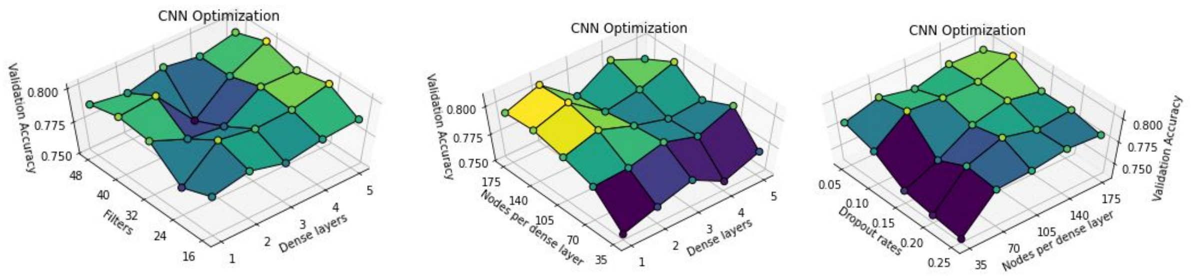
Sleep Stage Classification using Deep Learning Techniques

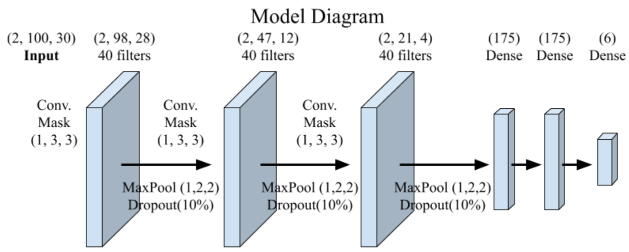
Prathna Vaswani (SNR: 1279690), *Noah Smeets* (SNR: 1279292), *Marco Wedemeyer* (SNR: 2001451)
Submission Account Name: *random_state_666*

Connection to Scientific Literature - Predicting sleep stages based on EEG readings can aid in diagnosing sleep disorders early, but is a challenging task because an individual's physiology affects EEG readings [2][3]. EEG sleep stage classification has been attempted with deep learning previously (e.g. [2][3][5]). So far, no dominant modelling architecture has been established; convolutional neural networks (CNN), hybrid CNNs [1], and LSTM models [5] have been used. Taking the run time of the algorithm into account, we decided to build a CNN for this assignment. We created an ensemble by combining the results of eleven identical but separately trained models, to increase generalizability, similar to the authors in [4].

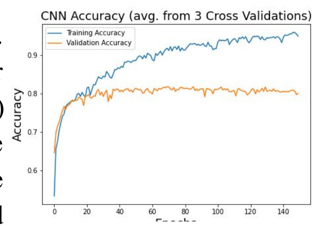
Model Tuning - Based on the literature, the CNN was chosen as the base algorithm. The amount of convolutional layers was maximised (3 layers) to allow for abstract feature representations to form, which was constrained by the shape of the data and the inclusion of max pooling layers. After several tests with different convolutional masks, the (1,3,3) mask performed the best. The hyperparameters optimized were the dropout rates, number of dense layers, dense layer node counts, and number of filters per convolutional layer. These hyperparameters were evaluated pairwise with 3 fold stratified cross-validation (see *CNN Optimization* graphs). The models were trained using 20 epochs, which resulted in the best bias-variance tradeoff (see *CNN Accuracy* and *CNN Loss* graphs).

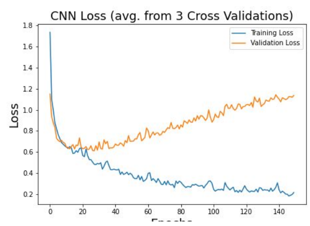


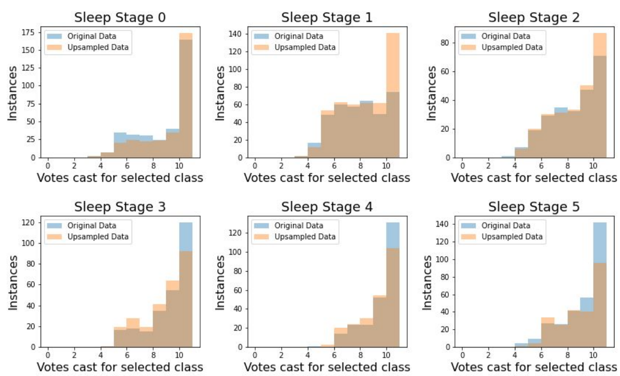
Final Model Architecture - The final model selected (see *Model Diagram*) is a CNN with three convolutional layers with masks of (1,3,3) and 40 filters. After each convolutional layer, max pooling with a mask of (1,2,2) and dropout of 10% is applied. This is followed by 2 dense layers with 175 nodes each and an output layer with 6 nodes. These parameters were chosen based off of the model tuning results. An ensemble of 11 such randomly initialized models was created to further improve generalizability. 11 was chosen as it is a prime number with a good balance between computational cost and generalizability.



Results - The best test accuracy achieved was 69.95%. Histograms of the votes casted for the selected class per class, with original (blue) and upsampled data (orange) (see histograms), showed that the ensemble was least sure about its votes for class 1, which aligns with theory. Stage 1 and 2 are the transition stages between the awake and deep sleep stages [6], making them difficult to classify. In an attempt to improve the accuracy, class 1 was upsampled to twice the number of instances. This did not improve performance, with a test accuracy of 67.9%. This change seemed to bias the ensemble towards class 1 rather than helping it learn, showing that changes in the model architecture are needed instead. Although our architecture outperformed the baseline, the use of hybrid CNNs or RNNs could achieve better performance, however, at the cost of computational complexity.







Literature

[1] Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of neural engineering*, *16*(3), 031001.

[2] Chambon, S., Galtier, M. N., Arnal, P. J., Wainrib, G., & Gramfort, A. (2018). A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(4), 758-769.

[3] Dong, H., Supratak, A., Pan, W., Wu, C., Matthews, P. M., & Guo, Y. (2017). Mixed neural network approach for temporal sleep stage classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(2), 324-333.

[4] Jin, L. P., & Dong, J. (2016). Ensemble deep learning for biomedical time series classification. *Computational intelligence and neuroscience*, *2016*.

[5] Michielli, N., Acharya, U. R., & Molinari, F. (2019). Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals. *Computers in biology and medicine*, 106, 71-81.

[6] Purves, D., Augustine, G., Fitzpatrick, D., Katz, L., LaMantia, A., McNamara, J., & Williams, S. (2001). *Neuroscience 2nd edition*. sunderland (ma) sinauer associates.