Applications of Genetic Algorithms, Fuzzy Logic and Swarm Intelligence in Data Analytics

| Rumesh Mohan  *Master of Data Analytics* | Sachini Jayawardena  *Master of Data Analytics* | Sajith Sanjeewa  *Master of Data Analytics* |
| --- | --- | --- |
| *Department of Industrial Management* | *Department of Industrial Management* | *Department of Industrial Management* |
| *University of Kelaniya* | *University of Kelaniya* | *University of Kelaniya* |
| *Dalugama, Kelaniya 11600, Sri Lanka* | *Dalugama, Kelaniya 11600, Sri Lanka* | *Dalugama, Kelaniya 11600, Sri Lanka* |
| rumeshmohan@gmail.com | sachinidjayawardena@gmail.com | sanjeewasajith@gmail.com |

***Abstract* - Currently data analytics gets more and more attention in analyzing massive raw data to make conclusions about information. It is proven that Genetic Algorithm, Fuzzy Logic and Swarm Intelligence based applications have given optimized solutions in data analytics rather than traditional methods due to large volume, dynamic changes and noise etc. Data Analysis is the art of extracting knowledge from a large set of data. This paper focuses on applications that are based on Genetic Algorithm, Fuzzy Logic and Swarm Intelligence and its related domains in data analytics which cater the needs and wants of mankind.**

*Keywords - Fuzzy Logic (FL), Genetic Algorithm (GA), Swarm Intelligence (SI), Convolutional Neural Networks (CNN), Linguistic Variables, Proportional Integral (PI), Multilevel Image Segmentation Model (MIS-XMACO).*

1. **INTRODUCTION.**

Genetic Algorithm, Fuzzy Logic and Swarm Intelligence are subfields of Artificial Intelligence (AI) which we will be emphasising and explaining throughout this research.

***A. Genetic Algorithm.***

Genetic Algorithm was developed by John Holland, his colleagues and students at the University of Michigan in 1975. GA and its variants have been used in a variety of fields for problem solving. It provides solutions for problems such as space research, economics, geography, remote sensing, agriculture, data mining and cancer detection etc. Genetic Algorithm is an application of heuristic search algorithms [1]. Heuristic Search methods are used to arrive at an approximate global solution. Heuristic search is a problem specific search method. It is further categorized into a meta-heuristic search optimization approach which is an independent solution method under global search method, which can be used to arrive at a global optimal solution.

***B. Fuzzy Logic.***

Fuzzy Logic is a mathematical technique that allows inclusion of vague human interpretations and assessments into computational problems. This innovative mathematical approach which is qualitative and not fully quantitative was introduced as a logic to accommodate partial truth or vagueness to find solutions that are impossible using binary logic. This logic was first introduced by Lotfi Aliasker Zadeh [16]. The key advantage of using these methods is its ability to deal with meaning-less systems and linguistic variables that allow faster and crisp development of controllers which assist in problem solving systems.

Fuzzy Logic Controller operates under 3 steps:

**Step 1: Fuzzification**

In this step external devices/sensors crisp variable inputs are converted into fuzzy (vagueness) variables using the membership functions.

**Step 2: Fuzzy Inference System (FIS).**

This process is simple and usually uses if-then rules with a set of propositional logic to function [13]. This inference system functions under the knowledge base which consists of fuzzy sets and fuzzy rules [11]. Basically mapping the outputs after fuzzification using fuzzy sets and fuzzy rules and sending its output for defuzzification is the process handled by the fuzzy inference system [10]. There are 2 types of fuzzy inference systems: Mamdani FIS, Sugeno and Tsukamoto FIS.

**Step 3: Defuzzification.**

Defuzzification converts fuzzified value into crisp value again so it can be processed by actuators. Defuzzification uses the centre of gravity method to find an accurate final crisp value to output. Mamdani uses the Center of Gravity technique for the defuzzification process; while Sugeno and Tsukamoto use Weighted Average to calculate the crisp output [12].

***C. Swarm Intelligence.***

Swarm Intelligence, nature -inspired metaheuristic algorithm was first introduced in 1989 by Gerardo Benny and Joon Wang. Since then SI has emerged in many fields including engineering, economics, health and computer science, etc. Simulating simple agents interacting locally with each other much like natural swarms. Scientists have been studying the behavior of social insects due to their efficiency of solving problems such as finding the best path from their nest to food sources or locating their nests. Even though the insects are ingenuous they are intelligent as a swarm by interacting with each other and the environment. These intelligent behaviors are transformed to swarm intelligence algorithms to find optimized solutions for complex problems [17], [30].

SI, which uses the knowledge of collective objects (people, ants, bees, etc.) together, defines an optimal solution for a complex problem. Collective objects are either natural or artificial. Impression of SI is that a solution given by an individual person may not be the best solution or it may not suit others. It might be a solution for that person only. To avoid this problem, getting solutions from a cluster of people (swarm) and finally computing all solutions together will be the best possible optimal solution which might suit everyone.

Example:

Prediction of 100 oranges in a particular box, suppose an individual prediction is 80. Instead of getting individual opinion, a group response would be more accurate.

| P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 60 | 95 | 100 | 84 | 70 | 100 | 85 | 80 | 95 | 72 |

Average no of oranges = (60+95+100+84+70+100+85+80+95+72)/10 = 84.1 = 84

Therefore the difference (error) from individual to swarm reduces from 20 to 16.

1. **LITERATURE REVIEW.**

The concept of GA derives from the biological concept of evolution, also known as survival of the fittest. It is an algorithm which follows the evolutionary process. The weakest elements are eliminated while the strongest elements grow stronger in this process. Finding an optimal solution using GA involves randomly finding a solution in the solution space from the strings of integers known as chromosomes. GA uses the objective function or the fitness function to arrive at a fitness value for each chromosome. The elements inside these strings are known as genes. According to Holland’s algorithm, genes will be assigned an integer value between 0 and 1 [2].

The population generated in GA will be controlled by genetic operators, selection, crossover and mutation. The first step is the random generation of a population of chromosomes. Then, the fitness function is applied to each chromosome to assess each value. The selection of parents for the cross-over will be based on the fitness value assessed for each chromosome. After that two parents are selected for reproduction using proportionate probability or using a selection method such as roulette wheel selection method, tournament selection method. The genes of uniformly selected two chromosomes will be swapped from each parent to produce offspring. In mutation, the genes in the produced offspring(chromosomes) will be again assigned a new gene value after selecting a gene randomly. The fitness value will be assessed for these offspring and included in the population. Then, the weakest chromosomes will be removed, and stronger chromosomes will be kept in the population. Finally, a chromosome is selected based on the fitness value in the population. The GA process will be iterated repeatedly and terminated once an approximate solution is found [1].

***Applications of Genetic Algorithms in Data Analytics.***

*A. Research on the Application of Genetic Algorithms in Physical Education.*

This section discusses how the GA has been used in developing a mathematical model to group intelligent paper grouping systems for students who use Massive Open Online Courses (MOOC) platforms to learn physical education in China. The modern and novel ways that physical education can be taught by expanding scope of the learning, different regions, other schools, other country sports other than one’s own course content. If a student wants to take up multiple courses, it will be possible since the courses are readily available any time. In this study, the improvement of genetic algorithms is realized by using race selection in factor operation.

Based on this, an improved adaptive GA selection mechanism is proposed to allow the target population to operate adaptively, allowing the GA to be used in the first step. It enables the GA to converge too soon in the early stages or too late in the late stages. An improved GA is proposed to avoid the risk of slow and partial convergence of the basic GA for intelligent grouping systems. By continuously eliminating similar individuals, the algorithm can rapidly expand the search space while ensuring the stability and diversity of the clusters. As a result, based on the improved GA, this study proposes a new method of intelligent grouping to create test papers for physical education. Using this model, an automatic paper is formed with the specified parameters [3].

*B. Automatically Designing CNN Architectures Using Genetic Algorithm for Image Classification.*

Image classification with automatic Convolutional Neural Networks (CNN) is another application of GA. GA can be effectively used in Image classification tasks where the users have less domain knowledge in CNN. Therefore, GA will help the users to obtain an approximate CNN Architecture design for the provided images. The studies show that the evolutionary algorithm based CNN architecture designs takes less time compared to reinforcement learning designs. Majority of the CNN blocks are designed based on the basic components of blocks of CNNs. Since the CNN has limited generalization ability, the use of GA will narrow the search space to easily identify the best performance. Predefined sets of building blocks with the required population size with maximum generation number and image dataset for classification will be used to get a best design of CNN for the image classification [4].

*C. Prediction of Indonesian Inflation Rate Using Regression Model Based on Genetic Algorithm.*

The inflation rate of a country acts as one of the main indicators which assures the stability of its currency. In Indonesia, Bank Indonesia is focused on maintaining Rupiah stability. The inflation rate also shows the general trends in development of prices. Consumer Price Index (CPI) is used to measure the inflation of a country. The Changes of CPI implies the increase or decrease in inflation rate of goods and services over time. Therefore, using historical CPI data, a country can predict its future inflation rates. The method used for this is the time series data predictions which is based on regression modelling. In this scenario, the Multiple Regression Analysis method is used to predict the value of one dependent variable based on two or more independent variables, which produces a regression model. The GA has been used in finding the optimal prediction model from historical CPI data to be used for the prediction of Indonesian inflation rate [5].

***Applications of Fuzzy Logic in Data Analytics.***

*A. Fuzzy Logic systems to detect the risk of being affected by Heart Disease.*

Today heart disease is common among humans be it any age, gender, family history and habits, etc. Change in behaviours, characters and situational issues make it even worse that drive it to chronic diseases. Heart diseases account for about 35% of the global death rate. This heart disease detection system was enhanced over the years and the recent invention in the year of 2021 using fuzzy logic along with improved C4.5 algorithm was able to collect abundant samples as data from target patients. This experimental work obtained 94.55% accuracy. At first patients’ data are input into the system and fuzzification is done, later Mamdani method is applied in a fuzzy inference engine of 87 rules and finally defuzzification is done by the centre of gravity method to obtain the accuracy of the system. This innovation is considered as a milestone in data analytics as well as in medicine [15].

*B. Fuzzy Logic systems used in Agriculture for Crop Management and Climate Predictions.*

Agriculture mainly requires a proper climate for better yields. Due to global warming and unpredictable climatic changes, it is necessary to have technology based analytical systems to help farmers predict future weather conditions to enhance crop yields and management. The Southern Oscillation Index (SOI) is one measure of the significant fluctuations in air pressure occurring between the western and eastern tropical pacific. In Indonesia, Scientists have collected these SOI data together with other climate related metrics and with the use of FIS have found ways to increase crop production [14]. Fuzzy image analysis with integration of date-grading systems developed using FIS allowed easy crop management that was accurate close to 0.9 probability providing better decision making than humans on storing and preserving crops. Also fuzzy based intelligent systems help improve decision making in supply-chain management [14].

*C. Fuzzy Logic systems used in Nuclear Power Plant Reactors.*

Nuclear energy is now the most reliable energy source in today’s world as its generation and fatalities of generation doesn’t affect and imply much complications compared to other energy sources but agreeable that it has its own disadvantages. Nuclear based power generation requires proper power plants maintained with serious protocols and regulations. Nowadays fuzzy logic controllers are used in designing and constructing nuclear power plants to control turbine power, reactor pressure, and reactor water level based on collected data from simulations with conventional PI controllers. Fuzzy Logic controllers which are used in nuclear power plants control high power range turbine, pressure and reactor water level (if it consists) simultaneously with the consideration of interactions among plant subsystems. Plant simulation data obtained show that these controllers perform better compared to the existing PI controllers, thus implying that the fuzzy logic control system based application in nuclear power plants is feasible and beneficial, enhancing its production by saving time and power consumption drastically.

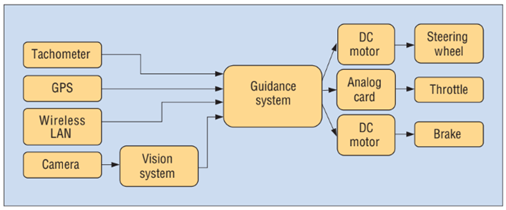
*D. Fuzzy Logic systems used in Production of Plastic Pipes.*

Water plays a vital role in plastic pipe production. Cooling water system is used to manage the temperature of water during this process. During the water cooling process freon compressor is centred within a refrigerator which does the water cooling process. As water moves between the cold and warm part for the cycle to continue in the procedure of manufacturing pipes, the compressor keeps on working always at maximum power without any downtime consuming power unnecessarily. In order to enhance and improve this freon compressor, a fuzzy logic system is integrated to the freon compressor. By this the load of the freon compressor is optimized through temperature monitoring and control system improvement using data obtained every time interval [9]. The system entirely collects the temperature of the freon compressor as well as the power it consumes, adds it to the Mamdani FIS and takes decisions that optimize both temperature and power utilized by freon compressor and make it process with less load without any disruptions [8].

*E. Fuzzy Logic systems used in Automated Vehicle Control.*

The automated vehicle consists of sensorial equipment which supplies all the collected data from the external environment to the fuzzy logic based guidance system. This guidance system controls signals to manage the vehicle by making use of its actuators like brakes, steering and throttle [10]. Using fuzzy logic in guidance systems as a part of an automated vehicle control system is one of the best solutions as it mimics how human drives, allows to convert human procedural knowledge into control algorithms and also has a history of proven tested methods for these systems. The vehicle consists of various other features which are controlled by the guidance system under heavy data monitoring and analytical approach like Adaptive Cruise Control (ACC) + Stop&Go controller, adaptive navigation, vehicle tracking and accident prevention, etc.

Fig.1: Guidance System used in Automated Vehicle Control



***Applications of Swarm Intelligence in Data Analytics.***

Table 1: Some Swarm Intelligence Algorithms [29].

| **Algorithm** | **Source of Inspiration** | |
| --- | --- | --- |
| Brainstorming Optimization – BSO  Teaching-Learning-Based Optimization – TLBO | Brainstorming process of human  How teachers influence learners | Human Society |
| Practical Swarm Optimization – PSO  Ant Colony Optimization – ACO  Artificial Bee Colony – ABC  Firefly Algorithm – FA  Glowworm Swarm Optimization – GSO  Cuckoo Search Optimization – CSO  Monarch Butterfly Optimization – MBO  Cockroach Swarm Optimization - CoSO | Bird flocking and foraging  Ants foraging mechanisms  Foraging behavior of honey bees  Bioluminescence of fireflies  Luciferin induced glow of glowworm  Obligate brood parasitism in cuckoos  Migration behavior of monarch butterfly  Food searching behavior of cockroach | Insects, Birds, etc. |
| Bat Sonar Algorithm – BSA  Grey Wolf Optimization – GWO  Monkey Algorithm – MA  Lion Optimization Algorithm - LOA  Elephant Herding Algorithm – EHA  Cat Swarm Optimization - CaSO | Echolocation behavior of microbats  Leadership and hunting of grey wolves  Climbing techniques used by monkeys  Cooperation characteristics of lions  Herding behavior of elephants  Tracing(inactive) and seeking(active) behavior of cats | Animals |
| Bacterial Foraging Optimization – BFO | Group foraging behavior of bacteria | Microscopic |
| Fireworks Algorithm - FA | Firework explosions | Other |

*A. Application on Ant Colony Optimization for Image*

*Segmentation and Internet of Things (IOT)*

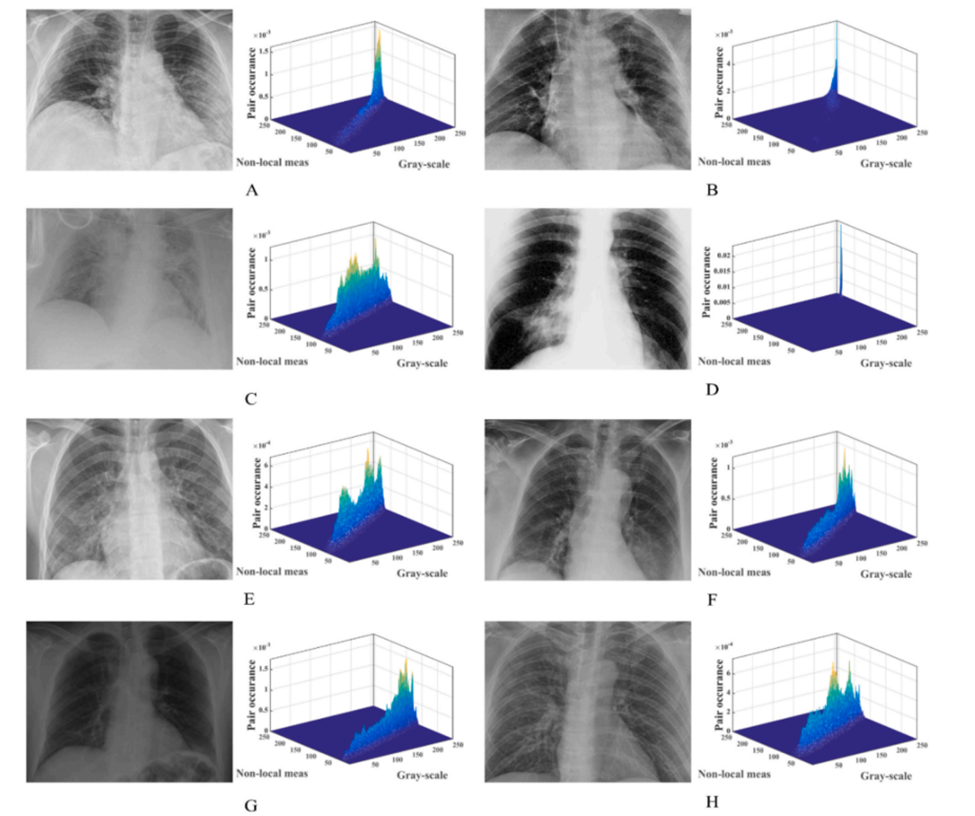
Ant Colony Optimization is a well-known swarm intelligence technique inspired by the indirect communication of real ants. ACO was preliminary designated for solving combinatorial optimization problems. Ants use pheromones to track other ants to the shortest path between food source and their nests. Ants randomly search their space for food sources and each ant leaves a pheromone trail along those paths. By tracing these higher volumes of pheromones within a certain period of time more ants start following it for food sources. Similarly, ACO uses virtual pheromones to find optimized solutions for many real world problems (Olaru et al., 2019).

Table 2: Some common applications of ACO algorithms.

| **Problem Type** | **Optimizing Problem** | **Description** |
| --- | --- | --- |
| Routing | Traveling Salesman  Vehicle Routing | One or more agents visit a predefined set of locations.  Objective function and constraints depend on the visiting order of locations. |
| Scheduling | Industrial Scheduling  Project Scheduling | Concern assignments of jobs to one or different machines over time. |
| Assignment and Layout | Graph Coloring  Course Timetabling  Ambulance Location  Supply Chain Management | Set of items assigned to a given number of resources subject to some constraints. |
| Bioinformatics | Protein Folding  Docking  DNA Sequencing | Applications for molecular biology. |
| Dynamic | Network Routing | Applications for communication network problems, find availability of links or the cost of traversing links. |

Image segmentation is a fundamental task of medical diagnosis. According to Ailiang Qi et al. (2022). ACO algorithm based application used for COVID-19 X-ray image segmentation. Figure 2 illustrates eight real COVID-19 X-rays and their corresponding image segmentation experiments. This is done by ant colony optimization with directional crossover strategy and the directional mutation strategy (XMACO, MIS-XMACO) , an advanced algorithm of ACO [18].

Fig. 2 : Original COVID – 19 X-ray image and the corresponding 2D histograma.(Qi et al., 2022)



SI has a great involvement in Internet of Things (IoT) based systems in logical control of their operations. IoT is dominating urban cities and transforming them to smarter cities. It uses cutting edge IT technology to improve the quality of services (QoS) in cities. To achieve this IoT build up with a large number of geographically distributed nodes with smart sensors. These sensors capture data, aggregate and analyse.

IoT applications are diversified and available in smart homes, traffic control, health care, smart transportation, smart education, sales and emergence response, etc. Cloud computing is the data storage and data analysis mechanism because of its large storage and processing strength. This is known as task offloading. Due to bandwidth limitation delays on data analysis occurs. The fog computing concept is introduced to overcome this overhead by bringing computing capabilities closer to IoT devices. The edge computing has some similarities to the fog computing. Edge provides computing and storage services using small data centers close to IoT devices via WiFi access points. Fog computing provides computing, networking, and storage services for IoT devices using network devices within LAN [19]. Quality of service (QoS) in fog computing needs effective finding of proper nodes, improving response time of the tasks and load balancing between nodes. ACO and PSO are used to optimize these QoS requirements. ACO based scheduler is commonly used in this matter [20].

*B. Cuckoo Search Algorithm for Power Generation Optimization*

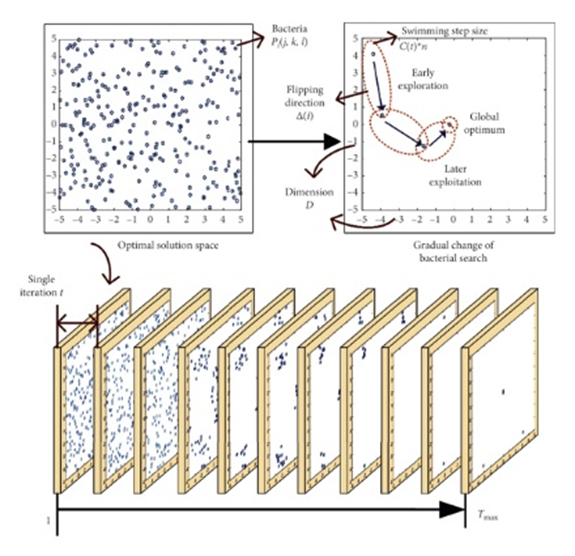
This nature inspired computation and SI method was introduced by Xin-She Yang and Suash Deb in 2009. CSA is based on the obligate brood parasitism of some cuckoo species and their egg laying strategy. Cuckoo birds lay their eggs in other birds’ nests (Host Birds of other species with similar texture, size, and colour of eggs). Each egg in a nest represents a solution and a Cuckoo egg represents a new solution. Cuckoo has the capability of selecting the nests containing recently laid eggs and also removing the existing eggs to increase the hatching probability of their own eggs. Some host birds are able to throw out these cuckoo eggs or build a new nest in a distinct location. By selecting the most suitable nests they maximize the cuckoo egg’s survival rate. High quality eggs which are more similar to host eggs bring next generation mature cuckoos. Snap Drift Cuckoo Search (SDCS), Dynamic Step Size Cuckoo Search Algorithm (DMQL-CS), Discrete binary Cuckoo Search, Neural based Cuckoo Search are some variants developed to enhance the Cuckoo Search [21], [22]. CSA has been used in different optimization fields including image processing, engineering optimization, real world applications, scheduling, feature selection, planning, forecasting and logistics distribution center locationing [23].

Photovoltaic based energy generation is becoming more popular around the world. International renewable energy agency has been forecasting that solar PV capacity may reach between 1760 and 2500 gigawatts (GW) by 2030 (Gielen et al., 2016 as cited in Jude & Jana, 2020). A PV cell is the primary element of PV systems. A PV module designed using several PV cells joined in series and parallel configuration. To estimate optimum power generation capacity from PV systems two popular circuit models proposed. Single Diode Model (SDM) and Double Diode Model (DDM) are the two models respectively preferable for outdoor irregular weather conditions and normal weather conditions. CSO and some of its variants have been used for parameter extraction of PV cells [24].

*C. Bacterial Foraging Optimization.*

Bacterial foraging optimization is a nature inspired optimization algorithm based on the behaviour of “Escherichia Coli” bacteria commonly known as E.coli. It simulates the foraging behavior of searching for nutrients. This is used to find approximate solutions to extremely difficult or impossible real world optimization problems which were proposed by Professor Passino in 2002. Compared to other optimization algorithms BFO is becoming popular due to its simple bacterial individual structure and behavior, varied group and characteristics and efficient life cycle. Probabilistic planning, load dispatch, reconstruction, loss minimization of power energy networks, robot automation, optimization of wireless networks are examples for BFO utilization [25].

Fig. 3 : Fundamental structure of the BFO algorithm [25].



BFO estimates cost function after each iterative step of the algorithm towards progressively better fitness. Optimized parameters represent the position of bacteria within the space. One bacteria is positioned at each point. After each progressive step bacteria move to new positions and cost function is calculated. Further movements are decided by decreasing direction of cost function and finally with highest fitness.

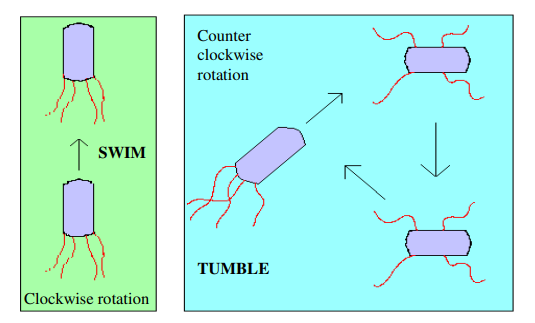
Four main processes of the BFO are Chemotaxis, Swarming, Reproduction, Elimination & Dispersal.

**Chemotaxis** is considered to be a great amount of swimming and flipping or tumbling motions of the bacteria. When a bacterium meets a better environment (nutrients and noxious free), it continues swimming in the same direction. When it encounters an unfavorable environment it changes its direction, known as tumbling.

**Swarming** defined the numerical relationship for the bacterial behavior based on attraction and repulsion.

During the **reproduction** phase fitness of all the bacteria is calculated and sorted.

In the final phase each bacterium is **dispersed** with a probability while the total remains the same. Once a bacterium is **eliminated**, it will be randomly dispersed to a new location.

Fig. 4 : Swim and Tumble of a bacterium [32].

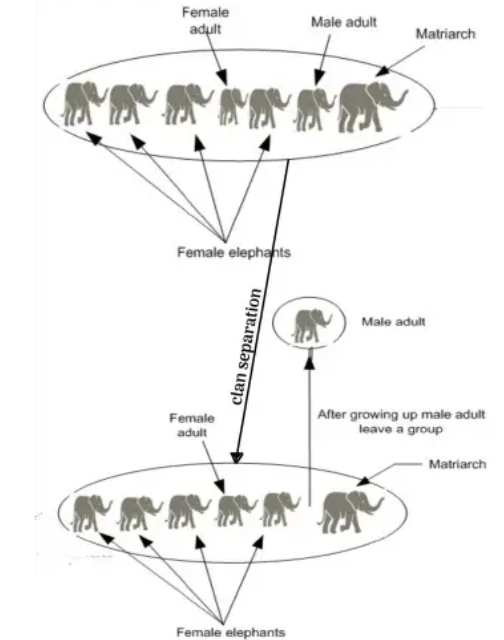
This BFO and its enhanced variants are used in airline scheduling, logistic delivery and vehicle routing planning (Gan & Xiao, 2019) and also predicting stock indices [26].

*D. Elephant Herding Optimization Algorithm for Cancer Prediction.*

Cancer is a kind of uncontrolled growth of abnormal body cells. Early prognosis is the way of better treatment. According to Nayak, Das, Bhanja & Senapati (2020), cancer is the second leading cause of human deaths in 2018. Predictions are possible by evaluating cancer data sets using these types of algorithms [27], [28].

Elephant Herding Optimization algorithm (EHO) developed several types of cancer prediction methods. This algorithm is inspired by the behavior of elephant groups. Elephants live as groups known as clans and the group leader is a Matriarch, the best female member. The grown up male elephants leave the group and live alone or live with male clan. Grown male elephants make decisions individually and move aimlessly. Based on this behavior two agents, the updating and separating operators develop to simulate this behavior.

Fig. 5 : Elephant Clan and Separation [31].



EHO is also used to optimize wireless sensor network localization problems, QoS aware web service composition, benchmark problems and also traveling salesman problem.

Elephant clan optimization algorithm (ECO) developed as a different version of EHO. Matriarch leads the group and moves based on the best experience of the population.

**III. METHODOLOGY OF FINDING THE FACTS**

One main part of this study is research methodology to collect information. This paragraph briefs the chosen methodology and the reasons behind choosing it. It includes design, methodology, data collection, justification for choice and ethical issues, etc. According to Chua (2006), the outcome of the research is determined by method(s) and the research design.

This study explores a phenomenon, generates ideas, investigating past, current and emerging concerns. There are a variety of methods of data collection. This paper is written based on research papers, journal articles, book chapters, case studies, videos and images available on the internet. Different findings have been combined to form more detailed information. Most of the data sources acquired from academic works and professional organizations who conduct researches, surveys and studies on AI, AI based applications, its subfields and adjoins.

Literature review is to discuss different theories, approaches, arguments, findings and analysis relevant to the study. Surface based search (Google Scholar) and deep web search (Tor) used to find information mainly for the last five years. Different keywords, combinations of phrases and references of findings are also used. Citation done for rewarding the authors as an ethical concern as well as a best practice.

**IV. RESULTS**

Most of the results that we see due to the usage of GA, FL, SI applications are discussed above including the changes these applications have brought into the technological and social arena that encompassed subtle simulation of things and living beings that we as a normal human forget to pay attention towards. It's really astonishing and scintillating to write up research based on things that are improving mankind with the learnings from nature under the subfields of AI. Some findings are still in the maturing stage and some are going around the world we live in to make drastic changes spanning years.

**V. DISCUSSION**

Access to the most valid and reliable data sources is difficult or costly. Therefore quality information gathering is likely to be difficult. There are some conflicts in between some research works too. Some researchers have defined some concepts as a whole unit while others as separate subunits. Based on initial algorithms, variants are developed to enhance the capabilities and accuracy of analysing processes. Also hybrid modules are better than individual modules. Behavior of most natural objects (Crows, Cockroach, Spiders, Serpentes, Termite etc.) could be transformed as problem solving and optimizing techniques based on SI. The complexity of insect colonies, beautiful sights of a large swarm of birds or elephant herd surprised with the simplicity of hidden rules. GA has been applied in solving optimization problems related to many areas of studies, it has also been used as a component in developing hybrid models based on either GA with Fuzzy logic or SI, as these three concepts can be integrated together in model development. Fuzzy Logic is enhancing how we work on the traditional old methods with new ideologies and aspects that might help mankind grow efficiently and effectively. Hence all three of these AI subfields are attributing to the betterment of the society and community inducing excellent technological advancement.

**VI. CONCLUSION**

Genetic Algorithm, Fuzzy Logic and Swarm Intelligence based problem solving and optimization algorithm based applications are playing a major role in data analytics to find better answers for the real world complex and complicated problems. This paper has reviewed related works that apply Genetic Algorithm, Fuzzy Logic and Swarm Intelligence in data analytics. The fundamentals and the findings are briefly summarized with several examples. The above three subfields of AI have been widely used in the field of data analytics in the past decades. Some findings have matured with the time and some are still enhancing for more accuracy. With the development of AI and Data Analytics, these three subfields have great opportunities in different aspects of development in the future world. Although it faces challenges, it serves mankind for the betterment of their lives in many ways.

**REFERENCES**

1. G. Garai, “Application of genetic algorithms in numerous scientific fields,” *Genetic Algorithms*, 2022.
2. J. Zhang, “3rd International Workshop on Materials Engineering and Computer Sciences (IWMECS 2018),” in *Advances in Computer Science Research*, 2018.
3. H. Wang, “Research on the application of genetic algorithms in Physical Education,” *Journal of Mathematics*, vol. 2022, pp. 1–8, 2022.
4. Y. Sun, B. Xue, M. Zhang, G. G. Yen, and J. Lv, “Automatically designing CNN architectures using the genetic algorithm for Image Classification,” *IEEE Transactions on Cybernetics*, vol. 50, no. 9, pp. 3840–3854, 2020.
5. F. l Dharma, Shabrina, A. Noviana, M. Tahir, N. H. Hendrastuti, and Wahyono, “Prediction of Indonesian Inflation Rate Using Regression Model Based on Genetic Algorithms,” *JOIN (Jurnal Online Informatika)* , vol. 5, no. 1, pp. 45–52, Jun. 2020.
6. J.E. Naranjo, C. González, R. García, and T.D. Pedro,’’*Using Fuzzy Logic in Automated Vehicle Control’’*, vol. 22, no. 1, pp. 36-45, 2017.
7. Poonam Gupta. Application of Fuzzy Logic in Daily Life. International Journal of Advanced Research in Computer Science. 2017; 8(5):1795-1800.
8. M.Kochev and M.Ivanova, “*Modernization of a Cooling Water Control System for the Production of Plastic Pipes Using Fuzzy Logic*”, 2021.
9. Z. Gao et al., “A Stable Self-Tuning Fuzzy Logic Control System for Industrial Temperature Regulation”, IEEE Industry Applications Conference. Thirty-Fifth IAS Annual Meeting and World Conference on Industrial Applications of Electrical Energy, vol. 2, Rome, Italy, 2000, 1232-1240, doi: 10.1109/IAS.2000.881990.
10. Zadeh, L.A. The role of fuzzy logic in modeling, identification and control. In Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A Zadeh; World Scientific: Singapore, 1996; pp. 783–795.
11. H.J. Zimmermann, (1985), Fuzzy Set Theory and Its Applications,

Kluwer.

1. [W. Eka Sari](https://aip.scitation.org/author/Sari%2C+Wahyuni+Eka), [O. Wahyunggoro](https://aip.scitation.org/author/Wahyunggoro%2C+Oyas) and [S.Fauziati](https://aip.scitation.org/author/Fauziati%2C+Silmi), “*A comparative study on fuzzy MamdaniSugeno-Tsukamoto for the childhood tuberculosis diagnosis*”, 2016.
2. R.R. Yager, (1982), Fuzzy Sets and Possibility Theory, Pergamon

Press, Oxford.

# Maramin and Mushthofa, “*Fuzzy Logic Systems and Applications in Agro-industrial Engineering and Technology*”, 2018.

# J.Kaur and B.S. Khehra, “*Fuzzy Logic and Hybrid based Approaches for the Risk of Heart Disease Detection: State-of-the-Art Review*”, 2022; 103(2): 681–697.

# Zadeh, L.A. (1965). Fuzzy Sets. Information and Control, 8: 338–353.

1. E. M.P.Saka, "Analysis of Swarm Intelligence–Based Algorithms for Constrained Optimization," *Swarm Intelligence and Bio-Inspired Computation,* pp. 25-48, 2013.
2. D. Z. Y. A. H. W. C. A. F. C. C. Ailiang Qi, "Directional mutation and crossover boosted ant colony optimization with application to COVID-19 X-ray image segmentation," *Computers in Biology and Medicine,* vol. 148, 2022.
3. A. D. P. K. P. R. P. Mamata Rath, "Swarm intelligence as a solution for technological problems associated with Internet of Things," in *Swarm Intelligence for Resource Management in Internet of Things*, 2020, pp. 21-45.
4. M. H. M. Mohamed K. Hussein, "Efficient Task Offloading for IoT-Based Applications in Fog Computing Using Ant Colony Optimization," 2020.
5. D.-d. X. L. Z. T. Juan Li, "Using Cuckoo Search Algorithm with Q-Learning and Genetic Operation to Solve the Problem of Logistics Distribution Center Location," Mathematics 2020, vol. 8, no. 149, 2020.
6. E. C. R.-O. Xin-She Yang, "A Cuckoo Search Algorithm for Multimodal Optimization," *Computational Intelligence and Metaheuristic Algorithms with Applications,* 2014.
7. M. A. Al-Abaji, "Cuckoo Search Algorithm: Review and its Application," *Tikrit Journal of Pure Science,* vol. 26 (2), pp. 137-141, 2021.
8. K. C. J. Srihari Gude, "Parameter extraction of photovoltaic cell using an improved cuckoo search optimization," *Solar Energy,* vol. 204, pp. 280-293, 2020.
9. L. W. D. P. Huang Chen, "Bacterial Foraging Optimization Based on Self-Adaptive Chemotaxis Strategy," *Computational Intelligence and Neuroscience,* 2020.
10. X. G. a. B. Xiao, "Improved Bacterial Foraging Optimization Algorithm with Comprehensive Swarm Learning Strategies," *Advances in Swarm Intelligence,* p. 325–334, 2020.
11. S. D. B. R. S. Monalisa Nayak, "Elephant herding optimization technique based neural network for cancer prediction," *Informatics in Medicine Unlocked,* vol. 21, 2020.
12. A. A. E. H. h. I. Ahmed Talaat, "Elephant herd optimization with neural networks: A case study on acute Lymphoblastic Leukemia diagnosis," 2017.
13. L. Q. S. S. C. Z. X. S. Jian Yang, "Swarm Intelligence in Data Science: Applications, Opportunities and Challenges," p. 3–14, 2020.
14. B. B. F. David Martens, "swarm intelligence for data mining," pp. 1-42, 2010.
15. R. R. A. W. S. Saman M AlMufti, "Review on Elephant Herding Optimization Algorithm Performance in Solving Optimization Problems," *International Journal of Engineering & Technology,* vol. 7, no. 4, 2018.
16. A. B. D. A. Swagatam Das, "Bacterial Foraging Optimization Algorithm: Theoretical Foundations, Analysis, and Applications," 2009.