Syst.* doi:10.1142/S0129065704001899

92. Rodriguez S., Tang X., Lien J.M. and Amato N.M. (2006). An obstacle-based rapidly-exploring random tree. In Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006, 895-900. IEEE. doi:10.1109/ROBOT.2006.1641823

93. Zhao Y. and Zhang Y. (2008). Comparison of decision tree methods for finding active objects. *Advances in Space Research*, 41(12), 1955-1959. doi:10.1016/j.asr.2007.07.020

94. Yu X., Hyyppä J., Vastaranta M., Holopainen M. and Viitala R. (2011). Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and remote sensing*, 66(1), 28-37. doi:10.1016/j.isprsjprs.2010.08.003

95. Hermens L.A. and Shlimmer J.C. (1994). A machine-learning apprentice for the completion of repetitive forms. *IEEE Expert*, 9(1), 28-33. doi:10.1109/64.295135

96. Kung J.Y. and Shu M.H. (2015). Some scheduling problems on a single machine with general job effects of position-dependent learning and start-time-dependent deterioration. *Asia-Pacific Journal of Operational Research*, 32(02). doi:10.1142/S0217595915500025

97. Langley P. and Simon H.A. (1995). Applications of machine learning and rule induction. *Communications of the ACM*, 38(11), 54-64.

98. Durso F.T. and Manning C.A. (2008). Air traffic control. *Reviews of human factors and ergonomics*, 4(1), 195-244. doi:10.1518/155723408X342853

99. Djokic J., Lorenz B. and Fricke H. (2010). Air traffic control complexity as workload driver. Transportation research part C: emerging technologies, 18(6), 930-936. doi:10.1016/j.trc.2010.03.005

100. Hanh T.T.B. and Van Hung D. (2007). Verification of an air-traffic control system with probabilistic real-time model-checking. UNU-IIST report.

101. Saba T., Sulong G. and Rehman A. (2011). Retracted Article: Document image analysis: issues, comparison of methods and remaining problems. *Artificial Intelligence Review*, 35(2), 101-118. doi:10.1007/s10462-010-9186-6

102. Murman E.M., Walton M. and Rebentisch E. (2000). Challenges in the better, faster, cheaper era of aeronautical design, engineering and manufacturing. *The Aeronautical Journal*, 104(1040), 481-489. doi:10.1017/S0001924000091983

103. Morrow D. (1997). Improving consultations between health-care professionals and older clients: Implications for pharmacists. *The International Journal of Aging and Human Development*, 44(1), 47-72. doi:10.2190/GQX9-F4UJ-5RQ2-N1YD

104. Parcell G. and Collison C. ( ). Learning to Fly Summary. 1-9.

105. Sengupta N., Sahidullah M. and Saha G. (2016). Lung sound classification using cepstral-based statistical features. *Computers in biology and medicine*, 75, 118-129. doi:10.1016/j.compbiomed.2016.05.013

106. Ream N. (1977). Discrete-time signal processing. *Electronics and Power*, 23(2), 157. doi:10.1049/ep.1977.0078

107. Al-Jumaily A.H.J., Sali A., Mandeep J.S. and Ismail A. (2015). Propagation measurement on earth-sky signal effects for high speed train satellite channel in tropical region at Ku-band. *Int. J. Antennas Propag.* doi:10.1155/2015/270949

108. Smith S.W. (1999). The Scientist and Engineer's Guide to Digital Signal Processing. *IEEE Signal Processing Magazine*. doi:10.1109/79.826412

109. Le Guernic P., Benveniste A., Bournai P. and Gautier T. (1986). Signal--A data flow-oriented language for signal processing. *IEEE transactions on acoustics, speech, and signal processing*, 34(2), 362-374. doi:10.1109/TASSP.1986.1164809

110. Zhang Q. and Benveniste A. (1992). Wavelet networks. *IEEE transactions on Neural Networks*, 3(6), 889-898. doi:10.1109/72.165591

111. Movaghar A., Jamzad M. and Asadi H. (2014). Artificial Intelligence and Signal Processing: International Symposium. *AISP 2013 Tehran, Iran*, December 25-26, 2013 Revised Selected Papers. *Commun. Comput. Inf. Sci.* (2014). doi:10.1007/978-3-319-10849-0

112. Hagan M.T., Demuth H.B. and Beale M.H. (1996). Neural Network Design. *Bost. Massachusetts PWS*. doi:10.1007/1-84628-303-5

113. Thrun S. (2010). Toward robotic cars. *Communications of the ACM*, 53(4), 99-106. doi:10.1145/1721654.1721679

114. Lassa T. (2013). The Beginning of the End of Driving. *Mot. Trend*.

115. Gehrig S.K. and Stein F.J. (1999). Cartography and dead reckoning using stereo vision for an\nautonomous car. *Proc. 1999 Int. Conf. Image Process. (Cat. 99CH36348)*, 4, 30-34. doi:10.1109/ICIP.1999.819461

116. Luettel T., Himmelsbach M. and Wuensche H.J. (2012).

Autonomous Ground Vehicles-Concepts and a Path to the Future. *Proceedings of the IEEE, 100(Centennial-Issue)*, 1831-1839. doi:10.1109/JPROC.2012.2189803

117. Hamid U.Z.A., Pushkin K., Zamzuri H., Gueraiche D. and Rahman M.A.A. (2016). Current Collision Mitigation Technologies for Advanced Driver Assistance Systems – A Survey. *Perintis eJournal,* 6(2). doi:10.1105/tpc.15.01050

118. Chen C., Seff A., Kornhauser A. and Xiao J. (2015). DeepDriving: Learning affordance for direct perception in autonomous driving. *Proceedings of the IEEE International Conference on Computer Vision*, 2722-2730. doi:10.1109/ICCV.2015.312

119. LeCun Y.A., Bengio Y. and Hinton G.E. (2015). Deep learning. *Nature,* 521, 436-444. doi:10.1038/nature14539.

120. Bird S., Klein E. and Loper E. (2009). Natural Language Processing with Python. *analyzing text with the natural language toolkit*. doi:10.1097/00004770-200204000-00018

121. Collobert R., Weston J., Bottou L., Karlen M., Kavukcuoglu K. and Kuksa P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, 12(Aug), 2493-2537. doi:10.1. 1.231.4614.

122. Soon W.M., Ng H.T. and Lim D.C.Y. (2001). A machine learning approach to coreference resolution of noun phrases. *Computational linguistics*, 27(4), 521-544. doi:10.1162/08912010175334 2653.

123. Farghaly A. and Shaalan K. (2009). Arabic natural language processing : Challenges and solutions. *ACM Trans. Asian Lang. Inf. Process.,* 8(4), 14. doi:10.1145/1644879. 1644881.http.

124. Westerdahl S. (2003). Organising and managing work. Organisational, managerial and strategic behaviour in theory and practice. *Scand. J. Manag.*, 3(19), 392-394. doi:10.1016/S0956-5221(03)00030-7.

125. Rokach L. and Maimon O. (2010). Decision Trees. in *Data Mining and Knowledge Discovery Handbook,* 154-192. doi:10.1007/ 0-387-25465-X_9.

126. Kuncheva L.I. (2004). Classifier ensembles for changing environments. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. doi:10.1007/978-3-540-25966-4_1.

127. Usmani M., Adil S.H., Raza K. and Ali S.S.A. (2016). Stock market prediction using machine learning techniques. *In Computer and Information Sciences (ICCOINS), 2016 3rd International Conference*, IEEE, 322-327. doi:10.1109/ICCOINS.2016.7783235

128. Coupelon O. (2007). Neural network modeling for stock movement prediction A state of the art. *Network*.

129. Yoo P.D., Kim M.H. and Jan T. (2005). Machine Learning Techniques and Use of Event Information for Stock Market Prediction: A Survey and Evaluation. *Int. Conf. Comput. Intell. Model. Control Autom. Int. Conf. Intell. Agents, Web Technol. Internet Commer,* 2, 835-841. doi:10.1109/CIMCA.2005. 1631572.

130. Glahn H.R. and Lowry D.A. (1972). The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *J. Appl. Meteorol,* 11(8), 1203-1211. doi:10.1175/1520-0450(1972)011<1203: TUOMOS>2.0.CO;2.

131. Holtslag A.A.M., De Bruijn E.I.F. and Pan H.L. (1990). A High Resolution Air Mass Transformation Model for Short-Range Weather Forecasting. *Mon. Weather Rev.,* 118(8), 1561-1575. doi:10. 1175/1520-0493(1990)118<1561:AHRAMT>2.0.CO;2

132. Nash J.E. and Sutcliffe J.V. (1970). River flow forecasting through conceptual models part I - A discussion of principles. *J. Hydrol.,* 10(3), 282-290. doi:10.1016/0022-1694(70)90255-6.

133. Gneiting T. and Raftery A.E. (2005). Weather forecasting with ensemble methods. *Science,* 310(5746), 248-249. doi:10.1126/ science .1115255.

134. Hapgood M. (2012). Astrophysics: Prepare for the coming space weather storm. *Nature,* 484(7394), 311-313. doi:10.1038/484311a.

135. Lee K., Cha Y. and Park J. (1992). Short-term load forecasting using an artificial neural network. *Power Syst. IEEE Trans.,* 7(1), 124-132. doi:10.1109/59.141695.

136. Jolliffe I.T. (1982). A Note on the Use of Principal Components in Regression. *J.R. Stat. Soc. Ser. C (Applied Stat.,* 31(3), 300-303. doi:10.2307/2348005

137. Holmstrom M., Liu D. and Vo C. (2016). Machine Learning Applied to Weather Forecasting. *Meteorol. Appl,* 1-5.

138. Salman A.G., Kanigoro B. and Heryadi Y. (2015). Weather forecasting using deep learning techniques. in *2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS),* 281-285. doi:10.1109/ ICACSIS.2015.7415154.

139. Li J., Ward J.K., Tong J., Collins L. and Platt G. (2016). Machine learning for solar irradiance forecasting of photovoltaic system. *Renew. Energy,* 90, 542-553. doi:10.1016/ j.renene. 2015.12.069

140. Ji H.F., Li X.J. and Zhang H.Y. (2009). Natural products and drug discovery: Can thousands of years of ancient medical knowledge lead us to new and powerful drug combinations in the fight against cancer and dementia?. *EMBO Rep.,* 10(3), 194-200. doi:10.1038/embor.2009.12.

141. Pan S.Y., Zhou S.F., Gao S.H., Yu Z.L., Zhang S.F., Tang M.K. and Ko K.M. (2013). New perspectives on how to discover drugs from herbal medicines: CAM'S outstanding contribution to modern therapeutics. *Evidence-based Complement. Altern. Med.* doi:10.1155/2013/627375

142. Mitchell J.B.O. (2014). Machine learning methods in chemoinformatics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.,* 4(5), 468-481. doi:10.1002/wcms.1183.

143. Patwardhan B. (2005). Ethnopharmacology and drug discovery. *Journal of Ethnopharmacology,* 100(1-2), 50-52. doi:10.1016/ j.jep.2005.06.006.

144. Lee J.A., Uhlik M.T., Moxham C.M., Tomandl D. and Sall D.J. (2012). Modern phenotypic drug discovery is a viable, neoclassic pharma strategy. *Journal of Medicinal Chemistry,* 55(10), 4527-4538. doi:10.1021/jm201649s

145. Murphy R.F. (2011). An active role for machine learning in drug development. *Nature Chemical Biology*, 7, 327-330. doi:10.1038/ nchembio.576.

146. Bernick J.P. (2015). The Role of Machine Learning in Drug Design and Delivery. *J. Dev. Drugs,* 4(3), 1-2.

147. Hopkins A.L. (2009). Drug discovery: Predicting promiscuity. *Nature,* 462*,* 167-168.

148. Wale N. (2011). Machine learning in drug discovery and development. *Drug Dev. Res.,* 72, 112-119.

149. Bowen A. and Casadevall A. (2015). Increasing disparities between resource inputs and outcomes, as measured by certain health deliverables, in biomedical research. *Proc. Natl. Acad. Sci.,* 112(36), 11335-11340. doi:10.1073/pnas.1504955112.

150. Williams C.K.I. (2003). Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. *J. Am. Stat. Assoc.*, 98, 489-489. doi:10.1198/jasa.2003.s269

151. Lavecchia A. (2015). Machine-learning approaches in drug discovery: Methods and applications. *Drug Discovery Today,* 20(3), 318-331. doi:10.1016/j.drudis.2014.10.012.

152. American Cancer Society (2016). Cancer Facts and Figures 2016. *Cancer Facts Fig. 2016*. doi:10.1097/ 01.NNR.0000289503.22414.79

153. World Health Organisation (2018). Cancer. *http://www.who.int/en/ news-room/fact-sheets/detail/ cancer*.

154. Alberts B. (2008). Molecular Biology of the Cell. *Amino Acids* doi:10.1002/1521-3773(20010316)40: 6<9823::AID-ANIE9823>3.3.CO;2-C.

155. Schneider K. (2001). Cell Biology and Cancer. *Couns. about cancer Strateg. Genet. Couns.,* 1-17.

156. Aktipis C.A. and Nesse R.M. (2013). Evolutionary foundations for cancer biology. *Evolutionary Applications,* 6(1), 144-159. doi:10.1111/eva.12034

157. Harris R.J.C. (1971). Cancer and the environment. *Int. J. Environ. Stud.* doi:10.1080/00207237008709397

158. Hanahan D. and Weinberg R.A. (2011). Hallmarks of cancer: The next generation. *Cell,* 144(5), 646-674. doi:10.1016/ j.cell.2011.02.013

159. Bughin J. (2017). Artificial Intelligence – The Next Digital Frontier?Artificial Intelligence. doi:10.1016/S1353-4858(17)30039-9

160. Kourou K., Exarchos T.P., Exarchos K.P., Karamouzis M.V. and Fotiadis D.I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal,* 13, 8-17. doi:10.1016/j.csbj.2014.11.005

161. Cruz J.A. and Wishart D.S. (2006). Applications of machine learning in cancer prediction and prognosis. *Cancer Informatics*, 2. doi:10.1177/117693510600200030

162. Sirintrapun S. Joseph, Ahmet Zehir, Aijazuddin Syed, JianJiong Gao, Nikolaus Schultz and Donavan T. Cheng (2015). Translational Bioinformatics and Clinical Research (Biomedical) Informatics. *Surgical pathology clinics* 8(2), 269-288. doi:10.1016/j.path.2015.02.015

163. Ataka K. and Heberle J. (2007). Biochemical applications of surface-enhanced infrared absorption spectroscopy. *Anal. Bioanal. Chem.,* 388(1), 47-54. doi:10.1007/s00216-006-1071-4

164. Hume D. (2000). A Treatise of Human Nature. *A Treatise Hum. Nat.* doi:10.2307/2216614

165. Binda C., Mattevi A. and Edmondson D.E. (2002). Structure-function relationships in flavoenzyme-dependent amine oxidations: A comparison of polyamine oxidase and monoamine oxidase. *Journal of Biological Chemistry*. doi:10.1074/jbc.R200005200

166. Doi E. (1993). Gels and gelling of globular proteins. *Trends in Food Science and Technology,* 4(1), 1-5. doi:10.1016/S0924-2244(05) 80003-2.

167. Fraley S.I., Feng Y., Krishnamurthy R., Kim D.H., Celedon A., Longmore G.D. and Wirtz D. (2010). A distinctive role for focal adhesion proteins in three-dimensional cell motility. *Nature cell biology*, 12(6), 598-604. doi:10.1038/ncb2062

168. Sippl M.J. (1993). Recognition of errors in three???dimensional structures of proteins. *Proteins Struct. Funct. Bioinforma,* 17(4), 355-362. doi:10.1002/prot.340170404.

169. Bowie J.U., Lüthy R. and Eisenberg D. (1991). A method to identify protein sequences that fold into a known three-dimensional stucture. *Science,* 253(5016), 164-170. doi:10.1126/science. 1853201.

170. Lazaridis T. and Karplus M. (2000). Effective energy functions for protein structure prediction. *Current opinion in structural biology*, 10(2), 139-145. doi:10.1016/S0959-440X(00)00063-4.

171. Rost B. and Sander C. (1993). Prediction of protein secondary structure at better than 70% accuracy. *Journal of* *molecular biology*, 232(2), 584-599. doi:10.1006/ jmbi.1993.1413.

172. Zhang Y. (2008). Progress and challenges in protein structure prediction. *Current Opinion in Structural Biology,* 18(3), 342-348. doi:10.1016/j.sbi.2008.02.004.

173. Ginsberg M.L. (1988). Multivalued logics: a uniform approach to reasoning in artificial intelligence. *Comput. Intell.,* 4(3), 265-316. doi:10.1111/j.1467-8640.1988. tb00280.x

174. Porollo A. and Meller J. (2007). Prediction-based fingerprints of protein–protein interactions. *Proteins: Structure, Function, and Bioinformatics*, 66(3), 630-645. doi:10.1002/prot.21248.

175. Pirovano W. and Heringa J. (2010). Protein secondary structure prediction. *Methods in molecular biology (Clifton, N.J.),* 327-348. doi:10.1007/978-1-60327-241-4_19.

176. Koehl P. and Levitt M. (1999). A brighter future for protein structure prediction. *Nature Structural Biology*, 6, 108-111. doi:10.1038/5794.