Prescient Image Classification of Malaria Cells Using Artificial Neural Network in Orange

Muhammad Sohail
Dept. of computer science
Riphah International University Lahore
Pakistan
maliksohail284@yahoo.com

Mehwish Shabbir
Dept. of computer science
Riphah International University Lahore
Pakistan
mehwish.shabbir22@gmail.com

Hafiz Muneeb Ahmed
Dept. of computer science
Riphah International University Lahore
Pakistan
ahmed.muneeb8470@gmail.com

Shahzaib Iqbal
Dept. of computer science
Riphah International University Lahore
Pakistan
shahzaibiqbal43@gmail.com

Kainat Noor
Dept. of computer science
Riphah International University Lahore
Pakistan
Nkainat801@gmail.com

Misbah Akbar
Dept. of computer science
Riphah International University Lahore
Pakistan
alis81535@gmail.com

***Abstract*— Intestinal sickness has been annihilated from the created nations yet at the same time influencing a huge piece of the world negatively. A bigger level of malaria fever is assessed to influence small kids. So as to lessen mortality from pediatric intestinal sickness, there ought to be a proficient and powerful expectation strategy. In medicinal services, information mining is one of the most crucial and inspiring regions of examination with the target of finding significant data from tremendous informational collections and gives a proficient systematic way to deal with distinguishing obscure and important data in social insurance information. In this investigation, the fundamental point is to anticipate the picture cell classification of Malaria fever utilizing profound learning model, for example, neural network. The dataset was downloaded from Kaggle. The detailed accuracy provided by the Orange tool using Artificial Neural Network is 93.50 %.**

***Keywords*—** *Malaria, Artificial neural Network (ANN), Orange.*

I. INTRODUCTION

Malaria is a typical infection and some of the time deadly as well and that is the reason it is considered as a genuine medical issue over globe. Malaria is brought about by Plasmodium parasites, which are most generally transmitted through the chomp of the Anopheles mosquito. The chromatography is a technique used to perform tests. In order to check the patients, have infected or not blood samples of the patients collected. If the patients have infected their blood cells will produce antibodies and the disease-causing agents are called antigens. Antigens may be bacteria virus germs or any other things else. The antibodies are produced as a result when antigens enter into the body but the number of antibodies is always low. To check the results of infected persons their blood put on slides and a solution containing the antibodies mixed it and this solution is called serum. This solution does not contain only antibodies but enzymes also. When antibodies and antigens interact with each other than enzymes are used to strengthen this interaction. During this procedure the shape and color of blood changed depending upon the level of disease and it was examined using microscope.

An artificial neural network is a deep learning technique that is used in this article and the image cell dataset of malaria was taken from Kaggle. Orange a data mining tool used to perform the predictive analysis of Malaria. Profound learning is a subset of AI in man-made brainpower (computer-based intelligence) that has systems fit for taking in unaided from information that is unstructured or unlabeled. Otherwise called profound neural learning or profound neural system.

II. LITERATURE REVIEW

This examination depended on the possibility that qualities utilized for different boundaries influencing malaria fluctuate according to geographical spread's model which is having a low mistake the rate has demonstrated to be helpful in malaria flare-up forecast among other expectation strategies. He proved SVM performed well than ANN [1].

This paper provided an information about the development and diagnosis of malaria disease in machine learning and image analysis. Information about automated diagnosis of malaria was also the part of this study [2].

The main objective of this paper was to predict malaria by using a big data approach to environmental and clinical data. They claimed that the prediction of malaria by using clinical and environmental data is a new approach to big data analytics [3].

In this examination, the malaria malady was predicated by utilizing some order calculations, for example, J48, HoeffdingTree, LMT, and REP Tree, and it was demonstrated that REP Tree gave preferred precision over different calculations [4].

Predicative analysis of malaria disease was done using classification algorithm in Weka and R. Naïve Bayes algorithm performed well in weka than R. SMOTE algorithms was also the part of this study [5].

In this article, the classification of malaria cells was done by using deep learning, and the Conventional Neural network used to obtain better results. The comparative analysis was also done on pre-trained CNN [6].

In this paper CUCKOO search algorithm was utilized for planning a heuristic scale, which is additionally

surveyed through different investigations to assess its exactness. Distinctive execution assessment estimates like sensitivity, accuracy, specificity, and precision are utilized to evaluate the vigor of the model towards early distinguishing proof of Malaria in the untimely stage [7].

In this article Conventional neural network was used to predict the malaria cell image and the patterns of normal and abnormal persons by using deep learning model in Python environment. The accuracy provided this model was 97% [8].

RapidMiner having simulation was utilized in this paper to foresee malaria illness. Profound learning calculations were executed in this exploration and the motivation behind this investigation to flare-up the malaria infection. A Historical dataset was executed in AI to anticipate the outcomes [9].

The KNIME tool was used to predict malaria disease in this paper. Four algorithms were used and Random Forest provided the best and remarkable results. All the algorithms belonged to the classifications [10].

III. METHODOLOGY

In order to classify and prediction of malaria image cells, deep learning technique is used in this research. Neural network model is used to classify the patients. The image dataset was downloaded from Kaggle. The dataset contained two type of images uninfected and parasitized. The dataset contained thousands of cell images but in this research few of cell images were selected for both training and testing. The normal cell images are those images having no malaria disease and parasitized cell images having malaria disease as showed in Fig 1&2.

The Orange a data mining tool was used in this paper. An orange tools is a free, an open source tool and people can use it easily. It had strong graphical user interface and provide best visualizations. It was selected because it provides greater benefits in medical science. Image analytics strategy used in this paper. By default, image analytics was not installed in orange tool interface and it was installed by using built in ads-on facility.

The training and testing dataset contained uninfected and parasitized images and target for both training and testing dataset was designed. CSV file of training target contained “yes” and “no” and testing target contained “?”. The “yes” for parasitized,” no” for uninfected and “?” for unknown category of cell images.

To start a practical work in Orange tool I selected “Import Images” from the directory. A file containing cell images of uninfected patients was loaded in it. To check the file successfully uploaded there “Image Viewer” was imported from the directory a double clicking on it provided the detail of images. The same process was used for parasitized cell images to train our data. “Image Embedding” reads the image and transfers them to a remote server or assess them locally. So as to join cell pictures of both uninfected and parasitized train cell images "Concatenate" menu was chosen and the association was made between them. A CSV file containing target values was loaded into the workflow and it was merged. Neural Network was applied to check the test and score of the training dataset. Activation function was also changed to check the classified accuracy.

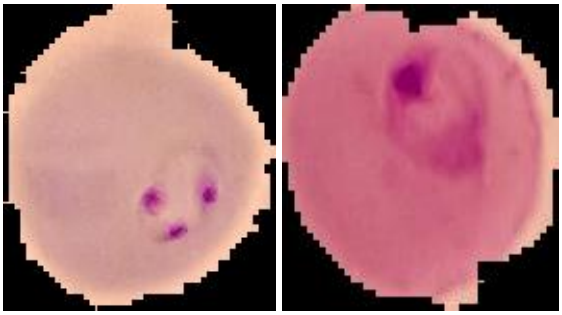


Fig 1. Parasitized Cell Images

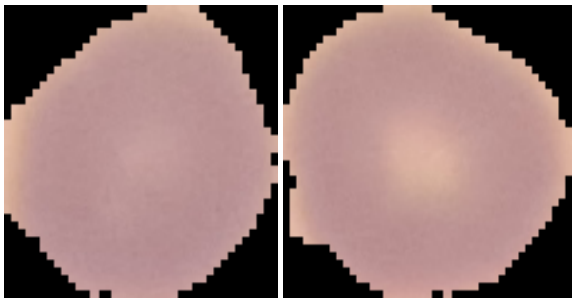


Fig 2. Uninfected Cell Images

A similar procedure was utilized for testing Image cells data and CSV document for testing information contained "?" values. The model was prepared and tried proficiently. The expectations were done and results were noted plainly.

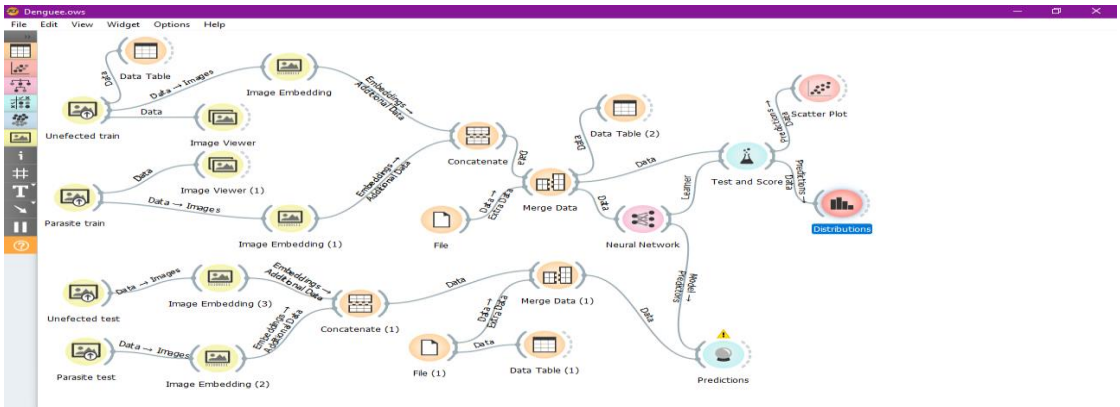


Fig 3. Workflow of Practical Work

IV. RESULTS

The results were noted after the model was set up for both preparing and testing cell images of malaria fever. The neural network model gave the 93.50% classified accuracy using logistic activation function which is exceptionally acknowledged . Precision was 93.90% and recall was 93.50%.

Evaluation Results

Model	AUC	CA	F1	Precision	Recall
Neural Network	0.988	0.932	0.932	0.935	0.932

Fig 4. Detailed Accuracy of (ANN) ReLu Activation

Evaluation Results

Model	AUC	CA	F1	Precision	Recall
Neural Network	0.986	0.917	0.917	0.924	0.917

Fig 5. Detailed Accuracy of (ANN) tanh Activation

Evaluation Results

Model	AUC	CA	F1	Precision	Recall
Neural Network	0.989	0.935	0.935	0.939	0.935

Fig 6. Detailed Accuracy of (ANN) Logistic Activation

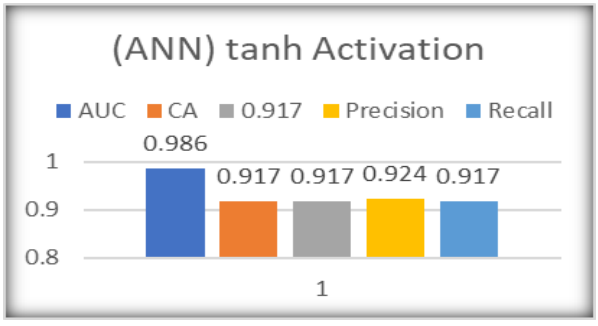


Fig. 7 Graphical Representation of tanh Activation

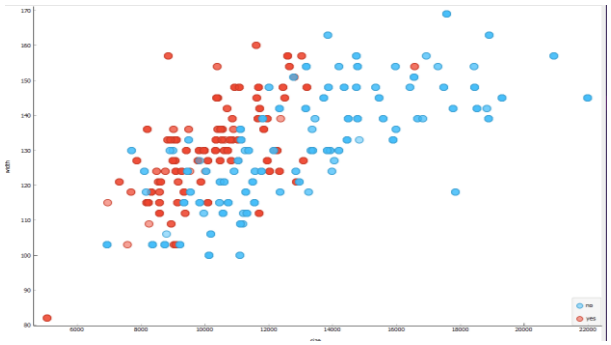


Fig 8. Scatter Plot visualization for uninfected and parasitized cell images.

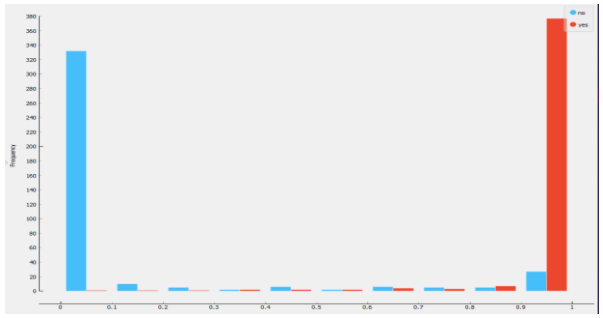


Fig. 9 Distribution for Neural Network

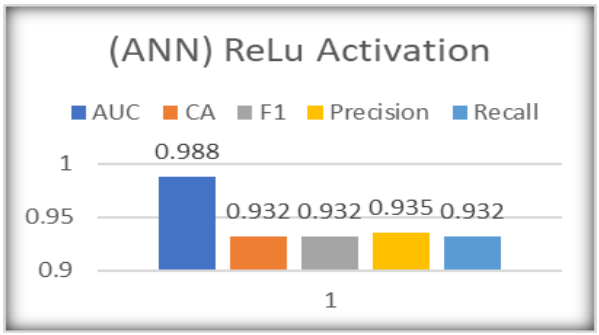


Fig. 10 Graphical Representation of ReLu Activation

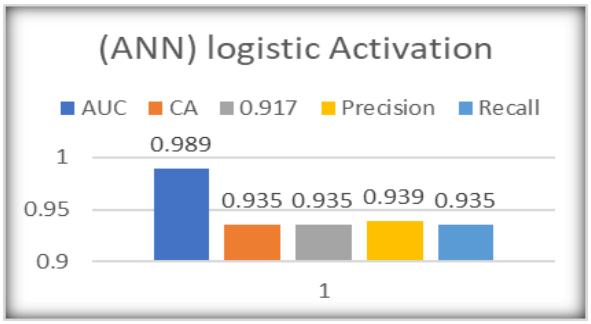
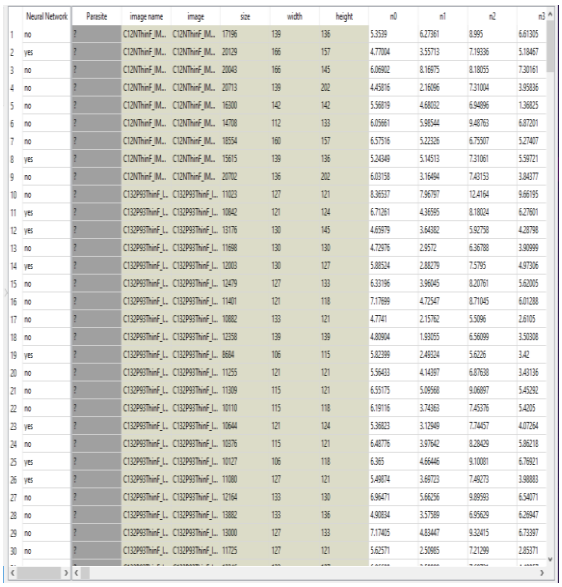


Fig. 11 Graphical Representation of logistic Activation



	Neural Network	Parasite	image name	image	size	width	height	n0	n1	n2	n3
1	no	?	C12NThinF_IM...	C12NThinF_IM...	17196	139	136	5.3539	6.27361	8.995	6.61305
2	yes	?	C12NThinF_IM...	C12NThinF_IM...	20129	166	157	4.77004	3.55713	7.19336	5.18467
3	no	?	C12NThinF_IM...	C12NThinF_IM...	23043	166	145	6.06902	8.16875	8.18055	7.30161
4	no	?	C12NThinF_IM...	C12NThinF_IM...	20713	139	202	4.45816	2.16096	7.31004	3.95836
5	no	?	C12NThinF_IM...	C12NThinF_IM...	16300	142	142	5.56819	4.68032	6.94896	1.38825
6	no	?	C12NThinF_IM...	C12NThinF_IM...	14708	112	133	6.05661	5.98544	9.48763	6.87201
7	no	?	C12NThinF_IM...	C12NThinF_IM...	18554	160	157	6.57516	5.22326	6.75507	5.27407
8	yes	?	C12NThinF_IM...	C12NThinF_IM...	15615	139	136	5.24349	5.14513	7.31061	5.59721
9	no	?	C12NThinF_IM...	C12NThinF_IM...	20702	136	202	6.03158	3.16494	7.43153	3.84377
10	no	?	C132P93ThinF_I...	C132P93ThinF_I...	11023	127	121	8.36537	7.96787	12.4164	9.66195
11	yes	?	C132P93ThinF_I...	C132P93ThinF_I...	10842	121	124	6.71261	4.36595	8.18024	6.27601
12	yes	?	C132P93ThinF_I...	C132P93ThinF_I...	13176	130	145	4.65979	3.64382	5.92758	4.28798
13	no	?	C132P93ThinF_I...	C132P93ThinF_I...	11698	130	130	4.72976	2.9572	6.36708	3.90999
14	yes	?	C132P93ThinF_I...	C132P93ThinF_I...	12003	130	127	5.88524	2.88279	7.5795	4.97306
15	no	?	C132P93ThinF_I...	C132P93ThinF_I...	12479	127	133	6.33196	3.96045	8.20761	5.62005
16	no	?	C132P93ThinF_I...	C132P93ThinF_I...	11401	121	118	7.17699	4.72547	8.71045	6.01288
17	no	?	C132P93ThinF_I...	C132P93ThinF_I...	10882	133	121	4.7741	2.15762	5.9266	2.6105
18	no	?	C132P93ThinF_I...	C132P93ThinF_I...	12358	139	139	4.80904	1.93055	6.56099	3.50308
19	yes	?	C132P93ThinF_I...	C132P93ThinF_I...	8684	106	115	5.82399	2.49324	5.6226	3.42
20	no	?	C132P93ThinF_I...	C132P93ThinF_I...	11255	121	121	5.56403	4.14397	6.87638	3.43136
21	no	?	C132P93ThinF_I...	C132P93ThinF_I...	11309	115	121	6.55175	5.08568	9.06897	5.45292
22	no	?	C132P93ThinF_I...	C132P93ThinF_I...	10110	115	118	6.19116	3.74363	7.45376	5.4205
23	yes	?	C132P93ThinF_I...	C132P93ThinF_I...	10644	121	124	5.36823	3.12949	7.74457	4.07264
24	no	?	C132P93ThinF_I...	C132P93ThinF_I...	10376	115	121	6.48776	3.97642	8.28429	5.86218
25	yes	?	C132P93ThinF_I...	C132P93ThinF_I...	10127	106	118	6.365	4.66446	9.10081	6.76921
26	yes	?	C132P93ThinF_I...	C132P93ThinF_I...	11080	127	121	5.49874	3.69723	7.49273	3.98883
27	no	?	C132P93ThinF_I...	C132P93ThinF_I...	12164	133	130	6.96471	5.66256	9.89593	6.54071
28	no	?	C132P93ThinF_I...	C132P93ThinF_I...	13882	133	136	4.90834	3.57589	6.95629	6.26947
29	no	?	C132P93ThinF_I...	C132P93ThinF_I...	13000	127	133	7.17405	4.83447	9.32415	6.73397
30	no	?	C132P93ThinF_I...	C132P93ThinF_I...	11725	127	121	5.62571	2.50865	7.21299	2.85371

Fig. 12 Prediction on Testing Cell Images

V. CONCLUSIONS

This research paper concluded that the logistic activation function for the neural network provided the best accuracy than others function. 93.50% classified accuracy obtained by using a logistic activation function. This research will be helpful to diagnose malaria fever by using many other deep learning algorithms and activation functions. This research can also be implemented in many other medical diseases to diagnose and lesser the work of doctors. As medical science is growing day by day so there will be many other advanced techniques that will help the researchers in the future.

REFERENCES

[1] V. Sharma, A. Kumar, D. Lakshmi Panat, and G. Karajkhede, "Malaria outbreak prediction model using machine learning," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 4, no. 12, 2015.

[2] M. Poostchi, K. Silamut, R. J. Maude, S. Jaeger, and G. Thoma, "Image analysis and machine learning for detecting malaria," *Translational Research*, vol. 194, pp. 36-55, 2018.

[3] S. Thakur and R. Dharavath, "Artificial neural network based prediction of malaria abundances using big data: A knowledge capturing approach," *Clinical Epidemiology and Global Health*, vol. 7, no. 1, pp. 121-126, 2019.

[4] T. Olayinka and S. Chiemeke, "Predicting Paediatric Malaria Occurrence Using Classification Algorithm in Data Mining," *Journal of Advances in Mathematics and Computer Science*, pp. 1-10, 2019.

[5] T. Sajana and M. Narasingarao, "Classification of Imbalanced Malaria Disease Using Naïve Bayesian Algorithm," *International Journal of Engineering & Technology*, vol. 7, no. 2.7, pp. 786-790, 2018.

[6] M. S. Alabdulrahman, "A Decision Support System (DSS) for Classification and Retrieval of Leukemia from Microscopic Blood Cell Images," Morgan State University, 2017.

[7] M. J. Sadiq and V. Balaram, "DTBC: decision tree based binary classification using with feature selection and optimization for malaria infected erythrocyte detection," *Int. J. Appl. Eng. Res*, vol. 12, no. 24, pp. 15923-15934, 2017.

[8] Y. Dong, Z. Jiang, H. Shen, and W. D. Pan, "Classification accuracies of malaria infected cells using deep convolutional neural networks based on decompressed images," in *SoutheastCon 2017*, 2017: IEEE, pp. 1-6.

[9] M. Masinde, "Africa's Malaria Epidemic Predictor: Application of Machine Learning on Malaria Incidence and Climate Data," in *Proceedings of the 2020 the 4th International Conference on Compute and Data Analysis*, 2020, pp. 29-37.

[10] S. Egieyeh, J. Syce, S. F. Malan, and A. Christoffels, "Predictive classifier models built from natural products with antimalarial bioactivity using machine learning approach," *PloS one*, vol. 13, no. 9, p. e0204644, 2018.