International Journal of Research in Circuits, Devices and Systems

E-ISSN: 2708-454X
P-ISSN: 2708-4531
IJRCDS 2023; 4(1): 29-34

www.circuitsjournal.com
Received: 14-10-2022
Accepted: 15-11-2022

Rupin Kirtan Harendra
School of Computer Science Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

Corresponding Author:
Rupin Kirtan Harendra
School of Computer Science Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

Bioinformatics and computational biology: An overview on the state of the art

Rupin Kirtan Harendra

Abstract
Bioinformatics is a multifaceted discipline that combines several scientific fields such as biology, computer science, chemistry, statistics, and mathematics. Biological data gets entered at an alarming rate and due to this vast amount of data many of the obstacles in biology have to do with computing. Bioinformatics has firmly established itself as a molecular biology specialty, including a wide range of topics ranging from structural biology to genomics and gene expression studies. Computational biology is a branch of biology that solves problems by combining computer science, statistics, and mathematics. Developing algorithms, theoretical models, computational simulations, and mathematical models for statistical inference all come under computational biology. This paper will provide an introduction to the topic, the types of information used, the studies being conducted and finally the applications of this field.

Keywords: Bioinformatics, computational biology, molecular biology, information technology

1. Introduction

Because bioinformatics is a new field that incorporates biology and technology, it was difficult for researchers to establish an accurate definition at first. The term Bio (Molecular Biology) informatics (Information Technology) refers to the tools and methods used to manage, analyse, and manipulate large sets of biological data. It also refers to the design, development, and use of software tools to generate, store, annotate, access, and analyse data and information related to molecular biology. Researchers have made numerous contributions to the definition and explanation of Bioinformatics, but all agree that it is a combination of Biology, Computer Science, Statistics, and Mathematics. It is the science of managing and analyzing voluminous amounts of biological data using advanced computing techniques and in turn understand and organize this information on a large scale. The origins of Bioinformatics can be traced back to Mendel's discovery of genetics inheritance in 1865 then the structure of DNA was determined by James Watson and Francis Crick and in later 1960s the protein and RNA structures determined by Dayhoff's led to the term Bioinformatics being used to describe the management of DNA, RNA, and protein sequence data analysis Later in 2000, the first draught of the Human Genome Sequence was announced, with the discovery of 20 to 25,000 human genes. Accessing this large volume of data would be challenging for biologists, and this ushered in a new era in modern biology, with the help of new computerized technology, resulting in the merging of biology and computer science to form Bioinformatics. The entire sequencing of the human genome through Bioinformatics has aided in the discovery of the genetic contribution to many diseases, and bioinformatics' applications include drug development, personalized medicine, preventive medicine, gene therapy, cancer therapy, and others. While the main target of bioinformatics is medicine it can also contribute to other fields like livestock, agriculture and even space explorations [1, 2, 3, 4, 5]. These applications will be expanded upon in the coming topics.

2. Bioinformatic Algorithms

The following are some of the most notable algorithmic trends in bioinformatics:
1. Finding common threads between strings (such as proteins of different organisms).
2. Recognizing certain patterns in strings (such as genes).
3. Identifying commonalities across spatial structural elements (such as motifs).
4. Tree construction (called phylogenetic trees expressing the evolution of organisms whose

DNA or proteins are currently known).
5. Using already tagged data sets to organize new data into clusters. Microarray data and the pathways' related behavior are used to make decisions.

3. Bioinformatic Aims [6, 7]

1. To organise data in such a way that scholars can access existing data and contribute new items as they become available.
2. To create and construct software tools to assist in data analysis and management.
3. To employ this biological data in a biologically meaningful way in the analysis and interpretation of the outcomes.
4. To assist pharmaceutical industry researchers in understanding the protein structures that lead to and aid in the development of the drug business.
5. To aid and assist medical professionals in understanding gene architecture that will aid in the detection and diagnosis of diseases such as cancer. Cancer is a genetic disease in which cells cannot follow the normal sequential cycle and divide in a normal manner which basically means they will mutate and divide uncontrollably and the chromosomes of the cancer cells will be arranged in an incorrect order or have large parts missing. Bioinformatics in Cancer known as Cancer Bioinformatics is used to produce precision medicinal procedures and products, which delivers the safest and most successful therapy method based on each subject's gene and protein variations, relies heavily on bioinformatics to monitor and anticipate its efficiency and effectiveness.

4. Data, Data types and Structures

Biological data is often distinguished by its massive size. DNA, RNA, Protein Sequences, and Micro Array Pictures are the four forms of data that are made and gathered. The first three columns are written data, and the fourth is a digital image. Because of the large amount of biological data collected, it's obvious that these data are represented in a variety of ways. There are four different kinds of data structures: DNA, RNA, and protein sequences are represented by strings; protein structures are represented by trees; metabolic and signalling processes are represented by graphs; and Strings (like words and phrases) are also employed by academics to express meaning. Substrings, subtrees, and subgraphs are also of importance to academics and biologists. Raw DNA sequences are 1,000-base-long strings of the four base letters that make up genes. There are currently 9.5 billion bases in 8.2 million entries in the GenBank database of nucleic acid sequences. Protein sequences, which are made up of 20 amino acid letters, are the next level. There are currently 300,00 protein sequences known. In terms of information, macromolecular structural data is more complex. The protein data bank now contains 13,000 entries.

Whole genome sequencing has recently been a focal point in bioinformatics. Because raw DNA sequences, or genomes, are made up of strings of base letters ranging from 1.6 million to 3 billion, analysing such large sequences is inefficient. As a result, a key feature of complete genomes is the distinction between coding and non-coding regions, with repetitive sequences accounting for the majority of the base sequences, particularly in eukaryotes (organisms with a nucleus and other membrane-bound organelles), such as animals, plants, fungi and algae as well. As a result, we can now monitor the expression levels of nearly every gene in a cell. Gene expression is the most fundamental level at which the genotype leads to the phenotype, or observable trait, in genetics. The genotype is the genetic information recorded in DNA, but the phenotype is the "interpretation" of that information. Therefore, there is a very apparent variety in the size and complexity of different datasets with repeating segments which leads us into the next topic where the redundancy and multiplicity of this data will be discussed.

5. Overabundance and multiplicity and grouping of data

Much of the data can be classified together based on biological commonalities, such as sequence segments that are frequently repeated at different places of genomic DNA, as detailed in the previous article [8]. Genes (a sequence of chemical events undergone by a molecule or class of compounds in a live organism) can be grouped together based on their specific roles or the metabolic pathway to which they belong [9]. Proteins have comparable sequences because they make up the majority of an organism's biological gene make up. This means that organisms have multiple copies of a specific gene due to duplication, and other species may have identical or similar proteins that were inherited when they diverged in evolution. Even if proteins differ in sequence they still adopt similar or equivalent structures this means that while the rate discovery of novel folds has decreased the number of structures in the Protein Data Bank has increased exponentially [10]. Novel folds can be represented by the inclusion of a relatively higher number of rarely occurring Structural motifs in their structures and, to a lesser extent, by a novel topological combination of commonly occurring Structural motifs. Folds are the basic building block of protein structures, and Structural Motifs are the super-secondary structural elements in protein structures.

There are several terms used to characterise the link between two proteins or the genes from which they are derived: comparable proteins have similar folds but different sequences, whereas homologous proteins are structurally and sequentially identical. It can be difficult to tell the difference between the two categories, especially if the link between the two proteins is distant [11, 12]. It's vital to distinguish between orthologues and paralogues, which are proteins connected by gene duplication within a genome that have developed from a common ancestral gene in distinct species [13].

As can be seen from the preceding discussion, creating methods for detecting similarities between different biomolecules and identifying those that are related is a critical component of managing such a massive amount of data. The important databases that enable access to original sources of data, as well as some secondary databases that categorise the data, will be discussed in the following section. These databases contain categories that make genome comparisons easier, allowing for the identification of similar themes among related genomes as well as highlighting traits that are unique to some.

6. Databases

6.1 Protein sequence databases

There are three types of protein sequence databases: primary, composite, and secondary. In main databases,

which serve as a repository for raw data, about 300,000 protein sequences are stored. SWISS-PROT [14] and PIRInternational [15] are two of the most popular archives, and they annotate the sequences as well as explain the functions, domain structure, and post-translational modifications of the proteins. OWL [16] and NRDB [17] are composite databases that aggregate and filter sequence data from many main databases to create non-redundant sets that are more complete than the component databases. Protein sequence data from translated coding regions in DNA sequence databases is also included. Secondary databases contain information derived from protein sequences that can be used to determine if a new sequence belongs to a known protein family. One of the most well-known is PROSITE [18], a database of short sequence patterns and profiles that characterise biologically significant sites in proteins. PRINTS [19], which compiles a database of protein fingerprints, which are collections of conserved motifs that define a protein family, expands on this concept. Motifs are generally segregated along a protein sequence, although they may be continuous in 3D space when the protein is folded. Because fingerprints contain several motifs, they can encode protein folds and functions more flexibly than PROSITE. These several secondary databases were recently merged into a single resource known as InterPro [20].

6.2 Structural Databases

Then we look at databases of macromolecular structure. The Protein Data Bank (PDB) [21, 22] is a central repository for all 3D macromolecule structures, such as proteins, RNA, DNA, and complexes. Although the bulk of the structures are solved using x-ray crystallography and NMR, some theoretical models are also offered. Because the information presented in individual PDB entries may be difficult to extract, PDBsum [23] provides a separate Web page for each structure in the PDB, giving thorough structural analysis, schematic diagrams, and data on interactions between different molecules in a specific entry. CATH, SCOP, and FSSP are three significant databases that classify proteins based on their structure in order to find structural and evolutionary connections [24, 25, 26]. Protein groups become more similar at lower levels of the classification tree, establishing a hierarchical structural taxonomy. A number of databases are also dedicated to specific macromolecules. The Nucleic Acids Database, NDB [27], contains nucleic acid structures, the HIV Protease Database, HIV-1, HIV-2, and SIV protease structures and complexes, and ReLiBase [28], receptor-ligand complexes.

6.3 Entrez Genome Database

As explained previously the biggest highlight currently lies with the fact that complete genome sequences for different organisms is available. DNA sequences for individual genes that encode protein and RNA products can be found in the GenBank [29], EMBL [30], and DDBJ databases. The Entrez nucleotide database, like the composite protein sequence database, collects sequence data from these basic databases. Individual genomes are released at different sites since whole-genome sequencing is frequently done through worldwide cooperation. The Entrez genome database, which presently contains over 1,000 organisms, brings together all full and partial genomes in one place (August 2000). A list of full genomes, all chromosomes in an organism, comprehensive representations of single chromosomes designating coding and non-coding areas, and specific genes are all available in addition to the raw nucleotide sequence. At each level, there are graphical presentations, pre-computed analyses, and linkages to additional Entrez sections, such as single gene annotations, which include the protein sequence and sequence alignments with similar genes in other genomes. The most basic application of the database is to predict the function of uncharacterized proteins based on their homology to known proteins, as well as to find evolutionary patterns of protein occurrence.

6.4 Gene Expression Level Data

Expression assays, which assess the expression levels of individual genes, have been the most recent sources of genomicscale data. These tests are used to determine the mRNA or protein products that are produced by the cell and hence measure the expression levels. For the measure of mRNA there are three main methods:

- The Affymatrix GeneChip and SAGE technologies, as well as the cDNA microarray. The first approach compares relative levels of mRNA abundance comparing samples, whereas the second and third methods compare absolute amounts.
- The majority of gene expression research has been done on yeast and human genomes, and there is currently no central archive for this data. The majority of studies examine mRNA levels throughout the yeast cell cycle, however some concentrate on a single stage. In people, the primary application has been to better understand expression in tumour and cancer cells. Data from microarray tests on human cancer cells is provided by the Molecular Portraits of Breast Tumours, Lymphoma, and Leukaemia Molecular Profiling projects.
- Currently, the only methods for determining protein abundance are 2D gel electrophoresis and mass spectrometry. Only the most abundant proteins can be shown because gels can only resolve roughly 1,000 proteins.

The next section will elaborate on how this data is applied and integrated in different fields and problems faced.

7. Integration of Bioinformatics

Integrating numerous sources of data is frequently the most profitable study in bioinformatics. When combined with information about a protein's function, occurrence in different genomes, and interactions with other molecules, for example, the 3D coordinates of a protein become more valuable. In this method, individual pieces of data are placed in context with other data. Unfortunately, due to differences in nomenclature and file formats, accessing and cross-referencing diverse sources of information is not always easy. This problem is frequently solved at a fundamental level by providing external links to other databases; for example, in PDBsum, web pages for specific structures direct users to matching entries in the PDB, NDB, CATH, SCOP, and SWISS-PROT databases. Attempts have been made at a higher level to combine access to several data sources. One is the Sequence Retrieval System, or SRS, which indexes nucleic acid, protein sequence, protein motif, protein structure, and bibliographic databases to allow users to retrieve, link, and access items from each. The Entrez facility, which was previously mentioned, is another option. A search in either database for a single gene returns the genome from which it originated, the protein sequence it

encodes, its structure, a bibliographic reference, and equivalent entries for all related genes.

7.1 DNA Sequences

Separating coding and non-coding sections, as well as identifying introns, exons, and promoter regions for annotating genomic DNA, are all steps in the investigation of raw DNA sequences.

7.2 Protein Sequences

Protein sequence studies include the creation of algorithms for sequence comparisons, methods for creating many sequence alignments, and the search for functional domains from conserved sequence patterns in such alignments.

7.3 Structural Data

Structural data investigations include the prediction of secondary and tertiary protein structures, the development of methods for 3D structural alignments, the analysis of protein geometries using distance and angular measurements, the computation of surface and volume shapes, and the analysis of protein interactions with other subunits, DNA, RNA, and smaller molecules. As a result of the growing availability of annotated genomic sequences, computational genomics and proteomics — large-scale analysis of complete genomes and the proteins they encode – have arisen. The research includes characterization of protein composition and metabolic pathways across genomes, identification of interacting proteins, gene product assignment and prediction, and large-scale assessments of gene expression levels.

In addition to establishing connections, much of the variation in proteins is due to the interactions between them. Bioinformatics is the study of biological data in the form of facts to deduce and comprehend observations for a different form of data. Statistical criteria derived from structures, such as the proclivity of certain amino acid sequences to make certain secondary structural components, are commonly used in these procedures, particularly the former. Another example is the use of structural data to understand a protein's function; research has looked into the relationship between different protein folds and functions, as well as the similarities between different binding sites in the absence of homology [39]. When paired with similarity tests, these findings provide us a clearer picture of how much biological information may be accurately communicated between homologous proteins. This leads us into the next section which is the ongoing research and applications of bioinformatics.

8. Research and applications of Bioinformatics

8.1 Identifying Homologous Properties Between Biomolecules

As previously said, one of the primary motives behind bioinformatics is the search for commonality among diverse biomolecules. Apart from facilitating systematic data organisation, protein homologue identification has various immediate practical benefits. The most obvious is information transfer between connected proteins. If a protein is poorly characterised, for example, it is possible to hunt for homologues that are well-studied and, with caution, apply some of the latter's information to the former. When it comes to structural data, theoretical models of proteins are usually based on experimentally solved structures of close homologues. Fold recognition employs similar methods, with tertiary structure predictions based on the discovery of distant homologous structures and analysing if the prediction is energetically viable. Investigations in low-level species such as yeast might be undertaken and the results used to homologues in higher-level organisms such as humans, where tests are more challenging, where biochemical or structural data is insufficient.

Homologue finding is routinely used to annotate individual genes, and functional data is frequently transmitted to confirm coding areas in newly sequenced genomes. It also simplifies the difficulty of understanding complex genomes on a wider scale by first analysing simple species and then applying the same ideas to more intricate ones. The same concept can be used in the opposite direction. Checking whether homologues of important microbial proteins are lacking in humans quickly identifies new treatment options. On a lesser scale, structural variations between similar proteins could be used to create therapeutic compounds that bind to one structure but not the other. This leads on to the next section, drug design.

8.2 Drug design

One of the early medical applications of bioinformatics was to help with rational medicine design. Through linkage studies and its similarities to MMR genes in mice, the gene has been associated to nonpolyposis colorectal cancer. Given the nucleotide sequence, translation tools can be used to calculate the expected amino acid sequence of the translated protein. The structure of the human protein can then be modelled using empirically defined structures based on sequence similarity, utilising sequence search methods to discover homologues in model organisms. Finally, docking algorithms could generate chemicals that bind to the model structure, allowing biochemical experiments to assess their biological activity on the genuine protein.

8.3 Generalisation of subject areas through Large-scale censuses

Although databases can efficiently store all of the information relating to genomes, structures, and expression datasets, it is beneficial to compress all of this data into clearly understandable trends and facts for customers. Broad generalisations aid in the identification of interesting subject areas for more extensive investigation, as well as the placement of fresh observations in the correct perspective. This allows them to see if they are out of the ordinary in any manner. One can answer a variety of evolutionary, biochemical, and biophysical questions with these large-scale censuses. Because specific protein folds are usually associated to specific biochemical functions, these studies demonstrate the diversity of metabolic pathways in various animals.

One of the most fascinating new sources of genetic information, as we described before, is expression data. We can query whether the high occurrence of a protein fold in a genome is predictive of high expression levels by combining expression data with structural and functional classifications of proteins. The subcellular localizations of proteins and their interactions with one another provide additional genomic size data that can be considered in largescale surveys. We may then start compiling a map of all protein-protein interactions in an organism using structural data.

8.4 Applications in medical science

Gene expression analysis has been the subject of current medical applications. This frequently requires collecting expression data for cells affected by diseases like cancer and aortic stenosis and comparing the results to normal expression levels. The identification of genes that are expressed differently in sick cells lays the groundwork for understanding disease aetiology and identifying potential treatment targets.

When provided a lead molecule, microarray experiments can be used to analyse reactions to pharmacological intervention and provide early tests to discover or predict the toxicity of experimental drugs. Further advancements in bioinformatics, paired with individual-level experimental genomics, are expected to transform the future of healthcare. Post-natal genotyping can be used to identify susceptibility or immunity to particular diseases and infections in a typical patient situation. With this knowledge, a one-of-a-kind vaccination combination might be offered, potentially saving healthcare costs by avoiding unnecessary treatments and averting problems later in life. Regular testing throughout one's life may lead to dietary recommendations and early disease identification. Additionally, drug-based therapy could be tailored to the patient's needs and disease.

Now the next section will briefly discuss Computational Biology.

9. Computational Biology: Definition and Application

First what is the difference between computational biology and bioinformatics while there are quite a few overlapping characteristics between them there are still some key differences. Bioinformaticians are biologists who specialise in using computational tools and systems to solve biological challenges. Computational biologists are computer scientists, mathematicians, statisticians, and engineers who specialise in designing theories, algorithms, and approaches for such tools and systems. This is a simple differentiation between the two.

Initially, computational biology was concerned with the sequencing and structure of biological molecules, which was often done in an evolutionary environment. However, starting in the 1990s, it began to focus more on function analysis. The sequence and structural similarities of an unknown and a known protein, as well as the proteins' interactions with other molecules, are used to make functional predictions. Because such analyses can be lengthy, computational biology has grown increasingly entwined with systems biology, which aims to understand how massive interacting networks of biological components, particularly biological processes, function. Because biochemical, regulatory, and genetic pathways are highly branched and interleaved, as well as dynamic, advanced computational approaches are required to describe and analyse them. In addition, new technology platforms for the rapid, automated (high-throughput) generation of biological data have permitted a change from traditional hypothesis-driven testing to data-driven analysis, permitting massive-scale computer studies on genome-wide databases. As a result, without the power of computers and computer science methodologies, many aspects of biological study would be impossible.

10. Conclusion

In light of the current data avalanche, computational approaches have become increasingly crucial in biological research. Bioinformatics was founded for the analysis of biological sequences and now covers a wide range of topics such as structural biology, genomics, and gene expression investigations. In this review, we provided an overview and assessment of the current state of the field. We looked at the many types of biological information and databases that are commonly used, as well as some of the current studies – with an emphasis on transcription regulatory systems – and a few practical applications of the issue.

11. References

1. Zadeh J. An undergraduate program in bioinformatics," Potentials, July-Aug. 2006; IEEE. 2006;25:3.
2. Chavan PR. Application of Bioinformatics in the Field of Cancer Research", 11th Workshop on Medical Informatics & CME on Biomedical Communication, 2008 November, 20-22.
3. Poe D, Venkatraman N, Hansen C, Singh G. Component-Based Approach for Educating Students in Bioinformatics," Education, IEEE Transactions on. 2009 Feb;52:1.
4. Umarji M, Seaman C, Koru AG. Hongfang Liu; "Software Engineering Education for Bioinformatics," Software Engineering Education and Training, 2009. CSEET '09. 22nd Conference on, 2009 Feb.
5. Jana R, Aqel M, Srivastava P, Mahanti PK. Soft Computing Methodologies in Bioinformatics, European Journal of Scientific Research. 2009;26:2.
6. Holm L, Sander C. Touring protein fold space with Dali/FSSP. Nucleic Acids Res, 1998.
7. Sai GH, Tripathi K, Tyagi AK. Internet of Things-Based e-Health Care: Key Challenges and Recommended Solutions for Future. In: Singh, P.K., Wierzchoń, S.T., Tanwar, S., Rodrigues, J.J.P.C., Ganzha, M. (eds) Proceedings of Third International Conference on Computing, Communications, and Cyber-Security. Lecture Notes in Networks and Systems. Springer, Singapore, 2023, 421. https://doi.org/10.1007/978-981-19-1142-2_37
8. Subasree S, Sakthivel NK, Khushboo Tripathi, Deepshikha Agarwal, Amit Kumar Tyagi. Combining the advantages of radiomic features based feature extraction and hyper parameters tuned RERNN using LOA for breast cancer classification, Biomedical Signal Processing and Control. 2022, 72Part A. 103354, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2021.103354.
9. Somisetti K, Tripathi K, Verma JK. Design, Implementation, and Controlling of a Humanoid Robot, 2020 International Conference on Computational Performance Evaluation (ComPE), 2020, p. 831-836. Doi:10.1109/ComPE49325.2020.9200020.
10. Midha S, Triptahi K. Extended TLS security and Defensive Algorithm in OpenFlow SDN," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2019, p. 141-146, Doi:10.1109/CONFLUENCE.2019.8776607.
11. Midha S, Tripathi K. Extended Security in Heterogeneous Distributed SDN Architecture. In: Hura, G., Singh, A., Siong Hoe, L. (eds) Advances in Communication and Computational Technology. Lecture Notes in Electrical Engineering. Springer,

Singapore, 2021, 668. https://doi.org/10.1007/978-981-15-5341-7_75
12. Mapanga I, Kumar V, Makondo W, Kushboo T, Kadebu P, Chanda W. Design and implementation of an intrusion detection system using MLP-NN for MANET, 2017 IST-Africa Week Conference (IST-Africa), 2017, p. 1-12. Doi: 10.23919/ISTAFRICA.2017.8102374.
13. Agarwal D, Tripathi K. A Framework for Structural Damage detection system in automobiles for flexible Insurance claim using IOT and Machine Learning, 2022 International Mobile and Embedded Technology Conference (MECON), 2022, p. 5-8. doi: 10.1109/MECON53876.2022.9751889.
14. Khushboo Tripathi, Deepshikha Agarwal, Kumar Krishen, In: Handbook of Research of Internet of Things and Cyber-Physical Systems, Apple Academic Press, 2022, p, 32. EBOOKISBN9781003277323.
15. Khushboo Tripathi, Neha Bhateja, Ashish Dhillon. Changing the Conventional Banking System through Blockchain, In: The Smart Cyber Ecosystem for Sustainable Development, Wiley, 2021 September, 12. https://doi.org/10.1002/9781119761655.ch20
16. Vishnuram G, Tripathi K, Kumar Tyagi A. Ethical Hacking: Importance, Controversies and Scope in the Future," 2022 International Conference on Computer Communication and Informatics (ICCCI), 2022, p. 01-06. Doi:10.1109/ICCCI54379.2022.9740860.
17. Nair MM, Tyagi AK, Sreenath N. The Future with Industry 4.0 at the Core of Society 5.0: Open Issues, Future Opportunities and Challenges, 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021, p. 1-7. Doi: 10.1109/ICCCI50826.2021.9402498.
18. Tyagi AK, Fernandez TF, Mishra S, Kumari S. Intelligent Automation Systems at the Core of Industry 4.0. In: Abraham A., Piuri V., Gandhi N., Siarry P., Kaklauskas A., Madureira A. (eds) Intelligent Systems Design and Applications. ISDA 2020. Advances in Intelligent Systems and Computing, Springer, 2021, 1351. Cham. https://doi.org/10.1007/978-3-030-71187-0_1
19. Goyal, Deepti, Tyagi, Amit. A Look at Top 35 Problems in the Computer Science Field for the Next Decade, 2020. 10.1201/9781003052098-40.
20. Madhav AVS, Tyagi AK. The World with Future Technologies (Post-COVID-19): Open Issues, Challenges, and the Road Ahead. In: Tyagi A.K., Abraham A., Kaklauskas A. (eds) Intelligent Interactive Multimedia Systems for e-Healthcare Applications. Springer, Singapore. 2022. https://doi.org/10.1007/978-981-16-6542-4_22
21. Mishra S, Tyagi AK. The Role of Machine Learning Techniques in Internet of Things-Based Cloud Applications. In: Pal S., De D., Buyya R. (eds) Artificial Intelligence-based Internet of Things Systems. Internet of Things (Technology, Communications and Computing). Springer, Cham. 2022. https://doi.org/10.1007/978-3-030-87059-1_4
22. Rekha G, Tyagi AK, Anuradha N. Integration of Fog Computing and Internet of Things: An Useful Overview. In: Singh P., Kar A., Singh Y., Kolekar M., Tanwar S. (eds) Proceedings of ICRIC 2019. Lecture Notes in Electrical Engineering. Springer, Cham, 2020, 597. https://doi.org/10.1007/978-3-030-29407-6_8
23. Sheth HSK, Tyagi AK. Mobile Cloud Computing: Issues, Applications and Scope in COVID-19. In: Abraham, A., Gandhi, N., Hanne, T., Hong, TP., Nogueira Rios, T., Ding, W. (eds) Intelligent Systems Design and Applications. ISDA 2021. Lecture Notes in Networks and Systems, 2022, 418. Springer, Cham. https://doi.org/10.1007/978-3-030-96308-8_55
24. Tyagi, Amit Kumar, Shamila M. Spy in the Crowd: How User's Privacy Is Getting Affected with the Integration of Internet of Thing's Devices (March 20, 2019). Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur - India, 2019 February 26-28.
25. Tyagi, Amit Kumar. Building a Smart and Sustainable Environment using Internet of Things (February 22, 2019). Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur - India, 2019 February 26-28.
26. Nair MM, Kumari S, Tyagi AK, Sravanthi K. Deep Learning for Medical Image Recognition: Open Issues and a Way to Forward. In: Goyal D., Gupta A.K., Piuri V., Ganzha M., Paprzycki M. (eds) Proceedings of the Second International Conference on Information Management and Machine Intelligence. Lecture Notes in Networks and Systems, 2021, 166. Springer, Singapore. https://doi.org/10.1007/978-981-15-9689-6_38
27. Amit Kumar Tyagi, Meghna Mannoj Nair, Deep. Learning for Clinical and Health Informatics, in the book "Computational Analysis and Deep Learning for Medical Care: Principles, Methods, and Applications", 2021 July 28. DOI: https://doi.org/10.1002/9781119785750.ch5
28. Kumari S, Vani V, Malik S, Tyagi AK, Reddy S. Analysis of Text Mining Tools in Disease Prediction. In: Abraham A., Hanne T., Castillo O., Gandhi N., Nogueira Rios T., Hong TP. (eds) Hybrid Intelligent Systems. HIS 2020. Advances in Intelligent Systems and Computing. Springer, Cham, 2021, 1375. https://doi.org/10.1007/978-3-030-73050-5_55
29. Gudeti B, Mishra S, Malik S, Fernandez TF, Tyagi AK, Kumari S. A Novel Approach to Predict Chronic Kidney Disease using Machine Learning Algorithms, 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, 2020, p. 1630-1635. Doi: 10.1109/ICECA49313.2020.9297392.
30. Patel NJ, Tripathi K. Trust Value based Algorithm to Identify and Defense Gray-Hole and Black-Hole attack present in MANET using Clustering Method. International journal of scientific research in science, engineering and technology, 2018;4:281-287.