Artificial Intelligence

Comprehensive Information

Collected by Sudha Mantri (Dean, Leesburg School of AI, VA, USA)

V 1.0
Dated 8/16/2018

Note: This will be continually updated and shared.

Table of Contents

Disclaimer

I thank Siraj, all the internet sources, experts, and School of AI team members. This information is not meant to violate any copyrights. Kindly know that I am an eternal student (always learning) and making notes (literally). However, I wish to add that this information is meant to be used by School of AI Deans/Wizards to help members and aspirants learn AI Free of charge. Please let me know of any copyright violations and I shall duly oblige.

- *School of AI Deans please use this information please feel free to use this document and/or create your own sessions/training as you deem fit

Below is the School of AI Guidelines for all the deans. Please follow them and share information if/as need be.

https://gallery.mailchimp.com/21a24cf96b7ec98b2144fca5c/files/c9a14833-10b2-42c0-a024-f5310cb395bc/Copy_of_School_of_AI_Meetup_Guidelines.pdf

School_of_AI_Meetup_Guidelines.pdf

Version History

V 1.0	8/16/2018	Initial Version (Draft) meant for School of AI Deans as reference – Use to teach others for FREE

History of AI

https://en.wikipedia.org/wiki/History_of_artificial_intelligence

Year	AI Significance
1936	The Turing Machine. Massive data analysis, recognition, and decoding.
1938	Learning is major feature of AI. Operant Conditioning (OC or Thorndike's law of effect). OC brought data science (Psychometry), making psychology a quantitative field. A being will learn a behavior if conditioned using OC techniques.
1940	Theory of Adaptive Process Control. Norbert Wiener. Both Reinforcement Learning and Operant Conditioning are based of Theory of Adaptive Process Control.
1947	Transistor
1960s	B.F Skinner took Operating Conditions to the next level. OCs were later used in AI: Example, OCs used for artificial spiking neural networks (ASNN) to act as robot brain controllers for learning.
1952	Spiking Neuron. Alan Lloyd Hodgkin and Andrew Huxley
	Computer could play Tic-Tac-Toe. OXO by Alexander Shafto "Sandy" Douglas
1955	Term 'AI' coined in a workshop. Dartmouth Professor John McCarthy
1956	General Problem Solver (GOFAI- Good Old Fashioned AI). Allen Newell. Herbert Simon Physical Symbol System Hypothesis
1958	Early version of artificial neural network. Frank Rosenblatt coined 'perceptrons'. Mark 1 perceptron. While Rosenblatt was working on his Mark 1, an MIT professor named Marvin Minsky pushed for symbolic systems approach. In 1969, Minsky co-authored a book with Seymour Papert, called "Perceptrons" arguing against Rosenblatt's approach to artificial neural networks. Perceptrons and artificial neural networks languished for nearly a decade. But Geoff Hilton developed a multi layered neural network.
1959	Arthur Samuel. Checkers against itself
1962	First Computer Science Department - Purdue
1964	First Neuro Science Department - aka 'Psycho Biology' - UC Irvine
1968	Intel founded
1970s	Reinforcement learning (RL) - has data, is quantitative and is unrelated to human behavior and is usually used in Machine Learning as an approach to training. Like OC, it is based on adaptive learning/process control. Richard S Sutton pioneering force.
1970s	Dean Pormerleau built a self-driving car
1980	Chinese Room Argument - - John Sorrell
1981	Advent of Personal Computers - IBM
Mid 1980s	Geoff Hinton created a new version of an artificial neural network, based on Rosenblatt's Perceptron. Except his version include several hidden layers. This allowed his artificial neural network to work on much more complicated patterns.
1990	Yann Lecun built OCR for handwritings
1995	Paper released 'Random fonts for the simulation of handwriting'
	http://luc.devroye.org/mcdougallpaper.pdf
1997	IBM DeepBlue beats Kasparov
2000s	First 1 billion transistor processor. As of 2017, the largest transistor count in a commercially available single-chip processor is 19.2 billion— AMD's Ryzen-based Epyc
2002	MyFont released Human like Handwriting Fonts
2011	IBM Watson beats Brad Rutter and Ken Jennings at Jeopardy

2014	Chatbot passes Turing Test. Russian chatterbot named "Eugene Goostman" passes the test if it is mistaken for a human more than 30% of the time. Competition at the Royal Society in London
2017	AlphaGo beats humans. Oct 2017: AlphaGo Zero learnt to play the game of Go simply by playing games against itself, starting from completely random play.
2000s	John Searle – consciousness in artificial intelligence. Presents Weak AI (symbolic systems approach) vs Strong AI. Famous for Chinese Room Argument.
2000s	Planning AI also called Automated planning and scheduling: Takes possible initial states of the world, desired goals, and set of possible actions. Planning AI attempts to synthesize a plan that guarantees a desired goal (i.e end in a goal state) starting from any of the initial states. This is applied in a dynamically changing or unknown environments. For example, unmanned vehicles.
From Mid 1940s	Heuristic reasoning continues to be central to AI. Best choice from start to goal nodes. Heuristic choice depends on domain/problem space: Google Maps, Directions, Legal contracts, logistics, video games and many more.

Definitions, Compare/Contrast

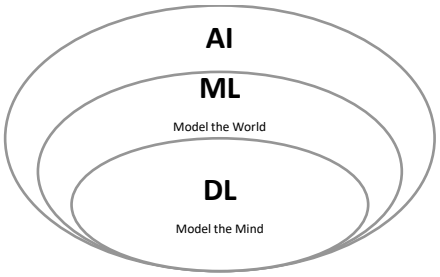
Humans are addictively fascinated with defining and categorizing everything. The benefit I see in that exercise is it makes communication easy (please note I said communication not understanding). There are many definitions for Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL). We can say that right now, AI field is still in its infancy and nomenclature is fluid. There is tremendous opportunity in the field.

Human VS Artificial Intelligence

The exponential increase-in/availability-of digital data provided the impetus for advancements in AI.

Human	Artificial Intelligence
Use 5 senses - visual, audio, skin feeling, smell, and taste – and arguably more senses.	Computer systems that can accomplish some tasks that require human intelligence
Use available information to take decisions (may be more biased and unpredictable/inconsistent – I may like ice cream one day but not the next)	Use available information for some decisions and provide more accurate and consistent results (which may be desirable in some fields but not in others)
Communicate	Interact with humans using their natural language (language processors)
Identify patterns in data (and act/decide) – Amazing to observe this.	Learn from the data and automate repetitive (neural network) tasks based on given algorithms/learning
Remember what people have said	Possibly not as advanced as humans
Adapt to new situations (infinite possibilities as humans are in general good with multiple things)	Learn from their mistakes and adapt to new environments (limited possibilities at the moment but for a specific task – Google AlphaGo – AI seems to be excelling)

AI vs ML vs DL



	Artificial Intelligence (AI)	Machine Learning (ML)	Deep Learning (DL)
Definition	Intelligent machines that are meant to imitate/think/act more like human beings	System learns things without being programmed to do so. Logistic Regression. System is able to make predictions or take decisions	System thinks and learns like human brain using artificial neural networks. Initial dataset is small with less testing time. Performance

		based on the past data. Model being trained based on features. This can be done on smaller amounts of data.	improves with more data and continual learning. Concept of weights, bias, variance. Problem solved in an end to end method. Best features are selected by the system
Progress	Era of Weak AI. Machines in their infancy. Any code or technique or algorithm that enables machines to mimic develop and demonstrate human cognition and behavior.	To move from weak AI to strong AI machines need to learn the ways of humans. Techniques and process of ML help machines in this endeavor Machines learn to predict outcomes on the go by recognizing patterns in the input data.	Machines demonstrate ability to learn deeply by drawing meaningful inferences from large data sets. DL requires Artificial Neural Networks (ANN) in layers which are like biological neural networks in humans. ANNs connect and communicate with each other to make sense of vast input data
Types (Just categorization)	1. Reactive machines (Say clothes are clumped on one side in a washing machine, then machine re-centers clothes). No past experience/memory for a new decision 2. Limited memory (Say neural network programmed to identify car). Machine may use short-lived or limited/past information 3. Theory of mind. Systems able to understand human emotions and adjust self-behavior according to the human needs (popularly known as user experience) 4. Self-awareness (self-fixing. Predicting which action is best for the humans – to make human life happy). Systems are aware of themselves and understand their own internal states. Systems	1. Supervised learning: Systems are able to predict future outcomes based on the past data (system learns from bunch of input images. Then it matches a new image). Requires both input and an output to be given to the model for it to be trained 2. Unsupervised learning: Systems are able to identify hidden patterns from the input data provided. By clustering etc., the patterns and similarities/anomalies become apparent to the system 3. Reinforcement Learning: Systems are given no training. It may learn on the basis of reward/punishment (Yes/No answers)	DL uses different approaches, algorithms to achieve Deeper learning. 1. Multi Layer Perceptrons (MLPs) 2. Convolutional neural net (CNNs)

	predict other people's feelings/actions.		
Examples	News generation – finds news articles based on news feeds, Google command – smart home devices. Amazon prime music, SAAVN, Philips Hue, OLA, NDTV, Zomata, TED, Goibibo, ESPN Cricinfo, Amazon Echo etc., KBS –DENDRAL, MYSIN, XCON/RI	Spam detection, Chats, Mail categories (All mail, Trash, important, spam), search engine result refining (weights - categorized) - Google search.	Sortie (translation) - neural network (different words and patterns) Chat boats – ask a question it gives answer. Specific tasks – IBM deep blue, Google DeepMind, Google AlphaGo. Convert a black and white image to, Automated library categorization of books – translation of handwritten/printed data to digital form
Future is happening	Detect crimes before they happen, humanoid AI helpers	Increase efficiency in healthcare, better marketing techniques, Finance, sports, trading, trading.	Increase personalization – hyper intelligent systems, decision making.

Internet Era	AI Era
A company has something and sells it on internet (implements ecommerce) – Surprise! This does not make it an internet company	Traditional tech company adds neural networks (ML or DL) – Surprise! This does not make it an AI company
Data is power. Internet companies can do pervasive A/B testing and improve - Short cycle time (e.g. shipping packages)	Strategic data acquisition. Unified data warehouse. Spotting Pervasive automation opportunities (theoretically anything a person take a second or less to think could be automated and combined into an automated workflow) – Data Centers, Cyber security etc.
Decisions usually made from top down	Push decision making from CEO to engineers/ product managers
Analysts/Managers may use wireframes etc., for communicating requirements	Product managers focus more on data and achieving 99% accuracy between initial and goal states. Traditional wireframes etc. not applicable

ML vs DL

Machine Learning	Deep Learning
Enables machines to take decisions on their own, based on the past data	Enables machines to take decisions with the help of artificial neural networks
May Need small or large training data (may be spreadsheet or tolls from end of the year or many images of cats)	May need small amount of initial data but Needs a large amount of data to train against.

Works well on low-end systems (laptop)	Needs high end system to work (cloud)
Most features need to be identified in advance and manually coded (lot of human work)	The machine learns the features from the data it is provided
The problem is divided into parts and solved individually and then combined	The problem is solved in an end to end manner may be one or two neural networks (we may need machine and deep learning together)
Testing takes longer	Testing takes less time
Crisp rules, explain why a certain decision was taken	Since the system takes decisions based on its own logic, the reasoning may be difficult to interpret (like a magic black box – we just know it just came up with right answer we don't know how it came there)

Suggested Prerequisites

Following knowledge will help but please do not focus too much on any one area. Mathematicians, Programmers, Visionaries, Domain experts are all needed to make a product successful. The more one knows, the better, but, there is no need to be expert of all. Please know what the goals are and pick the right technology/algorithm/team to make it happen. Start with an example and complete it

1. Math needed for ML and BL
    a. Linear algebra
    b. Calculus,
    c. Probability Theory/ Statistics,
    d. Matrix calculation
2. Domain knowledge (in addition to simply having Datasets)
3. Algorithms (Know what your needs are to pick the best algorithm
4. Technology (Start with what you are comfortable – Python is popular)
    a. Python
    b. Advanced Python – NumPy
    c. Angular JS
    d. JavaScript
5. Datasets

Popular Algorithms

We need to understand 'Bias' and 'Variance' in the context of AI and algorithms and learning/testing Models can be compromised from high bias and/or high variance. One needs to evaluate the model against the data sets (including training and cross validation). But features, relevancy, noisy then training a model get a zero (or close to zero) error will be misleading.

In other words – Simply choosing algorithms and data sets, and training for zero error may not necessarily guarantee perfect outcomes. Domain knowledge is helpful.

Algorithms may be categorized as below (Usually Economic value created goes down the list below)

1. Supervised Learning (A to be B mapping)
    a. Transfer Learning (Learn from one problem to the other)
        i. For example, system learns to recognize cats/dogs/people. There may be lot of pictures of cats and other terrestrial objects. Hence, transfer learning uses what it has learned with terrestrial objects and tries to (for example) beat x ray and radiology images. There is lot of terrestrial data but much less data in radiology images for a given condition.
2. Unsupervised Learning
    a. Google Brain project (System watches you tube and tries to figure out concept of cat) i.e. figure things out by themselves without requiring a lot of label data
    b. autoencoder
3. Reinforced Learning
    a. AlphaGo and also Video games (self-learn by playing against itself)
    b. This needs a lot more data than unsupervised learning but it can be self-learning. E.g. Robotic application that can self-learn by simulating a robot.

Some Examples of System Learning

Going from input A to output B (A -> B)

Input (A)	Output (B)
Picture	Is it you? (Y/N)
Loan Application	Will it repay (Y/N)
A video or a link	Will you click it (Y/N)
Audio	Output text transcript
English	To French (for example)
Image/Radar etc.	Position of other cars (example)

Following are some commonly used algorithms. Please refer to https://docs.microsoft.com/en-us/azure/machine-learning/studio/algorithm-choice which good suggestions as to what to use when.

1. Linear Regression
2. Logistic Regression
3. Decision Tree
4. Naive Bayes
5. SVM
6. kNN

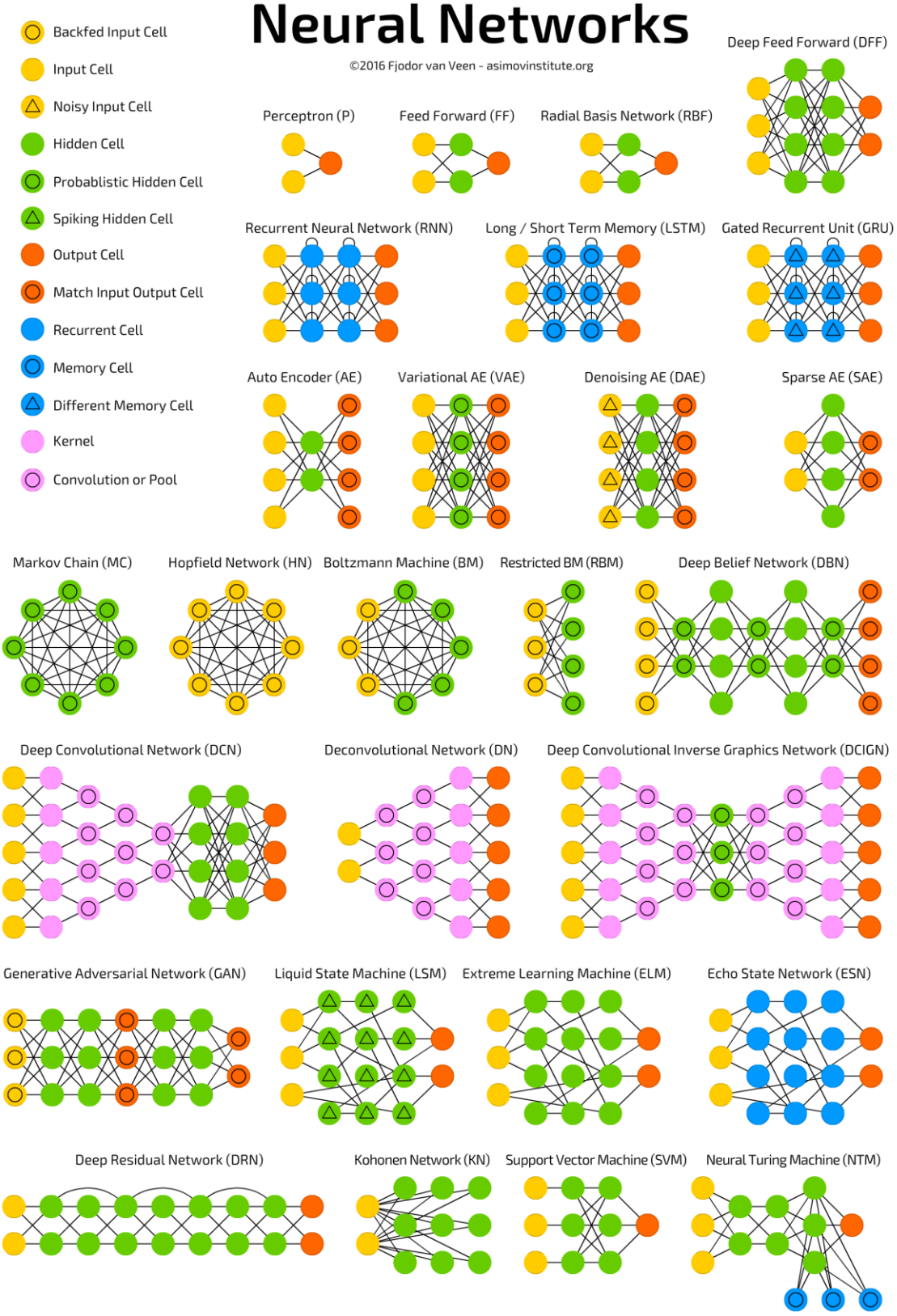
7. Random Forest
8. K-Means
9. Dimensionality Reduction Algorithms
10. Gradient Boosting algorithms
11. GBM

Please refer to the following article for additional information on the neural networks.

1. http://www.asimovinstitute.org/neural-network-zoo/

A mostly complete chart of

Neural Networks



Common neural networks: Following are 6 Types of Artificial Neural Networks Currently Being Used

1. Feedforward Neural Network – Artificial Neuron
2. Radial basis function Neural Network
3. Kohonen Self Organizing Neural Network
4. Recurrent Neural Network (RNN) – Long Short-Term Memory
5. Convolutional Neural Network
6. Modular Neural Network

Datasets

Free dataset (please remember – domain knowledge goes far!). Please search and you will find many more. Below are just a few examples

- https://github.com/awesomedata/awesome-public-datasets
- https://www.kaggle.com/datasets
- USA Census Data: https://www2.census.gov/acs2013_1yr/summaryfile/
- Cause of Death - https://wonder.cdc.gov/
- Airbnb, Walmart etc.

Environment

1. Python, anaconda, IDLE
2. Jupiter note book – Windows
3. Cloud
    a. Anaconda cloud
    b. AWS cloud
    c. Azure

Hands On Training

TBD. Please feel free to provide suggestions. This will be progressively filled as we move forward. Deans, please feel free to do your own sessions/training as you deem fit.

Project 1

TBD. Please feel free to provide suggestions. This will be progressively filled as we move forward Deans, please feel free to do your own sessions/training as you deem fit.

Resources

- School of AI owes it to Siraj Raval.
    - https://www.youtube.com/channel/UCWN3xxRkmTPmbKwht9FuE5A
- YouTube and WWW are good resources. There are many experts but below are just a few individuals:
    - Geoffrey Hinton, Andrew Ng, Demis Hassabis, Richard Sutton, Dag Kittlous, Ray Kurzweil, Yann Lecun, Stuart Russel, Jurgen Schmidhuber, Nathan Benaich, Alex Champandard, Jana Eggers, Adam Coates, John C Havens.
- Below are a few institutions/ free courses
    - http://openclassroom.stanford.edu/MainFolder/CoursePage.php?course=DeepLearning
    - https://ai.google/education/
    - https://www.class-central.com/course/udacity-intro-to-deep-learning-5681
    - https://www.coursera.org/learn/machine-learning
    - https://www.class-central.com/course/edx-machine-learning-7231
    - https://www.class-central.com/course/fundamentals-of-deep-learning-for-computer-vision-10730
    - https://selfdrivingcars.mit.edu/