WHITEPAPER

STAY AHEAD WITH ARTIFICIAL NEURAL NETWORK

TABLE OF **CONTENTS**

FAST-PACED BUSINESS ENVIRONMENT

Till a few years ago, help features were the life support of tedious and complicated applications. As businesses move into the fast lane with time being a major constraint, the clichéd click-based directions provided by help files have become passé. Today, **voice-based searches** and **voice-activated assistance** has become the new normal. These appendages are powered by technologies such as **Artificial Intelligence (AI) / Machine Learning (ML).**

Similarly, as data proliferation is increasing, it is becoming highly important to seek out **smarter ways** of **data analysis** for continued progress to keep pace with the fast movers in the industry. In recent years, ML has emerged as the vital solution for skimming and analyzing humongous amounts of data to derive **quick, actionable insights.**

ML algorithms allow machines to recognize patterns, construct prediction models, or classify images or videos through learning. ML algorithms can be implemented using a wide variety of methods such as decision trees, random forest, clustering, neural network, and more. **Artificial Neural Networks (ANNs)** come under **Deep Learning,** which in nothing but **ML,** and ML is a subfield of **AI.**

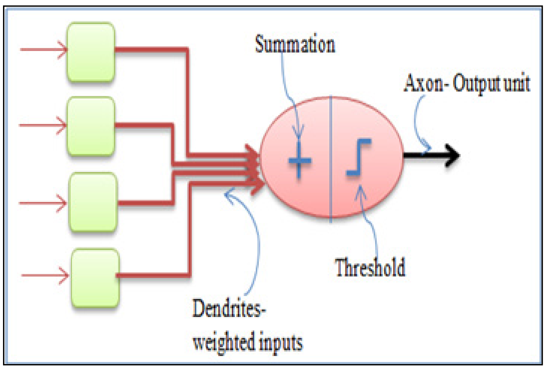
Fast analysis and **interpretation of unstructured data** is what ANNs bring to the table in a fast paced, complex business environment.

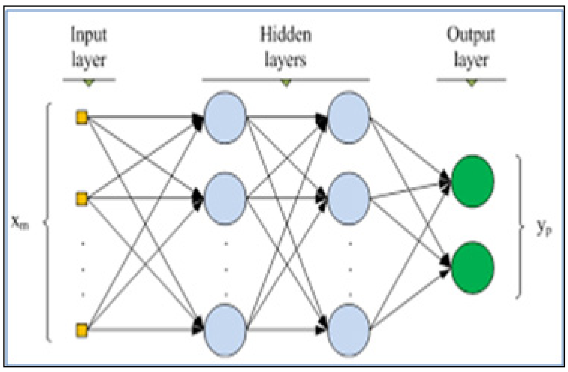
ANNS VERSUS BIOLOGICAL NEURAL NETWORKS

Numerous researches in AI and ML have led to the discovery about the presence of **brain-inspired neural networks** in many real-life applications. The purpose of this paper is to discuss the design of **ANN architecture,** which has gathered a powerful momentum in the recent years as one of the most efficient and **fast learning methods** to solve **complex problems.**

ANNs are **mathematical models** that are inspired by the way **biological neural networks** in the human brain **process information,** which is capable of self-learning as well as pattern recognition. ANNs are generally presented as systems of highly interconnected **"neurons",** which can compute values from inputs.

The following diagrams are basically a comparison between the biological neural network and artificial neural network. This particular ANN has three layers of networks: **the input layer, the middle layer,** and **the output layer.** We can consider each circle to be a neuron, so each neuron accepts input, processes it, and produces the output. When the output produced in the first neuron is passed to the next neuron, a weight gets multiplied to it. Then, the new value gets accepted by the next neuron as its input. The **input** and **output** of the ANN can be compared to **dentrites** and **axons** of biological neural networks. **Weight multiplication** can be compared to **synapses** of a biological neural network.





As in a biological neural system, the **smallest unit** of the ANN is the **neuron.** These neurons are connected with each other with immense **parallel processing power** and try to replicate the behavior of a real brain. ANNs adopt and modify their architecture in order to learn similar to how a human brain would function, use input from the outside world, and modify the information that is already collected or modify their internal structure. To put it simply, the ANNs **change weights of connections** based on **input** and **desired output.**

If we see the structure of an ANN, there are a few components, where we can do changes if we want to modify the architecture; for example:

- Create new connections among neurons
- Add and delete neurons
- Modify input functions or activation function or
- Change the weights of connections

Here, we can see that **changing weights** is the **most practical approach.**

THE WORKING OF AN ANN IN COMPLEX BUSINESS ENVIRONMENT

In real business problems, we usually have some collected data based on which we need to create our model for predictions, or classification, or any other processing. This data is called the **knowledge base.** In fact, based on **behavior during the training** and the **nature of knowledge base,** learning is divided into three classes:

01

Unsupervised Learning

Knowledge base contains only inputs. The network attempts to identify similar inputs and to put them into categories.

02

Reinforcement Learning

Knowledge base contains inputs, but the network is also provided with additional pieces of information during the training. What happens is that once the network calculates the output for one of the inputs, we provide information that indicates whether the result was right or wrong and possibly, the nature of the mistake that the network made.

03

Supervised Learning

Knowledge base contains inputs and desired outputs. This way, the network can check if its calculated output is the same as desired output and take appropriate actions based on that.

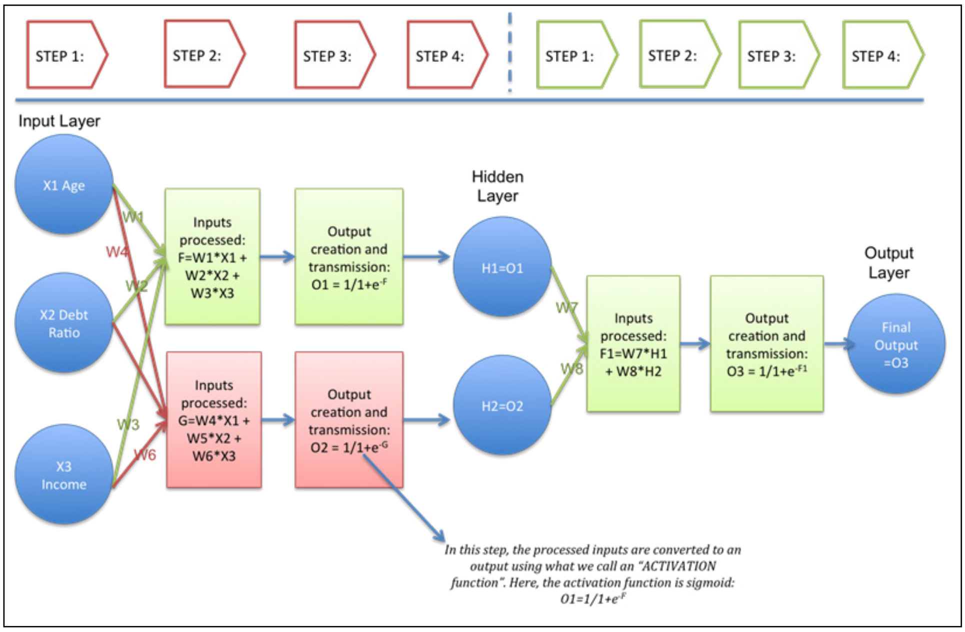
Supervised Learning is the most commonly used form of ANN. Here, we have a **knowledge base** that contains a **vector of input values** and a **vector of desired output values.** Once the network calculates the output for one of the inputs, cost function (for example: mean squared error function) calculates the **error vector.** This **error** indicates how close our guess is to the **desired output.** Now, this error is sent back to the **neural network,** and **weights** are modified accordingly. This process is called **Backpropagation**. It is an **advanced mathematical algorithm,** using which the ANN **adjusts all weights at once.**

The entire point of training is to set the **correct values** to the **weights,** so we get the **desired output** in our **neural network.** This means that we are trying to make the **value of the error vector** as small as possible, i.e., to find a global minimum of the cost function. For this purpose, we use technique called **gradient descent.**

To summarize:

1. The whole processes starts by randomly initializing weights in the **neural network connections**
2. Then we send the **first set of input values** to the neural network and propagate the values through it to get the **output value.**
3. We compare **output value** to the **expected output value** and calculate the **error** using **cost functions.**
4. We propagate the **error** back to the **network** and set the **weights** according to that **information**
5. Repeat **steps** from **2** to **4** for **every input value** we have in the **training set.**
6. When the **entire training set** has been sent through the **neural network,** we have finished **one epoch.** After that, we repeat **more epochs.**

The following figure shows the common steps in a single layer neural network:



ANN IN DIFFERENT BUSINESS SCENARIOS

Some of the real-time business applications of ANN are given below:

- Voice search and voice-activated assistants
- Classification, pattern and sequence recognition, pattern detection and sequential decision making
- Data processing, including filtering, clustering, blind signal separation and compression
- Search engine design
- Automatic translation of documents
- Many security applications, for example: for access control, use face recognition as one of its components
- Facial recognition by systems that learn about features, which are relevant for identifying a person
- Translation between two languages
- Time series prediction and modeling
- The problem of identifying entities, such as places, titles, names, actions, etc. from documents
- Futures Market Trading, which helps in trading in financial instruments and commodities
- Speech recognition
- Automatic coloration of black and white images
- Direct marketing (companies use past purchase behavior to guesstimate whether you might be willing to purchase even more)
- Image recognition to help recognize people and objects in images as well as understanding context and content; for example: Social Media, Tourism, and Retail
- Image caption generation that helps to describe the contents of an image

VARIANTS OF NEURAL NETWORKS

There is a "zoo" of different neural network paradigms. They have completely different structure, different learning philosophies, and hence completely different areas of application. The most common neural networks are:

- Single layer perceptions
- Multi-Layer Perceptions (MLP), often trained with Back propagation of Error
- Radial-Basis Function Networks (RBF)
- Recurrent MLPs
- Long short-term memory (LSTM)
- Self Organizing Feature Maps (SOM)
- Adaptive Resonance Theory (ART)
- Hopfield Networks
- Boltzmann Machines, Restricted Boltzmann Machines
- Convolutional Networks and Deep-Learning networks
- Generative Adversarial Networks (GANs)

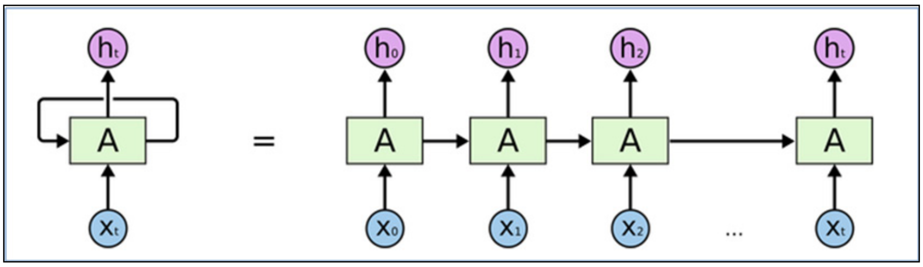
And there are a lot of other less common, specialized approaches.

SOME MOST COMMON NEURAL NETWORKS AND THEIR APPLICATIONS

01

Recurrent Networks

Recurrent Net dynamically captures the structure of the data and uses it to predict the next element in the series. Those elements could be anything like next words in a sentence or next letter in a word or next number in a data series. **Recurrent Nets** are mainly used for natural language generation. The basic architecture of RNN is as follows:



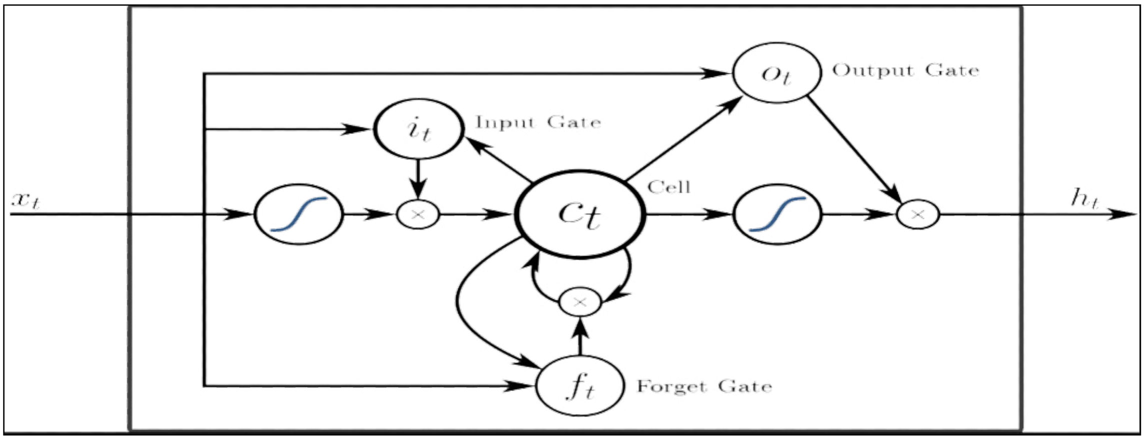
02

Long Short-Term Memory

Long Short-Term Memory (LSTM) is a special type of RNN that is designed to **model temporal sequences** and their **long-range dependencies** more accurately than usual RNNs.

LSTM has different architecture as it does not use activation function within its recurrent components, the stored values are not modified, and the gradient does not tend to vanish during training. Usually, LSTM units are implemented in “blocks” with several units.

LSTM is used for language modeling, image captioning (with and without attention), hand writing generation, image generation using attention models, question answering, video to text.

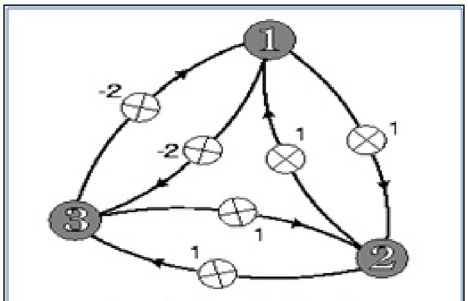


03

Hopfield Networks

Hopfield Network is initially trained to store a number of patterns or memories. It is then able to recognize any of the learnt patterns by providing only partial or even some corrupted information about that pattern. It is single-layered and recurrent network i.e., each neuron is connected to every other neuron in the network.

In Hopfield Network, all the nodes are fully interconnected, i.e. each node is an input for all the other nodes in the network. The following figure shows the architecture of a 3-node Hopfield Network:

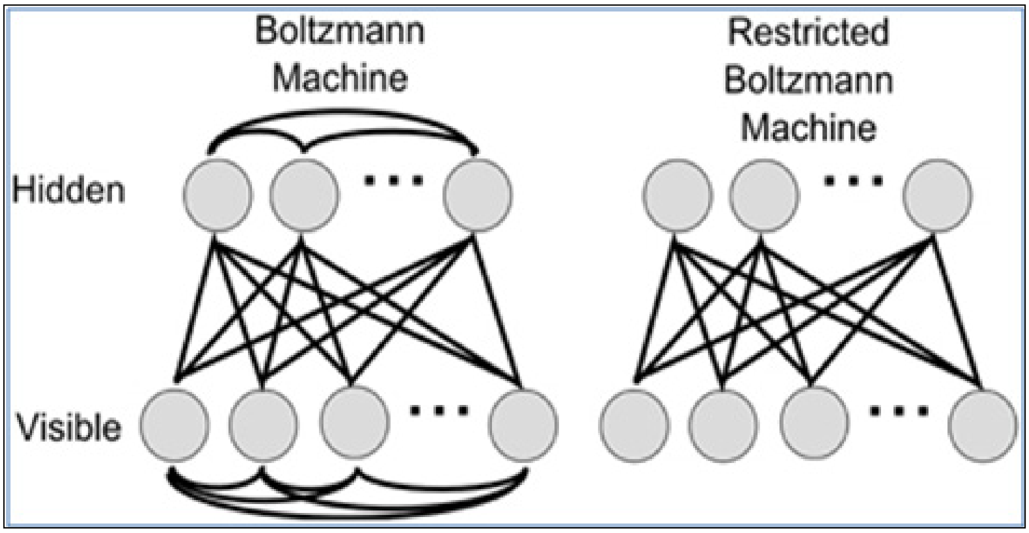


3 node Hopfield net

04

Restricted Boltzmann Machine

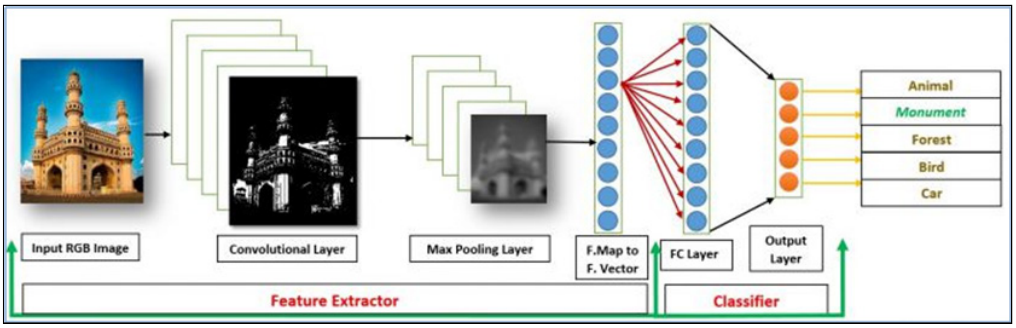
Restricted Boltzmann Machine (RBM) is an unsupervised deep stochastic neural network (also known as an undirected graph model) essentially performs a binary version of factor analysis. In RBM, we map the training data into different dimensional space and use constructive divergence algorithm to update the neurons.
RBM is used for dimensionality reduction, classification, regression, collaborative filtering (Recommendation Engine), feature learning, and topic modeling.



05

Convolutional Neural Networks

Convolutional Neural Networks come under deep artificial neural networks that are used to classify images and videos. Convolution is the integral measuring of how much two functions overlap as one passes over the other. There are numerous architectures in the field of Convolutional Networks. The most common are: LeNet, AlexNet, ZF Net, GoogLeNet, VGGNet, and ResNet.



06

Sequence-to-Sequence Models

Sequence-to-Sequence model is usually a combination of **two recurrent neural networks** – an **encoder** for processing the input and a **decoder** to produce the output. The encoder and decoder can use the same or different sets of parameters.

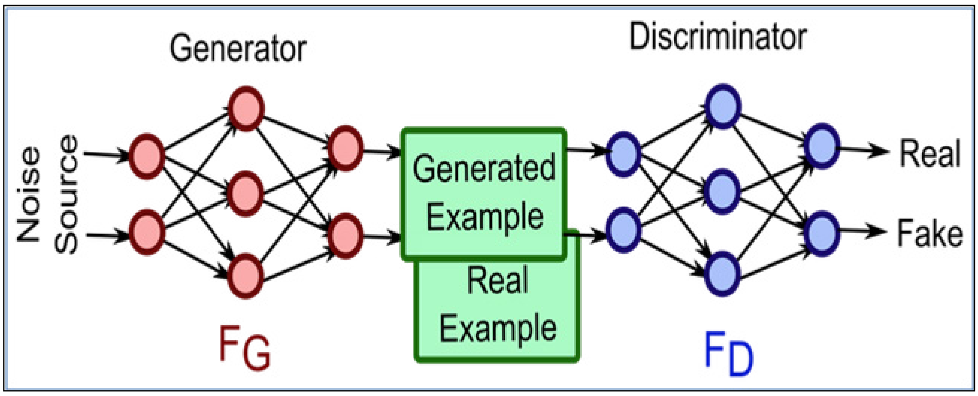
Sequence-to-Sequence models are mainly used in question answering systems, chat bots, and machine translation.

07

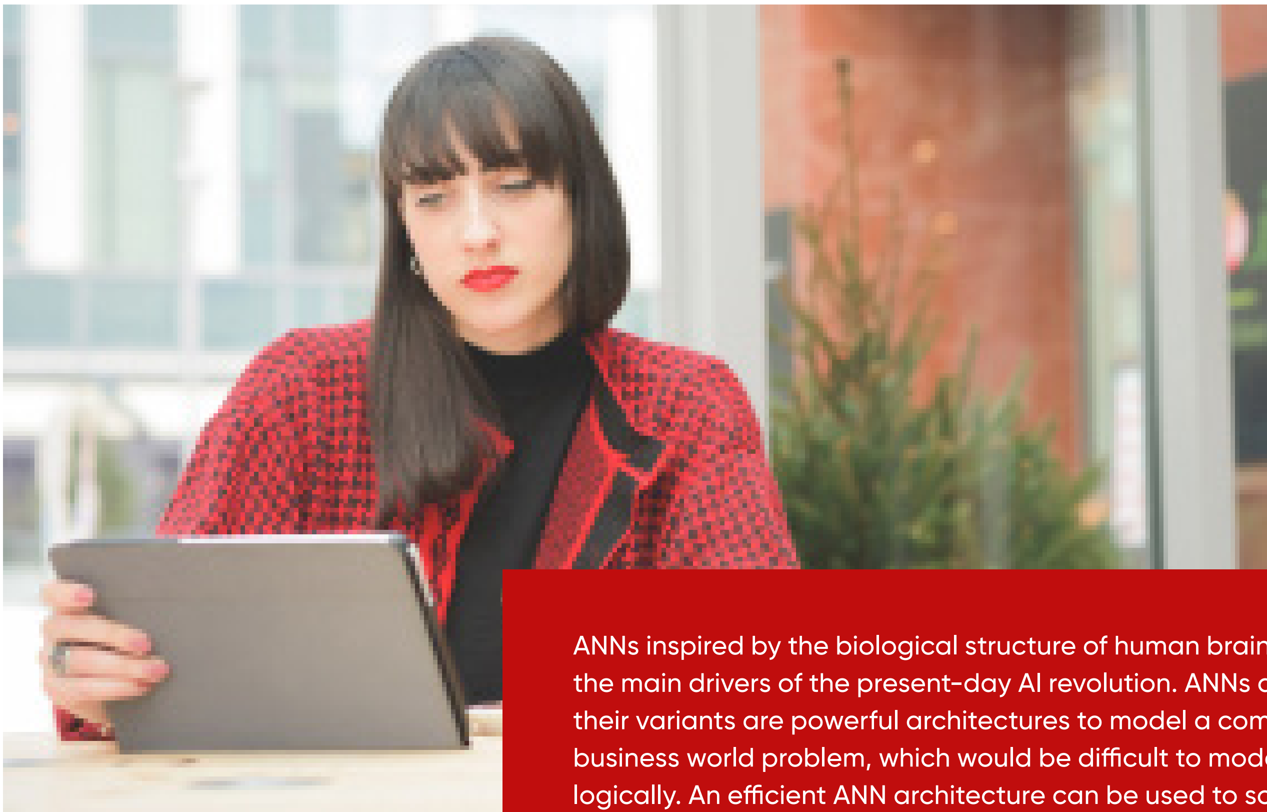
Generative Adversarial Networks (GANs)

In GAN Model, we simultaneously train two models – a **generator** that learns to output fake samples from an unknown distribution and a **discriminator** that learns to distinguish fake from real samples.

If we are in process of developing a complete AI system, the system shouldn't just interpret the world; it also should generate something. GANs are designed to provide the solution for this kind of problems.



CONCLUSION



ANNs inspired by the biological structure of human brain are the main drivers of the present-day AI revolution. ANNs and their variants are powerful architectures to model a complex business world problem, which would be difficult to model logically. An efficient ANN architecture can be used to solve complex non-linear problems instead of defining functions with high degree polynomials, which may lead to over fitting. Parallel processing using ANNs expedites analysis and saves time, effort, and money as compared to other conventional approaches.

REFERENCES

Prof. Leslie Smith, “An Introduction to Neural Networks”, University of Stirling, 1996, 98, 2001, 2003

Vincent Cheung and Kevin Cannons, “An Introduction of Neural Networks”, Manitoba, Canada, May 27, 2002.

Anil K Jain, Jianchang Mao and K.M Mohiuddin, “Artificial Neural Networks: A Tutorial”, Michigan State university, 1996.

Introduction to Artificial Neural Networks, Data Science foundation,2017

https://www.iisc.ac.in/from-expert-to-amateur-a-simple-method-to-fool-your-deep-ne...,2018

https://medium.com/@tharanignanasegaram/artificial-neural-network-a-brief-introduction-572d462666f1

ABOUT
THE AUTHOR

DR. ASHUTOSH KUMAR PANDEY

Lead Consultant – Artificial Intelligence & Cognitive Sciences

Ashutosh is a M.Sc. (Mathematics) from the Indian Institute of Technology Roorkee and holds a Ph.D. in the area of Optimization Techniques from the Banaras Hindu University.

Ashutosh has over 11+ years of experience in IT industry. He has implementation experience in machine learning, predictive modeling, forecasting, optimization, and transportation across industry verticals including Banking, Retail/ CPG, Insurance, Healthcare and Telecom. He has a good expertise in unsupervised and supervised machine learning techniques, for example artificial neural network, Random forest, KNN and has the know-how to translate real-time business issues into suitable machine learning solutions.

Ashutosh is a M.Sc. (Mathematics) from the Indian Institute of Technology Roorkee and holds a Ph.D. in the area of Optimization Techniques from the Banaras Hindu University.

ABOUT **DATAMATICS**

Datamatics provides intelligent solutions for data-driven businesses to increase productivity and enhance the customer experience. With a complete digital approach, Datamatics portfolio spans across Information Technology Services, Business Process Management, Engineering Services and Big Data & Analytics all powered by Artificial Intelligence.

It has established products in Robotic Process Automation, Intelligent Document Processing, Business Intelligence and Automated Fare Collection.

Datamatics services global customers across Banking, Financial Services, Insurance, Healthcare, Manufacturing, International Organizations, and Media & Publishing.

The Company has presence across 4 continents with major delivery centers in the USA, India, and Philippines. To know more about Datamatics, visit www.datamatics.com

FOLLOW US ON

_012022

website: datamatics.com | email: business@datamatics.com

USA | UK | UAE | India | Philippines