Graduate Certificate in AI for Renewable Energy Forecasting

Introduction to Artificial Intelligence

Artificial Intelligence (AI): The simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving.

Reinforcement Learning: A type of machine learning where an agent learns to behave in an environment, by performing certain actions and observing the results/rewards. The agent aims to maximize the total reward over time.

Supervised Learning: A type of machine learning where the model is trained on a labeled dataset, i.e., a dataset that has both the input features and the correct output labels. The model learns to predict the output labels for new, unseen data.

Unsupervised Learning: A type of machine learning where the model is trained on an unlabeled dataset, i.e., a dataset that only has the input features but no output labels. The model learns to identify patterns and structures in the data.

Deep Learning: A subset of machine learning that is based on artificial neural networks with representation learning. It can process a wide range of data resources, requires less data preprocessing by humans, and can often produce more accurate results than traditional machine learning approaches.

Neural Network: A computing model whose layered structure resembles the neural network of the human brain. It is composed of input and output layers, as well as (in most cases) a hidden layer consisting of units that transform the input into something the output layer can use.

Activation Function: A function used in artificial neural networks to transform the summed weighted input from the node into an activation signal which is then sent to the next layer. Common activation functions include the sigmoid, tanh, and ReLU functions.

Backpropagation: A method used in artificial neural networks to calculate the gradient that is needed in the calculation of the weights to be used in the network. It is commonly used to train deep learning models.

Gradient Descent: An optimization algorithm used to minimize some function by iteratively moving in the direction of steepest descent as defined by the negative of the gradient. In machine learning, we use gradient descent to update the parameters of our model.

Overfitting: A situation where a statistical model describes the random error in the training data to such a degree that it negatively impacts the performance of the model on new data. This results in a model that has poor predictive accuracy.

Underfitting: A situation where a statistical model cannot capture the underlying trend of the data. This

results in a model that has poor predictive accuracy on both the training data and new data.

Regularization: A technique used to prevent overfitting in machine learning models by adding a penalty term to the loss function. This encourages the model to have smaller weights, making it less likely to fit the noise in the training data.

Cross-Validation: A technique used to evaluate the performance of a machine learning model. The data is divided into k subsets, or folds. The model is trained on k-1 of the folds, and the remaining fold is used for testing. This process is repeated k times, with a different fold used for testing each time. The average performance across the k runs is then used as the overall performance of the model.

Feature Engineering: The process of creating new features or modifying existing features in a dataset to improve the performance of a machine learning model. This can include techniques such as one-hot encoding, binning, and polynomial features.

Natural Language Processing (NLP): A field of AI that focuses on the interaction between computers and humans through natural language. The ultimate objective of NLP is to read, decipher, understand, and make sense of the human language in a valuable way.

Named Entity Recognition (NER): A subtask of NLP that seeks to locate and classify named entities in text into predefined categories such as person names, organizations, locations, medical codes, time expressions, quantities, monetary values, percentages, etc.

Part-of-Speech Tagging (POS): A subtask of NLP that involves identifying the part of speech for each word in a given text, i.e., whether a word is a noun, pronoun, adjective, verb, adverbs, etc.

Sentiment Analysis: A subtask of NLP that involves determining the emotional tone behind words to understand the attitudes, opinions, and emotions of a speaker or writer.

Support Vector Machines (SVMs): A type of supervised machine learning algorithm that can be used for both classification or regression challenges. However, SVMs are mostly used in classification problems.

Decision Trees: A type of supervised machine learning algorithm that is mostly used in classification problems. It works for both categorical and continuous input and output variables.

Random Forest: An ensemble learning method that operates by constructing multiple decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

Naive Bayes: A family of probabilistic algorithms based on applying Bayes' theorem with strong (naive) independence assumptions between the features.

K-Nearest Neighbors (KNN): A type of instance-based learning algorithm that is used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space.

Principal Component Analysis (PCA): A technique used to emphasize variation and bring out strong patterns

in a dataset. It is often used to make data easy to explore and visualize.

Linear Regression: A statistical model that is used to analyze the relationship between two continuous quantitative variables.

Logistic Regression: A statistical model that is used to analyze the relationship between one dependent binary variable and one or more nominal, ordinal, interval, or ratio-level independent variables.

Time Series Analysis: A statistical technique that deals with time series data, or trend analysis. It is used to identify patterns in data collected over time, which can then be used to make predictions about future data points.

Autoregressive Integrated Moving Average (ARIMA): A forecasting technique that projects the future values of a series based entirely on its own inertia.

Long Short-Term Memory (LSTM): A type of recurrent neural network capable of learning long-term dependencies, which makes it perfect for tasks such as understanding language and speech recognition.

Convolutional Neural Network (CNN): A type of deep learning algorithm that is commonly applied to image processing, natural language processing, and speech recognition.

Generative Adversarial Network (GAN): A class of machine learning systems invented by Ian Goodfellow and his colleagues in 2014. Instead of following the standard approach of supervised learning, these systems use two neural networks contesting with each other in a zero-sum game framework.

Recommender Systems: A subclass of information filtering systems that seek to predict the 'rating' or 'preference' a user would give to an item.

Markov Decision Processes (MDPs): A mathematical framework for decision making in situations where outcomes are partly random and partly under the control of a decision-maker.

Q-Learning: A value-based algorithm in reinforcement learning that seeks to find the optimal policy by iteratively determining the expected utility of each action at each state.

Deep Q-Network (DQN): A deep learning approach to the Q-learning algorithm, which is used to solve Markov decision processes.

Monte Carlo Tree Search (MCTS): A heuristic search algorithm used for decision making in perfect and imperfect information games.

Genetic Algorithm: A search heuristic that is inspired by Charles Darwin's theory of natural evolution. This algorithm reflects the process of natural selection where the fittest individuals are selected for reproduction in order to produce the offspring of the next generation.

Simulated Annealing: A probabilistic technique for approximating the global optimum of a given function. It uses a random search instead of a gradient descent and can avoid getting stuck in local minima.

Swarm Intelligence: A form of artificial intelligence based on the collective behavior of decentralized, self-organized systems.