**MACHINE LEARNING (MTCS 031)**

**UNIT –V**

**SYLLABUS:**

***GENETIC ALGORITHMS*** – an illustrative example, Hypothesis space search, Genetic Programming, Models of Evolution and Learning; Learning first order rules-sequential covering algorithms-General to specific beam search-FOIL;

***REINFORCEMENT LEARNING*** - The Learning Task, Q Learning.

**PYQ:**

**7 MARKS**

* Discuss any four examples of machine learning applications.
* How is Candidate Elimination algorithm different from Find-S Algorithm? Explain
* Describe Reinforcement Learning? Classify the various reinforcement learning algorithms.
* Explain the operations involved in Hypothesis Space searching through Genetic Algorithm.
* What are the Genetic operators? What are the role of Genetic operators in Genetic algorithm?
* Name and Describe the main features of Genetic Algorithm with proper diagram.
* Explain the Q function and Q Learning Algorithm.

**2 MARKS**

* List the issues in Decision Tree Learning.

**Examples of Machine Learning Applications:**

Machine learning has become ubiquitous in our daily lives. Here are four significant examples:

**1. Image Recognition**

* **Application:** Identifying objects, faces, scenes, and activities in images and videos.
* **Examples:** Image search, self-driving cars, medical image analysis (e.g., detecting tumors), facial recognition for security systems.
* **Underlying Techniques:** Convolutional Neural Networks (CNNs) are commonly used for image recognition tasks.

**2. Natural Language Processing (NLP)**

* **Application:** Understanding, interpreting, and generating human language.
* **Examples:** Machine translation, sentiment analysis, chatbots, text summarization, speech recognition.
* **Underlying Techniques:** Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), and Transformer models are often employed.

**3. Recommendation Systems**

* **Application:** Suggesting products, movies, music, or other items based on user preferences and behavior.
* **Examples:** Amazon product recommendations, Netflix movie suggestions, Spotify music playlists.
* **Underlying Techniques:** Collaborative filtering, content-based filtering, and hybrid approaches are used.

**4. Fraud Detection**

* **Application:** Identifying fraudulent transactions or activities.
* **Examples:** Credit card fraud detection, insurance claim fraud detection, detecting fake accounts on social media platforms.
* **Underlying Techniques:** Anomaly detection, classification, and clustering algorithms are commonly used.

**Candidate Elimination vs. Find-S Algorithm**

**Find-S** and **Candidate Elimination** are two fundamental algorithms in machine learning used for learning a concept from a set of training examples. While both aim to find a hypothesis that fits the training data, they differ significantly in their approach.

**Find-S Algorithm**

* **Goal:** To find the most specific hypothesis that is consistent with all positive training examples.
* **Process:** Starts with a maximally specific hypothesis and generalizes it as new positive examples are encountered.
* **Limitations:** Highly sensitive to noise in the data. If there's a noisy positive example, the final hypothesis might be overly specific.

**Candidate Elimination Algorithm**

* **Goal:** To find a set of hypotheses consistent with all training examples.
* **Process:** Maintains two sets: a set of most general hypotheses (G) and a set of most specific hypotheses (S). These sets are updated as new examples are processed.
* **Advantage:** More robust to noise than Find-S as it considers a set of hypotheses instead of a single one.
* **Complexity:** More complex than Find-S due to the management of two sets.

**Key Differences:**

|  |  |  |
| --- | --- | --- |
| **Feature** | **Find-S** | **Candidate Elimination** |
| Hypothesis Space | Single most specific hypothesis | Set of consistent hypotheses |
| Noise Handling | Sensitive to noise | More robust to noise |
| Complexity | Simpler | More complex |
| Output | Single hypothesis | Set of hypotheses |

**Genetic Algorithms**

* **Genetic Algorithms (GAs)** are optimization and search techniques inspired by the principles of natural selection and genetics.
* They are used to find approximate solutions to complex problems.
* These algorithms mimic the principle of natural genetics and natural selection to construct search and optimization procedure.
* Following are the benefits of Genetic algorithm :
  + **Robustness**: GAs are robust and can handle complex, nonlinear optimization problems.
  + **Optimization over Large State Spaces**: They are effective in searching large and complex solution spaces.
  + They do not break on slight change in input or presence of noise.

**Operations involved in Hypothesis Space searching through Genetic Algorithm.**

**Or**

**Different phases of genetic algorithm**

**1. Initial population** :

* The process begins with a set of individuals which is called a population.
* Each individual is a solution to the problem we want to solve.
* An individual is characterized by a set of parameters (variables) known as genes.
* Genes are joined into a string to form a chromosome (solution).
* In a genetic algorithm, the set of genes of an individual is represented using a string.

**2. FA (Factor Analysis) fitness function :**

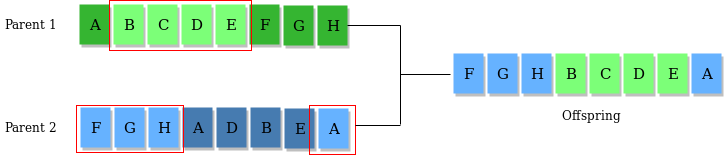
* The fitness function determines how fit an individual is (the ability of all individual to compete with other individual).
* It gives a fitness score to each individual.
* The probability that an individual will be selected for reproduction is based on its fitness score.

**3. Selection :**

* The idea of selection phase is to select the fittest individuals and let them pass their genes to the next generation.
* Two pairs of individuals (parents) are selected based on their fitness scores.
* Individuals with high fitness have more chance to be selected for reproduction.

**4. Crossover :**

* Crossover is the most significant phase in a genetic algorithm.
* For each pair of parents to be mated, a crossover point is chosen at random from within the genes.
* Offspring are created by exchanging the genes of parents among themselves until the crossover point is reached.

****

**5. Mutation :**

* When new offspring formed, some of their genes can be subjected to a mutation with a low random probability.
* This implies that some of the bits in the bit string can be flipped.
* Mutation occurs to maintain diversity within the population and prevent premature convergence.

****

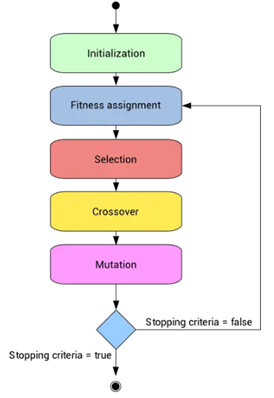
**6. Termination :**

* The algorithm terminates if the population has converged (does not produce offspring which are significantly different from the previous generation).
* Then it is said that the genetic algorithm has provided a set of solutions to our problem.

**Summary of Genetic Algorithm Operations**

1. **Initialization**: Create a diverse initial population.
2. **Evaluation**: Assess fitness of solutions.
3. **Selection**: Choose the best solutions to act as parents.
4. **Crossover**: Combine parent solutions to create offspring.
5. **Mutation**: Introduce variability by randomly altering offspring.
6. **Replacement**: Update the population with new offspring.
7. **Termination**: End the search based on criteria.

These operations collectively help in searching through the hypothesis space to find optimal or near-optimal solutions by simulating evolutionary processes.



**Role of Genetic Operators in Genetic Algorithms**

 **Selection:**

* Determines which individuals in the current population will be selected as parents for the next generation.
* Based on fitness scores, individuals with higher fitness have a higher probability of being selected.
* Common methods include roulette wheel selection, tournament selection, and rank selection.

 **Crossover:**

* Combines genetic material from two parent individuals to create offspring.
* Introduces new combinations of genes, increasing the diversity of the population.
* Common types include one-point crossover, two-point crossover, and uniform crossover.

 **Mutation:**

* Randomly alters the genetic material of an individual to introduce new genetic variations.
* Helps prevent premature convergence and maintains diversity in the population.
* Common types include bit-flip mutation, swap mutation, and inversion mutation