OPTIMAL APPROACH TO THE PANCREATIC CANCER PREDICTION BASED ON ARTIFICIAL NEURAL NETWORK COMPARED WITH DECISION TREE ALGORITHM FOR IMPROVED ACCURACY

D. Yoga lasyaa[1], P. Nirmala[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The aim of the work is to evaluate the accuracy in predicting pancreatic cancer using Artificial Neural Network(ANN) and Decision Tree Algorithms.
Materials and Methods: In this work there are two groups in which each group has 20 sample sizes and total sample size is found to be 40, with pre-test power of 80% (G-power), α=0.05, confidence interval 95%. The effectiveness in identifying the pancreatic cancer by the algorithms is evaluated.
Results: It has been observed that the artificial neural network is much better than the decision tree algorithm in terms of accuracy in predicting the cancer. Each algorithm gives different accuracies where ANN has better mean accuracy of 90.06% which is better than the decision tree algorithm 85.9%. The statistical results are also provided where Artificial Neural Network and Decision tree algorithms have statistical significance different values i.e. $p<0.01$ (independent sample T-test).
Conclusion: The results show that the proposed artificial neural network is efficient when compared to decision tree algorithms.

Keywords: Pancreatic cancer, Artificial Neural Network (ANN), Cancer Prediction(CP), Prognosis, Innovative Pancreatic Cancer Detection, Machine Learning.

[1]Research Scholar, Department of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical & Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode: 602105.
[2*]Project Guide, Department of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical & Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode: 602105.

1. Introduction

This research employs sophisticated detection methods to predict pancreatic cancer as an advanced diagnosis method, which is challenging to analyze because the symptoms only appear in the later stages of life. Cancer-specific symptoms appear in most patients at an advanced stage. Considering the low incidence rate, pancreatic cancer remains the fourth greatest cause of cancer-related death in both men and women. Pancreatic cancer has a higher risk of mortality (Muhammad et al., 2019). The pancreatic cancer risk was determined at the individual level due to a lack of effective screening techniques or early stage disease diagnosis is difficult. Endoscopic ultrasonography, magnetic resonance imaging, computed tomography, or positron emission tomography have also been demonstrated in different clinical investigations. The disease itself usually results in a limited patient survival rate, and therapy generally results in a recovery, though some patients do experience long-term remissions (Saxena et al., 2021; Tong et al., 2020)1). Even with the best treatment, physicians must assess the possible survival advantages against the impact of treatment problems on the quality of life of their patients. Artificial neural networks, a widely used method in machine learning, a non-linear mode and biological neural model system, both structurally and functionally ((Saxena et al., 2021; Tong et al., 2020). Artificial neural networks have shown to be effective in handling a variety of problems in machine learning. ANN can train and generalize the complicated and non-linear properties of the input data (Chalumuri et al., 2020). ANN has been effectively employed as a nonlinear pattern recognition technology to make diagnostic and prognostic predictions in a variety of clinical situations which are considered difficult. Artificial neural networks are nonlinear systems that may be better at describing interactions between health risk variables than conventional statistical models. In difficult medical decision-making, ANNs are becoming more widely used, it has been used to prognostic patient outcomes in a variety of cancers (Ansari et al., 2013)

Many articles have been published on detection of pancreatic cancer using artificial neural network algorithms for 5 years. PUBMED published 80 articles and in science direct 45 articles were published from the last 5 years. It is critical to investigate novel technologies for early detection of pancreatic cancer, in order to enhance the prognosis of people with the disease (Yu et al., 2005). Nonlinear regression computational devices, or ANNs, have been implemented in classification and survival prediction in a variety of biomedical systems, including cancer, for over 45 years. Artificial neural networks (ANNs) are regression devices with layers that operate as computer nodes and have exceptional data processing capabilities. They can identify nonlinearities that are just not stated clearly as inputs, giving them the ability to learn and adapt (Ahmed, 2005). ANNs have the advantage of being able to deal with "uncertain" data and taking advantage of the availability of multiple components of data that will not show logically to influence the solution of a problem. They're good for solving, classifying and predicting problems since they determine a value for each individual in a group (Marchevsky, 2007). In medical diagnosis, prognosis, and survival analysis, artificial neural networks have now been found to be successful as general non-linear models. This work opens with an overview of artificial neural networks that are employed for non-linear regression models in cancer patient survival analyses. These methods are appealing because they allow for the simulation of time-dependent risks in the case of complex non-linear and non-additive covariate effects. First, the relevance of neural networks in cancer prognosis is discussed in terms of statistical techniques and parametric procedures (Arsene & Lisboa, 2007).

Our institution is passionate about high quality evidence based research and has excelled in various domains (Anupong et al., 2022; Bharathiraja et al., 2022; Jothi et al., 2022; Kale et al., 2022; Palanisamy et al., 2022; Ram et al., 2022; Sumathy et al., 2022; Thanigaivel et al., 2022; Vickram et al., 2022; Yaashikaa et al., 2022).The accuracy in detecting the pancreatic cancer prediction using existing algorithms is found to be very low. This made it possible to do research under this work. Hence the aim of work is to propose an innovative pancreatic cancer detection using machine learning model to predict pancreatic cancer with more accuracy.

2. Materials and Methods

This work is done in the specialization lab of the Biomedical Engineering department in Saveetha School of Engineering. The sample data sets are taken from Kaggle.com. Pre-test analysis was also done by keeping g-power at 80%, threshold at 0.05% & confidence interval at 95%. Number of groups used in this work is two. The total number of samples is 40, each group has a sample size of 20 (Muhammad et al., 2019)

The proposed algorithm is improvised for artificial neural networks for the pancreatic cancer detection. This algorithm is a way to promote machine learning through the innovative pancreatic cancer detection method. It enables non-linear process modeling and has become a very popular and useful tool for the artificial neural network for solving many problems such as classification, clustering, regression, pattern recognition, dimension reduction, structured prediction, machine translation, anomaly detection, decision making. Artificial Neural Networks are used to model nonlinear problems and predict the output value. This study concludes that the artificial neural network algorithm has better accuracy than the existing algorithm decision tree by using the innovative pancreatic cancer detection method. The effective approach for the Prognosis prediction of pancreatic cancer with artificial neural networks has better accuracy.

Google uses artificial neural networks in recurring connections to support voice search. Microsoft also claims to have developed a speech recognition system, using neural networks, that can transcribe conversations with a little more precision than humans. ANNs have the ability to learn and model complex and non-linear relationships, which is very important because in real life many of the relationships between inputs and outputs are complex and non-linear.

Working station in proposing the new algorithm for better accuracy for treating pancreatic cancer is mostly in a personal computer in software named matlab with all the required add-ons installed. Both the algorithms have been simulated using the software Matlab. The code is simulated for artificial neural algorithms and decision tree algorithms for detection. Performance of the algorithm in pancreatic cancer prediction is shown and values are noted in Table 1 and Table 2 also shows the statistical analysis response values. The resultant graph in Fig 2. shows the simple bar mean and mean of loss in the proposed algorithm.

Overall testing setup was done on Matlab 2021b software. Datasets were obtained from kaggle.com and used in the testing methods. The output data values were noted.

Statistical Analysis

To validate the outcomes of both algorithms, statistical algorithm evaluation was carried out using IBM-SPSS software (Wagner & III, 2019). The mean accuracy is found by performing independent samples t-test. The data set is trained and tested with various percentages, and the output is noted.

3. Results

The pancreatic cancer prediction with the proposed algorithm was programmed using the matlab. Although training the datasets seek to minimize errors in neural networks and provide better accuracy for proposed algorithms, the artificial neural networks in common use to address this until an acceptable accuracy is achieved.

Table 1. shows the accuracy of 20 samples in comparison of customized artificial neural network and decision tree algorithms. Each sample gives different accuracies where customized artificial neural networks have better accuracy 90.05% which is greater than decision tree algorithm.

In Table 2. the statistical analysis of group 1 artificial neural network accuracy is 90.06 standard deviation 1.17127 and standard error 0.26190 whereas group 2 decision tree algorithm mean standard deviation is 1.16524 and standard error is 0.26056. Table 3 shows independent samples tested by equal variance on the basis of assumed and not assumed. On the basis of significance $p=0.01$ i.e $p<0.05$

In Fig.1 the graph shows performance analysis of ANN which is obtained by using Microsoft Excel. Performance analysis compares the accuracy values of ANN and Decision Tree algorithm.

In Fig.2 pancreatic cancer prediction using ANN is shown in a bar graph. Predicted accuracy values can be done by using SPSS software. Performance analysis compares the mean accuracy values of artificial neural networks and decision tree algorithms.

4. Discussion

The results of the study shows that the Artificial Neural Network is better than the decision tree algorithm with p value less than 0.01. The purpose of this work is to study the predicted survival time of pancreatic cancer patients using data from patients.

In the general population, patient clinical data has been shown to be useful in identifying therapy and patient outcome trends. Models based on patient-level data can provide insight into the best course of treatment for individual patients as well as their long-term prospects. The goal of this review is to critically evaluate the use of machine learning for analyzing, prognostic and managing resectable pancreatic cancer with prediction types (Bradley et al., 2019). Artificial neural networks have traditionally been used in the medical community to develop predictive models based on explanatory

variables or traits. This research findings demonstrate that using machine learning algorithms may increase significantly the survival of patients using prognosis prediction and feature selection (Floyd, 2007). The Cancer Prediction ability of artificial neural networks to model land cover changes as a powerful tool for pancreatic cancer. However, the development of algorithms for cancer prediction and accurate prognosis prediction is practically possible through innovative pancreatic cancer detection method. Therefore, part of pancreatic cancer was accurately predicted in this investigation using artificial neural networks. Artificial neural networks have been developed as an effective statistical technique over the past 40 years. They have been used in many areas and have become established as practicable calculation methods in informatics, biochemistry and medicine (Caron, 2022). This article proposes a model using ANN for pancreatic ductal adrenalcarcinoma. Several attributes were considered and the concordance index was found to be 0.79. This article successfully predicted patients survival after surgery (Ansari et al., 2013). The authors in this article proposed a biomarker method using decision tree which is also a machine learning model.From several control samples a sensitivity of about 88.9% was achieved. When compared to Decision Tree, ANN uses nonlinear statistics and consists of a strongly interconnected set of processing units and weighted connections(Yu et al., 2005). There are also other approaches selected by the authors for predicting pancreatic cancer. One of them is using data mining techniques (Floyd, 2007). Different data mining tools and techniques are implemented by the authors in this work and a maximum of 99% accuracy has been achieved using these techniques(T., 2019). Limited data sources, which limit generalizability and bias, a lack of external validation, and the requirement for transparency in internal validation, sequential sampling, and candidate predictor selection were the key methodological difficulties highlighted (Diederich, 2008). The future of research depends on increasing the concept of a multidisciplinary team to include specialists from computing and data science, with algorithms built in collaboration with physicians and considered as supplements rather than replacements for traditional clinical decision-making (Nougnignon Comlan Deguenonvo & Thiam, 2021)

5. Conclusion

This study concludes that Artificial Neural Network-based techniques for predicting pancreatic cancer have a higher accuracy and it is found that proposed algorithm artificial neural network has better accuracy than the existing decision tree algorithm.

Declaration

Conflict of interests

No conflict of interest in this manuscript.

Author contribution

Author YL was involved in Methodology creation, simulation, data collection, data analysis, Manuscript writing. Author PN was involved in conceptualization, guidance and critical review of manuscript.

Acknowledgement

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding: We thank the following organizations for providing financial support that enabled us to complete the study.

1. Malathi Polyclinic, Banglore.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

6. References

Ahmed, F. E. (2005). Artificial neural networks for diagnosis and survival prediction in colon cancer. *Molecular Cancer*, *4*, 29.

Ansari, D., Nilsson, J., Andersson, R., Regnér, S., Tingstedt, B., & Andersson, B. (2013). Artificial neural networks predict survival from pancreatic cancer after radical surgery. *American Journal of Surgery*, *205*(1), 1–7.

Anupong, W., Yi-Chia, L., Jagdish, M., Kumar, R., Selvam, P. D., Saravanakumar, R., & Dhabliya, D. (2022). Hybrid distributed energy sources providing climate security to the agriculture environment and enhancing the yield. In *Sustainable Energy Technologies and Assessments* (Vol. 52, p. 102142). https://doi.org/10.1016/j.seta.2022.102142

Arsene, C. T. C., & Lisboa, P. J. G. (2007). Artificial Neural Networks Used in the Survival Analysis of Breast Cancer Patients: A Node-Negative Study. In *Outcome Prediction in Cancer* (pp.

191–239). https://doi.org/10.1016/b978-044452855-1/50010-6

Bharathiraja, B., Jayamuthunagai, J., Sreejith, R., Iyyappan, J., & Praveenkumar, R. (2022). Techno economic analysis of malic acid production using crude glycerol derived from waste cooking oil. *Bioresource Technology*, *351*, 126956.

Bradley, A., van der Meer, R., & McKay, C. (2019). Personalized Pancreatic Cancer Management: A Systematic Review of How Machine Learning Is Supporting Decision-making. *Pancreas*, *48*(5), 598–604.

Caron, P. (2022). MANAGEMENT OF THYROTOXICOSIS AND PREGNANCY: Review of the current literature and an update of the care pathway. *Annales D'endocrinologie*. https://doi.org/10.1016/j.ando.2022.01.006

Chalumuri, A., Kune, R., & Manoj, B. S. (2020). Training an Artificial Neural Network Using Qubits as Artificial Neurons: A Quantum Computing Approach. In *Procedia Computer Science* (Vol. 171, pp. 568–575). https://doi.org/10.1016/j.procs.2020.04.061

Diederich, J. (2008). *Rule Extraction from Support Vector Machines*. Springer Science & Business Media.

Floyd, S. (2007). *Data Mining Techniques for Prognosis in Pancreatic Cancer*.

Jothi, K. J., Jeeva Jothi, K., Balachandran, S., Mohanraj, K., Prakash, N., Subhasri, A., Santhana Gopala Krishnan, P., & Palanivelu, K. (2022). Fabrications of hybrid Polyurethane-Pd doped ZrO2 smart carriers for self-healing high corrosion protective coatings. In *Environmental Research* (Vol. 211, p. 113095). https://doi.org/10.1016/j.envres.2022.113095

Kale, V. N., Rajesh, J., Maiyalagan, T., Lee, C. W., & Gnanamuthu, R. M. (2022). Fabrication of Ni–Mg–Ag alloy electrodeposited material on the aluminium surface using anodizing technique and their enhanced corrosion resistance for engineering application. In *Materials Chemistry and Physics* (Vol. 282, p. 125900). https://doi.org/10.1016/j.matchemphys.2022.125900

Marchevsky, A. M. (2007). The Use of Artificial Neural Networks for the Diagnosis and Estimation of Prognosis in Cancer Patients. In *Outcome Prediction in Cancer* (pp. 243–259). https://doi.org/10.1016/b978-044452855-1/50011-8

Muhammad, W., Hart, G. R., Nartowt, B., Farrell, J. J., Johung, K., Liang, Y., & Deng, J. (2019). Pancreatic Cancer Prediction Through an Artificial Neural Network. *Frontiers in Artificial Intelligence*, *2*, 2.

Nougnignon Comlan Deguenonvo, G., & Thiam, I. (2021). [Multivisceral pancratic cytosteatonecrosis found at autopsy]. *Annales de pathologie*. https://doi.org/10.1016/j.annpat.2021.09.004

Palanisamy, R., Karuppiah, D., Rengapillai, S., Abdollahifar, M., Ramasamy, G., Wang, F.-M., Liu, W.-R., Ponnuchamy, K., Shim, J., & Marimuthu, S. (2022). A reign of bio-mass derived carbon with the synergy of energy storage and biomedical applications. In *Journal of Energy Storage* (Vol. 51, p. 104422). https://doi.org/10.1016/j.est.2022.104422

Ram, G. D., Dinesh Ram, G., Praveen Kumar, S., Yuvaraj, T., Babu, T. S., & Balasubramanian, K. (2022). Simulation and investigation of MEMS bilayer solar energy harvester for smart wireless sensor applications. In *Sustainable Energy Technologies and Assessments* (Vol. 52, p. 102102). https://doi.org/10.1016/j.seta.2022.102102

Saxena, A., Brault, N., & Rashid, S. (2021). *Big Data and Artificial Intelligence for Healthcare Applications*. CRC Press.

Sumathy, B., Kumar, A., Sungeetha, D., Hashmi, A., Saxena, A., Kumar Shukla, P., & Nuagah, S. J. (2022). Machine Learning Technique to Detect and Classify Mental Illness on Social Media Using Lexicon-Based Recommender System. *Computational Intelligence and Neuroscience*, *2022*, 5906797.

Thanigaivel, S., Vickram, S., Dey, N., Gulothungan, G., Subbaiya, R., Govarthanan, M., Karmegam, N., & Kim, W. (2022). The urge of algal biomass-based fuels for environmental sustainability against a steady tide of biofuel conflict analysis: Is third-generation algal biorefinery a boon? In *Fuel* (Vol. 317, p. 123494). https://doi.org/10.1016/j.fuel.2022.123494

Tong, Z., Liu, Y., Ma, H., Zhang, J., Lin, B., Bao, X., Xu, X., Gu, C., Zheng, Y., Liu, L., Fang, W., Deng, S., & Zhao, P. (2020). Development, Validation and Comparison of Artificial Neural Network Models and Logistic Regression Models Predicting Survival of Unresectable Pancreatic Cancer. *Frontiers in Bioengineering and Biotechnology*, *8*, 196.

T., S. (2019). A Study on Implementation of different Data Mining Techniques on Healthcare. In

International Journal of Computer Applications (Vol. 178, Issue 31, pp. 13–17). https://doi.org/10.5120/ijca2019919116

Vickram, S., Rohini, K., Anbarasu, K., Dey, N., Jeyanthi, P., Thanigaivel, S., Issac, P. K., & Arockiaraj, J. (2022). Semenogelin, a coagulum macromolecule monitoring factor involved in the first step of fertilization: A prospective review. *International Journal of Biological Macromolecules*, *209*(Pt A), 951–962.

Wagner, W. E., & III. (2019). *Using IBM® SPSS® Statistics for Research Methods and Social Science Statistics*. SAGE Publications.

Yaashikaa, P. R., Keerthana Devi, M., & Senthil Kumar, P. (2022). Algal biofuels: Technological perspective on cultivation, fuel extraction and engineering genetic pathway for enhancing productivity. In *Fuel* (Vol. 320, p. 123814). https://doi.org/10.1016/j.fuel.2022.123814

Yu, Y., Chen, S., Wang, L.-S., Chen, W.-L., Guo, W.-J., Yan, H., Zhang, W.-H., Peng, C.-H., Zhang, S.-D., Li, H.-W., & Chen, G.-Q. (2005). Prediction of pancreatic cancer by serum biomarkers using surface-enhanced laser desorption/ionization-based decision tree classification. *Oncology*, *68*(1), 79–86.

Tables and Figures

Table 1: The accuracy of samples of artificial neural network and decision tree algorithms. The maximum accuracy using ANN is found to be 90.6% and 85.9% by decision tree algorithm.

ARTIFICIAL NEURAL NETWORK ACCURACY(%)	DECISION TREE ACCURACY(%)
91.3	88.3
90.9	86.7
89.3	87.7
88.7	88.0
87.7	86.1
90.9	86.4
90.6	86.7
89.6	85.4
90.6	85.4
88.3	86.1
89.6	85.4
89.3	85.4
90.3	85.4
89.6	84.8

89.6	85.8
91.6	86.4
91.9	83.8
91.6	86.7
89.0	84.8
90.9	84.5

Table 2 : Statistical analysis of artificial neural network and decision tree algorithm. The group statistics provides mean, standard deviation, standard error mean of the accuracy for a total sample of 40. The innovative artificial neural network has better accuracy than the decision tree algorithm.

	GROUP	N	MEAN	STANDARD DEVIATION	STANDARD ERROR MEAN
ACCURACY	ANN	20	90.0650	1.17127	.26190
	DECISION TREE	20	85.9900	1.16524	.26056

Table 3 : Independent sample t-test on pancreatic cancer prediction using SPSS software. The mean standard deviation and significance difference of customized artificial neural network has better accuracy consistency with decision tree algorithm. There is significance difference between the two groups.

		Levene's Test for Equality of Variances		T - Test for Equality of Mean						95% Confidence Interval of the Difference	
		F	Sig.	t	df	One sided p	Two - Sided p	Mean diff.	Std Error diff.	Lower	Upper
ACCURACY	Equal	.14	.70	11.03	38	<0.0	<0.0	4.07500	.3694	3.33271	4.822

	variances assumed	1	9	0		1	1	0	4	2	8
	Equal Variances not assumed			11.030	37.999	<0.01	<0.01	4.075000	.36944	3.32712	4.8228

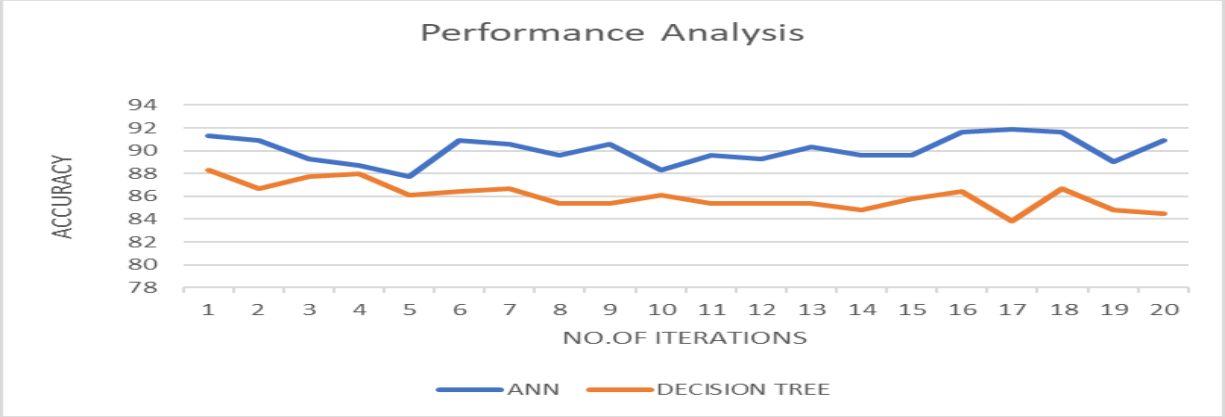

Fig.1. Performance analysis of ANN and Decision Tree denoted by its accuracy. Performance analysis compares the accuracy values of ANN and Decision Tree algorithm.

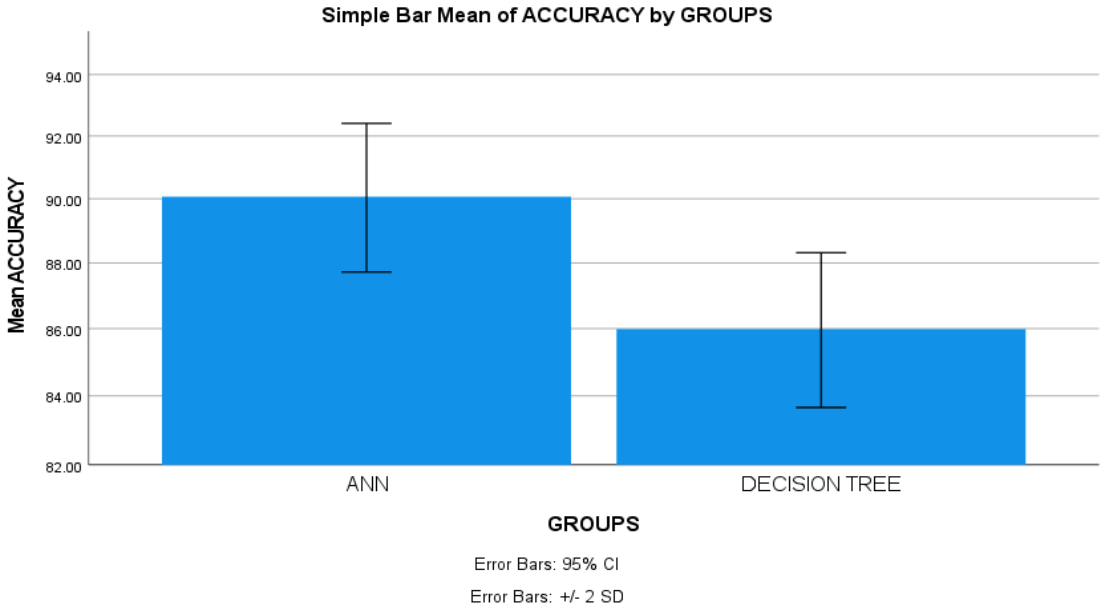

Fig. 2. By comparing artificial neural network and decision tree algorithm on diagnosis of pancreatic cancer prediction, ANN has better accuracy.The accuracy values are represented in a bar graph to predict the the best algorithm using SPSS software.The above figure shows artificial neural network and decision tree algorithm as groups in X-axis and the mean accuracy in Y -axis with +/-2SD.