A Multi-Layer Perceptron (MLP) Neural Networks for Stellar Classification: A Review of Methods and Results

Ayman H. Abdel-aziem*, Tamer H. M. Soliman

Information System Department; October 6 University; Cairo, Egypt

Emails: Ayman.Hasanein.comp@o6u.edu.eg; tamer.hasan.comp@o6u.edu.eg

Abstract

The remarkable capacity of artificial intelligence (AI) to analyze enormous quantities of information and create precise forecasts has led to its growing prominence in the field of scientific Astrophysics. Stellar categorization is the process by which stars are sorted according to the characteristics revealed by their spectra. To analyze the star's electromagnetic radiation, a diffraction or prism screen separates it into a spectrum with an assortment of hues and spectral lines used to categorize the star. Star wavelengths are an extremely important piece of data for space-based photography studies. Employing data from over 100,000 cases and a variety of AI models, this study demonstrates how to categorize stellar properties as either a Galaxy or a Star. This paper used the multi-layer perceptron (MLP) neural network (NN) for stellar classification. The MLP is applied in 18 features. This paper showed the correlation between these features. This paper achieved 97% accuracy from the MLP model. This study compared various optimizers to show the best optimizer. The Adagrad optimizer is the best optimizer due to getting the highest validation accuracy.

Keywords: Neural Networks; Multi-Layer Perceptron (MLP); Stellar; Classification.

1. Introduction

Several sky survey innovations, including the Sloan Digital Sky Survey(SDSS), have been finished and put into service thanks to advancements in science and engineering, making it simpler to acquire data as well as knowledge regarding numerous celestial objects. The vast amounts of data gathered via observation and collecting may now be efficiently classified thanks to advances in artificial intelligence technologies[1], [2].

We need to employ the lowest feasible resolution (while still giving the necessary categorization accuracy) to detect dim objects. Therefore, we can wonder what the best resolution is for various spectral categorization jobs[3], [4]. Investigations of the Galaxy that attempt to learn more about its structure and history need a high degree of categorization precision. However, in certain cases, just very rough spectral types are required for the categorization job at hand. Exact astrometric observations in broad CCD fields at high zenith angles are a nice illustration of such a challenge. Correcting star locations for color-dependent reflections from the atmosphere may be done using just a crude spectroscopic type[5]–[7].

Categorization accuracies have been shown to vary with wavelength resolution in a variety of spectral categorization investigations. Seitter, using the Hoher List Observatory's three target prisms, obtained the most outstanding data from observations by doing qualitative MK assessments[4], [8].

Doi: https://doi.org/10.54216/IJAACI.030203
Received: August 18, 2022 Revised: December 19, 2022 Accepted: February 16, 2023

The importance of researching stellar characteristics and expanding our knowledge of the cosmos cannot be overstated. There have been numerous recent successes in this area of academic study. For instance, SVM, random forest, and other algorithms are used to create the basic classifier model in Stacking ensemble learning-based star or galaxy categorization study[9], [10].

The feed-forward NN is supplemented by the MLP. The input layer, the output layer, and the hidden layer make up its structure. The signal to be analyzed enters the network at the input layer. The output layer is responsible for completing the necessary action, such as forecasting or categorization. The real computing engine of the MLP is a finite number of hidden layers positioned across the input and output layers. Like a feed-forward network, an MLP has a single route of data flow, from the input to the output layer[11], [12]. The MLP uses the backpropagation learning technique to hone the abilities of its neuronal components. MLPs may estimate any continuous function and provide solutions to problems that cannot be separated into linear subproblems. Pattern categorization, acknowledgment, forecasting, and estimation are the main applications of MLP[13]–[16].

2. Steller Classification

Astronomers rely heavily on stellar classification because it gives them "systems" to use when comparing new sorts of stars. Cleverly separating "peculiar" items and learning more about the processes that produce "normal" objects may be accomplished with the help of a solid categorization system. There will come a time when the number of items in a certain "peculiar" class will be large enough to justify revising the definition of "normal" to include them. The long task of analysing all stars in detail may be shortened by isolating prototypes via categorization. As a result, astronomy relies heavily on the continual operations of star categorization and reference frame maintenance.

3. Multi-Layer Perceptron Neural Networks (MLP)

High skills in modeling the nonlinear behavior of complicated structures are found in MLP designs, which are built on the principles of the natural nervous system. In addition, these frameworks are well-suited for solving nonlinear forecasting issues. This approach gets its desired results by first learning the steps involved in addressing the issue and then discovering the underlying link between those steps. To do this, a large amount of data is employed during the training phase, and the resulting relationships are exploited to arrive at the correct output. There are several varieties of neural networks, but the back-propagation net is by far the most common. The neurons in each layer of the structure function in parallel with one another. Every successive layer is fully interconnected with its predecessor and successor[17]–[20].

The MLP is a supervised learning NN that uses the back-propagation algorithm. As can be seen in Figure 1, MLP functions best with a three-layer architecture consisting of an input layer, a hidden layer or layers, and an output layer or layers, where each neuron is linked to all the neurons in the layer that follows. The usefulness of MLP in non-linear issues has been widely discussed[21], [22].

Doi: https://doi.org/10.54216/IJAACI.030203
Received: August 18, 2022 Revised: December 19, 2022 Accepted: February 16, 2023

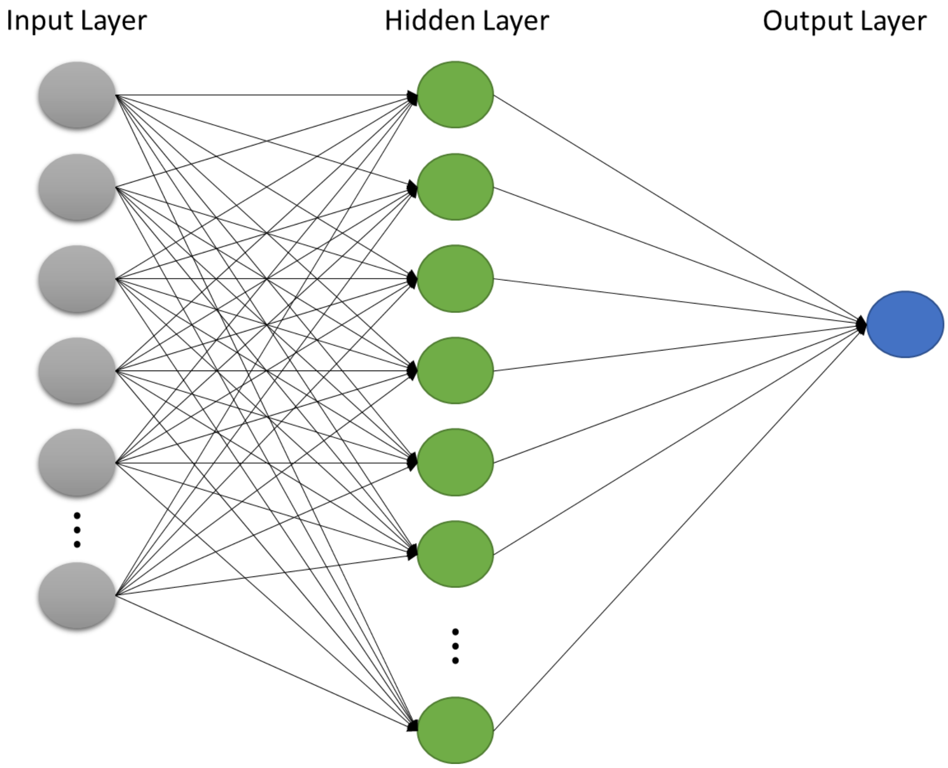


Figure 1: The design of MLP.

The calculation of output, input, and bias variables as:

$$C_i = \sum_{i=1}^{n} E_{ij} IN_i + B_i \quad (1)$$

Where $E_{ij}, IN_i, and\ B_i$ refer to the weights, input variables and bias variable.

The sigmoid function is an activation function for two target classes can be computed as:

$$S_i = \frac{1}{1+e^{c_j}} \quad (2)$$

The ultimate of output variable can be computed as:

$$O_i = S_i\left(\sum_{i=1}^{n} E_{ij} IN_i + B_i\right) \quad (3)$$

4. Application

The MLP is applied in the stellar classification dataset. This dataset has 18 features and 100000 observations. Table 1 shows the instance of first and last five records from this dataset. The first row shows the observations and first column shows the 18 features.

Table 1: Instance of staller dataset.

	STA1	**STA2**	**STA3**	**STA4**	**STA5**	**...**	**STA99995**	**STA99996**	**STA99997**	**STA99998**	**STA99999**
STC1	1.24E+18	1.24E+18	1.24E+18	1.24E+18	1.24E+18	...	1.24E+18	1.24E+18	1.24E+18	1.24E+18	1.24E+18
STC2	135.6891	144.8261	142.1888	338.741	345.2826	...	39.62071	29.49382	224.5874	212.2686	196.8961

Doi: https://doi.org/10.54216/IJAACI.030203
Received: August 18, 2022 Revised: December 19, 2022 Accepted: February 16, 2023

STC 3	32.49463	31.27419	35.58244	-0.40283	21.18387	...	-2.59407	19.79887	15.70071	46.66037	49.46464
STC 4	23.87882	24.77759	25.26307	22.13682	19.43718	...	22.16759	22.69118	21.16916	25.35039	22.62171
STC 5	22.2753	22.83188	22.66389	23.77656	17.58028	...	22.97586	22.38628	19.26997	21.63757	21.79745
STC 6	20.39501	22.58444	20.60976	21.61162	16.49747	...	21.90404	20.45003	18.20428	19.91386	20.60115
STC 7	19.16573	21.16812	19.34857	20.50454	15.97711	...	21.30548	19.75759	17.69034	19.07254	20.00959
STC 8	18.79371	21.61427	18.94827	19.2501	15.54461	...	20.73569	19.41526	17.35221	18.62482	19.28075
STC 9	3606	4518	3606	4192	8102	...	7778	7917	5314	3650	3650
STC 10	301	301	301	301	301	...	301	301	301	301	301
STC 11	2	5	2	3	3	...	2	1	4	4	4
STC 12	79	119	120	214	137	...	581	289	308	131	60
STC 13	6.54E+18	1.18E+19	5.15E+18	1.03E+19	6.89E+18	...	1.06E+19	8.59E+18	3.11E+18	7.60E+18	8.34E+18
STC 14	GALAXY	GALAXY	GALAXY	GALAXY	GALAXY	...	GALAXY	GALAXY	GALAXY	GALAXY	GALAXY
STC 15	0.634794	0.779136	0.644195	0.932346	0.116123	...	0	0.404895	0.143366	0.45504	0.542944
STC 16	5812	10445	4576	9149	6121	...	9374	7626	2764	6751	7410
STC 17	56354	58158	55592	58039	56187	...	57749	56934	54535	56368	57104
STC 18	171	427	299	775	842	...	438	866	74	470	851

This paper performed some descriptive statics on this dataset. These statistics are mean, standard deviation, minimum and maximum values in dataset. Table 2 shows these statistics.

Table 2: Some descriptive statistics on the stellar dataset.

	Count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
STC1	1.00E+05	1.24E+18	8.44E+12	1.24E+18	1.24E+18	1.24E+18	1.24E+18	1.24E+18
STC2	100000	177.6291	96.50224	0.005528	127.5182	180.9007	233.895	359.9998
STC3	100000	24.13531	19.64467	-18.7853	5.146771	23.64592	39.90155	83.00052
STC4	100000	21.98047	31.76929	-9999	20.35235	22.17914	23.68744	32.78139
STC5	100000	20.53139	31.75029	-9999	18.96523	21.09984	22.12377	31.60224
STC6	100000	19.64576	1.85476	9.82207	18.13583	20.12529	21.04479	29.57186
STC7	100000	19.08485	1.757895	9.469903	17.73229	19.40515	20.3965	32.14147
STC8	100000	18.66881	31.72815	-9999	17.46068	19.0046	19.92112	29.38374
STC9	100000	4481.366	1964.765	109	3187	4188	5326	8162
STC10	100000	301	0	301	301	301	301	301
STC11	100000	3.51161	1.586912	1	2	4	5	6
STC12	100000	186.1305	149.0111	11	82	146	241	989
STC13	1.00E+05	5.78E+18	3.32E+18	3.00E+17	2.84E+18	5.61E+18	8.33E+18	1.41E+19
STC14	100000	0.576661	0.730707	-0.00997	0.054517	0.424173	0.704154	7.011245

Doi: https://doi.org/10.54216/IJAACI.030203
Received: August 18, 2022 Revised: December 19, 2022 Accepted: February 16, 2023

STC15	100000	5137.01	2952.303	266	2526	4987	7400.25	12547
STC16	100000	55588.65	1808.484	51608	54234	55868.5	56777	58932
STC17	100000	449.3127	272.4984	1	221	433	645	1000

Then performed some analysis on the dataset to show the distribution of the dataset. Figure 2 shows the distribution of plate feature. The target class in this paper named class feature. Figure 3 shows the distribution of the target class. From Figure 3 there are 3 categories. The galaxy has observation more than QSO, and STAR.

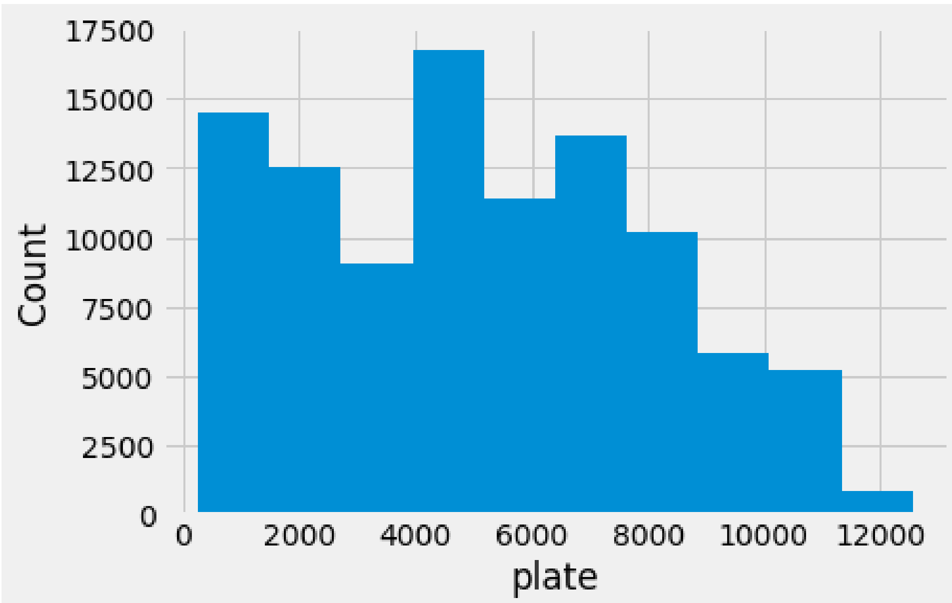


Figure 2: The distribution of the plate feature.

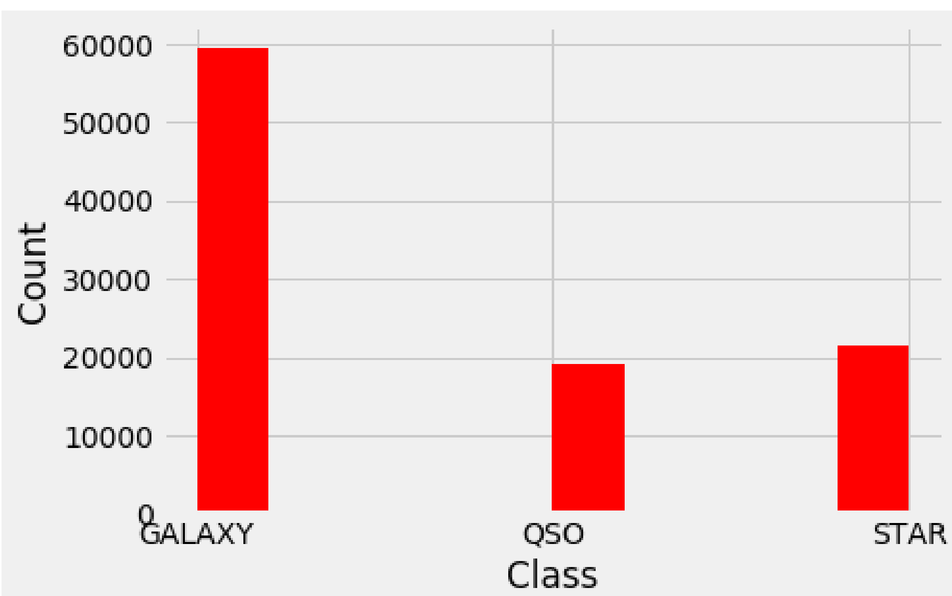


Figure 3: The distribution of the target class.

Doi: https://doi.org/10.54216/IJAACI.030203
Received: August 18, 2022 Revised: December 19, 2022 Accepted: February 16, 2023

Then the heatmap is obtained from the dataset to show the correlation between 18 features as shown in Figure 4. If the value of correlation is greater than 0.5 this is strong correlation.

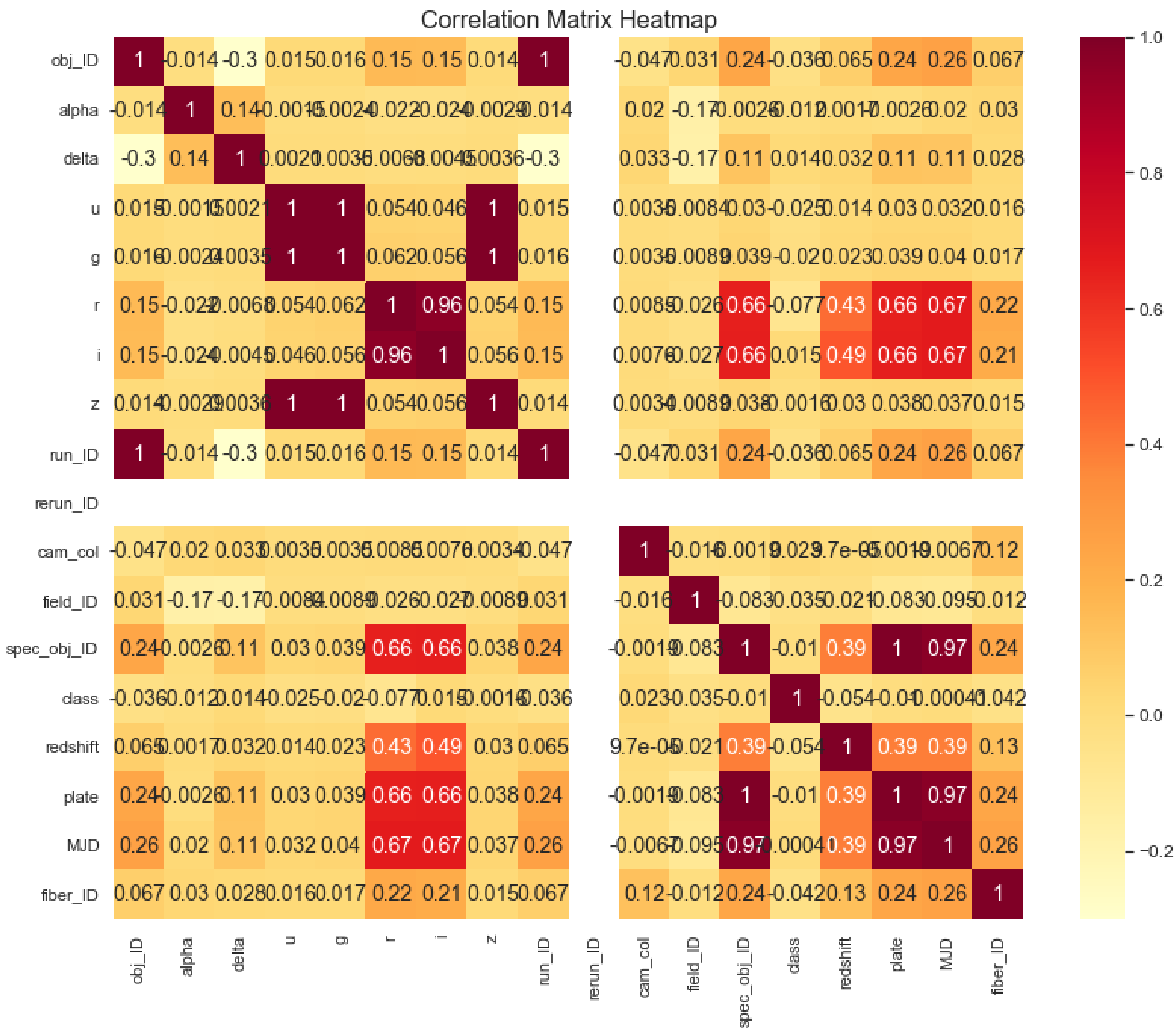


Figure 4: The correlation between 18 features.

Then MLP model is applied on the stellar classification dataset. Figure 5 shows the steps of the proposed MLP model. First, we divide the dataset into an 80% train size and 20% test size. Then we built three layers of dense and added three softmax activation function in three layers. Then we used the Adam optimizer and added 10 epochs, and batch size 64. Table 3 shows the accuracy of validation and train by using Adan optimizer.

Doi: https://doi.org/10.54216/IJAACI.030203
Received: August 18, 2022 Revised: December 19, 2022 Accepted: February 16, 2023

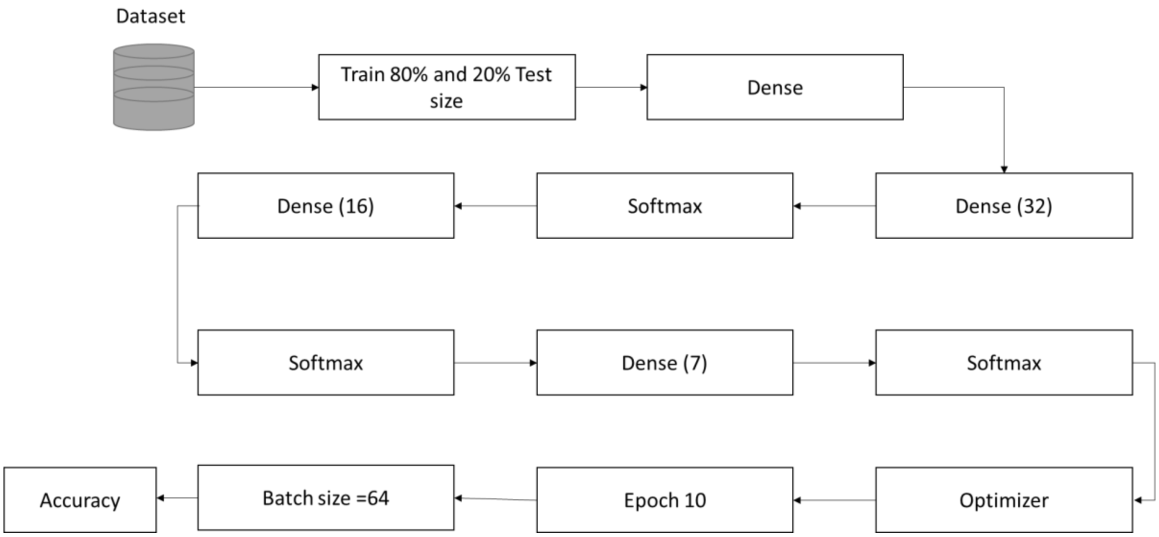


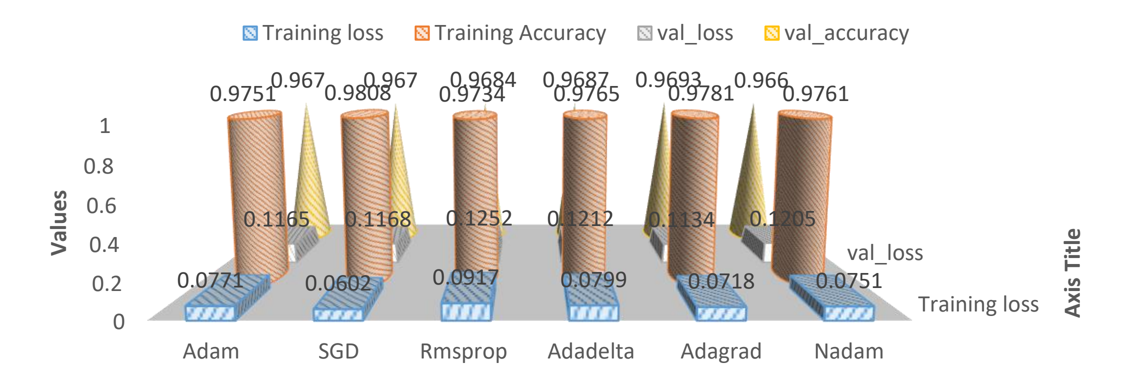
Figure 5: The steps of the MLP.

Then we changed the optimizer to show the accuracy of the model. We used six optimizers and get the results as shown in Table 3. In SGD optimizer the training loss is the lowest loss in all optimizer. The training accuracy in SGD optimizer is the greatest in all optimizers. The validation loss in Adagrad optimizer is the lowest loss. The greatest validation accuracy in all optimizer in a Adagrad optimizer. Figure 6 shows this comparison.

Table 3: Comparison between various optimizers.

Optimizers	Training loss	Training Accuracy	val_loss	val_accuracy
Adam	0.0771	0.9751	0.1165	0.967
SGD	**0.0602**	**0.9808**	0.1168	0.967
Rmsprop	0.0917	0.9734	0.1252	0.9684
Adadelta	0.0799	0.9765	0.1212	0.9687
Adagrad	0.0718	0.9781	**0.1134**	**0.9693**
Nadam	0.0751	0.9761	0.1205	0.966

Doi: https://doi.org/10.54216/IJAACI.030203
Received: August 18, 2022 Revised: December 19, 2022 Accepted: February 16, 2023



	Adam	SGD	Rmsprop	Adadelta	Adagrad	Nadam
Training loss	0.0771	0.0602	0.0917	0.0799	0.0718	0.0751
Training Accuracy	0.9751	0.9808	0.9734	0.9765	0.9781	0.9761
val_loss	0.1165	0.1168	0.1252	0.1212	0.1134	0.1205
val_accuracy	0.967	0.967	0.9684	0.9687	0.9693	0.966

Optimizers

Figure 6: Comparison between optimizers.

5. Conclusion

AI models have challenges, such as how to choose the best weights for each layer of a NN so that useful characteristics may be extracted from the input data and used to build a precise prediction. The best forecasting algorithm requires input data, which is an essential and helpful resource for estimating stellar classification. This paper applied the MLP model to a stellar classification dataset. This dataset has 18 features and 100000 samples. The MLP shows 97% with three layers. This paper compared the Adam optimizer with various optimizers to show the best optimizer in this dataset. From this comparison, the SGD is the best optimizer in training loss and training accuracy. The Adagrad optimizer is the best in the validation loss and validation accuracy.

References

[1] T. M. Brown, D. W. Latham, M. E. Everett, and G. A. Esquerdo, “Kepler input catalog: photometric calibration and stellar classification,” *Astron. J.*, vol. 142, no. 4, p. 112, 2011.

[2] T. Kuntzer, M. Tewes, and F. Courbin, “Stellar classification from single-band imaging using machine learning,” *Astron. Astrophys.*, vol. 591, p. A54, 2016.

[3] C. Damiani *et al.*, “Stellar classification of CoRoT targets,” *Astron. Astrophys.*, vol. 595, p. A95, 2016.

[4] J. W. Richards *et al.*, “Active learning to overcome sample selection bias: application to photometric variable star classification,” *Astrophys. J.*, vol. 744, no. 2, p. 192, 2011.

[5] C. A. L. Bailer-Jones, “Automated stellar classification for large surveys: a review of methods and results,” *arXiv Prepr. astro-ph/0102223*, 2001.

[6] D. I. Hoffman, T. E. Harrison, and B. J. McNamara, “Automated variable star classification using the Northern Sky Variability Survey,” *Astron. J.*, vol. 138, no. 2, p. 466, 2009.

[7] Z. Hosenie, R. Lyon, B. Stappers, A. Mootoovaloo, and V. McBride, “Imbalance learning for variable star classification,” *Mon. Not. R. Astron. Soc.*, vol. 493, no. 4, pp. 6050–6059, 2020.

[8] I. Becker, K. Pichara, M. Catelan, P. Protopapas, C. Aguirre, and F. Nikzat, “Scalable end-to-end recurrent neural network for variable star classification,” *Mon. Not. R. Astron. Soc.*, vol. 493, no. 2, pp. 2981–2995,

Doi: https://doi.org/10.54216/IJAACI.030203
Received: August 18, 2022 Revised: December 19, 2022 Accepted: February 16, 2023

2020.

[9] F. R. Lima-Junior and L. C. R. Carpinetti, "Predicting supply chain performance based on SCOR® metrics and multilayer perceptron neural networks," *Int. J. Prod. Econ.*, vol. 212, pp. 19–38, 2019.

[10] K.-C. Ke and M.-S. Huang, "Quality prediction for injection molding by using a multilayer perceptron neural network," *Polymers (Basel).*, vol. 12, no. 8, p. 1812, 2020.

[11] M. Hosseinzadeh *et al.*, "A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things," *J. Supercomput.*, vol. 77, pp. 3616–3637, 2021.

[12] N. Tamim, M. Elshrkawey, G. Abdel Azim, and H. Nassar, "Retinal blood vessel segmentation using hybrid features and multi-layer perceptron neural networks," *Symmetry (Basel).*, vol. 12, no. 6, p. 894, 2020.

[13] M. Desai and M. Shah, "An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN)," *Clin. eHealth*, vol. 4, pp. 1–11, 2021.

[14] E. B. Moustafa and A. Elsheikh, "Predicting Characteristics of Dissimilar Laser Welded Polymeric Joints Using a Multi-Layer Perceptrons Model Coupled with Archimedes Optimizer," *Polymers (Basel).*, vol. 15, no. 1, p. 233, 2023.

[15] W. Qiao, M. Khishe, and S. Ravakhah, "Underwater targets classification using local wavelet acoustic pattern and Multi-Layer Perceptron neural network optimized by modified Whale Optimization Algorithm," *Ocean Eng.*, vol. 219, p. 108415, 2021.

[16] S. Ghimire *et al.*, "Hybrid convolutional neural network-multilayer perceptron model for solar radiation prediction," *Cognit. Comput.*, vol. 15, no. 2, pp. 645–671, 2023.

[17] A. A. Ewees, M. Abd Elaziz, Z. Alameer, H. Ye, and Z. Jianhua, "Improving multilayer perceptron neural network using chaotic grasshopper optimization algorithm to forecast iron ore price volatility," *Resour. Policy*, vol. 65, p. 101555, 2020.

[18] P. Du *et al.*, "Hydrodynamic detection based on multilayer perceptron and optimization using dynamic mode decomposition," *Ocean Eng.*, vol. 278, p. 114258, 2023.

[19] Alber S. Aziz,Hoda K. Mohamed,Ahmed Abdelhafeez, Unveiling the Power of Convolutional Networks: Applied Computational Intelligence for Arrhythmia Detection from ECG Signals,International Journal of Advances in Applied Computational Intelligence, Vol. 1 , No. 2 , (2022) : 63-72 (Doi : https://doi.org/10.54216/IJAACI.010205)

[20] R. Ramezanian, A. Peymanfar, and S. B. Ebrahimi, "An integrated framework of genetic network programming and multi-layer perceptron neural network for prediction of daily stock return: An application in Tehran stock exchange market," *Appl. Soft Comput.*, vol. 82, p. 105551, 2019.

[21] N. I. M. Yusoff, D. I. Alhamali, A. N. H. Ibrahim, S. A. P. Rosyidi, and N. A. Hassan, "Engineering characteristics of nanosilica/polymer-modified bitumen and predicting their rheological properties using multilayer perceptron neural network model," *Constr. Build. Mater.*, vol. 204, pp. 781–799, 2019.

[22] J. Zhang, C. Li, Y. Yin, J. Zhang, and M. Grzegorzek, "Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer," *Artif. Intell. Rev.*, vol. 56, no. 2, pp. 1013–1070, 2023.

Doi: https://doi.org/10.54216/IJAACI.030203
Received: August 18, 2022 Revised: December 19, 2022 Accepted: February 16, 2023