The Best of the Machine Learning Algorithms Used in Artificial Intelligence

Indrasen Poola

Abstract: Artificial Intelligence is the best answer for tomorrow as our belief in intelligence is losing naturally and gradually. With high confidence, we will observe multiple roles taken over by machines in the next few years: customer service representatives, legal assistants, medical assistants, even primary care physicians and many others. It will start with human augmentation but move pretty rapidly towards human replacement. In this paper, we are discussing different machine learning algorithms used in Artificial Intelligence.

Keywords: Artificial Intelligence, Big Data, Machine Learning

1. INTRODUCTION

Digital world is taking us for joy ride. We will be able to guide only some part of the Artificial Intelligence revolution but disruption and natural evolution through machine learning algorithms will take care of the rest. We can combine nature inspired algorithm with machine learning to improve accuracy level of our data. For example, genetic algorithms (Brownlee J. , 2011) can help turning hyper parameters or choosing features.

"Artificial Intelligence" is no match for "Natural Stupidity". Dependency of humans will reduce for mundane tasks. Initially, Artificial Intelligence would handle routine repetitive tasks in organizations. We are few years away when we will slowly be integrated with corporate applications. This will eliminate many manual back end processing jobs with the help of supervised and unsupervised machine learning algorithms.

Humanity still has a long ways to go before machine artificial intelligence can take on anything close to resembling sentience. Hardware being the limiting factor. A generic learning pattern requires an incredible amount of hardware resources but software.

2. ARTIFICIAL INTELLIGENCE

Machines just need a shorthand way to do things like checking the current weather or adding an event to your calendar. The technique with which machines achieve such results is called Artificial Intelligence.

2.1 Machine Learning:

Machine learning is a top strategic trend for 2016, according to Gartner[1]. And Ovum[2] predicts that machine learning will be a necessary element for data preparation and predictive analysis in businesses moving forward. *Machine learning (ML) is a discipline where a program or system can learn from existing data and dynamically alter its behavior based on the ever-changing data. Therefore, the system has the ability to learn without being explicitly programmed*[3]. Machine Learning algorithms can be broadly categorized as classification, clustering, regression, dimensionality reduction and anomaly detection etc.

---

[1] Gartner Identifies the Top 10 Strategic Technology Trends for 2016, *Visit: http://www.gartner.com/newsroom/id/3143521*

[2] Ovum reveals the reality of Big Data for 2016: Cloud and appliances will drive the next wave of adoption, with Spark the fastest growing workload, *Visit: https://www.ovum.com/press_releases/ovum-reveals-the-reality-of-big-data-for-2016-cloud-and-appliances-will-drive-the-next-wave-of-adoption-with-spark-the-fastest-growing-workload/*

[3] AI is about Machine Reasoning – Or when Machine Learning is just a Fancy Plugin, *Visit: http://www.reasoning.world/ai-is-about-machine-reasoning-or-when-machine-learning-is-just-a-fancy-plugin/*
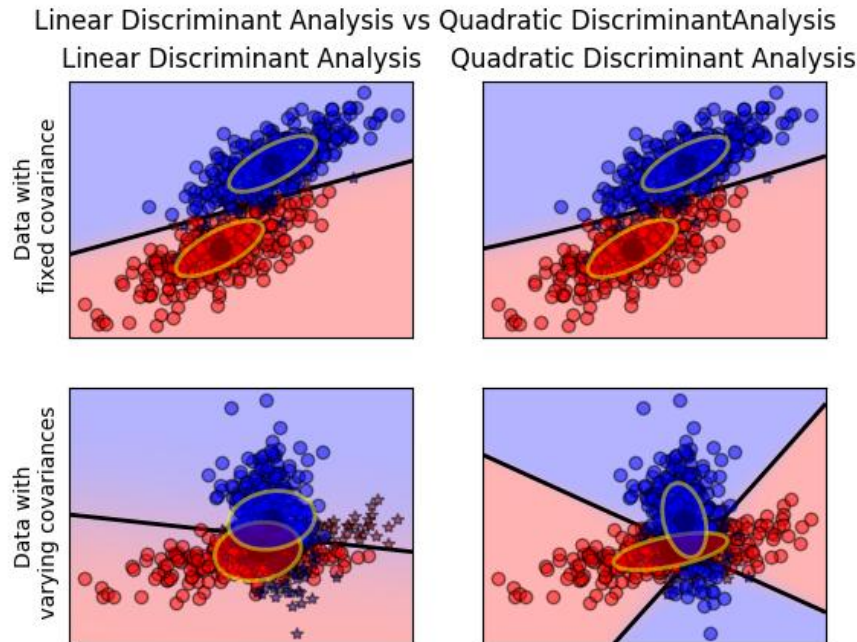
Figure 1Linear and Quadratic Discriminant Analysis with covariance ellipsoid - A typical example of classification

2.2 Machine Learning and Cognitive Systems

Cognitive Computing has interesting use cases catering to multiple industries and functions (Kelly, 2013). Routine repetitive applications initially will get automated through *Workplace Artificial Intelligence*. Next would come corporate applications starting to add value to the business. The machine Learning module acts as the core computing engine, which using algorithms & techniques helps *Cognitive Systems* to identify patterns, perform complex tasks like prediction, estimation, forecasting and anomaly detection. Pen source machine learning libraries like Mahout, Spark ML have made machine learning algorithms accessible to a wider audience (Kohavi & Provost, 1998). With larger and more complex data sets entering the health care field, machine learning models and AI will become table stakes, and health care incumbents will have to find ways to use these algorithms as well. Apple, Google, Microsoft, Intel and IBM played a key role in making deep learning capabilities accessible to the developer community through their Cognitive Services & APIs which could be easily embedded into other applications.

Artificial Intelligence is impossible until we master *Quantum Computing* since these systems are not 'self-aware'. These are technologies that deal with what are known as 'difficult' problems where conventional programming is not suited. Artificial Intelligence would not sprung up from nothing. Even if it self-replicates itself as in *Science Fiction* movie like *Terminator* according to Jon Von Neumann's self-replication machine[4], the materials for robots to build robots themselves would not materialize out of thin air. What intimates our study is four things:

1)   Who will control and monitor the daily transactions[5] of this database and for what purpose?
2)   What if the system crashes after we become depending on it?
3)   Artificial Intelligence will always be as intelligent as you let it be. If artificial intelligence does not learn, will it be useless?
4)   How would the artificial intelligence manager respond to an emergency situation, a one-in-a-million incident that it has no data on?
5)

---

[4] John von Neumann was a Hungarian-American mathematician, physicist, and computer scientist who first gave the concept of self-replication.

[5] Agent technologies deals with similar monetary transactions making use of machine learning algorithms

2.1.1 The Human Factor

The fact that Artificial Intelligence exists is because of capitalism. If capitalism ends, then Artificial Intelligence ends. What will happen to the people who become unemployed due to this technological factor is a matter of concern. Economics means of production still exists. Someone has to produce the raw materials, be it a mining company that's listed on Stock market, which buys trucks from another vehicle manufacturer etc. and so forth in order for productions to take place. Artificial Intelligence scientists believe there is a 50% chance that artificial intelligence will reach HLMI (Hi Level Machine Intelligence) by 2040-2050. That increases to 90% by 2075. High Level Machine Intelligence is defined as being able to carry out most human professions at least as well as a typical human.

In the U.S., the bottom 25-33 percent[6] of income earners could basically vanish without significantly disrupting the national economy. As automation increases, our focus should be on tasks that utilize creativity and emotion' resonates strongly. Artificial intelligence is guided to be able to sift through massive amounts of data. There is, however; one concern that people will use machines as an excuse not to learn the skills to be able to critically think in their fields. For example, the last thing we want is people to make decisions based on "The robot told us" instead of understanding why it got to that conclusion.

Although having Artificial Intelligence is helpful, we believe, it cannot replace experts. There is certainly a sense of fear around the impact of artificial intelligence making support functions redundant. That said, it's an exciting time to think about the jobs of the future and how we can best utilities the qualities we possess as humans that cannot be mimicked.

In Australia, CEDA (Committee for Economic Development of Australia) predicts[7] that technology will replace 40% of the workforce within 20 years. PWC predict 44% in the same time frame (PricewaterhouseCoopers, 2017). It is not just support functions that will be replaced by artificial intelligence, higher value, professional jobs are also being targeted.

Technology is a tool and we should drive it than letting it drive us. We see people unable to read maps who will blindly follow SATNAV, TOMTOM, GARMIN and they don't know where they are so if the system fails. European Union has already considered an electronic persons identifier for jobs that once were done physically by humans in order to replace the taxable income depletion or to develop a fund in case a human sues an electronic person due to a negative interaction. That is why, there are so many factors to consider because Information Technology people or services are not just plug and play like many decision makers may think. The human factor 'must' prevail over artificial judgement.

2.1.1.1 The movie Sully is a good example[8]

While the main character Captain Sully makes the decision of landing in the river Hudson which was right, the simulated version showed he could easily return back to the runway which was absolutely incorrect. Later, it was the very humans who accepted that the technology that they relied upon was not right and Captain Sully's decision was absolutely right. So we are all hands up for an expert over ARTIFICIAL INTELLIGENCE. The best practical approach to find the best or good algorithms for a given problem is trial and error. Heuristics provide a good guide, but sometimes/often you can get best results by breaking some rules or modeling assumptions.

2.1.2 How Will Artificial Intelligence Transform The Workplace

People are afraid of artificial intelligence machines. People misuse, abuse, and overuse such items. Are these artificial intelligence machines disposable or serviceable or simply replaceable? Artificial intelligence in the workplace makes the working environment dull and boring. So far, economic activity at large almost invents grunt work so large populations can be employed fully, whether an entrepreneur or the umpteenth levels of social hierarchy below them. What happens when Workplace Artificial Intelligence removes grunt work or cannot distinguish social impact and productivity impact is a question answered with the help of machine learning algorithms. What happens to all the people who actually depend on grunt work and how would they purpose themselves is another debate. Likely, these people would not be able to re-purpose themselves without the use of machine learning algorithms. When artificial intelligence machine fails, the bad results are eventually traced to human error. That of the one who created it.

---

[6] https://www.technologyreview.com/s/519241/report-suggests-nearly-half-of-us-jobs-are-vulnerable-to-computerization/

[7] http://www.abc.net.au/news/2015-06-16/technology-could-make-almost-40pc-of-jobs-redundant-report/6548560

[8] The story of Chesley Sullenberger, an American pilot who became a hero after landing his damaged plane on the Hudson River in order to save the flight's passengers and crew. http://www.imdb.com/title/tt3263904/

2.2 Supervised Learning Algorithms

Most of what people consider *Workplace Artificial Intelligence* is closer to programmer-assisted learning. Meaning that the algorithms and goals are predefined ahead of time forecasting algorithms like ARIMA, TBATS, Prophet. Most of applied machine learning (e.g. predictive modeling) is concerned with supervised learning algorithms. These algorithms are divided into following classifications (Brownlee D. J., 2017):

2.2 Regression

Regression is concerned with modelling the relationship between variables that is iteratively refined using a measure of error in the predictions made by the model. Regression methods are a work horse of statistics and have been cooped into statistical machine learning (Smith, 2016). This may be confusing because we can use regression to refer to the class of problem and the class of algorithm. For instance, a solution may be dependent on the outcome at multiple 'nodes' where each node may continue to vary for each event – creating an invoice for a logistics service that may have many variables to resolve for each transaction. Really, regression is a process.

Regression helps predicting a continuous-valued attribute associated with an object. Its usage includes applications such as Drug response, Stock prices. The most appropriated algorithms to this branch are SVR, ridge regression, Lasso, Ordinary Least Squares, Logistic Regression, Stepwise Regression, Multivariate Adaptive Regression Splines (MARS), and Locally Estimated Scatterplot Smoothing (LOESS).

2.2.1. Regularization Methods

An extension made to another method (typically regression methods) that penalizes models based on their complexity, favoring simpler models that are also better at generalizing. Regularization methods are popular, powerful and generally simple modifications made to other methods. Examples of such algorithms include Ridge Regression, Least Absolute Shrinkage and Selection Operator (LASSO) and Elastic Net.

2.2.2 Instance-based Methods

Instance based learning model a decision problem with instances or examples of training data that are deemed important or required to the model (Daelemans, 2005). Such methods typically build up a database of example data and compare new data to the database using a similarity measure in order to find the best match and make a prediction. For this reason, instance-based methods are also called winner-take all methods and memory-based learning. Focus is put on representation of the stored instances and similarity measures used between instances. Some vital algorithms are k-Nearest Neighbour (kNN), Learning Vector Quantization (LVQ) and Self-Organizing Map (SOM).

2.3 Classification

Classification algorithms are to identify which category an object belongs to. Its applications include Spam detection, Image recognition. The most popular algorithms in this classification are SVM, nearest neighbors, random forest, Classification and Regression Tree (CART), Iterative Dichotomiser 3 (ID3), C4.5, Chi-squared Automatic Interaction Detection (CHARTIFICIAL INTELLIGENCED), Decision Stump, Multivariate Adaptive Regression Splines (MARS), and Gradient Boosting Machines (GBM).

2.3.1 Decision Tree Learning

Decision tree methods construct a model of decisions made based on actual values of attributes in the data. Decisions fork in tree structures until a prediction decision is made for a given record. It will change how taxes will be collected. Artificial Intelligence derived layoffs and taxes are dependent on decision trees that are trained on data for classification and regression problems.
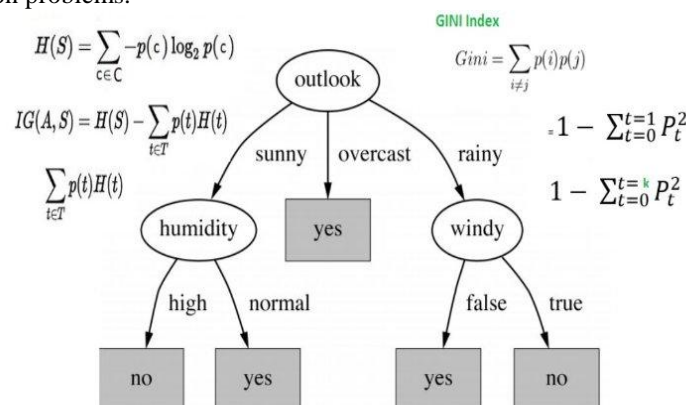
$$H(S) = \sum_{c \in C} -p(c) \log_2 p(c)$$

$$IG(A,S) = H(S) - \sum_{t \in T} p(t)H(t)$$

$$\sum_{t \in T} p(t)H(t)$$

GINI Index

$$Gini = \sum_{i \neq j} p(i)p(j)$$

$$= 1 - \sum_{t=0}^{t=1} P_t^2$$

$$1 - \sum_{t=0}^{t=k} P_t^2$$

Figure 4 Decision Trees Algorithms with Math

2.3.2 Bayesian
Bayesian methods are those that are explicitly apply Bayes' Theorem for problems such as classification and regression. The best suited algorithms are Naive Bayes classifier, Averaged One-Dependence Estimators (AODE) and Bayesian Belief Network (BBN).

2.3.3 Kernel Methods
Kernel Methods are best known for the popular method Support Vector Machines which is really a constellation of methods in and of itself. Kernel Methods are concerned with mapping input data into a higher dimensional vector space where some classification or regression problems are easier to model. The machine learning algorithms are included Support Vector Machines (SVM), Radial Basis Function (RBF) and Linear Discriminate Analysis (LDA).
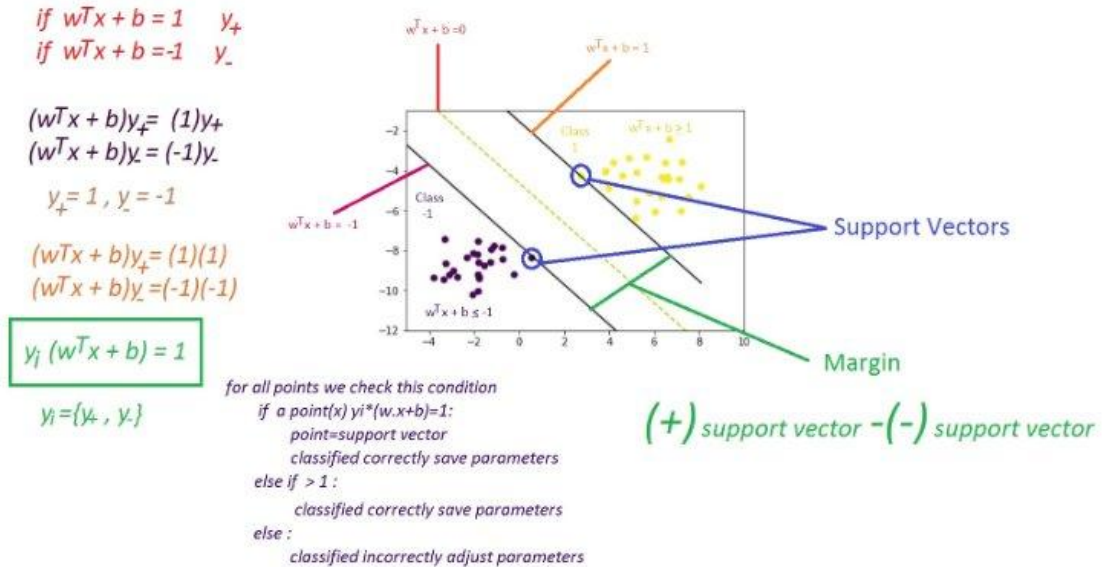

Figure 5 Support Vector machine with Math

2.3 Clustering Methods
Clustering, like regression describes the class of problem and the class of methods. Clustering methods are typically organized by the modelling approaches such as centroid-based and hierarchal. All methods are concerned with using the inherent structures in the data to best organize the data into groups of maximum commonality. Automatic grouping of similar objects into sets is called clustering. Applications such as Customer segmentation, Grouping experiment outcomes are achieved by clustering. The most useful algorithms in this domain are k-Means, spectral clustering, mean-shift, and Expectation Maximization (EM).

2.3.1 Association Rule Learning
Association rule learning are methods that extract rules that best explain observed relationships between variables in data. These rules can discover important and commercially useful associations in large multidimensional datasets that can be exploited by an organization. These include Apriori algorithm, and Eclat algorithm.

2.4 Dimensionality Reduction
Like clustering methods, *Dimensionality Reduction* seek and exploit the inherent structure in the data, but in this case in an unsupervised manner or order to summarize or describe data using less information. This can be useful to visualize dimensional data or to simplify data which can then be used in a supervised learning method. Reducing the number of random variables to consider helps in visualization, and increased efficiency. The most opted algorithms in this branch are Principal Component Analysis (PCA), feature selection, non-negative matrix factorization, Partial Least Squares Regression (PLS), Sammon Mapping, Multidimensional Scaling (MDS) and Projection Pursuit.

2.4.1 Ensemble Methods
Ensemble methods are models composed of multiple weaker models that are independently trained and whose predictions are combined in some way to make the overall prediction. Much effort is put into what types of weak learners to combine and the ways in which to combine them. This is a very powerful class of techniques and as such is very popular because of Boosting, Bootstrapped Aggregation (Bagging), AdaBoost, Stacked Generalization (blending), Gradient Boosting Machines (GBM) and Random Forest machine learning algorithms.
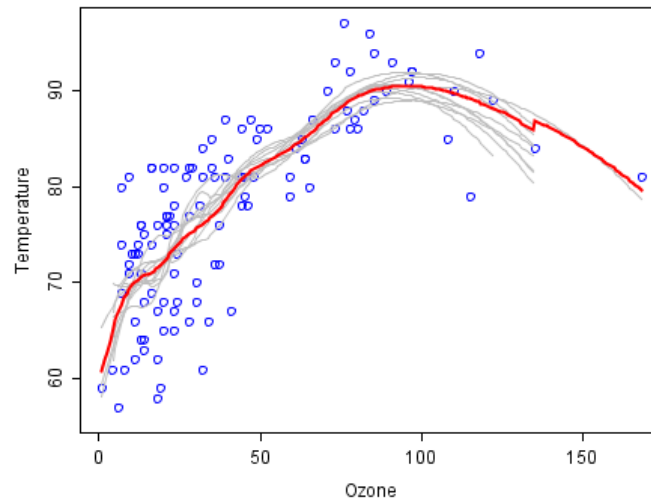
Figure 6 Ensemble of Lines of Best Fit - Weak members are grey, the combined prediction is red - Plot from Wikiperdia, licensed under public domain

3. Artificial Neural Networks

Artificial Neural Networks are models that are inspired by the structure and/or function of biological neural networks. They are a class of pattern matching that are commonly used for regression and classification problems but are really an enormous subfield comprised of hundreds of algorithms and variations for all manner of problem types.
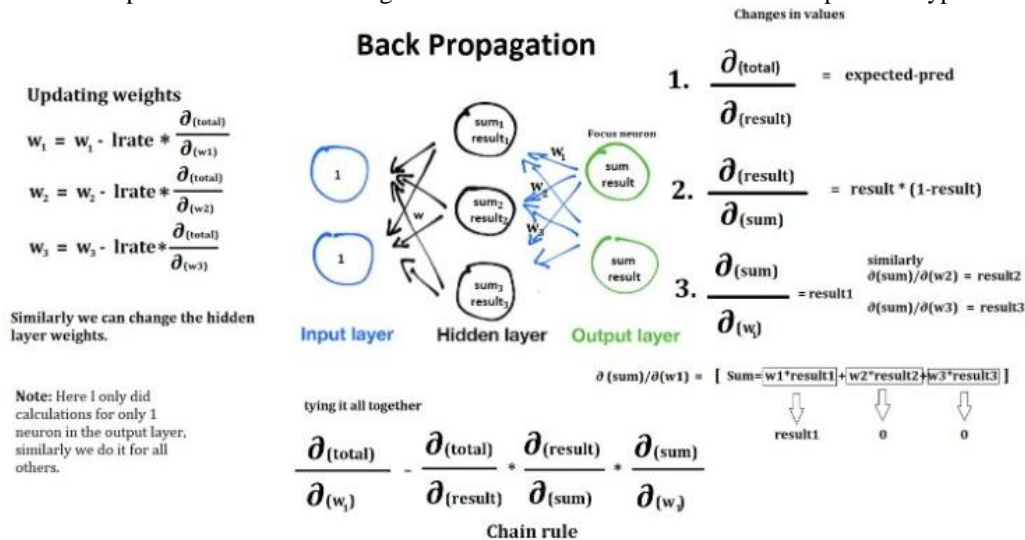

Figure 7 Artificial neural networks with Math

Biological individuals have a logical series of an algorithm, which is regulates their commands and response. These algorithms we may call as a 'Genetic Material' as 'DNA or RNA'; but I would like to see them as an "ALGORITHMS" which is regulates there all activities like responses and commands. Because, particular DNA or RNA sequences have special type of code, which can be used by different performers, here performers are Enzymes.

Reinforcement learning models are a reward or punishment way of learning. This allows to explore and memorize the states of an environment or the actions with a way very similar on how the actual brain learns using the pleasure circuit (TD-Learning). It also has a very useful ability: blocking, which naturally allows a reinforcement learning model to only use the stimuli and information that is useful to predict the reward, the useless stimuli will be "blocked" that is filtered out. This is currently being used in combination with deep learning to model more biologically plausible and powerful neural networks that can for example maybe solve the Go game problem[9].

Genetic algorithms are best used when facing a very high dimension problem or multimodal optimizations - where you have multiple equally good solutions, aka multiple equilibriums. Also, a big advantage is that genetic algorithms are

---

[9] DeepMind AlphaGo, Visit: https://deepmind.com/research/alphago/

derivative-free cost optimization methods, so they are very generic and can be applied to virtually any problem and find a good solution - even if other algorithms may find better ones.

Especially, using Machine Learning (ML) techniques to probe statistical regularities and build impeccable data-driven products around neural networks. Recently, there has been an upsurge in the availability of many easy-to-use machine and deep learning packages[10] such as scikit-learn, Weka, Tensorflow etc. Some of the classically popular methods include Perceptron, Back-Propagation, Hopfield Network, Self-Organizing Map (SOM), and Learning Vector Quantization (LVQ).

3.1 Deep Learning

Deep Learning methods are a modern update to Artificial Neural Networks that exploit abundant cheap computation. These are concerned with building much larger and more complex neural networks, many methods are concerned with semi-supervised learning problems where large datasets contain very little labelled data (See Below Table 1: A typical table showing formatted research data available for download).

Deep learning is a growing field of artificial intelligence (Marcus, 2012). It can provide opportunities for a challenging and lucrative career. Especially, if you are interested in computer vision challenges like image recognition, it can provide you the skills you need. The suit of deep learning algorithm comprises on Restricted Boltzmann Machine (RBM), Deep Belief Networks (DBN), Convolutional Network, and Stacked Auto-encoders.

4. Algorithms and Complex Optimizations

The complexity of Human brain is to process data and information if could be expressed in mathematical terms or any set of systems, then we can almost mimic it using machines. Machine learning algorithms suffer variance of some degree. This is important for understanding the computational efficiency (Krishna, 2017) and scalability of our Machine Learning Algorithm and for exploiting sparsity in our datasets. *Knowledge of data structures such as Binary Trees, Hashing, Heap, Stack etc., Dynamic Programming, Randomized & Sublinear Algorithm, Graphs, Gradient/Stochastic Descents and Primal-Dual methods are needed[11].*

4.1 Convex Optimization

There are many reasons why the mathematics of Machine Learning is important. Machine Learning theory is a field that intersects statistical, probabilistic, computer science and algorithmic aspects arising from learning iteratively from data using logistic regression and finding hidden insights which can be used to build intelligent applications.
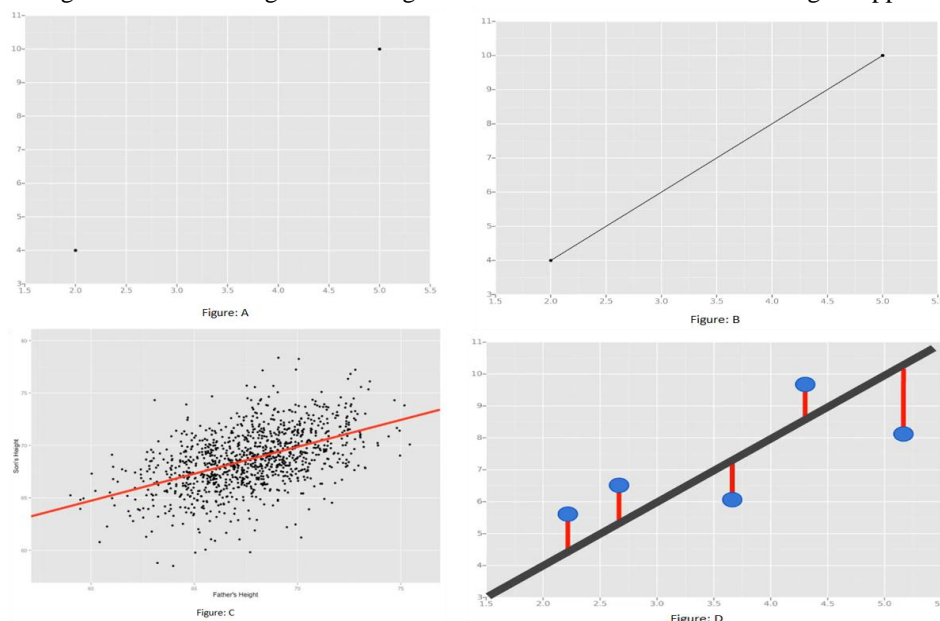

Figure 8 Linear Regression (Machine Learning Algorithm - Implementation in Python Jupyter Notebook)

[10] A Simple Checklist to a lucrative path of becoming an AI Engineer (Deep Learning), Visit:
https://www.linkedin.com/pulse/simple-checklist-lucrative-path-becoming-ai-engineer-deep-kashyap

[11] The Mathematics of Machine Learning by Wale Akinfaderin, Visit:
https://www.linkedin.com/pulse/mathematics-machine-learning-wale-akinfaderin

Probabilistic models, for example, Monte-Carlo, Markov Chains, Markovian processes, Gaussian mixtures, etc. and probabilistic graphical models, for example, Bayesian networks, Credal networks, Markov graphs, etc. are great for uncertain situations since with the help of these we can manipulate uncertain values and hidden variables. Graphical models are kind of close to deep learning, but they are more flexible. It's easier to define a PGM from a semantic of what you want to do than a deep learning network.

Despite the immense possibilities of Machine and Deep Learning, a thorough mathematical understanding of many of these techniques is necessary for a good grasp of the inner workings of the algorithms and getting good results. There's a long way before we can control it and then to master it. Artificial personality is still the one to watch out for. We always need to change our skills with optimism and willingness to change. Let's hope that the people required to make important decisions get more intelligent. When shall machine learning algorithm replace the humans? It will happen. The only question is when.

5. CONCLUSION

Artificial Intelligence revolution is going to happen in phases, in existing scenario we are looking at tasks within the current socio- economic structure which are going to get automated. Beyond this phase next level of evolution is going to happen in changed structure which is again dependent on current level of Artificial Intelligence penetration and direction it is going to take. And nobody can say what new economic social structure would be. And sure it will make lot of things redundant and throw up new opportunities as well. Just like it happened over entire century, only difference would be the pace and cycle of changes will be short. At this stage we're only trying to grasp quantum computing (Neven, 2013) capabilities, much less understand it.

REFERENCES

1. Brownlee, D. J. (2017). *Machine Learning Mastery With R.* Melbourne: Jason Brownlee.
2. Brownlee, J. (2011). *Clever Algorithms: Nature-Inspired Programming Recipes.* Texas Austin: Jason Brownlee.
3. Daelemans, W. (2005). Memory-Based Language Processing. *Cambridge University Press*.
4. Isenkul, M., Sakar, B., & Kursun, O. (2014). Improved spiral test using digitized graphics tablet for monitoring Parkinson's disease. *The 2nd International Conference on e-Health and Telemedicine (ICEHTM-2014)* (pp. pp. 171-175). Istanbul: International Conference on E-Health and TeleMedicine.
5. Kelly, J. a. (2013). Smart Machines: IBM's Watson and the Era of Cognitive Computing. *Columbia Business School Publishing*.
6. Kohavi, R., & Provost, F. (1998). Machine Learning. *Machine Learning*, 271–274.
7. Krishna, D. (2017). *Managing algorithmic risks.* New York: Deloitte Development LLC.
8. Marcus, G. (2012). Is "Deep Learning" a Revolution in Artificial Intelligence? *The New Yorker*.
9. Neven, H. (2013). Launching the Quantum Artificial Intelligence Lab. *Google Research Blog*.
10. PricewaterhouseCoopers. (2017). Will robots steal our jobs? *PricewaterhouseCoopers*, 1-19.
11. Smith, T. C. (2016). Statistical Genomics: Methods and Protocols . In T. C. Frank, *Introducing Machine Learning Concepts with WEKA* (pp. pages 353-378). New York, NY, 2016.: Springer.