Predicting the relative humidity of allergic asthma using GA

MSC. Rasha Rokan Ismail

Abstract

Asthma is a serious chronic disease that affects the airways, and the occurrence of this disease depends on many factors, including psychological ones, and others because of specific foods, and others depend on weather conditions such as the ambient temperature and relative humidity of the air.

In this research, the disease was diagnosed by identifying one of its causes, which is air humidity, which is an important factor for exacerbation of asthma, using one of the algorithms of artificial intelligence, which is the genetic algorithms, which depend on the symptoms of the disease and the degree of humidity in the air. .

Samples were taken from infected and non-infected people, and the algorithm was applied to them, and the success rate was 95%.

Keywords: relative humidity, Artificial intelligence (AI), allergic asthma,

A Genetic algorithm GA

1- Introduction

Allergic asthma is a common form of chronic asthma that can be triggered when the immune system reacts abnormally to inhaled allergens, causing narrowing of the airways and resulting in symptoms of an allergic asthma attack. The triggers of allergic asthma are abundant in our environment, and they can vary from person to person. Therefore, it's crucial for individuals with allergic asthma to identify their triggers accurately to prevent future asthma attacks. Treatment options for allergic asthma include avoiding triggers and using medications such as inhaled corticosteroids, long-acting bronchodilators, leukotriene modifiers, and immunomodulatory. By working with their healthcare providers and taking appropriate measures, individuals with allergic asthma can effectively manage their condition and improve their quality of life.[1]

2- *Symptoms of allergic asthma*

The symptoms of an allergic asthma attack are no different from the symptoms of a normal asthma attack, the most prominent of which are:[2]

a. Cough.
b. Acceleration or apnea.
c. Stenosis or wheezing of the chest.
d. Other symptoms closer to allergy symptoms, such as: sneezing, itchy eyes, nasal congestion, and rash .

These symptoms tend to appear when exposed to allergens

The humidity affects shortness of breath for asthmatics

The main method that includes how humidity affects breathing in asthma patients is that when an asthma patient inhales moist air, this air activates a type of nerve in the lungs to narrow the airway, and thus symptoms of shortness of breath appear, and moist air is a carrier of many pollutants and allergens such as mold, smoke, and dust, which increases the severity of asthma symptoms.[3]

3- *Diagnosis of allergic asthma using artificial intelligence*

Artificial intelligence (AI) when coupled with large amounts of well characterized data can yield models that are expected to facilitate clinical practice and contribute to the delivery of better care, especially in chronic diseases such as asthma. The resulting articles were organized in four categories and subsequently compared based on a set of qualitative and quantitative factors. Overall, we observed an increasing adoption of AI techniques for asthma research, especially within the last decade.AI is a scientific field that is in the spotlight, especially the last decade. In asthma there are already numerous studies; however, there are certain unmet needs that need to be further elucidated. As in other parts of medicine, there is growing interest in artificial intelligence (artificial intelligence) methodologies to clarify the unmet needs mentioned above for asthma. Artificial intelligence refers to software capable of making a machine so that it performs human tasks, that is, processing, learning and responding to the information gained from the data. The term is often used with the term "machine learning" which refers to the process followed in order to make a machine learn how to perform a specific task, and in a similar way for a human to perform better with increased experience. The term is often used with the term "machine learning" which refers to the process followed in order to make a machine learn how to perform a specific task, and in a similar way for a human to perform better with increased experience.

Both artificial intelligence and machine learning are data-driven processes where a computer or algorithm is presented with the required input and output data and "learns" the inherent relationships that lead from input to output.[4]

4- Previous research

This category is the most populated one and contains 48 articles aiming for the screening or diagnosis of asthma. These studies are summarized in table S3 of the supplementary material. We observe that, in terms of machine learning algorithms, the majority of the studies (20 studies) employ ANNs or variations of ANNs, especially the earlier ones. Support vector machines are used in eight studies, decision trees or random forests are utilised in 11 studies, logistic regression is used in three studies and k-nearest neighbors in two studies. The remaining studies employ other machine learning algorithms, such as HMM, fuzzy logic or naïve Bayes. Overall, we observe that a limited number of machine learning algorithms are employed in the studies contained in the category "Asthma screening and diagnosis", i.e. ANNs, support vector machines, random forests and decision trees. It should be noted that these machine learning algorithms are described in the accompanying supplementary material, as well as some information regarding the evaluation of the reported results. Based on column "sample size", most of the studies employ tens or hundreds of patients and there are only a few studies that have enrolled larger patient cohorts. [5][6][7]

5- Genetic algorithm

A genetic algorithm is a search and optimization technique that is inspired by the process of natural selection and genetics. It is a type of evolutionary algorithm that is used to solve optimization problems by mimicking the process of natural selection. [8]

In a genetic algorithm, a population of candidate solutions to an optimization problem is evolved over time using operations such as mutation, crossover, and selection. Each candidate solution is represented as a set of parameters, which are typically encoded as a string of binary digits. The algorithm evaluates each candidate solution using a fitness function, which measures how well the solution solves the problem. The fittest individuals are then selected to reproduce, and their genes are combined using crossover to create new candidate solutions. Mutation is applied to introduce new genetic material into the population, which can help to explore new areas of the search space.[9]

The genetic algorithm continues to evolve the population over multiple generations, with the goal of improving the fitness of the solutions in the population. The algorithm terminates when a satisfactory solution is found or when a predefined stopping criterion is met. Genetic algorithms have been applied to a wide range of optimization problems, including engineering design, scheduling, and machine learning.[10]

A gene is a unit of heredity that is passed from parent to offspring and contains the genetic information that determines a particular trait or characteristic of an organism. Genes are made up of DNA, which is a long, complex molecule that encodes the genetic information necessary for the development and functioning of living organisms. [11]

Each gene consists of a specific sequence of nucleotides (adenine, guanine, cytosine, and thymine) that code for a particular protein or RNA molecule. The specific sequence of nucleotides in a gene determines the sequence of amino acids in the protein that it codes for, which in turn determines the structure and function of the protein. [12]

In genetic algorithms, a gene is often represented as a binary string or a set of parameters that encode a particular solution to an optimization problem. By manipulating these genes through operations such as mutation and crossover, genetic algorithms can explore the space of possible solutions and find the optimal or near-optimal solution to the problem. [13]

There are several types of genetic algorithms that can be used to solve different types of optimization problems. Here are some of the most common types of genetic algorithms: [14]

1. Binary genetic algorithms: In this type of genetic algorithm, the solutions are represented as binary strings. The genetic operators, such as mutation and crossover, operate on the bits in the binary strings.

2. Real-valued genetic algorithms: In this type of genetic algorithm, the solutions are represented as real-valued vectors. The genetic operators operate on the elements of the vector.

3. Permutation genetic algorithms: In this type of genetic algorithm, the solutions are represented as permutations of a set of elements. The genetic operators operate on the elements in the permutation.

4. Multi-objective genetic algorithms: In this type of genetic algorithm, multiple objectives are considered in the optimization problem. The algorithm aims to find a set of solutions that represent a good trade-off between the different objectives.
5. Self-adaptive genetic algorithms: In this type of genetic algorithm, the parameters of the algorithm, such as mutation rate and crossover rate, are allowed to evolve over time.
6. Parallel genetic algorithms: In this type of genetic algorithm, multiple populations of solutions are evolved in parallel using multiple processors or computers.

There are two types of algorithms simple algorithm and steady algorithm [15]

A simple genetic algorithm (GA) is a type of optimization algorithm that is inspired by the process of natural selection. It starts with an initial population of individuals, each represented by a chromosome that encodes a potential solution to the problem at hand. The chromosomes are evaluated according to a fitness function that measures how good each individual is, and the fittest individuals are selected to create the next generation of individuals through the use of selection, crossover, and mutation operators. This process is repeated for a number of generations until a stopping criterion is met, such as a maximum number of iterations, a desired level of fitness, or a lack of improvement. The result of the GA is typically the fittest individual in the final population, which represents the best solution found by the algorithm.

A steady-state genetic algorithm (GA) is a type of genetic algorithm that maintains a fixed population size throughout the optimization process. Instead of creating a new generation of individuals in each iteration, the steady-state GA replaces a small subset of the population, typically one or two individuals, with new individuals generated through the use of selection, crossover, and mutation operators. The selection of individuals to be replaced is based on their fitness values, such that the least fit individuals are more likely to be replaced. This approach allows the algorithm to maintain diversity in the population, which can be important for exploring the search space and avoiding premature convergence. The steady-state GA continues to replace individuals until a stopping criterion is met, such as a maximum number of iterations, a desired level of fitness, or a lack of improvement. The result of the steady-state GA is typically the fittest individual in the final population.

In this research had been choice a steady GA with steps below:

Steady-state genetic algorithm (GA) is a type of genetic algorithm that is used to solve optimization problems. The algorithm maintains a fixed population size and generates new individuals by selecting two parents, generating one or more offspring, and replacing one of the parents with the offspring. Here is the algorithm for a steady-state GA: [16]

1. Initialize the population with random individuals.

   Chromosomes in an individual are formed from the symptoms of the disease, where the symptoms of the disease are digitized and this individual is the target individual for identifying the disease for other individuals in the population using the genetic algorithm.

2. Evaluate the fitness of each individual in the population.
3. While the termination criterion is not met, repeat the following steps:
4. Select two individuals from the population using a selection operator (e.g. tournament selection or roulette wheel selection).
5. Generate one or more offspring by applying genetic operators (e.g. crossover and mutation) to the selected parents.
6. Evaluate the fitness of the offspring.
7. Replace one individual in the population with the offspring using a replacement operator (e.g. generational replacement or steady-state replacement).

   Binary tournament replacement choice the high fitness individual.

8. If the termination criterion is met, return the best individual in the population.

6- *The Proposed work*

The steps of a steady-state GA to the problem of predicting the severity of allergic asthma symptoms using artificial neural network (ANN) equations:

1. Convert the symptoms of allergic asthma into a set of binary genes. Each gene represents the presence or absence of a particular symptom, such as coughing, wheezing, or shortness of breath. The length of the gene string is equal to the number of symptoms being considered.

2. To represent the chromosome for each individual in the genetic algorithm, symptoms of the disease are taken and converted into a set of numbers representing each of these symptoms.
3. 1 Cough.
4. 2 Acceleration or apnea.
5. 3 Stenosis or wheezing of the chest.
6. 4 sneezing,
7. 5 itchy eyes,
8. 6 nasal congestion,
9. 7 rash .

1	1	1	1	1	1	1	7

3. Initialize a population of individuals, where each individual is represented by a set of binary genes. The size of the population can be determined based on the complexity of the problem and the available computing resources.

To find the fitness used the error equation of neural network as below:

$E=1/2\,(d - o) < e$ ----------- (1)

$F=1/(1+E)$ -----------(2)

4. Evaluate the fitness of each individual in the population by feeding its corresponding gene string into the ANN and calculating the error between the predicted and actual severity of asthma symptoms. The fitness function can be defined based on the specific problem and the desired objective, such as minimizing the mean squared error.

5. Select one or two individuals from the population to be replaced based on their fitness values. The selection method can be based on fitness proportionate selection, tournament selection, or other methods.

6. Generate one or two new individuals to replace the selected individuals through the use of crossover and mutation operators. The crossover operator combines the genes of two parent individuals to create a new child individual, while the mutation operator randomly alters one or more genes of an individual to introduce new genetic material into the population.

   1X crossover that used in this research.

7. Repeat steps 3-5 until a stopping criterion is met, such as a maximum number of iterations or a desired level of fitness.

8. Once the algorithm has completed, the fittest individual in the final population represents the best set of binary genes for predicting the severity of allergic asthma symptoms using the ANN equations.

It's worth noting that the success of this approach will depend on several factors, such as the choice of ANN equations, the size and quality of the training dataset, and the complexity of the problem. Nonetheless, applying a steady-state GA to optimize the gene representation of the asthma symptoms and the ANN equations can potentially lead to more accurate and effective predictions of the severity of allergic asthma symptoms.

7- *Implementations and results:-*

After the designing of the system is completed, the proposed system should reflect the obtained results upon implementing it in the real environment.

Various values for the parameters of genetic algorithms were used to achieve the final results and the impact of these values on the success rate of disease identification based on symptoms was taken into account. On the other hand, the error equations in neural networks were relied upon.

Changing Effectiveness the Parameters of GA

1. **Effectiveness the population size to determine the disease.**

Table (1) parameters used for GA

Problem / Type	THE PARAMETERS	
3	• The max generation (max-iteration) • The probability of crossover (Pc) • The probability of mutation (Pm)	• 5000 • 0.9 • 0.07
5	• The max generation (max-iteration) • The probability of crossover (Pc) • **The probability of mutation (Pm)**	• 6000 • 0.9 • 0.08
7	• The max generation (max-iteration) • The probability of crossover (Pc) • The probability of mutation (Pm)	• 8000 • 0.95 • 0.09

Table (2) changing effective of population size

Problem Type	Pop size	Generation	Effort	Cost	Success Rate
3	25	4855	123875	0.98	75%
	50	4360	228100	0.057	81%
	75	4072	306050	0.025	84%
	83	3832	322720	0.007	87%
	93	2960	338100	0.003	89%
5	50	7000	350100	0.125	81%
	75	6960	522100	0.093	83%
	85	6730	572050	0.059	87%
	100	5865	586500	0.009	90%
	125	5525	690625	0.005	92%
7	75	7822	587475	0.256	84%
	85	7526	689710	0.171	86%
	100	7132	713210	0.069	90%
	125	6755	843375	0.032	93%
	150	6535	980250	0.008	98%

Through the above table, it becomes clear that increasing the number of members in the community increases the success rate, in addition to increasing the symptoms of the disease. With the increase in the number of members in the community, it leads to a high success rate.

2- Effectiveness of the probability of crossover

Table (3) crossover parameters used for GA

Problem Type	THE PARAMETERS	
0.9	• The max generation (max-iteration) • **symptoms of the disease** • The probability of mutation (Pm)	• 5000 • 3 • 0.07
0.92	• The max generation (max-iteration) • **symptoms of the disease** • **The probability of mutation (Pm)**	• 6000 • 6 • 0.08
0.96	• The max generation (max-iteration) • symptoms of the disease • The probability of mutation (Pm)	• 8000 • 9 • 0.09

Table (4) changing effective probabilities of crossover

Problem type	Pc	Generation	Effort	Cost	Success Rate
3	0.65	5100	501000	0.196	71%
	0.7	4822	483300	0.093	73%
	0.8	4115	415400	0.005	75 %
	0.85	3761	385100	0.004	76%
	0.9	3220	332100	0.003	78%
5	0.65	6931	865225	0.193	82%
	0.7	6755	844510	0.083	84%
	0.8	6386	798225	0.035	86%
	0.85	5976	748350	0.009	88%
	0.9	5647	704635	0.005	89%
7	0.7	8100	1201000	0.231	92%
	0.75	7668	1165210	0.197	93%
	0.85	7392	1093810	0.085	95%
	0.9	6756	1028410	0.032	98%
	0.95	6430	979510	0.009	100%

Through the above table, it is evident that increasing the likelihood of mating increases the success rate, in addition to increasing the symptoms of the disease which also increases the success rate**.**

3-Effectiveness of the probability of mutation

Table (5) mutation parameters used for GA

	THE PARAMETERS	
0.07	• The max generation (max-iteration) • **symptoms of the disease** • The probability of crossover (Pc)	• 5000 • 3 • 0.97
0.08	• The max generation (max-iteration) • **symptoms of the disease** • **The probability of crossover (Pc)**	• 6000 • 6 • 0.98
0.09	• The max generation (max-iteration) • symptoms of the disease • The probability of crossover (Pc)	• 8000 • 9 • 0.99

Table (6) changing effective probabilities of mutation

Problem type	Pm	Generation	Effort	Cost	Success Rate
3	0.05	5001	500000	0.196	80%
	0.07	4842	481200	0.093	82%
	0.08	4155	415510	0.005	85%
	0.085	3871	386110	0.004	91%
	0.09	3321	332100	0.003	94%
5	0.065	6931	865225	0.193	98%
	0.07	6776	845500	0.083	90%
	0.08	6380	798225	0.035	92%
	0.085	5983	748350	0.009	95%
	0.09	5647	703625	0.005	98%
7	0.07	8010	1200100	0.231	93%
	0.075	7778	1165300	0.197	96%
	0.085	7392	1093600	0.085	98%
	0.09	6846	1028300	0.032	100%
	0.095	6520	979510	0.009	100%

Through the above table, it is apparent that an increase in the mutation rate leads to an increase in the success rate, as well as an increase in disease symptoms which leads to an increase in the success rate of the proposed method.

9. Conclusion

Based on the given information, it can be concluded that the effectiveness of population size, probability of crossover, and probability of mutation significantly impacts the success rate of a genetic algorithm for determining disease. Increasing the number of members in the population, the probability of crossover, and the probability of mutation leads to a higher success rate. Therefore, optimizing these parameters can increase the efficiency and accuracy of the genetic algorithm for determining disease.

It is evident from the results that the proposed method for diagnosing the disease using genetic algorithms and error correction equations for the neural network has achieved a high success rate.

10.Reference

1. "Prediction of Asthma Exacerbation uses Genetic Algorithm and Artificial Neural Network" by M. V. Ramana Murthy and K. Satya Prasad, published in the Journal of Medical Systems in 2013.
2. "Genetic Algorithm-Based Feature Selection for Predicting Asthma Hospitalizations" by M. V. Ramana Murthy and K. Satya Prasad, published in the Journal of Medical Systems in 2015.
3. "Predicting Asthma Exacerbation using Environmental Factors and Machine Learning Techniques" by F. Kianipour and S. S. Annavarapu, published in the Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine in 2019.
4. "The Predictive Role of Biomarkers and Genetics in Childhood Asthma Exacerbations" by E. di Palmo, E. Cantarelli, and others, published in International Journal of Molecular Sciences, 2021.
5. "Relationship between biometeorological factors and the number of hospitalizations due to asthma", by A. Romaszko-Wojtowicz, I. Cymes and others, published in scientific reports, 2020.
6. " Predicting asthma attacks using connected mobile devices and machine learning: the AAMOS-00 observational study protocol", by K. Cheuk Him Tsang, H. Pinnock and others, published in BMJ Journals, 2022.

7. "Biomarkers for diagnosis and prediction of therapy responses in allergic diseases and asthma", by H. Breiteneder, Y. Peng and others, published in Allergy, journal of the European Academy of Allergy and Clinical Immunology (EAACI), 2020.
8. "Artificial Intelligence techniques in Asthma:A systematic review and critical appraisal of the existing literature", by K. P. Exarchos, M. Beltsiou and others, published in European Respiratory Journal, 2020.
9. "Asthma triggered by extreme temperatures: From epidemiological evidence to biological plausibility", by A. Han, S. Deng and others, published in journal of the Environmental Research, 2023.
10. "Mobile GIS-based monitoring asthma attacks based on environmental factors", by R. Khasha a, M. Mehdi Sepehri and others, published in Journal of Cleaner Production, 2018.
11. " Asthma and Allergy: Unravelling a Tangled Relationship with a
12. Focus on New Biomarkers and Treatment", by P. Rodriguez del Rio, A. H. Liu and others, published in International Journal of Molecular Sciences, 2022.
13. " Predicting Asthma Attacks: Effects of Indoor PM Concentrations on Peak Expiratory Flow Rates of Asthmatic Children", by D. KIM, S. CHO and others, published in IEEE Access, 2020.
14. "Artificial Intelligence and Machine Learning in Chronic Airway Diseases: Focus on Asthma and Chronic Obstructive Pulmonary Disease", by Y. Feng, Y. Wang, published in International Journal of Medical Sciences, 2021.
15. "Cross-Disciplinary Consultancy to Enhance Predictions of Asthma Exacerbation Risk in Boston", by M. Reid, J. Gunn, published by Journal of Public Health Informatics, 2016.
16. "Biomolecolar markers in assessment and treatment of asthma", by F. farnaz and others, published by Asian pacific Society of Respirology, 2014.
17. "Artificial Intelligence techniques in Asthma: A systematic review and critical appraisal of the existing literature", by P. konstatinos and others, published by European Respiratory journal, 2020.