About the Authors

Dr.K.Dharmarajan is with the Department of Computer Applications, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, as Professor and Associate Director-Digital Infrastructure. He has Doctorate in Computer Science at Bharathiar University, Coimbatore. He has Published 12 Patents, 10 Books and more than 80 Research Articles in Various Refereed Journals and Conferences. He has Produced 6 Ph.D Candidates, more than 10 M.Phil. Candidates and guiding 7 Ph.D. Research Scholars. He has Received Various Awards likes Distinguished Lifetime Leadership Award, Outstanding Academic Leadership Award, Bhishma Acharya Award, Outstanding Director Award, Faculty Excellence Awards, 2025, 2024, 2023, 2022 and 2021, Lifetime Achievement Award for Education, Best Teacher Award, Innovation Launched Award, Young Scientists Award and Best Social Worker Award. He is the Editor and Reviewer for Web of Science and Scopus Journals. He has Organized and Participated in many Conferences, Workshops and Seminars. Some of the other Key Roles he is Administering include Chief Superintendent of - UPSC, TNPSC and Exam Committee Members of ICAI, ICSI. He has served as an NSS Programme Officer and Lions Clubs Member of Society. He has 19+ years of Experience out of 15+ in Teaching and 3+ years in the Software Industry as a software Professional. His area of Specialization is Web Mining, Machine Learning, Cloud Computing and Big Data.

Dr.K.Abirami is with the Department of Advanced Computing & Analytics, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, as Assistant Professor and Research Coordinator. She has his Doctorate in Computer Science at Vels University, Pallavaram. She has Published 5 Patents, 5 Books and more than 30 Research Articles in Various Refereed Journals and Conferences. She has Guiding 6 Ph.D. Research Scholars. She has Received Various Awards likes Young Achiever Award, Best Innovative Award, Best Researcher Award, Faculty Excellence Awards 2024, 2023, 2022 and 2021, Lifetime Achievement Award for Education, Best Teacher Award, Young Scientists Award and Best Social Worker Award. She has Participated in many Conferences, Workshops and Seminars. She has 20 Years of Experience Out of 16 in Teaching and 4 in the Industry. His Area of Specialization is Web Mining, AI, Machine Learning, Cloud Computing and Big Data.

T.Haripriya is with the Department of Computer Applications, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, as Ph.D. Research Scholar. She has his M.Phil. in Computer Science at Ambedkar College, Chennai. She has Participated in some Conferences, Workshops and Seminars. His Research Work is Based on the Medical Field using Image Processing on the Liver Cancer Detection.

Leilani Katie Publication and PRESS
International Publishing Partner
142, Periyar Nagar, Madakulam,
Madurai - 625003, Tamil Nadu, India.
+91 9894820237 | +91 9790120237
Since 2023
SCAN NOW
ISBN 978-93-6348-379-8
9 789363 483798

Machine Learning: Concepts, Algorithms, and Real-World Applications
Dr.K.Dharmarajan | Dr.K.Abirami | T.Haripriya

Machine Learning:
CONCEPTS, ALGORITHMS, AND REAL-WORLD APPLICATIONS

Authors
Dr.K.Dharmarajan
Dr.K.Abirami
T.Haripriya

ISBN: 978-93-6348-379-8

Machine Learning: Concepts, Algorithms and Real-World Applications

Dr.K.Dharmarajan
Professor,
Associate Director-Digital Infrastructure,
Department of Computer Applications,
School of Computing Sciences,
Vels Institute of Science, Technology & Advanced Studies (VISTAS),
Chennai, Tamil Nadu, India.

Dr.K.Abirami
Assistant Professor & Research Coordinator,
Department of Advanced Computing & Analytics,
School of Computing Sciences,
Vels Institute of Science, Technology & Advanced Studies (VISTAS),
Chennai, Tamil Nadu, India.

T.Haripriya
Research Scholar,
Department of Computer Applications,
School of Computing Sciences,
Vels Institute of Science, Technology & Advanced Studies (VISTAS),
Chennai, Tamil Nadu, India.

Published by
Leilani Katie
Publication and PRESS
International Publishing Partner
142, Periyar Nagar, Madakulam,
Madurai - 625003, Tamil Nadu, India.
+91 9894820237 | +91 9790120237
leilanikatiepublication@gmail.com
www.leilanikatiepublication.org
Since 2023

Title: Machine Learning: Concepts, Algorithms and Real-World Applications

Authors: Dr.K.Dharmarajan
Dr.K.Abirami
T.Haripriya

Published by: Leilani Katie Publication and PRESS
Madurai – 625003, Tamil Nadu, India.

Edition Details: I

ISBN: 978-93-6348-379-8

Month & Year: February, 2026

Pages: 313

Price: ₹850/-

Social Media & E-Commerce Promotion

AICTE UNIVERSITY SYLLABUS

INTRODUCTION TO MACHINE LEARNING

UNIT I: INTRODUCTION TO MACHINE LEARNING

Introduction: Definition, History and Application of Machine Learning, Types of Machine Learning: Supervised, Unsupervised, Semi-Supervised, and Reinforcement Learning. Labelled and Unlabelled Dataset. Supervised Learning Tasks: Regression vs. Classification, Learning Framework: Training, Validation and Testing of ML models. Performance Evaluation Parameters: Confusion matrix, Accuracy, Precision, Recall, F1 Score, and AUC.

UNIT II: SUPERVISED LEARNING AND UNSUPERVISED LEARNING

Regression: Linear and non-linear Regression, Logistic Regression. Classification: Naïve Bayes, K-Nearest Neighbours, Decision Trees. Linear model: Introduction to Artificial Neural Networks, Perceptron Learning Algorithm, Single Layer Perceptron, Introduction to Support Vector Machine for linearly separable data. Clustering: K-Means, Hierarchical Clustering, DBSCAN, Clustering Validation Measures. ML Applications: Ethical Considerations in Machine Learning, Case study and Real-world Applications.

UNIVERSITY OF MADRAS

ARTIFICIAL INTELLIGENCE

UNITS I: Introduction–Definition – Future of Artificial Intelligence – Characteristics of Intelligent Agents–Typical Intelligent Agents – Problem Solving Approach to Typical AI problems.

UNITS II: Problem solving Methods – Search Strategies- Uninformed – Informed – Heuristics – Local Search Algorithms and Optimization Problems -Searching with Partial Observations – Constraint Satisfaction Problems – Constraint Propagation – Backtracking Search – Game Playing – Optimal Decisions in Games – Alpha – Beta Pruning – Stochastic Games

UNITS III: Knowledge Representation First Order Predicate Logic – Prolog Programming – Unification – Forward Chaining-Backward Chaining – Resolution – Knowledge Representation –Categories and Objects – Events – Mental Events and Mental Objects – Reasoning Systems for Categories -Reasoning with Default Information

UNITS IV: Software Agents Architecture for Intelligent Agents – Agent communication – Negotiation and Bargaining – Argumentation among Agents – Trust and Reputation in Multi-agent systems.

UNITS V: Representing knowledge using rules: Procedural Vs Declarative knowledge – Logic programming – Forward Vs Backward reasoning – Matching – Control knowledge Brief explanation of Expert Systems- Definition- Characteristics-architecture- Knowledge Engineering- Expert System Life Cycle-Knowledge Acquisition Strategies- Expert System Tools.

UNIVERSITY OF MADRAS

MACHINE LEARNING

UNITS I: Introduction: Machine Learning Foundations – Overview – Design of a Learning System – Types of Machine Learning – Supervised Learning and Unsupervised Learning – Applications of Machine Learning – Tools Overview for ML.

UNITS II: Supervised Learning – I: Simple Linear Regression – Multiple Linear Regression – Polynomial Regression – Ridge Regression – Lasso Regression – Evaluating Regression Models – Model Selection – Bagging – Ensemble Methods.

UNITS III: Supervised Learning – II: Classification – Logistic Regression – Decision Tree Regression and Classification – Random Forest Regression and Classification – Support Vector Machine Regression and Classification - Evaluating Classification Models.

UNITS IV: Unsupervised Learning: Clustering – K-Means Clustering – Density-Based Clustering – Dimensionality Reduction – Collaborative Filtering.

UNITS V: Association Rule Learning: Association Rule Learning – Concepts related to ARL – ARL Algorithms - Apriori – Eclat – Concepts and Algorithms.

VELS UNIVERSITY

INTRODUCTION TO ML

UNIT I - INTRODUCTION TO MACHINE LEARNING

Introduction to Machine learning: Type of Learning and Examples, basic concepts in machine learning, Computational Learning theory, Introduction to Parametric Models – Non-Parametric Models–Probability Basics.

UNIT II - SUPERVISED LEARNING

Supervised Learning Algorithms Supervised Machine Learning Algorithms, working of supervised machine learning algorithm, Naive Bayes algorithm, decision tree, Support Vector Machines, KNN, Random Forest algorithm.

UNIT III - UNSUPERVISED LEARNING

Clustering- K-means -EM Algorithm- Mixtures of Gaussians - Dimensionality Reduction – Factor analysis – Feature Selection - Principal Component Analysis - Probabilistic PCA – Independent components analysis - Singular Value Decomposition.

UNIT IV - REINFORCEMENT LEARNING

Reinforcement Learning Algorithms Reinforcement Machine Learning Algorithms, working of reinforcement machine learning algorithm, Finite Markov Decision Processes, Dynamic Programming, Monte Carlo Methods

UNIT V - APPLICATIONS AND USE CASE

Ranking: Priority Inbox - Ordering Email Messages by Priority - Writing a Priority Inbox – Spam Filtering - Analyzing Social Graphs - Social Network Analysis - Hacking Twitter Social Graph Data - Analyzing Twitter Networks – Case Study.

BHARATHIAR UNIVERSITY

MACHINE LEARNING- BASICS

UNIT I INTRODUCTION TO LEARNING

Algorithmic models of learning, Learning classifiers, functions, relations, grammars, probabilistic models, value functions, behaviors and programs for experience. Bayesian, maximum some posterior, and minimum description length frameworks.

UNIT II ML-MODELS

Parameter Estimation, sufficient statistics, decision trees, neural networks, support vector machines, Bayesian networks, bag of words classifiers, N-gram models; Markov and Hidden Markov models, probabilistic relational models, association rules, nearest neighbour classifiers, locally weighted regression, ensemble classifiers.

UNIT III COMPUTATIONAL LEARNING

Computational Learning theory, mistake bound analysis, sample complexity analysis, VC dimension, Occam learning, accuracy and confidence boosting, Dimensionality reduction: Principal component Analysis, feature selection and visualization.

UNIT IV UNSUPERVISED LEARNING

Unsupervised Learning: Clustering, mixture models, k-means clustering, hierarchical clustering, distributional clustering, Reinforcement learning; Learning from heterogeneous, distributed, data and knowledge.

UNIT V APPLICATIONS IN DATA MINING

Selected applications in data mining, automated knowledge acquisition, pattern recognition, program synthesis, text and language processing, internet-based information systems, human computer interaction, semantic web, and Bio informatics and computational biology.

BHARATHIAR UNIVERSITY

MACHINE LEARNING TECHNIQUES

UNIT I INTRODUCTION TO MACHINE LEARNING

Introduction – Types of Machine Learning – Supervised Learning – The Brain and the Neuron – Design a Learning System – Perspectives and Issues in Machine Learning – Concept Learning Task – Concept Learning as Search – Finding a Maximally Specific Hypothesis –Version Spaces and the Candidate Elimination Algorithm–Linear Discriminants–Perceptron–Linear Separability – Linear Regression.

UNIT II MACHINE LEARNING MODELS

Linear Models–Multi-Layer Perceptron–Going Forwards–Going Backwards: Back Propagation Error – Multi-Layer Perceptron in Practice – Examples of using the MLP–Overview–Deriving Back-Propagation–Radial Basis Functions and Splines–Concepts–RBF Network–Curse of Dimensionality– Interpolations and Basis Functions–Support Vector Machines.

UNIT III TREE & PROBABILISTIC MODEL

Tree and Probabilistic Models–Learning with Trees–Decision Trees–Constructing Decision Trees – Classification and Regression Trees – Ensemble Learning – Boosting – Bagging – Different ways to Combine Classifiers – Probability and Learning – Data into Probabilities–Basic Statistics– Gaussian Mixture Models–Nearest Neighbour Methods–Unsupervised Learning–Kmeans Algorithms–Vector Quantization–Self Organizing Feature Map.

UNIT IV DIMENSIONALITY REDUCTION AND EVOLUTIONARY MODELS

Dimensionality Reduction and Evolutionary Models-Dimensionality Reduction–Linear Discriminant Analysis–Locally Linear Embedding–Iso map–Least Squares Optimization–Evolutionary Learning– Genetic Algorithms–Genetic Offspring–Genetic Operators– Using Genetic Algorithms–Reinforcements Learning–Overview–Getting Lost Example–Markov Decision Process.

UNIT V GRAPHICAL MODEL

Graphical Models – Markov Chain Monte Carlo Methods– Sampling – Proposal Distribution – Markov Chain Monte Carlo – Graphical Models – Bayesian Networks – Markov Random Fields–Hidden Markov Models–Tracking Methods.

MANONMANIAM SUNDARANAR UNIVERSITY

MACHINE LEARNING

Unit I: Introduction: Machine Learning – Examples of Machine Learning Applications. Supervised Learning: Learning a Class from Examples – Vapnik - Chervonenkis (VC) Dimension – Probably Approximately Correct (PAC) Learning – Noise – Learning Multiple Classes – Regression – Model Selection and Generalization – Dimensions of a Supervised Machine Learning Algorithm. Bayesian Decision Theory: Introduction – Classification – Losses and Risks – Discriminant Functions – Association Rules.

Unit II: Parametric Methods: Maximum Likelihood Estimation – Evaluating an Estimator: Bias and Variance – The Bayes' Estimator – Parametric Classification – Regression – Tuning Model Complexity: Bias/Variance Dilemma – Model Selection Procedures. Nonparametric Methods: Nonparametric Density Estimation – Generalization to Multivariate Data – Nonparametric Classification – Condensed Nearest Neighbor – Distance- Based Classification – Outlier Detection – Nonparametric Regression: Smoothing Models

Unit III: Linear Discrimination – Generalizing the Linear Model – Geometry of the Linear Discriminant – Pairwise Separation – Gradient Descent – Logistic Discrimination – Discrimination by Regression – Learning to Rank. Multilayer Perceptrons: The Perceptron – Training a Perceptron –Learning Boolean Functions – Multilayer Perceptrons – MLP as a Universal Approximator – Backpropagation Algorithm

Unit IV: Combining Multiple Learners: Generating Diverse Learners – Model Combination Schemes – Voting – Bagging – Boosting – Stacked Generalization – Fine-Tuning an Ensemble – Cascading Reinforcement Learning: Elements of Reinforcement Learning – Model-Based Learning – Temporal Difference Learning – Generalization – Partially Observable States

Unit V: Machine Learning with Python: Data Pre-processing, Analysis & Visualization - Training Data and Test Data – Techniques – Algorithms: List of Common Machine Learning Algorithms- Decision Tree Algorithm- Naïve Bayes Algorithm - K-Means- Random Forest- Dimensionality Reduction Algorithm- Boosting Algorithms – Applications: Social Media-Refinement of Search Engine Results- Product Recommendations-Detection of Online frauds.

VELLORE INSTITUTE OF TECHNOLOGY

MACHINE LEARNING

Module:1 Introduction PAC Learning-Consistent and inconsistent hypothesis, FIND-S, Candidate Elimination, deterministic and stochastic generalities, error, VC Dimensions, lower bounds-Convex optimization review- Probability review

Module:2 Dimensionality Reduction Feature representation in different domains: text, image, video and audio, Feature selection: Filter, wrapper and embedded models, Feature Reduction: PCA, t-SNE

Module:3 Model Selection and Validation Estimation and approximation errors: ERM-SRM- Validation- Regularization-based algorithms

Module:4 Classification Models Supervised Learning, Perceptron – Single layer & Multi-layer – Linear SVM – Hard, Soft Margins, kernel Methods, Lazy SVM for Instance Based Learning, Handling imbalanced data: One Class SVM

Module:5 Ensemble Learning Bagging-Committee Machines and Stacking-Boosting-Ranking based aggregation

Module:6 Clustering Unsupervised Learning, Partitional Clustering-K-Means-Linkage-Based Clustering Algorithms-Birch Algorithm-CURE Algorithm-Density-based Clustering- Spectral Clustering.

Module:7 Online Learning Online Classification in the Realizable Case- Online Classification in the Unrealizable Case-Online Convex Optimization- The Online Perceptron Algorithm- On-line to batch conversion – Federated Learning

UNIVERSITY OF MADRAS

MACHINE LEARNING

UNIT I: BASIC MATHEMATICS FOR MACHINE LEARNING: Regression Correlation and Regression, types of correlation – Pearson's, Spearman's correlations –Ordinary Least Squares, Fitting a regression line, logistic regression, Rank Correlation Partial and Multiple correlation- Multiple regression, multicollinearity. Gradient descent methods, Newton method, interior point methods, active set, proximity methods, accelerated gradient methods, coordinate descent, cutting planes, stochastic gradient descent. Discriminant analysis, Principal component analysis, Factor analysis, k means.

UNIT II: INTRODUCTION TO MACHINE LEARNING: Introduction, Examples of various Learning Paradigms, Perspectives and Issues, Version Spaces, Finite and Infinite Hypothesis Spaces, PAC Learning, VC Dimension.

UNIT III: SUPERVISED LEARNING ALGORITHMS Learning a Class from Examples, Linear, Non-linear, Multi-class and Multi-label classification, Decision Trees: ID3, Classification and Regression Trees (CART), Regression: Linear Regression, Multiple Linear Regression, Logistic Regression. Neural Networks: Introduction, Perceptron, Multilayer Perceptron, Support vector machines: Linear and Nonlinear, Kernel Functions, K-Nearest Neighbours

UNIT IV: ENSEMBLE LEARNING: Ensemble Learning Model Combination Schemes, Voting, Error-Correcting Output Codes, Bagging: Random Forest Trees, Boosting: Adaboost, Stacking: Unsupervised Learning: Introduction to Clustering, Hierarchical: Agnes, Diana, Partitional: K-means clustering, K-Mode Clustering, Self Organizing Map, Expectation Maximization, Gaussian Mixture Models, Principal Component Analysis (PCA), Locally Linear Embedding (LLE), Factor Analysis

UNIT V: PROBABILISTIC LEARNING: Bayesian Learning, Bayes Optimal Classifier, Naïve Bayes Classifier, Bayesian Belief Networks, Mining Frequent Patterns: Machine Learning In Practice: Design, Analysis and Evaluation of Machine Learning Experiments, Other Issues: Handling imbalanced data sets

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to the management of Vels Institute of Science, Technology, and Advanced Studies (VISTAS) for their best support. I express my sincere thanks to my family for their endless support, encouragement, and constant belief in me throughout this challenging journey. I extend my sincere deepest heartfelt gratitude to my Research Supervisor **Prof.Dr.K.Dharmarajan,** for his unwavering guidance, blessings, and for providing me with the opportunity to embark on this challenging and innovative work. I also express my heartfelt gratitude to my dear best friend and constant supporter, **Dr.K.Abirami,** for her timely constant support and creative ideas. This book is dedicated to them, whose influence has profoundly shaped my life. I extend my sincere thanks to the Leilani Katie Publication and PRESS for her kindly co-operation and for publishing this book. Lastly, I express my sincere gratitude to the readers who will engage with this work. Your interest and involvement make this effort meaningful, and I hope the ideas and knowledge within these pages resonate with you. Any suggestions for enhancing the book further will be greatly appreciated.

With sincere regards
Dr.K.Dharmarajan
Dr.K.Abirami
T.Haripriya

CONTENT

CHAPTER II **MACHINE LEARNING FUNDAMENTALS AND DATA ENGINEERING** 2.1 Introduction to Machine Learning – Definition and Motivation 2.2 Types of Machine Learning – Supervised, Unsupervised, Semi-supervised and Reinforcement Learning 2.3 Machine Learning Life Cycle 2.4 Data Collection Techniques and Sources 2.5 Data Quality Issues and Data Cleaning Methods 2.6 Handling Missing Values and Noisy Data 2.7 Outlier Detection and Treatment 2.8 Data Normalization and Standardization 2.9 Feature Scaling Techniques 2.10 Encoding Techniques for Categorical Data 2.11 Feature Engineering and Feature Selection Methods 2.12 Dimensionality Reduction – Concepts and Techniques 2.13 Data Visualization Techniques 2.14 Exploratory Data Analysis (EDA) 2.15 Statistical Data Summarization 2.16 Data Balancing Techniques 2.17 Dataset Splitting – Training, Validation and Testing 2.18 Cross Validation Techniques 2.19 Machine Learning System Design Pipeline	**50**

CHAPTER III

SUPERVISED AND UNSUPERVISED LEARNING MODELS

CHAPTER IV

DEEP LEARNING, NATURAL LANGUAGE PROCESSING AND COMPUTER VISION

CHAPTER I

INTRODUCTION TO ARTIFICIAL INTELLIGENCE AND INTELLIGENT SYSTEMS

1.1 Meaning and Definition of Artificial Intelligence

Artificial Intelligence refers to the capability of a machine, computer, or system to perform tasks that typically require human intelligence. These tasks include learning from experience, reasoning, problem-solving, understanding natural language, recognizing patterns, making decisions, and adapting to new situations. Essentially, AI enables machines to mimic cognitive functions similar to those of the human mind. In simple terms, AI is about creating machines that can think, learn, and act intelligently, rather than just following pre-programmed instructions.

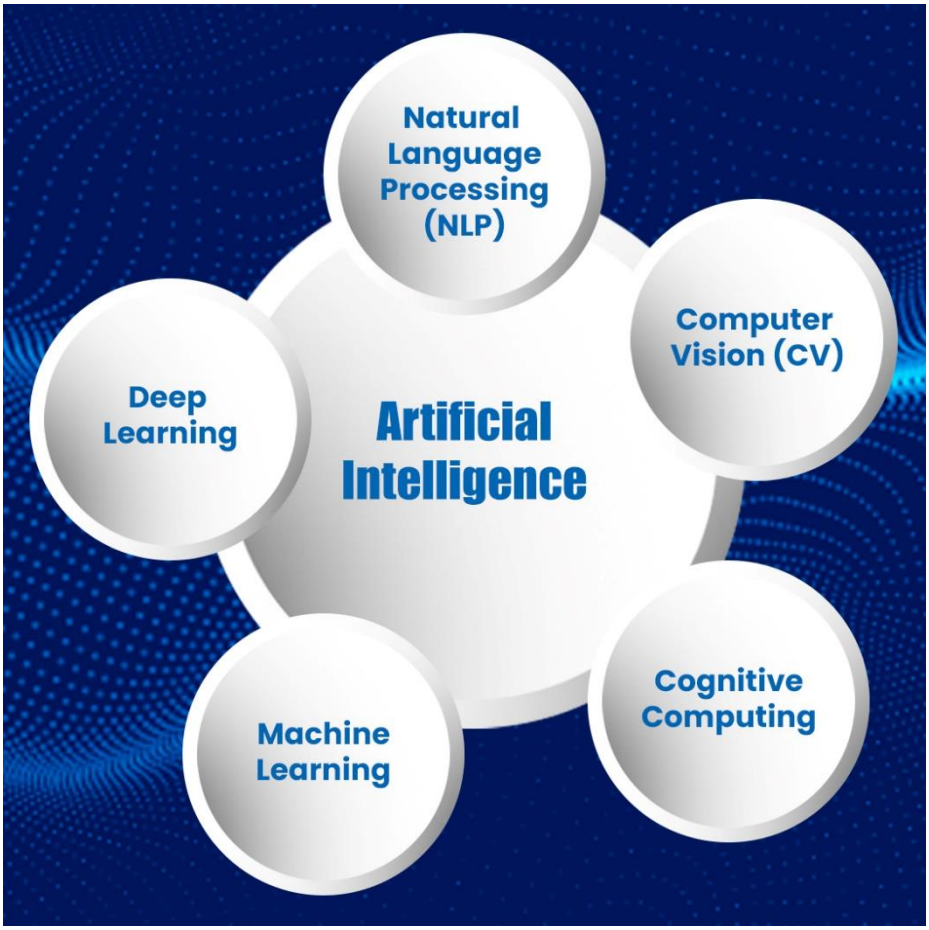

Figure 1.1: Artificial Intelligence

Definitions of Artificial Intelligence

Several Scholars and Organizations have defined AI in Various Ways:

- **John McCarthy (1955):** Known as the "Father of AI," he defined Artificial Intelligence as: "The science and engineering of making intelligent machines, especially intelligent computer programs." This emphasizes AI as both a scientific discipline and an engineering practice aimed at developing machines capable of intelligent behavior.
- **Alan Turing (1950):** Proposed the concept of machine intelligence and the famous Turing Test: "A machine can be said to think if it can successfully imitate human responses in a way indistinguishable from a human." This definition focuses on machines exhibiting human-like intelligence.
- **Russell and Norvig (2010):** In their book Artificial Intelligence: A Modern Approach, they define AI as: "The study of agents that receive percepts from the environment and perform actions." Here, AI is seen as creating autonomous agents that perceive, reason, and act.

Encyclopedia Britannica – AI is defined as

"Artificial intelligence is the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings."

1.2 Historical Evolution and Generations of AI

Here's a detailed explanation of the Historical Evolution and Generations of Artificial Intelligence in an academic-style format suitable for study or assignment purposes

Historical Evolution of Artificial Intelligence

Artificial Intelligence (AI) has evolved over decades, driven by advances in mathematics, computer science, and neuroscience. The history of AI can be divided into key stages

1. Early Concepts (Pre-1950s)

- The idea of intelligent machines dates back to ancient myths and early mechanical devices.
- Mathematical logic and automata theory laid the groundwork for AI.
- Alan Turing (1936–1950) introduced the concept of a machine that could simulate any human intelligence through the *Turing Machine* and later proposed the Turing Test for evaluating machine intelligence.