Automating Music Transcription with Neural Networks

PRESENTERS:
Ashton Doane, Robert Spencer, Sunny Wang

BACKGROUND

Our project presents a new method of transcribing music through the use of neural networks. The results of our findings serve as both a useful tool for musicians and a proof of concept for the revolutionary results that could be obtained from machine learning.

METHODS

We used the MAPS Dataset in order to train two models, the first of which was a deep neural network consisting of four fully connected layers and using the sigmoid activation function. The second model was a recurrent neural network using two LSTM layers and two fully connected layers using softmax. Each model would take in a size 229 array created from mel spectrogram data and output a size 88 one hot array representing every note on a piano.Multiple iterations of each model were created, and the best performing architectures were selected.

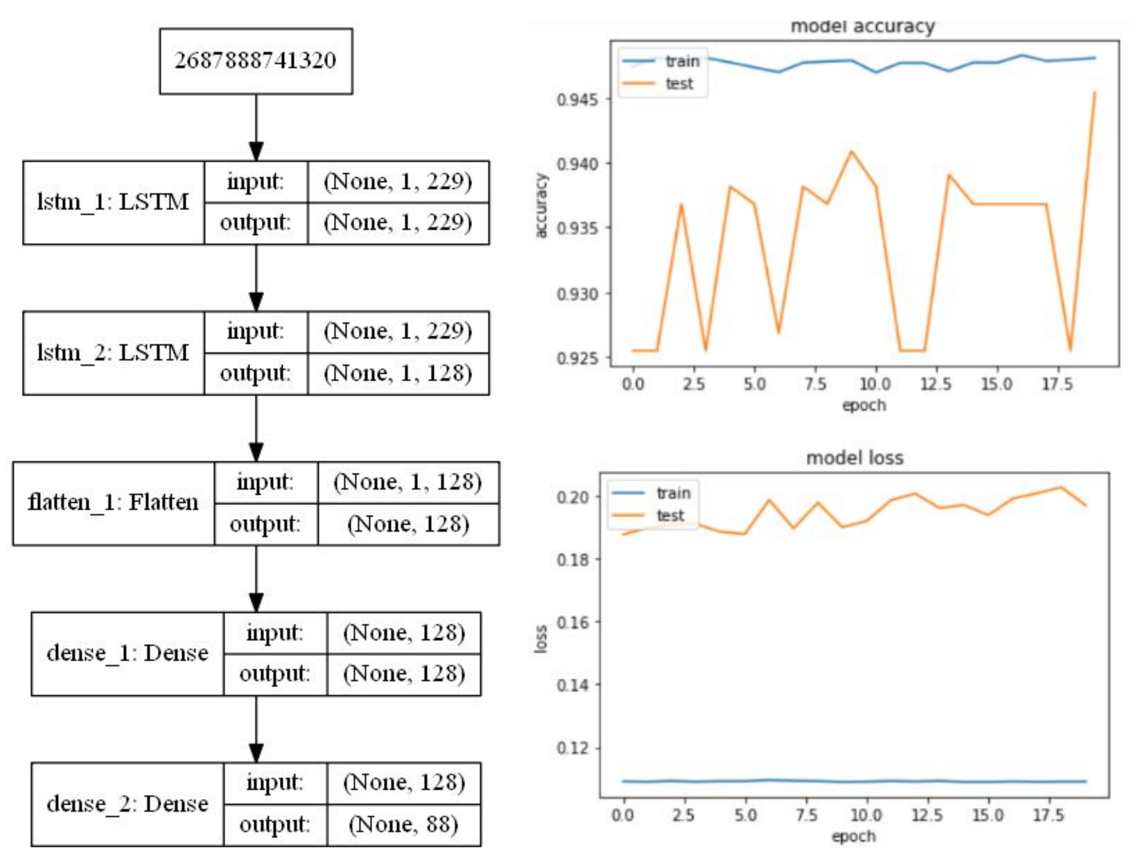
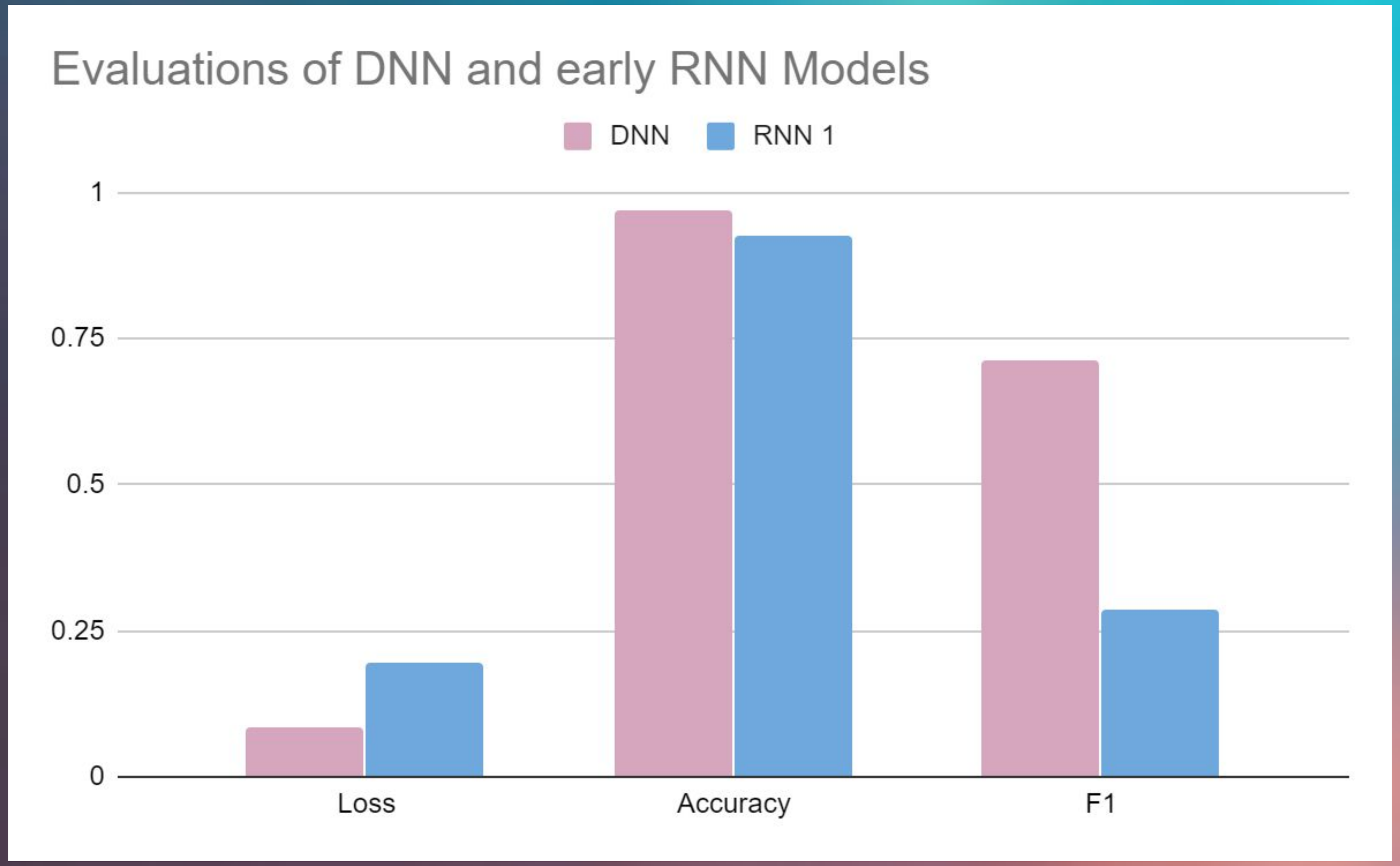


Figure 1: Training and Validation data (right) of our RNN model (left)

RESULTS

After both neural network models were trained through 20 epochs, the DNN received an overall F1 score of 71%, while the RNN received a score of 29%. When comparing per note accuracy on the test set, the DNN correctly identified 96.8% of the notes, while the RNN identified 95.5%.

Music transcription can be done **efficiently** and **accurately** with the use of **neural networks.**



Take a picture to **download** the **full paper**

DISCUSSION

Currently, the RNN has performed with subpar results and lies below the base level DNN model in terms of accuracy and f1 score. However, odd fluctuations during our validation set lead us to believe that there is an error to be fixed in our new model that will greatly improve it.

CONCLUSION

Our results have shown that a conventional deep neural network yields superior results when compared to a recurrent LSTM neural network when trained and tested on the same dataset. Overall, machine learning shows promise in the field of music transcription given the relatively high accuracies of our neural network models.

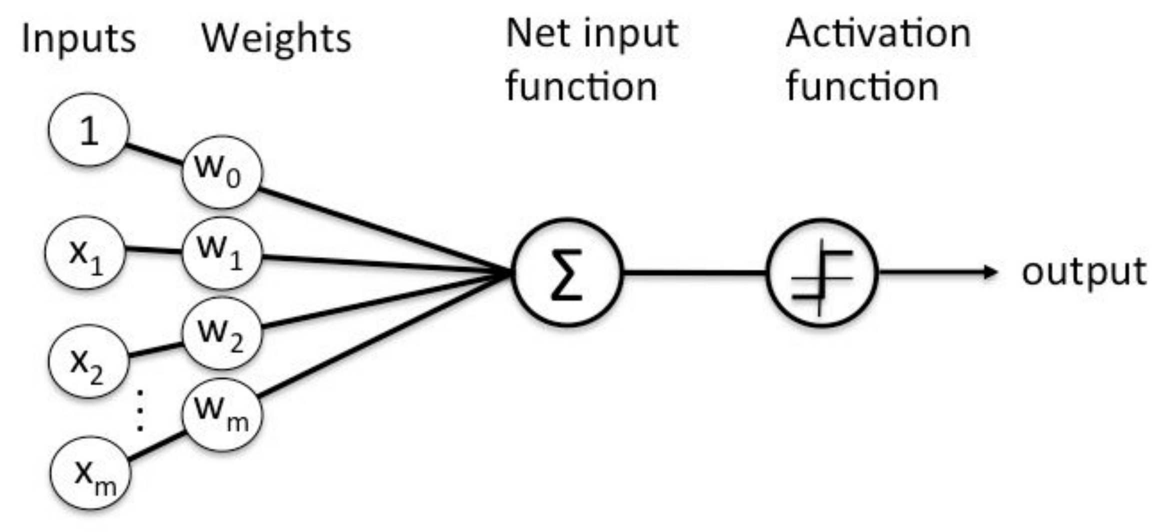
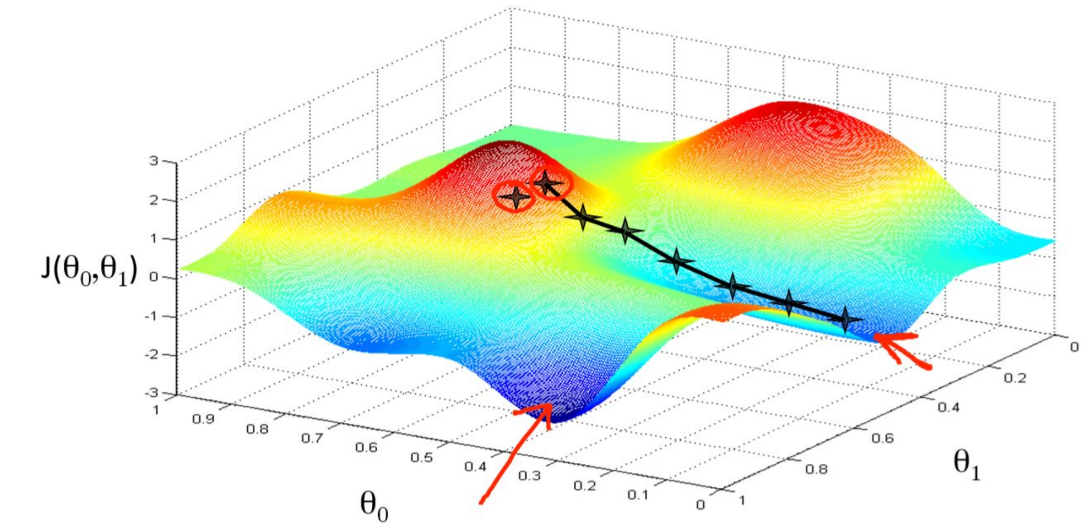


Figure 2: Visualization of computation within a neuron. Retrieved from https://skymind.ai/wiki/neural-network



$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \quad (\text{for } j = 0 \text{ and } j = 1)$$

Figure 4: Visualization of gradient descent. Retrieved from https://www.researchgate.net/figure/Gradient-descent-algorithm-direction-Equation-1-Image-courtesy-of-Andrew-Ng_fig13_298786734.

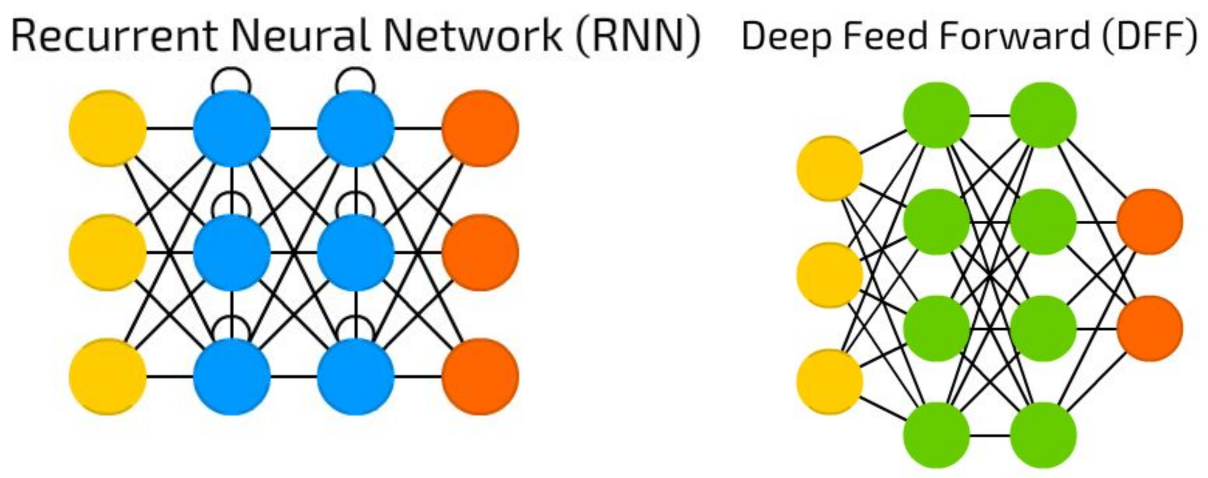


Figure 5: Visualization of recurrent neural network. Retrieved from https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464