

Objectives:

- Explain why algorithms are widely used in Biotechnology.
- Enumerate and give a basic understanding of some of the popular algorithms that are used in Biotechnology.
- Show how algorithms are crucial a crucial aspect of Biotechnology and how it positively impacts the field.

Scope:

This paper will focus on the algorithms that are used in biotechnology as well as discuss the benefits that these algorithms bring in the said field.

Presentation of the chosen technology:

a) . Uses and function –

1. **Sequence alignment algorithms** - Sequence alignment algorithm are computational methods used to identify regions of similarity between two or more biological sequences. Vladimir Levenstein performed one of the earliest attempts to align two sequences in 1965, dubbed "edit distance" and now commonly referred to as Levenshtein Distance. In biotechnology, these algorithms are commonly employed for tasks such as detecting homologous genes and predicting protein structure and function.

Example: The BLAST (Basic Local Alignment Search Tool) algorithm is a widely used sequence alignment algorithm that compares a query sequence to a database of known sequences to identify similar sequences. BLAST can be used to identify the function of a protein or to find homologous genes in different organisms.

2. **Hidden Markov Model** - A Hidden Markov Model (HMM) is a statistical model for representing probabilistic sequences. In biotechnology, HMMs are commonly employed for tasks such as gene prediction and protein sequence analysis. They offer a conceptual framework for creating complicated models simply by drawing an intuitive picture of biological items of interest.

Example: HMMs can be used to predict the secondary structure of a protein. The HMM model is trained on a set of known protein structures to learn the probability distribution of each amino acid and its position in the

protein structure. The trained model can then be used to predict the secondary structure of a new protein sequence.

3. **Clustering Algorithm** - Clustering algorithms are a sort of unsupervised machine learning method that is frequently used to evaluate gene expression data in order to discover clusters of genes that are co-expressed. Clustering algorithms operate by grouping together comparable data points based on a measure of similarity or distance. Clustering algorithms are frequently used in biotechnology to evaluate gene expression data in order to discover groupings of genes that are co-expressed. This can aid researchers in understanding the underlying biological processes and identifying possible drug development targets.

Example: k-means clustering is a widely used clustering algorithm that groups data into k clusters based on their similarity. In gene expression analysis, k-means clustering can be used to group genes with similar expression patterns into clusters. This can help to identify genes that are co-regulated or involved in the same biological pathway.

4. **Machine Learning Algorithms** - Machine learning algorithms are a subset of artificial intelligence that can use data to create predictions or judgments. To classify biological data, such as gene expression data or medical imaging, machine learning algorithms such as support vector machines, random forests, and neural networks are utilized. These algorithms can be trained to recognize patterns or relationships in data and generate predictions depending on fresh information. Machine learning algorithms are employed in biotechnology for a variety of tasks, including gene expression analysis, drug development, and personalized medicine.

Support vector machines (SVMs) can be used to classify cancer patients based on gene expression data. The SVM model is trained on a set of known cancer patients and healthy individuals to learn the patterns of gene expression that are associated with cancer. The trained model can then be used to classify new patients as either cancer or healthy based on their gene expression data.

5. **Genetic Algorithms** - Genetic algorithms are a sort of optimization technique that is inspired by natural selection. Genetic algorithms are employed in biotechnology for a variety of tasks, including protein structure prediction, gene expression analysis, and drug development. Genetic algorithms mimic the process of natural selection. They begin with

a population of potential solutions to a problem and then employ selection, crossover, and mutation operators iteratively to evolve the population towards superior solutions. Genetic algorithms are employed in biotechnology for a variety of tasks, including protein structure prediction, gene expression analysis, and drug development. A genetic algorithm, for example, may be used to find the best medicine combination to treat a specific ailment.

Example: Genetic algorithms can be used to optimize metabolic pathways in microbes for the production of biofuels or pharmaceuticals. The genetic algorithm starts with a population of potential pathway designs and evolves them over multiple generations by selecting the best designs and breeding them to create new designs. The goal is to find the optimal pathway design that maximizes the production of the desired product.

6. **Beehive algorithm** - The Beehive method, sometimes known as the Bees Algorithm, is a population-based search method that replicates honey bee colonies' food gathering behavior. Pham, Ghanbarzadeh, and colleagues created it in 2005. In its most basic form, the algorithm combines neighborhood search with global search and can be used for both combinatorial and continuous optimization. The algorithm begins with a population of bees spread randomly in the search space. Each bee next searches its immediate surroundings for a food supply (i.e., a solution to the optimization problem). The best food sources are then chosen, and more bees are sent to look for them. This method is repeated until either a good solution or a stopping criterion is reached.

Example: The beehive algorithm can be used to optimize the design of protein structures. The algorithm starts with a population of potential protein structures and evolves them over multiple generations by selecting the best structures and exchanging information between bees. The goal is to find the optimal protein structure that maximizes its stability or activity.

b.) Importance and Benefits

The use of algorithms in biotechnology is significant because it can aid in the advancement of research and development in the field. It can aid in the creation of new biotechnology products and processes, as well as the identification of potential therapeutic targets. Algorithms can also be used to investigate and model complicated biological systems, resulting in a greater knowledge of their behavior and application.

Algorithms are significant in biotechnology because they can help scientists discover new drug targets, simulate complicated biological processes, and increase the efficiency and effectiveness of research and development.

References:

- <https://open.oregonstate.education/appliedbioinformatics/chapter/chapter-3/>
- <https://www.nature.com/articles/nbt1205-1499/>
- <https://www.nature.com/articles/nbt1004-1315>
- <https://www.nature.com/articles/s41587-021-01187-w>
- <https://www.nature.com/articles/s41587-021-01187-w>
- https://en.wikipedia.org/wiki/Bees_algorithm
- <https://onlinelibrary.wiley.com/doi/abs/10.1002/int.20376>