

# An Intelligent Prenatal Screening System for the Prediction of Trisomy-21

**Amir Jamshidnezhad**

Ahvaz Jundishapur University of Medical Sciences: Ahvaz Jondishapur University of Medical Sciences

**Seyed Mohsen Hosseini** (✉ [mohsen77749@gmail.com](mailto:mohsen77749@gmail.com))

Ahvaz Jundishapur University of Medical Sciences: Ahvaz Jondishapur University of Medical Sciences

<https://orcid.org/0000-0002-4570-789X>

**Javad Mohammadi-Asl**

Ahvaz Jundishapur University of Medical Sciences: Ahvaz Jondishapur University of Medical Sciences

**Mandana Mahmoudi**

Ahvaz Jundishapur University of Medical Sciences: Ahvaz Jondishapur University of Medical Sciences

---

## Research note

**Keywords:** Trisomy-21, Prenatal Screening System, artificial neural network (ANN), genetic algorithm (GA)

**Posted Date:** October 23rd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-94963/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Informatics in Medicine Unlocked on January 1st, 2021. See the published version at <https://doi.org/10.1016/j.imu.2021.100625>.

# Abstract

**Objective:** Early and accurate diagnosis of genetic diseases such as Down syndrome can lead to correct action and prevent irreversible events. Therefore, the use of accurate and low-risk diagnostic methods to detect these diseases in the fetal period is important. In this study, a combined artificial neural network (ANN) and genetic algorithm (GA) was used to predict Down syndrome through first trimester screening test. In order to examine the proposed model, sample data were collected on 381 pregnant women who referred to a private screening reference laboratory for the first trimester screening test and NT ultrasound between 11 and 13 weeks of gestation. The proposed model in this study was a feedforward neural which its structure and input parameters were determined by (GA).

**Results:** The average of 10 times experiments showed that the developed model can accurately identify cases of Down syndrome with specificity of 99.72% and sensitivity of 90.91%, and a mean square error (MSE) of 0.61%. The results of this study showed that the use of GA in optimizing the structure of the neural network technique can increase the accuracy in diagnosing Down syndrome through the information of first trimester screening tests.

## 1- Introduction

Down syndrome is caused by the presence of a trisomy on chromosome 21 of humans and with prevalence rate of approximately 1 in every 800 births, which has the largest share of genetic abnormalities(1). Children with this disease often have varying degrees of physical-mental problems such as mental retardation, communication skills problems, congenital heart disease, and thyroid disease(2, 3).

Nowadays, it is possible to diagnose Down syndrome in the early stages of pregnancy by performing various tests. Routine tests that are usually done to diagnose such problems are done at first trimester screening tests. Although this test is known as a non-invasive, low-risk procedure and can only be performed using a maternal blood test and an ultrasound of the fetus, it has low accuracy (at best with 90% accuracy) and is only able to assess the likelihood of disease occurrence(4, 5).

Another non-invasive method is the measurement of free DNA in the mother's plasma (Cell free DNA OR non-invasive prenatal test (NIPT)), which is done by examining the DNA released from placental cells, which can be extracted from the mother's blood. The advantages of this method are its low risk and high accuracy, but this test requires special equipment and its high cost, delay in answering the test and the difficulty of isolating fetal DNA from the mother's blood are some of its disadvantages (6, 7).

Amniocentesis and Chorionic villus sampling (CVS) are two invasive procedures that are performed using amniotic fluid and fetal villi sampling, respectively. These two methods, despite their high accuracy (98–99%), increase the risk of miscarriage (1–2%) or harm to the mother, and also because of its high cost and fear of postpartum consequences for the baby, performing this test is not considered reasonable for

all pregnant mothers and is only recommended for those who are at risk based on screening test results(5, 8, 9).

ANNs, as one of the branches of machine learning, are a collection of intelligent computer calculations that have been widely used in the principles of classification and prediction for the last 25 years, especially in the field of medicine. These networks are in fact mathematical algorithms that are trained by data and their knowledge is obtained by the relationships between data. These networks, if properly trained, can act like the human brain and, in some cases, identify complex nonlinear relationships between dependent and independent variables that are not detectable by the human brain (8, 10).

GA is an evolutionary algorithm that is widely used in the optimization of machine learning systems and can provide an almost optimal solution with random search technique (11, 12). This algorithm works by modeling chromosome recombination capabilities in a process similar to what occurs in the meiotic stage of cell proliferation. An initial set of solutions, called a population, is created randomly to solve a problem. Each solution is considered as a chromosome. Chromosomes create a new generation each time the recombination and cycle is repeated. Chromosomes whose phenotypes perform best in the neural network are considered to be the parents of the next generation. After several generations, the population converges to the best chromosome, which can provide the desired architecture of the neural network(13, 14).

In this study, an intelligent model was achieved by using neural network technique and GA optimization capacities, which was obtained by examining various factors with higher accuracy and less risk in diagnosing screening tests.

## **2- Methodology**

### **2-1 Dataset Collection**

In order to create the required data, information about 381 pregnant women were collected, who referred to a private screening laboratory in Iran, Ahvaz, for the first trimester screening test and NT ultrasound between 11 and 13 weeks of pregnancy between February 20, 2018 and February 20, 2019.

### **2-2 Hybrid ANN-GA**

Usually a neural network consists of an input layer, an output layer, and one or more hidden (middle) layers. The input layer contains raw information that is given to the neural network for processing. The initial values in the input layer are received by the middle layer. The middle layer changes the received values by a series of weights and sends the new values to the next layer. Finally, the output layer processes the received information and produces the output(15). The neural-genetic network was designed by MATLAB software version 2015a. The ANN used in this study is the Feedforward back-propagation (BP) algorithm. Feedforward BP networks are networks that include an output layer, an input

layer, and at least one layer of processing neurons (middle layer) in which one-way information flows from the input layer neurons to the output layer neurons. In other words, the output of each layer becomes the input of the next layer (16, 17). The ANN used in this study has an input layer, two hidden layers and an output layer with a neuron. The number of input layer neuron (s) is from 1 to 7 neurons and is equivalent to the study factors. the number of intermediate layer neuron (s) is from 1 to 20 neurons determined by the GA. The transfer functions of each layer are also selected by the GA. For the training process, 12 types of training functions and 2 types of learning functions are used to select by the GA. The output of the ANN indicates healthy cases from the cases of Trisomy 21 (T21).

## **2-3 Genetic Algorithm**

The parameters that organize the ANN structure and the input factors were defined by haploid chromosomes containing 14 genes. The first two genes on each chromosome were related to the number of neurons in the first and second hidden layers with the values of 1 to 20. The third to fifth genes also determined the transfer function in the first and second layers as well as the output layer, which could accept values from 1 to 3. The sixth gene was related to the selection of the training function, which could contain values from 1 to 12.

The seventh gene also determined the learning function, which accepted values 1 and 2. The eighth to fourteenth genes, each indicating the presence or absence of one of the seven ANN input factors. Figure 1 shows the structure of the GA and its operation to optimize the ANN prediction performance in the hybrid model.

In order to create the initial population, chromosomes were created randomly. After evaluating the expression of each chromosome, the superior chromosomes were selected for the recombination process and the next generation were created and paired in pairs.

Chromosomes were considered as higher-score which their decoding resulted in the networks with lower MSE and greater accuracy. With a certain number of iterations of the generation cycle, the ANN with the lowest MSE was finally determined as the final optimal network.

## **2-4 Data analysis**

The performance of the generated model was evaluated by three indicators; sensitivity, specificity and accuracy, which are calculated using confusion matrix data(18-20).

## **3- Results**

### **3 -1 Study Data**

Out of 381 patients, 73 were excluded from the study due to lack of information, and finally 308 cases were included in the study, of which 286 were healthy and the remaining 22 had Down syndrome.

## 3-2 Designed model

In order to train the designed model, 80% of the data were selected, randomly and the remaining 20% of the data were allocated for the model testing. The population of each generation was 120 chromosomes in the GA. The probability ratios of one-point intersection and two-point intersection were 0.4 and 0.6, respectively, and in 0.8 of the total population, the intersection process took place. The mutation process also occurred randomly in 0.6 populations. The remaining 0.2 to create the next generation's population was also selected from the chromosomes of each population as well as mutants that performed better.

## 3-3 Classification results

After 10 times of performing the hybrid model of ANN-GA using the training dataset, this model was able to identify Down syndrome cases with the average accuracy of 96.32% and specificity and sensitivity of 99.72% and 90.91%, respectively, and with MSE equal to 0.61% in All data and with 82.27% accuracy as well as 98.62% and 75% specificity and sensitivity, respectively, and with MSE equal to 2.17% among the testing dataset. Sensitivity, specificity, accuracy and MSE related to 10 times the implementation of the hybrid ANN model and GA are listed in Table 1.

Table 1  
Hybrid artificial neural network performance outcomes in 10 times implementations

Iteration	Sensitivity (%)	Specificity (%)	Precision (%)	MSE (%)
1	90.91	100	100	0.56
2	86.36	100	100	0.65
3	86.36	100	100	0.73
4	90.91	99.65	95.24	0.71
5	95.46	99.65	95.45	0.37
6	95.45	99.65	95.45	0.54
7	86.36	99.30	90.48	0.89
8	90.91	99.65	95.24	0.64
9	95.46	99.30	91.30	0.47
10	90.91	100	100	0.61

## 4- Discussion

In recent years, machine learning techniques and especially ANNs in the field of diagnosis of fetal abnormalities have been provided robust results(19, 21, 22). But what can accurately demonstrate the capacity of a ANN for early detection of a disease is the optimal ANN architecture (13). This study showed that using GA, an optimal ANN architecture can be achieved for early detection of Down syndrome. In recent years, similar studies have been performed to detect aneuploidy abnormalities through ANNs, the results of which are shown in Table 2.

Table 2  
Comparison of current techniques with the proposed model

Writer	No of population	Machine learning method	Sensitivity	Specificity	Accuracy
Catic(22)	2500	Feedforward ANN	92%	88%	89.6%
		Feedback ANN	99%	98.67%	98.8%
Hannah (23)	Not reported	Decision Trees (J48)	88.5%	-	-
		Decision Trees (LMT)	83.3%	-	-
		Decision Trees (random trees)	80.8%	-	-
Current study	308	Hybrid ANN-GA	90.91%	99.72%	99.91%

By comparing the results of this study and similar studies, it can be showed that optimization algorithms such as GAs can improve the results of ANNs. Although some studies, such as the study by Catic et al., have shown that manual search can achieve acceptable results in ANN performance, it is not possible to search in a wide space of alternatives to create an optimal ANN architecture. In this study, a feedforward and feedback neural network were used, and unlike the Specificity and Accuracy parameters, sensitivity by both networks was more desirable in compare with the present study. Considering the high population used in Catic study, the achieved results could be expected.

Hannah et al. also predicted Down syndrome using three types of decision trees (36). A comparison of the results of this study with the present study showed that in the field of diagnosing chromosomal diseases such as Down syndrome, ANNs can perform better performance than the decision tree. Moreover, comparing the three methods of decision trees and ANN in predicting fetal distress syndrome, Huang et al. reported that the accuracy of diagnosis using discriminant analysis, decision tree and ANN were 82%, 86.4% and 97.8%, respectively (24).

Usually the screening tests are interpreted by software due to the ambiguous nature of them and the mother's age and weight are also used to improve the interpretation. Since Down Syndrome affects the

fetus's appearance, NT sonography also has a significant effect on the interpretation of such tests(25). The hybrid model can provide a clear, simple, fast and acceptable interpretation of this tests.

## 5- Conclusion

The results of this study showed that the use of GAs in optimizing the structure of the ANN can increase the accuracy in diagnosing Down syndrome through the information of first trimester screening tests. GA and its search strategies improve the results of these networks by finding the optimal structures of ANNs. The results showed that improving ANN conditions using GA is more efficient and cost-effective than manual methods.

## 6- Limitation

The main limitation of this study was the limited number of abnormal cases. It is possible to expect more definitive results from the model if there were more abnormality cases.

## List Of Abbreviations

List of abbreviations used in article

Abbreviation	Full form
ANN	Artificial Neural Network
GA	Genetic Algorithm
MSE	Mean Squad Error
NT	Nuchal Translucency
CRL	Crown-Rump Length
FHR	Fetal Heart Tate
HCG	Human chorionic Gonadotropin
PAPP-A	Pregnancy-Associated Plasma Protein A
NIPT	Non-Invasive Prenatal Test
CVS	Chorionic Villus Sampling
AIS	Artificial Immune System
LMT	Logistic Model Tree

## Declarations

Ethics approval and consent to participate

This study was approved in ethics committee of Ahvaz Jundishapur University of Medical Sciences (ethical code: IR.AJUMS.REC.1397.826).

### **Consent for publication**

In order to preserve the privacy and Confidentiality of individuals whose information is used in this research, personal information wasn't available to the researchers and only laboratory, clinical and social information was accessible.

### **Availability of data and materials**

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

### **Competing interests**

The authors declare that they have no competing interests

### **Funding**

this paper was financially supported by grant from vice chancellor of Research Affairs of Ahvaz Jundishapur University of Medical Sciences. This funding source had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results.

### **Authors' contributions**

The author's responsibilities were as follows: AM suggested the study, AM and SMH designed the study and conducted independent literature searches, SMH extracted the data, AM and SMH performed the statistical analysis, interpreting the findings, and wrote the manuscript. JM-A and MM helped in study designing and interpreting the findings. All four authors read and approved the final manuscript.

### **Acknowledgements**

The authors would like to appreciate the vice chancellor of Research Affairs of Ahvaz Jundishapur University of Medical Sciences for their financial and administrative support to undertake this project.

### **Ethics approval and consent to participate**

This study was approved by research ethics committee of *Ahvaz Jundishapur* University of Medical Sciences (IR.AJUMS.REC.1397.826). informed consent is not applicable in this study.

## **References**

1. Santoro SL, Bartman T, Cua CL, Lemle S, Skotko BG. Use of Electronic Health Record Integration for Down Syndrome Guidelines. *Pediatrics*. 2018;142(3):e20174119.
2. Channell MM, Hahn LJ, Rosser TC, Hamilton D, Frank-Crawford MA, Capone GT, et al. Characteristics Associated with Autism Spectrum Disorder Risk in Individuals with Down Syndrome. *Journal of Autism and Developmental Disorders*. 2019;49(9):3543-56.
3. Nguyen TL, Duchon A, Manousopoulou A, Loaëc N, Villiers B, Pani G, et al. Correction of cognitive deficits in mouse models of Down syndrome by a pharmacological inhibitor of DYRK1A. *Disease models & mechanisms*. 2018;11(9):dmm035634.
4. Santorum M, Wright D, Syngelaki A, Karagiotti N, Nicolaides KH. Accuracy of first-trimester combined test in screening for trisomies 21, 18 and 13. *Ultrasound in Obstetrics & Gynecology*. 2017;49(6):714-20.
5. Carlson LM, Vora NL. Prenatal Diagnosis: Screening and Diagnostic Tools. *Obstetrics and Gynecology Clinics*. 2017;44(2):245-56.
6. Neocleous AC, Neocleous C, Schizas CN, Biehl M, Petkov N, editors. Marker selection for the detection of trisomy 21 using generalized matrix learning vector quantization. 2017 International Joint Conference on Neural Networks (IJCNN); 2017 14-19 May 2017.
7. Taylor-Phillips S, Freeman K, Geppert J, Agbebiyi A, Uthman OA, Madan J, et al. Accuracy of non-invasive prenatal testing using cell-free DNA for detection of Down, Edwards and Patau syndromes: a systematic review and meta-analysis. *BMJ Open*. 2016;6(1):e010002.
8. Neocleous AC, Nicolaides KH, Schizas CN. First trimester noninvasive prenatal diagnosis: a computational intelligence approach. *IEEE journal of biomedical and health informatics*. 2015;20(5):1427-38.
9. Akolekar R, Beta J, Picciarelli G, Ogilvie C, D'Antonio F. Procedure-related risk of miscarriage following amniocentesis and chorionic villus sampling: a systematic review and meta-analysis. *Ultrasound in Obstetrics & Gynecology*. 2015;45(1):16-26.
10. Jayalakshmi T, Santhakumaran A, editors. A Novel Classification Method for Diagnosis of Diabetes Mellitus Using Artificial Neural Networks. 2010 International Conference on Data Storage and Data Engineering; 2010 9-10 Feb. 2010.
11. Ding S, Su C, Yu J. An optimizing BP neural network algorithm based on genetic algorithm. *Artificial intelligence review*. 2011;36(2):153-62.
12. Leung FH-F, Lam H-K, Ling S-H, Tam PK-S. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *IEEE Transactions on Neural networks*. 2003;14(1):79-88.
13. Heckerling PS, Gerber BS, Tape TG, Wigton RS. Use of genetic algorithms for neural networks to predict community-acquired pneumonia. *Artificial Intelligence in Medicine*. 2004;30(1):71-84.
14. Guo P, Wang X, Han Y, editors. The enhanced genetic algorithms for the optimization design. 2010 3rd International Conference on Biomedical Engineering and Informatics; 2010: IEEE.

15. Cilimkovic M. Neural networks and back propagation algorithm. Institute of Technology Blanchardstown, Blanchardstown Road North Dublin. 2015;15.
16. Montana DJ, Davis L, editors. Training feedforward neural networks using genetic algorithms. IJCAI; 1989.
17. Chen F-C. Back-propagation neural networks for nonlinear self-tuning adaptive control. IEEE control systems Magazine. 1990;10(3):44-8.
18. Grover D, Toghi B, editors. MNIST Dataset Classification Utilizing k-NN Classifier with Modified Sliding-window Metric. Science and Information Conference; 2019: Springer.
19. Troisi J, Sarno L, Martinelli P, Di Carlo C, Landolfi A, Scala G, et al. A metabolomics-based approach for non-invasive diagnosis of chromosomal anomalies. Metabolomics. 2017;13(11):140.
20. Al-Maqaleh BM, Abdullah AMG. Intelligent predictive system using classification techniques for heart disease diagnosis. International Journal of Computer Science Engineering (IJCSE). 2017;6(6):145-51.
21. Katlan DC, Tanacan A, Orgul G, Leblebicioglu K, Beksac MS. An Intelligent Prenatal Screening System for the Prediction of Trisomy-21 Using Triple Test Variables: The Hacettepe System. Gynecology Obstetrics & Reproductive Medicine. 2019;25(2):67-9.
22. Catic A, Gurbeta L, Kurtovic-Kozaric A, Mehmedbasic S, Badnjevic A. Application of Neural Networks for classification of Patau, Edwards, Down, Turner and Klinefelter Syndrome based on first trimester maternal serum screening data, ultrasonographic findings and patient demographics. BMC medical genomics. 2018;11(1):19.
23. Hannah E, Raamesh L. Early Prenatal Diagnosis of Down's Syndrome-A Machine Learning Approach. Soft Computing for Problem Solving: Springer; 2020. p. 467-77.
24. Huang M-L, Hsu Y-Y. Fetal distress prediction using discriminant analysis, decision tree, and artificial neural network. 2012.
25. Beksaç MS, Durak B, Özkan Ö, Çakar AN, Balci S, Karakaş Ü, et al. An artificial intelligent diagnostic system with neural networks to determine genetical disorders and fetal health by using maternal serum markers. European Journal of Obstetrics & Gynecology and Reproductive Biology. 1995;59(2):131-6.

## Figures

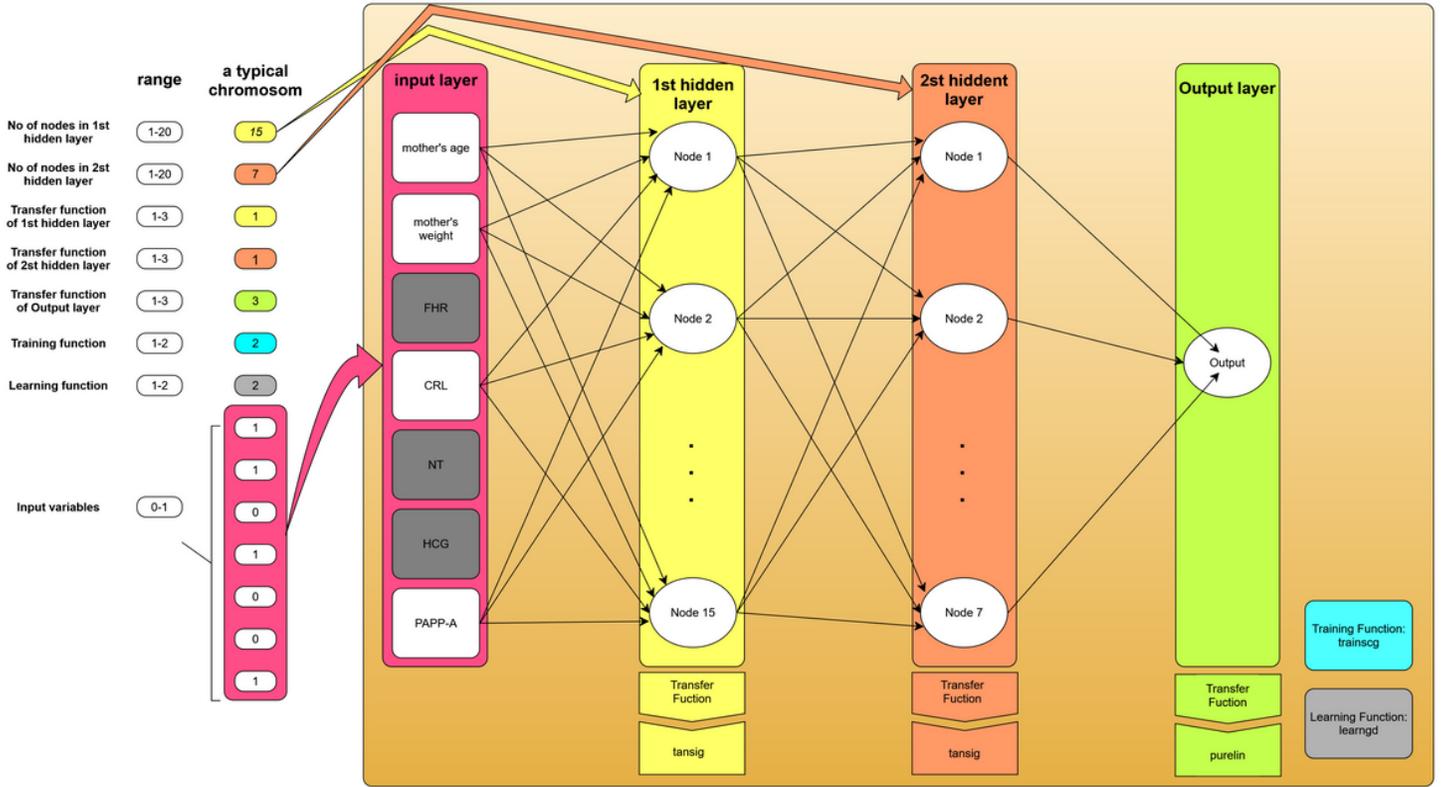


Figure 1

The structure of the proposed hybrid ANN-GA model