VIT UNIVERSITY
VELLORE INSTITUTE OF TECHNOLOGY

VIT®
BHOPAL
www.vitbhopal.ac.in

Introduction to Machine Learning

(Basics of Machine Learning)

Dr. Virendra Singh Kushwah

Assistant Professor Grade-II

School of Computing Science and Engineering

Virendra.Kushwah@vitbhopal.ac.in

7415869616

About the Course

- Course Code: CSA4008
- Course Title: Applied Machine Learning
- Course Type: LTP
- Course Credit: 04

Course Objectives

- To understand the concepts of supervised and unsupervised learning techniques.
- To analyze the regression, classification and clustering techniques and to implement their algorithms.
- To evaluate the performance of various machine learning techniques and to select appropriate features for training machine learning algorithms.

Course Outcomes

- Identify the characteristics of machine learning that makes it useful to solve real-world problems.
- Provide solution for classification and regression approaches in real-world applications.
- Choose an appropriate clustering technique to solve real world problems.
- Understand the methods to reduce the dimension of the dataset used in machine learning algorithms.
- Indicate a suitable machine learning model, implement and examine the performance of the chosen model for a given real world problems.

UNIT I
Introduction to Machine Learning

- Preface of Machine Learning, Types of Learning : Supervised - Unsupervised Learning- Reinforcement- theory of learning – feasibility of learning – error and noise – training versus testing – theory of generalization – generalization bound – approximation-generalization tradeoff – bias and variance – learning curve - Finite and Infinite Hypothesis Spaces, Probably Approximately Correct (PAC) Learning-Bayes theorem ,MDL principle.

UNIT II
Supervised Learning

- Learning a Class from Examples, Linear, Non-linear, Multi-class and Multi-label classification, Generalization error bounds: VC Dimension, Decision Trees: ID3, Classification and Regression Trees, Regression: Linear Regression, Multiple Linear Regression, Logistic Regression. Neural Networks: Introduction, Perceptron, Multilayer Perceptron, Support vector machines: Linear and Non-Linear, Kernel Functions, K-Nearest Neighbors.

UNIT III
Ensemble Learning and Unsupervised Learning

VIT®
BHOPAL
www.vitbhopal.ac.in

- Ensemble Learning Model Combination Schemes, Voting, Error-Correcting Output Codes, Bagging: Random Forest Trees, Boosting: Adaboost. Introduction to clustering: Hierarchical: Partitional: K-means clustering, K- Mode Clustering, Self-Organizing Map, Expectation Maximization, Gaussian Mixture Models- Principal components analysis (PCA), Locally Linear Embedding (LLE), Factor Analysis.

UNIT IV
Reinforcement Learning

- Passive reinforcement learning – Direct utility estimation - Adaptive dynamic programming- Temporal difference learning- Active Reinforcement learning –Exploration –learning an action-utility function -Generalization in reinforcement learning – policy search - Applications.

VIT®
BHOPAL
www.vitbhopal.ac.in

Unit V
Neural Networks and Machine Learning in Practice

- Introduction to Neural Networks - Fundamental concepts: neuron models and basic learning rules-Multilayer neural networks and back-propagation - Machine Learning in Practice Design, Analysis and Evaluation of Machine Learning Experiments.

Books / e-Books

1. Ethem Alpaydin, Introduction to Machine Learning , MIT Press, Prentice Hall of India, Third Edition 2014.
2. Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar ”Foundations of Machine Learning, MIT Press, 2012.
3. Tom Mitchell, Machine Learning, McGraw Hill, 3rd Edition, 1997.
4. Charu C. Aggarwal, Data Classification Algorithms and Applications, CRC Press, 2014.
5. Kevin P. Murphy ”Machine Learning: A Probabilistic Perspective”, The MIT Press, 2012.

List of Experiments as per Syllabus

VIT BHOPAL
www.vitbhopal.ac.in

Implement Decision Tree learning.
Implement Logistic Regression.
Implement classification using Multilayer perceptron.
Implement classification using SVM.
Implement Adaboost Algorithm.
Implement Bagging using Random Forests
Implement K-means Clustering to Find Natural Patterns in Data.
Implement Principle Component Analysis for Dimensionality Reduction.
Evaluating ML algorithm with balanced and unbalanced datasets.
Comparison of Machine Learning algorithms.

Let us Start the Course

Preface of Machine Learning

- Machine learning is an integral part of many commercial applications and research projects today, in areas ranging from medical diagnosis and treatment to finding your friends on social networks. Many people think that machine learning can only be applied by large companies with extensive research teams.
- you can build your own system for finding out how people feel on Twitter, or making predictions about global warming. The applications of machine learning are endless and, with the amount of data available today, mostly limited by your imagination.

- Arthur Samuel (1901-1990) was a pioneer of artificial intelligence research. From 1949 through the late 1960s, he did the best work in making computers learn from their experience. His vehicle for this was the game of checkers. Programs for playing games often fill the role in artificial intelligence research that the fruit fly Drosophila plays in genetics. Drosophilae are convenient for genetics because they breed fast and are cheap to keep, and games are convenient for artificial intelligence because it is easy to compare computer performance with that of people.

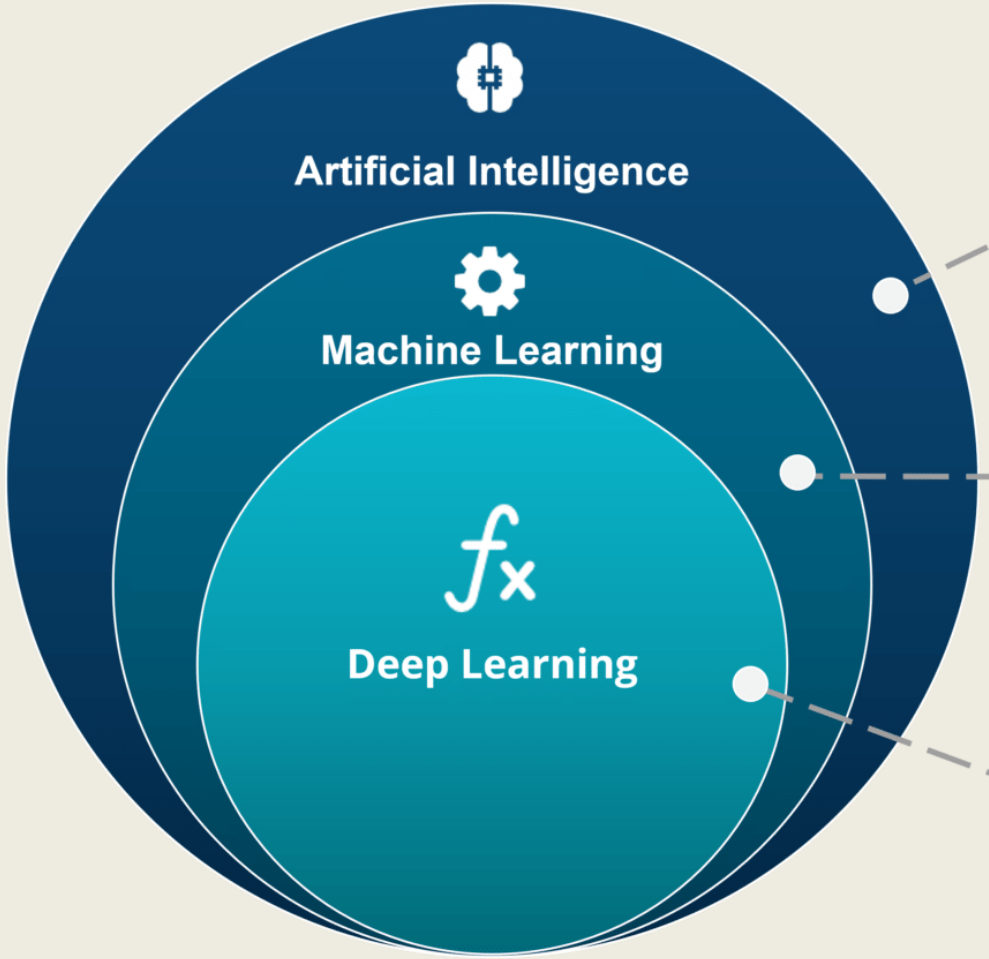
- The term Machine Learning was first coined by Arthur Samuel in the year 1959. Looking back, that year was probably the most significant in terms of technological advancements.

- If you browse through the net about ‘what is Machine Learning’, you’ll get at least 100 different definitions. However, the very first formal definition was given by Tom M. Mitchell:
- *“A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E.”*
- In simple terms, Machine learning is a subset of Artificial Intelligence (AI) which provides machines the ability to learn automatically & improve from experience without being explicitly programmed to do so. In the sense, it is the practice of getting Machines to solve problems by gaining the ability to think.

What is Machine Learning?

- Machine Learning is a concept which allows the machine to learn from examples and experience, and that too without being explicitly programmed. So instead of you writing the code, what you do is you feed data to the generic algorithm, and the algorithm/ machine builds the logic based on the given data.

- Machine Learning is a subset of artificial intelligence which focuses mainly on machine learning from their experience and making predictions based on its experience.



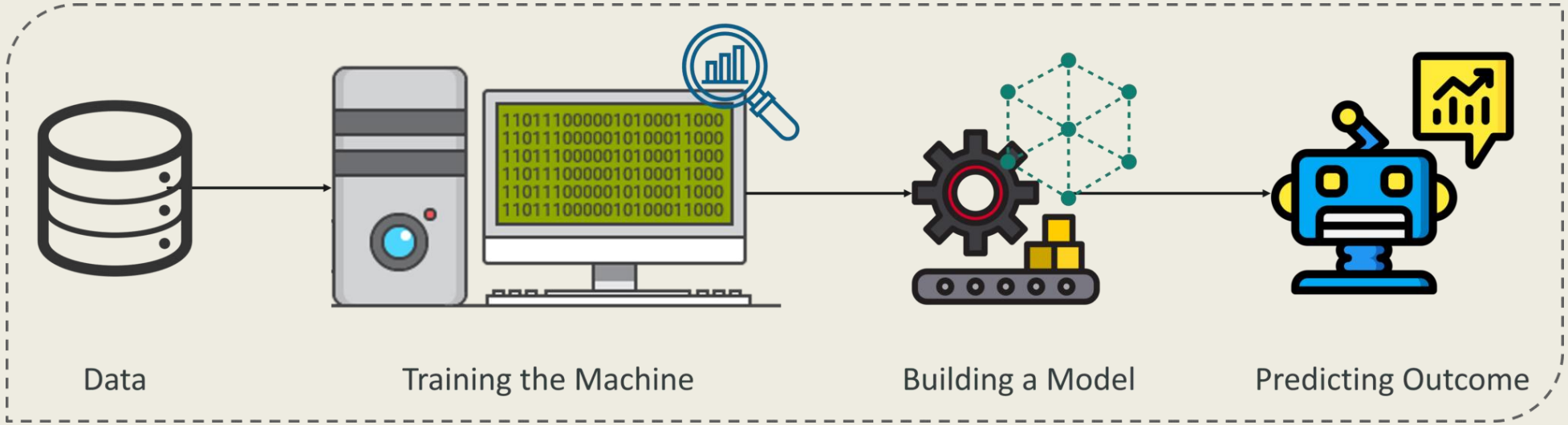
ARTIFICIAL INTELLIGENCE
A technique which enables machines to mimic human behaviour

MACHINE LEARNING
Subset of AI technique which use statistical methods to enable machines to improve with experience

DEEP LEARNING
Subset of ML which make the computation of multi-layer neural network feasible

What does it do?

- It enables the computers or the machines to make data-driven decisions rather than being explicitly programmed for carrying out a certain task. These programs or algorithms are designed in a way that they learn and improve over time when are exposed to new data.



Let us understand by examples

VIT®

BHOPAL

www.vitbhopal.ac.in