TOPIC: ROLE AND IMPORTANCE OF AI ON BIOLOGY

<u>ASSIGNMENT BY:</u>

NAME: SINGAM SETTY S A DESHIK

ROLL NO: 2101AI32

BRANCH: Artificial Intelligence and Data Science

GROUP: G8

Introduction:

Artificial Intelligence (AI) is the favoured way for data mining and large data processing in today's society. It uses effective learning and adaptability to provide answers to a number of engineering applications. Artificial Neural Network modelling, Reasoning-based decision algorithms, Simulation models, DNA computing, and Quantum computing are just few of the approaches that are available. AI's application in biological research has considerably reduced the fuzziness and randomness in managing such data.

This work provides an overview of machine learning and AI computer models, sophisticated data analytics, and optimization methodologies utilised in Bioengineering applications such as drug design and analysis, medical imaging, biologically inspired learning and adaptation for analytics, and so on.

The number of biological data has risen dramatically over time, as seen by ProteomicsDB's coverage of 92 percent of all known human genes annotated in the Swiss database. Millions of patient records have already been collected and kept in databases around the world. Medical imaging creates large volumes of data with much more complicated properties, and analysing this data will not only improve health-care services but also lead to big advances in research.

The level of analysis for biological data has evolved to a molecular level as computing and storage capacities have improved. In medicine, information technology techniques are used to examine patient records and reports, and intricate relationships and correlations are discovered through the analysis of large amounts of data. Oncology, liver pathology, thyroid disease diagnostics, rheumatology, dermatology, cardiology, neuropsychology, gynaecology, and perinatology are all examples of practical application. The 'Internet of things' (IoT) is a popular new system in which numerous objects are interconnected and share relevant sensory data and commands, allowing gadgets to understand and respond to the external environment.

In genomics: Artificial intelligence has made great progress in the world of computer science during the previous 10 years, and it is still one of the fastest developing fields. By 2014, there were several published research papers in which machine learning was used to analyse genome data. The advantage of computers is that you can give them a lot of training and teach them what is wrong and right while they learn from their mistakes and eventually begin to recognise patterns in data. The growing influence of AI in medicine and its global deployment is assisting in making machine decision making more accurate, personalised, and faster, but the current healthcare system is unable to keep up with the rapid improvements. The implementation of Electronic Health Records has advanced the clinical setup a bit and is being looked into as the maiden step towards revolutionizing modern day healthcare.

AI in proteomics: Protein-based studies, also known as 'proteomics,' have grown in popularity, and with more data available, machine learning has found increased applications in prediction, feature selection, pattern recognition, and a variety of automated tasks. The first step in understanding protein sequences is to generate a profile for the unknown protein based on its sequence using homology modelling techniques. The query sequence is aligned multiple times with existing databases that contain non-redundant records of protein structure and evolutionary information.

AI in proteome informatics: One of the first applications of machine learning was protein structure prediction and dynamic analysis of the predicted structures. Later, these became known as AI and have now evolved into an even broader interdisciplinary technology known as Deep Learning. Because determining protein structures is critical for understanding biological processes and cell functioning. Protein structure and fold prediction have had a significant impact on our understanding of how proteins work. Many new protein sequences have been stored in various databases around the world in recent years. Understanding how proteins acquire their three-dimensional structure remains one of biology's most perplexing mysteries. To summarise, machine learning-based AI is being used successfully in protein fold prediction and structure prediction applications. Its application is expected to spread in the coming days to areas such as disease-based genome modification prediction, protein-protein binding site prediction, and protein-protein network prediction. Phylogenetic AI: Computational Biology, also known as Bioinformatics, has grown by leaps and bounds in recent years. Phylogeny is explained using trees, where the roots represent the origin of evolution, the leaves represent the species/organism or a genomic sequence, the branches represent the relationship between the leaves, and the branch length represents the evolutionary time required for a particular evolution.

Next-generation sequencing and artificial intelligence:Biological databases are massive collections of biological data that have been collected, curated, and stored according to a predefined schema. These databases contain information from a wide range of fields, including proteomics, metabolomics, genomics, microarray data analysis, and the most recent trend, Next-Generation Sequence (NGS) data. NGS data now accounts for whole genome analysis and has undoubtedly helped with many genomics and proteomics issues. These databases can be divided into two types: structure databases and sequence databases. Sequence databases, also known as primary databases, store nucleic acid and protein sequences, while structure databases, also known as secondary databases, store protein structures. GenBank is a rapidly expanding database of known genetic sequences. GenBank files contain information such as accession numbers and gene names, as well as phylogenetic classification and references, in addition to sequence data. Over the years, many software tools have been developed to retrieve, analyse, and visualise data from such databases. These tools perform a variety of useful operations such as homology modelling, similarity and functional analysis, and so on. One such enthralling tool that is currently trending is "MethBank," which involves whole genome sequences and provides configurable and interactive data analysis.

Integration of Biological Databases with AI: We know that we will only be able to achieve optimal results if we can integrate data from a variety of different sources and then design an automated learning algorithm to analyse and predict based on previous learning experiences. Biological databases, which used to store compound sequences and structures, have now evolved to store more complex and bulk data. The majority of Microarray Profiling studies are based on a subset of the total expression dataset. We recognise that reaching full potential requires the integration and unification of all available data. ‘ONCOMINE’ is one such tool for rapid interpretation of gene’s potential role in a particular disease. Expression sets from multiple sources can be retrieved and analysed along while integrating it with multiple other resources such as gene ontology annotations, target gene data, etc.

SOURCE:

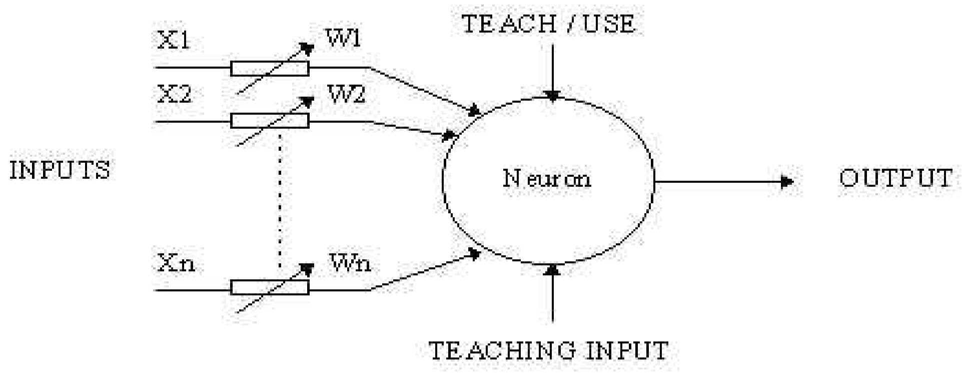
1. https://www.researchgate.net/publication/319623622_Artificial_Intelligence_in_Biological_Data

With the maturation of AI, artificial intelligence is expected to dominate biological data science in the near future. The majority of cutting-edge AI techniques are derived from computer vision, image recognition, or natural language processing. Although some efforts are being made, it is not easy to migrate existing AI techniques to the field of biological data science. The unique characteristics of massive data generated in biological data science necessitate the development of their own AI theory, methods, and systems. The maturity of AI in biological data science will, to some extent, indicate the realisation of precision medicine.

Now let us discuss regarding the use of AI in **Artificial Neural Networks (ANN):**

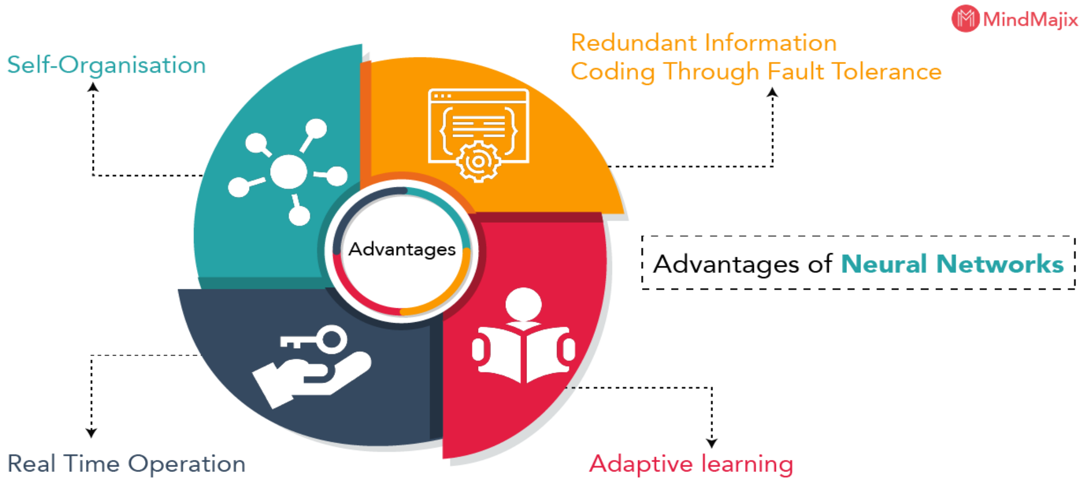
Let us go through one simple idea of ANN. One major use of AI in biology is in Artificial Neural Networks or ANNs. ANNs are computing systems similar to animal (and human) brain or (biological neural networks). These systems learn from examples without being given any task-specific rules or code. Image recognition is a prime example as ANN learns to identify images that contain dogs by examining sample images. These sample images are manually labelled as "dog" or "no dog" and use these results to identify dogs in other images.

Now let us discuss it briefly about ANN. ANN are the key tool of machine learning. They are systems developed from the idea of how neuron functions in the brain, the exact way we humans learn. Neural networks (NN) constitute both the input & output layers, as well as a hidden layer containing units that change input into the output so that the output layer can utilize the value. These are used for identifying patterns for data scientists and programmers to recognise and develop the machine to recognize patterns that are many and complex. The Function of Neural Networks in AI is that these technologies are used in the majority of business applications and commercial enterprises. Their primary goal is to solve complex problems such as pattern recognition or facial recognition, but they also have several other applications such as speech-to-text transcription, data analysis, handwriting recognition for check processing, weather prediction, and signal processing. The structure of ANN is described. ANN functions very similarly to the human brain. We can replicate the functioning of the brain using silicon and wires that act similarly to dendrites and neurons by making the necessary connections. Because dendrites accept external stimuli in the same way, the input generates electric impulses that travel through the neural network. ANN is made up of nodes that act like neurons. Links (wires) connect the nodes so that they can communicate with one another. Nodes take input data and perform small operations on trained data, with the results being passed to other nodes (neurons). The node's output is referred to as its node value. The image below depicts the basic structure of a neuron.



ANN is made up of nodes that act like neurons. Links (wires) connect the nodes so that they can communicate with one another. Nodes take input data and perform small operations on trained data, with the results being passed to other nodes (neurons). The node's output is referred to as its node value. The image below depicts the basic structure of a neuron.

ADVANTAGES OF NEURAL NETWORKS



Self-organization: During learning, an ANN can generate its own representation of the information that it receives.

Real-Time Operation: ANN calculations can be performed concurrently, and some special (hardware) devices are available to take advantage of this capability.

Adaptive learning: The ability to learn how to solve tasks based on the training set data

Redundant Information Coding Through Fault Tolerance: A network's semi-destruction reduces its corresponding performance. Furthermore, some networks will be able to retain data even if major network damage occurs.

SOURCE:

1. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-019-3225-3#Ack1

2. https://mindmajix.com/neural-network-in-artificial-intelligence

We are now going to discuss regarding the use of AI in some **biotechnology fields**

1. AI in **AGRICULTURAL BIOTECHNOLOGY:**
   Agricultural biotechnology creates genetically modified plants to boost crop yields or add new traits to existing plants. It includes traditional plant breeding, tissue culture, micropropagation, molecular breeding, and plant genetic engineering. Biotechnology companies are now using Artificial Intelligence and Machine Learning techniques to create and programme autonomous robots that can perform important agricultural tasks such as crop harvesting at a much faster rate than humans. Data captured by drones is processed and analysed using computer vision and deep learning algorithms. Machine Learning algorithms aid in the tracking and prediction of environmental changes such as weather changes that affect crop yield.
2. AI in **MEDICAL TECHNOLOGY**:
   Medical biotechnology uses living cells for the betterment of human health by producing drugs and antibiotics. In drug discovery, artificial intelligence and machine learning are widely used. Machine Learning is widely used in disease diagnosis because it uses true results to improve diagnostic tests, i.e., the more diagnostic tests run, the more accurate the results. AI is also assisting in the reduction of the radiation therapy planning process, which saves time and improves patient care. Another promising application of Artificial Intelligence and Machine Learning is the enhancement of EHRs with evidence-based medicines and clinical decision support systems. Apart from the applications mentioned above, these technologies are widely used in gene editing, radiology, and personalised medicine.
3. AI in **ANIMAL BIOTECHNOLOGY:**
   This branch employs molecular biology techniques to genetically engineer/modify animals for pharmaceutical, industrial, or agricultural purposes. Animal breeding is one area where Artificial Intelligence models can be useful. Selective breeding is a common practise in which animals with the most desirable characteristics are bred together so that their offspring have the same traits. This practise is also carried out at the molecular level, where genetic characteristics of animals are chosen and such animals are bred.
4. AI in **INDUSTRIAL BIOTECHNOLOGY:**
   This branch employs molecular biology techniques to genetically engineer/modify animals for pharmaceutical, industrial, or agricultural purposes. Selective breeding is a common practise in which animals with the most desirable characteristics are bred together so that their offspring have the same traits. This practise is also carried out at the molecular level, where genetic characteristics of animals are chosen and such animals are bred.

   Conclusion:

   Throughout history, humans have continued to adapt to new techniques and develop better technology. As we reviewed AI, we discovered that it had been evolving since its introduction in 1943, when McCullouch and Pitts introduced the world to the concept of artificial neurons. Since then, it has grown rapidly, sometimes unexpectedly, as new cutting-edge technology has come into play. Machine learning-based methods are more accurate, robust, and reliable than traditional methods. Conversely, because AI has flaws, we must constantly seek ways to improve its design and application. AI will continue to find extensive applications in a wide range of previously unexplored areas in the coming years. The measure of its success would eventually be measured by the amount of change it causes in people's lives. The ease with which people are adapting to AI technologies today suggests that the future looks bright.

Source:

1. https://qualetics.com/role-applications-of-ai-in-biotechnology/
2. https://www.researchgate.net/publication/319623622_Artificial_Intelligence_in_Biological_Data