Classification of ECG data using deep neural networking algorithms

Thomas Harrison - 35584689

***Abstract*—Using ResNet-50, long-short term memory and inception to perform classification on two ECG datasets demonstrated that there was not an overall best model. Each model was implemented with the 'Adam' optimiser and the 'SGD' optimiser to identify whether one optimiser produced higher classification scores than the other. Evaluating the first dataset showed that ResNet-50 (using SGD optimiser) and LSTM (using the Adam optimiser) had the highest prediction accuracy computing classification scores of 0.99 for precision, recall and f1-score. Res-Net had the lowest prediction accuracy on the first dataset using the Adam optimising producing classification metrics of 0.43, 0.74 and 0.26 for recall, precision and f1-score respectively showing potential signs of being 'stuck' on the local minima. Evaluating the second dataset demonstrated that the 'Adam' optimiser was more suited to this data, the classification metrics generated by the three algorithms using the 'Adam' optimiser were all higher than the classification metrics produced when the models used the 'SGD' optimiser. ResNet-50 and the Inception model both produced classification metrics of 0.98 for recall, precision and f1-score.**

***Index Terms*—ResNet, Inception, LSTM, Electrocardiogram (ECG) classification, Neural networks, Machine learning**

I. Introduction

Electrocardiogram (ECG) classification is a critical process for diagnosis in abnormal heartbeats. Detecting abnormal heartbeats is key in diagnosing early onset heart disease; these readings may be difficult to a medical practitioner to detect without any aid from a program [1]. This leads to the introduction of deep neural networks (DNNs), these tend to perform with higher accuracy for predictions and classification than traditional machine learning algorithms for larger amounts of data by embedding multi-layer architecture and eliminates the need for any feature extraction methods.

To determine the most suitable neural network for the ECG dataset three different types of neural networks were implemented, each had their own benefits and are structured differently to approach different issues. The first neural network was long-short term memory (LSTM) which is a type of reccural neural network which addressed the vanishing gradient problem, this type of neural network is effective at 'remembering' information of a series of timesteps making it a suitable choice for ECG data as it is a measurement of electrical activity of the heart over time. The second neural network explored is ResNet-50, a convolutional neural network which introduced the 'skip' connection which addressed the issue with decreased performance over a large amount of layers. The last neural network was Inception, this model addressed the issue with overfitting by creating each layer with multiple filters. Each of these algorithms were implemented using two different optimisers, 'Adam' and 'SGD' to explore how these effect the final accuracy. The accuracy and loss were measured for each algorithm using each optimiser to perform a comparison between the optimisers to determine the optimal choice for the dataset.

Two different datasets were used to test the model, both datasets contained one-dimensional samples with a final label of '1' and '0' to represent abnormal and normal heartbeats respectively. This data is trained using a 30:70 test train slit on the above models and classified as abnormal and normal ECG reading. Three different classification metrics were evaluated (recall, precision and f1-score) for each algorithm to compare the performance of each neural network as well as the approximate training time for each epoch.

II. Literature review

Previous studies on ECG data by Cláudia Brito et al used a learning model based of Res-Net architecture which achieved 96% accuracy using the stochastic gradient descent optimiser as opposed to an 83% accuracy using the 'Adam' optimiser [2] . These results suggest that ResNet architecture has high performance in classifying ECG data and suitable optimiser choices will be 'Adam' and 'SGD'. Similar results were seen in Antônio H. Ribeiro et al , which used an adapted convolution neural network similar to a residual networks, which achieved classification metric scores of above 80% for the f1-score utilising the 'Adam' optimiser [3]. Other convolutional neural networks have been utilised in Edward B.Panganiban et al , they used Google's Inception V3 model to train the final CNN layer and tested there data using binary and quinary classification achieving high accuracy scores of 98.73% and 97.33% which suggest Inception is a neural network that should be considered in this experiment [4]. It is important to chose the correct optimiser to achieve the highest performing model, the optimiser links the loss function to the model parameters by updating the model based on the output of the loss function to minimise the loss.

Further research into convolutions neural networks produce high quality ECG classification results [5] [6]. Sherin M.Mathews et al utilised an eleven layered convolution neural network achieving 93% accuracy across the ECG dataset [7]. S.Kiranyaz et al developed a a three layer convolutional neural network to study patient specific ECG to measure the R-peak wave, this produced results with high accuracy but a long training time [8]. P.Rajpurkar et al trained a 34-layered convolutional neural network to map a sequence of ECG samples to a sequence of rhythm classes which produced evaluation metrics of precision and recall which exceeded the average cardiologist performance [9].

Reccural networks have an advantage over convolutional neural networks (CNN) since CNNS are more suited to objects identification and classification; reccural neural networks (RNN) can easily process 1-D data such as sensor signals. The main issues with RNNs found when training time-based data was the vanishing gradient problem. However, the long-short term memory algorithm resolved this issue making it suitable for ECG classification [10] [11]. Similar studies have been done relating the use of LSTM based ECG classification for continuous real-time monitoring on patients, the use of a RNN was selected since the ECG waveform reflects temporal dependencies in sequential data which are captured more easily in LSTM compared to other neural networks [12].

III. Description of the dataset

Two different ECG datasets were used, the first dataset is the one provided and the second data set contains 14552 samples taken at a sampling frequency of 125Hz with each sample representing a set of ECG readings [13]. Each of these have an associated label of '0' and '1' which corresponds to normal heart rate and abnormal heart rate respectively.Each dataset was split into a 30:70 test train split

Each algorithm contained a final layer using the activation function 'softmax' to calculate predicted probabilities for each labels using a dense state with two output neurons.

IV. Deep neural networks

A. Optimisation parameters

To fully evaluate the possible performance of each data set two different optimasation parameters were used, 'Adam' and 'SGD'.

'Adam' contains the benefits of 'Adaptive gradient descent (AdaGrad)' which means that it is effective in settings with sparse gradients but struggles to train the model effectively in non-convex optimisations. 'Adam' also contains the benefits of RMSprop which accounts for some issues with Adagrad and is effective with on-line settings.

$$m_t = (\beta_1 \times m_{t-1}) + (1-\beta_1) \times g_t \tag{1}$$

$$v_t = \beta_2 \times v_{t-1} + (1-\beta_2) \times g_t^2 \tag{2}$$

$$w_t = w_{t-1} - n\frac{\hat{m_t}}{\sqrt{\hat{v_t}} + \epsilon} \tag{3}$$

'Adam' calculates the individual learning rates for different parameters by using the first and second moments (m_t and v_t) of the gradient to adjust the learning rate for each moment [14]. The weight update (w_{t+1}) is computed using the first moment which is normalised by the second moment [15].

Whilst 'Adam' is a strong performing optimisation parameter there are some disadvantages associated with it, in some areas 'Adam' does not converge to an optimal solution where stochastic gradient descent with momentum performs strongly. Based on these results the second optimisation parameter 'SGD' was included and evaluated alongside 'Adam' [16].

Stochastic gradient descent is given by

$$\theta_{k+1} = \theta_k - (\alpha \times \nabla l_{ik}(\theta_k)) \tag{4}$$

For a sample (x_{ik},y_{ik}) at iteration k. The stochastic gradient ($\nabla l_{ik}(\theta_k)$) is used to update θ SGD optimise objective functions based on their gradient or curvature information using only a few samples. Each update in SGD is computed using a small sample of training data or a minibatch as opposed to a single example since this reduces the variance in the parameter update which will lead to a more stable convergence.

B. Long short-term memory (LSTM) networks

Reccural neural networks (RNN) have downsides with short-term memory, if a sequence is too long they can not effectively carry information from earlier time steps. This occurs when the gradient value increases or decreases exponentially. LSTM is a type of RNN which is developed to learn long term dependencies. Long-short term memory networks are well suited to classifying and making predictions on time-series data and were developed to counter the vanishing gradient problem when training traditional reccurent networks through the introduction of the 'input', 'output' and 'forget' gate. LSTM allows the algorithm to forget according to the dependencies in the problem. This makes LSTM a suitable choice to classify the ECG dataset.

The structure of the LSTM used to classify the ECG datasets was a single layer neural network with 64 neurons, A dropout of 0.3 was added after the first layer.

C. ResNet-50

ResNet-50 is a fifty layer deep convolution neural network. Residual networks were proposed to overcome the problems of VGG styled CNNs where stacking convolutions layers to increase the depth of the model does not guarantee an increase in accuracy but after a certain depth the accuracy will decrease. To approach this problem; ResNet was developed which introduced a 'skip connection' [17], instead of stacking convolutional layers on top of one another the original input is added to the output of the convolution block. By introducing this 'skip connection' it mitigates the vanishing gradient problem by allowing an alternate shortcut path for the gradient to flow through and allows the model to learn an identity function which ensures the higher layer of the algorithm will perform better than the lower layer.

Convolution layer	output size	ResNet-50
conv1	112 × 112	1 × 7, 64 strides =2
Conv2	56 × 56	3 × 3 max pool, stride = 2 $3 \times \begin{bmatrix} 1\times 1 & 64 \\ 3\times 3 & 64 \\ 1\times 1 & 256 \end{bmatrix}$
Conv3	28 × 28	$4 \times \begin{bmatrix} 1\times 1 & 128 \\ 3\times 3 & 128 \\ 1\times 1 & 512 \end{bmatrix}$
Conv4	14 × 14	$6 \times \begin{bmatrix} 1\times 1 & 256 \\ 3\times 3 & 256 \\ 1\times 1 & 1024 \end{bmatrix}$
Conv5	7 × 7	$3 \times \begin{bmatrix} 1\times 1 & 512 \\ 3\times 3 & 512 \\ 1\times 1 & 2048 \end{bmatrix}$

TABLE I: Baseline architechture used for performing ResNet-50.

D. Inception

Standard convolutional neural networks are constructed by stacking multiple convolution layers causing the model to become deeper however, these types of models are prone to overfitting.

Neural networks are constructed with multiple layers; as the networks becomes progressively deeper they are prone to overfitting causing it to be difficult to pass gradient updates throughout the entire network. To overcome this issue, the inception model is constructed with filters of multiple sizes alongside max pooling. The outputs of these filters are concatenated and sent to the next inception model. However, using the outputs of the pooling layer alongside the outputs of the convolution layers would lead to increasingly large computation times as the number of layers increases. To overcome this issue, an inception model with dimensionalilty reduction was used; prior to computing the 3 × 3 and 5 ×5 convolution layers there is a 1 × 1 convolution layer.

V. Classification results

Dataset	Neural network	Recall	precision	f1-score
Ecg	LSTM	0.99	0.99	0.99
	ResNet-50	0.43	0.74	0.26
	Inception	0.99	0.99	0.99
Ecg 2	LSTM	0.95	0.95	0.95
ECG 2	ResNet-50	0.98	0.98	0.98
ECG 2	Inception	0.98	0.98	0.98

TABLE II: Weighted classification metrics for each model performed on the datasets using adam optimiser.

Dataset	Neural network	Recall	precision	f1-score
Ecg	LSTM	0.97	0.97	0.97
	ResNet-50	0.099	0.99	0.99
	Inception	0.78	0.78	0.78
Ecg 2	LSTM	0.0.47	0.49	0.48
ECG 2	ResNet-50	0.94	0.94	0.94
ECG 2	Inception	0.78	0.78	0.77

TABLE III: Weighted classification metrics for each model performed on the datasets using SGD optimiser.

Dataset	Model	Optimiser	Computation time per epoch/s
ECG	LSTM	Adam	4
ECG	LSTM	SGD	4
ECG	ResNet	Adam	280
ECG	ResNet	SGD	280
ECG	Inception	Adam	290
ECG	Inception	SGD	290
ECG 2	LSTM	Adam	17
ECG 2	LSTM	SGD	17
ECG 2	ResNet	Adam	867
ECG 2	ResNet	SGD	800
ECG 2	Inception	Adam	800
ECG 2	Inception	SGD	830

TABLE IV: Computation time per epoch in seconds for each algorithm and the corresponding optimiser.

A. Performance of the inception model

Figure 1a and figure 1b demonstrate the difference between the performance of the inception model when trained with the 'Adam' optimiser and 'SGD' optimiser. Figure 1a shows that the 'Adam' optimiser had a consistantly higher accuracy than the 'SGD' optimiser across all epochs however, after the first epoch the ECG dataset utilising the 'Adam' optimiser levelled meaning it showed minimal alteration in accuracy; the ECG dataset using 'SGD' continued to increase suggesting that if further epochs were included they may have converged with the same overall accuracy score. A range of 10 epochs were used for computational purposes.

Figure 2a and figure 2b demonstrate the performance of each optimiser which implemented with the second ECG dataset. These results are similar to those found with the first ECG dataset; the 'Adam' optimiser had the highest accuracy across all epochs.

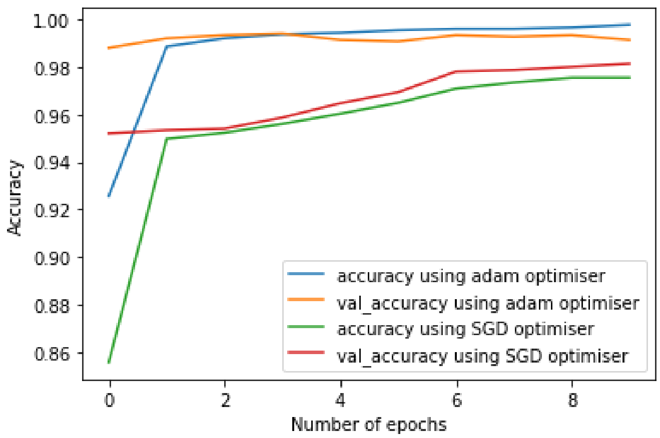
B. Performance of the ResNet model

Figure 3a and figure 3b demonstrate that the 'Adam' optimiser was not a suitable choice since the validation accuracy and loss remain constant across all epochs. This suggests that the 'Adam' optimiser is not suited for the first ECG dataset. The validation accuracy remaining constant throughout all epochs is a potential sign that the model became 'stuck' on the local minima and cannot accurately train the dataset. To further test this theory the learning rate should be adjusted since a learning rate that is too low causes weight updates that are too small to sufficiently train the model. Those issues are not seen with the 'SGD' optimiser, The validation accuracy and accuracy converge after 6 epochs and produced high classification results as shown in table III which demonstrates that ResNet predicts accurately with the first ECG data when accompanied with the SGD optimiser.

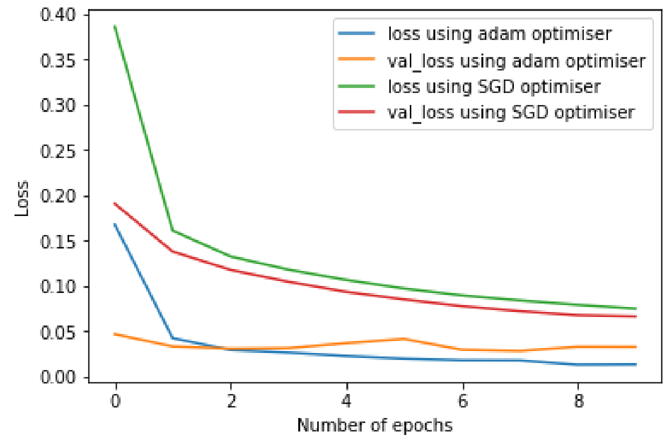
The opposite was seen when ResNet-50 was trained on the second ECG dataset, the use of the 'Adam' did not demonstrate the same issues seen with the first ECG data set and performed with a higher accuracy the 'SGD' as show in figure 4a.

C. Performance of the Lost-short term memory model

Figure 5a shows that, for the first dataset, the 'Adam' optimiser outperforms the SGD optimiser. The accuracy score across all of the epochs is higher when using the 'Adam'

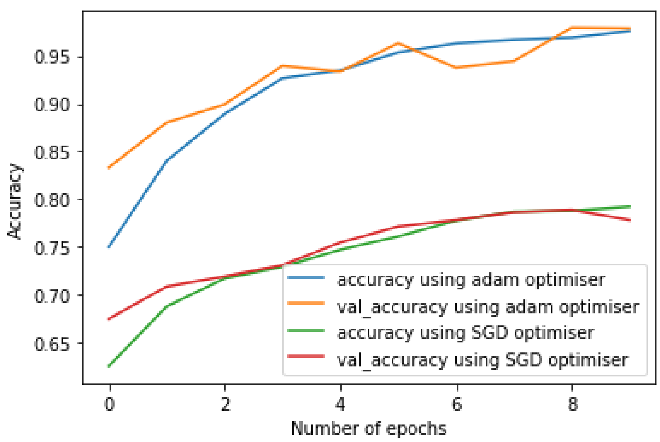


(a) Accuracy and validation accuracy of the inception model using the ECg dataset over 10 epochs.

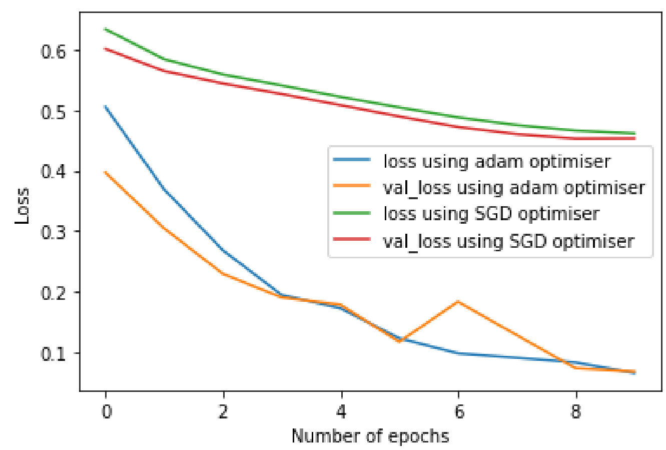


(b) Validation loss and loss of the inception model using the ECG dataset over 10 epochs.

Fig. 1: Accuracy and loss values for the inception model for the first ECG dataset over 10 epochs.



(a) Accuracy and validation accuracy of the inception model using the second ECG dataset over 10 epochs.



(b) Validation loss and loss of the inception model using the second ECG dataset over 10 epochs.

Fig. 2: Accuracy and loss values for the inception model for the second ECG dataset over 10 epochs.

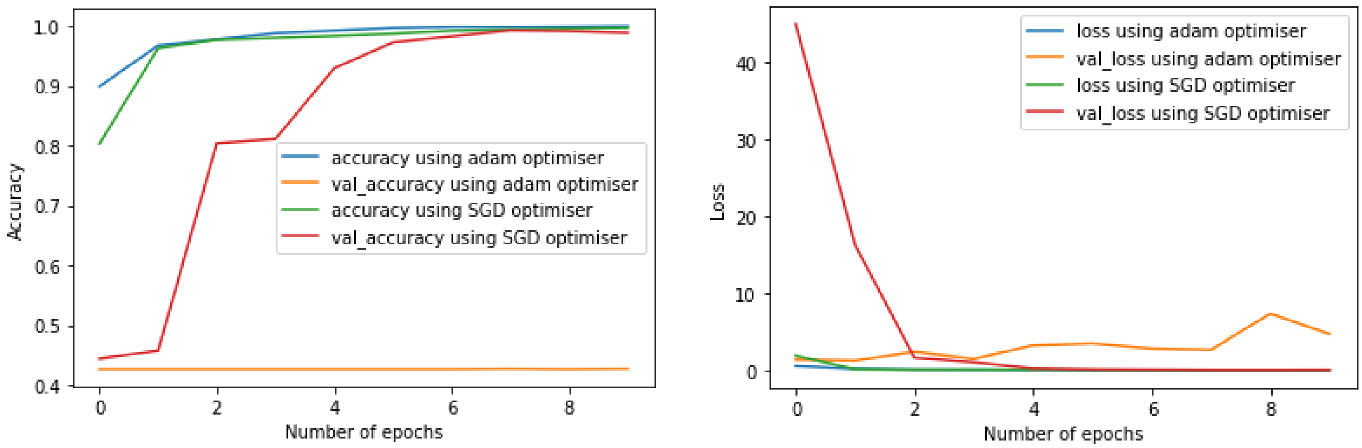
optimiser over the ‘SGD’ optimiser. Similar results were seen when comparing the loss values in figure 5b, when utilising the ‘Adam’ optimiser there was a significant decrease in loss across all epochs.

Figure 6a and figure 6b demonstrate that the stochastic gradient descent optimiser is not a suitable choice for the second dataset. Figure 6a shows no change in accuracy using the SGD across all epochs; it remains at a constant accuracy score of 0.6 likewise figure 6b shows no change in loss. This issue may be a result of the depth of the LSTM model, if the model is too ’shallow‘ (a small number of layers) then the model can not sufficiently learn the data and become ’stuck‘ on a local minima.

Comparing the classification metrics for these models as shown in table II and table III demonstrates that the ‘SGD’ optimiser performed slightly worse than the ‘Adam’ optimiser on the first dataset however, it produced significantly lower classification metrics on the second dataset.

VI. Comparison of Neural Network Models

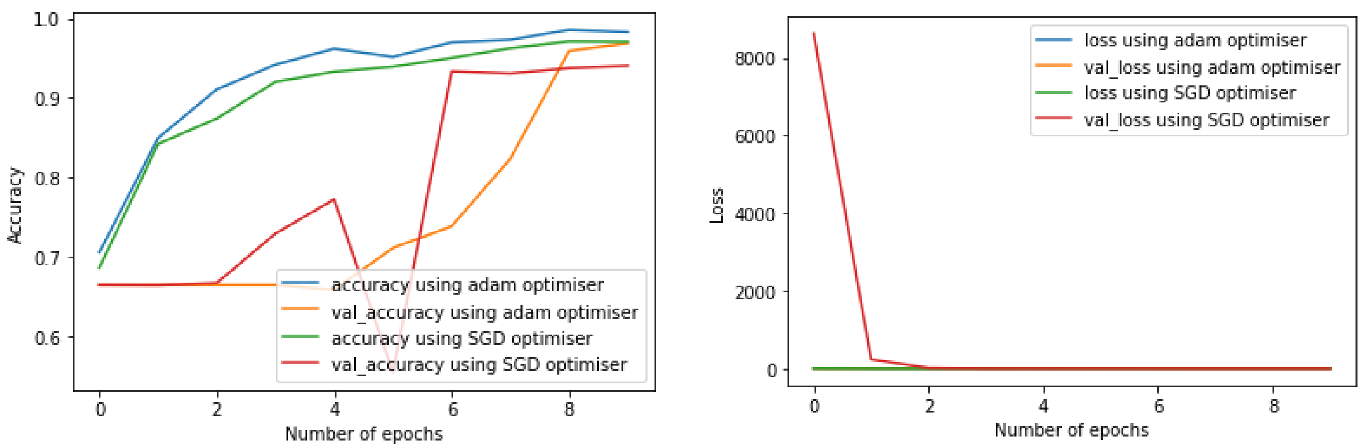
To gain an insight as to which would be the most efficient neural network to model ECG data, the overall accuracy was compared alongside the computation time per epoch. The approximate computation time per epoch is shown in table IV; ResNet-50 and the Inception model are the most computationally heavy with the inception model taking approximately 4% longer per epoch. The LSTM was the most computationally efficient however, this model contained far less layers than ResNet-50 and LSTM. Comparing the computation times to the classification scores in table II and table III shows that all three models were inconsistent on the classification metrics and depended strongly on the optimiser. The most consistent model, whilst being the most computationally heavy, was the



(a) Accuracy and validation accuracy of the ResNet-50 model on the ECG dataset over 10 epochs.

(b) Loss and validation loss of the ResNet-50 model on the ECG dataset over 10 epochs.

Fig. 3: Accuracy and loss values for the ResNet model for the first ECG dataset over 10 epochs.



(a) Accuracy and validation accuracy of the ResNet-50 model on the ECG dataset over 10 epochs.

(b) Loss and validation loss of the ResNet-50 model on the ECG dataset over 10 epochs.

Fig. 4: Accuracy and loss values for the ResNet model for the second ECG dataset over 10 epochs.

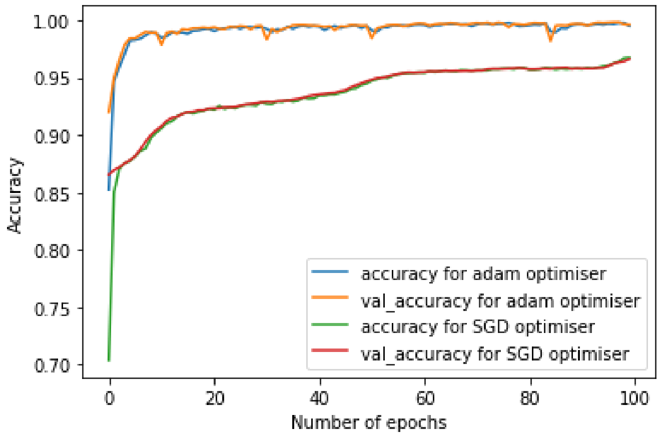
inception model which produced reasonably high classification metrics across all cases.

VII. Conclusion

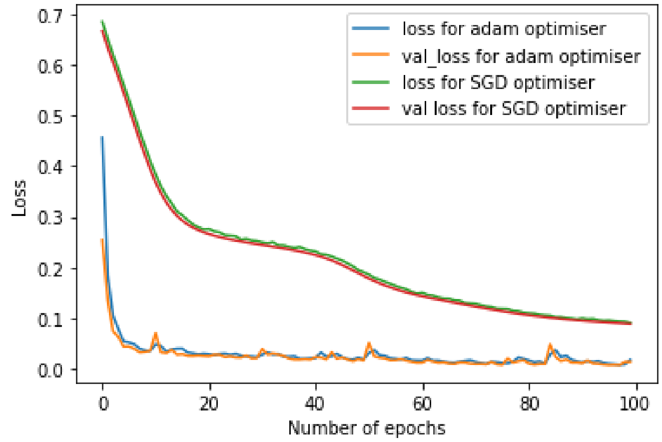
Three different neural network algorithms were constructed to perform classification on two different ECG datasets to classify heartrates into normal and abnormal cases; these algorithms were a reccural neural network LSTM, a residual neural network ResNet-50 and a convolutional neural network Inception. Each of these algorithms were evaluated using two different optimsation parameters (‘Adam’ and ‘SGD’) over a pre-set number of epochs (100 epochs for LST and 10 for Res-net and LSTM).

After performing classification on the first dataset, LSTM and Inception produced the highest classification metrics for recall, precision and F1-score with values of 0.99 utilising the ‘Adam’ optimiser. ResNet was not able to perform classification accurately with the results resembling possibilities of being ‘stuck’ on a local minima. A second set of classification results were taken on the first ECG dataset again but each algorithm implemented the ‘SGD’ opitimiser and the classification results showed accurate predictions throughout all three algorithms. LSTM produced accurate results of 0.97 for precision, recall and f1-score; these results are promising given the simple architecture and diminutive computation time in comparison to ResNet and Inception.

Performing classification on the second ECG dataset showed that ResNet was the best suited model in this case; ResNet produced classification metrics of 0.99 for recall, precision and f1-score utilising the ‘Adam’ optimiser and classification metrics of 0.94 utilising the ‘SGD’ optimiser. LSTM could not produce accurate results for the ‘SGD’ optimiser

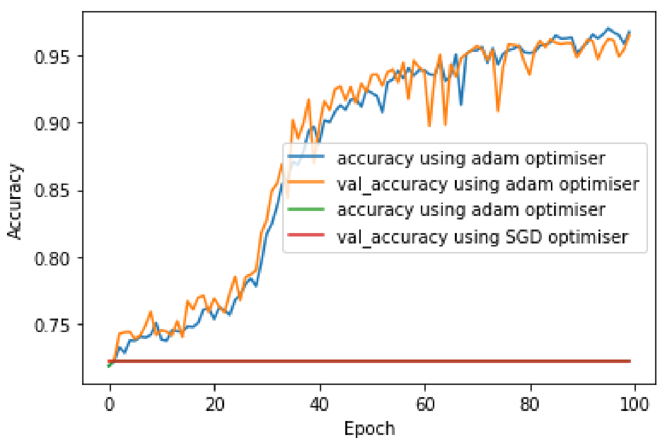


(a) Validation accuracy and accuracy of the LSTM model on the ECG dataset over 100 epochs.

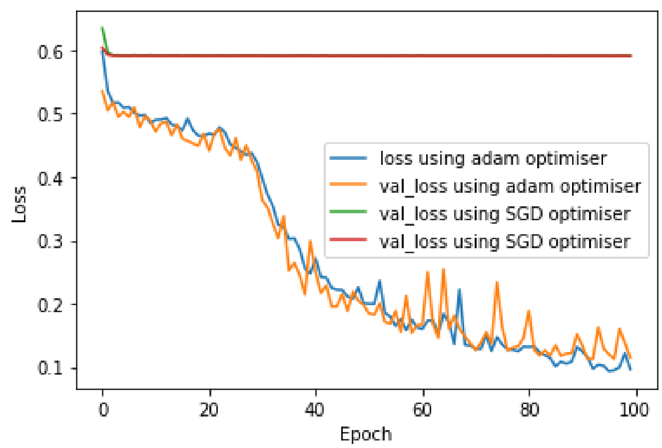


(b) Loss and validation loss of the LSTM model on the ECG dataset over 100 epochs.

Fig. 5: Accuracy and loss values for the LSTM model for the first ECG dataset over 100 epochs.



(a) Validation accuracy and accuracy of the LSTM model on the second ECG dataset over 100 epochs.



(b) Loss and validation loss of the LSTM model on the second ECG dataset over 100 epochs.

Fig. 6: Accuracy and loss values for the LSTM model for the second ECG dataset over 100 epochs.

demonstrating potential issues with being 'stuck' on a local minima. LSTM produced high classification metrics utilising the 'Adam' optimiser generating classification metrics of 0.95 for precision, recall and f1-score.

Despite the simplicity of the developed LSTM model the classification results show that it is a strong performing model in comparison to the more complex developed ResNet and Inception model. Further work on these would include adding addition layers to LSTM to test for further accuracy and implement larger higher initial learning rates to attempt to account for the local minima issue found in some circumstances.

References

[1] S. Chauhan and L. Vig, "Anomaly detection in ecg time signals via deep long short-term memory networks," in *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2015, pp. 1–7.

[2] C. Brito, A. Machado, and A. Sousa, "Electrocardiogram beat-classification based on a resnet network," vol. 264, 08 2019, pp. 55–59.

[3] A. H. Ribeiro, M. H. Ribeiro, G. M. M. Paixão, D. M. Oliveira, P. R. Gomes, J. A. Canazart, M. P. S. Ferreira, C. R. Andersson, P. W. Macfarlane, W. Meira, and et al., "Author correction: Automatic diagnosis of the 12-lead ecg using a deep neural network," *Nature Communications*, vol. 11, no. 1, 2020.

[4] E. B. Panganiban, A. C. Paglinawan, W. Y. Chung, and G. L. S. Paa, "Ecg diagnostic support system (edss): A deep learning neural network based classification system for detecting ecg abnormal rhythms from a low-powered wearable biosensors," *Sensing and Bio-Sensing Research*, vol. 31, p. 100398, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214180421000039

[5] M. Zubair, J. Kim, and C. Yoon, "An automated ecg beat classification system using convolutional neural networks," in *2016 6th International Conference on IT Convergence and Security (ICITCS)*, 2016, pp. 1–5.

[6] Y. Liang, S. Yin, Q. Tang, Z. Zheng, M. Elgendi, and Z. Chen, "Deep learning algorithm classifies heartbeat events based on electrocardiogram signals," *Frontiers in Physiology*, vol. 11, 2020.

[7] S. M. Mathews, C. Kambhamettu, and K. E. Barner, "A novel application of deep learning for single-lead ecg classification," *Computers in Biology and Medicine*, vol. 99, pp. 53–62, 2018. [Online]. Available:

https://www.sciencedirect.com/science/article/pii/S0010482518301264

[8] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ecg classification by 1-d convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2016.

[9] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," *arXiv preprint arXiv:1707.01836*, 2017.

[10] H. Palangi, L. Deng, Y. Shen, J. Gao, X. He, J. Chen, X. Song, and R. Ward, "Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 4, p. 694–707, 2016.

[11] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder–decoder for statistical machine translation," *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014.

[12] S. Saadatnejad, M. Oveisi, and M. Hashemi, "Lstm-based ecg classification for continuous monitoring on personal wearable devices," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 2, p. 515–523, 2020.

[13] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "Ecg heartbeat classification: A deep transferable representation," *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, Jun 2018. [Online]. Available: http://dx.doi.org/10.1109/ICHI.2018.00092

[14] Y. Wang, P. Zhou, and W. Zhong, "An optimization strategy based on hybrid algorithm of adam and sgd," *MATEC Web of Conferences*, vol. 232, p. 03007, 01 2018.

[15] B. Ginsburg, P. Castonguay, O. Hrinchuk, O. Kuchaiev, R. Leary, V. Lavrukhin, J. Li, H. Nguyen, Y. Zhang, J. M. Cohen, and et al., *Training Deep Networks with Stochastic Gradient Normalized by Layerwise Adaptive Second Moments*, Feb 2020.

[16] N. S. Keskar and R. Socher, "Improving generalization performance by switching from adam to sgd," 2017.

[17] H. Ahn and C. Yim, "Convolutional neural networks using skip connections with layer groups for super-resolution image reconstruction based on deep learning," *Applied Sciences*, vol. 10, no. 6, p. 1959, 2020.