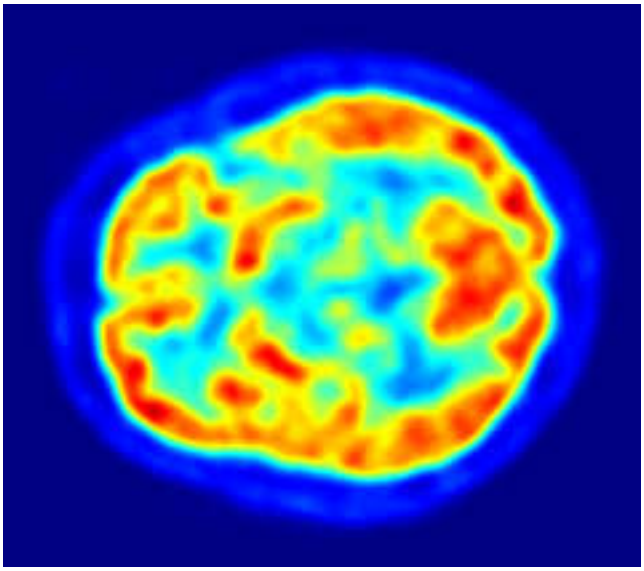
# Practice unit 2: Hyperparameter tuning of the initial practice

Statistical learning with deep artificial neural networks

## Activity

### Description

Parkinson's disease is a chronic and progressive neurodegenerative disorder. As dopamine-generating neurons in parts of the brain become damaged or die, people begin to experience difficulty speaking, writing, walking, or completing other simple tasks. These symptoms worsen over time, resulting in increased severity in patients.



In this activity, we propose the implementation of a dense deep neural network for the prediction of the severity of Parkinson's disease from data captured by telemonitoring of the voice of patients. Supplementary information about the data can be found at: https://archive.ics.uci.edu/ml/datasets/Parkinsons+Telemonitoring. from Grover et al. (2018).

This dataset is comprised of a variety of biomedical voice measurements from 42 people with early-stage Parkinson's disease recruited for a six-month trial of a telemonitoring device for remote monitoring of symptom progression. The recordings were automatically captured at the patients' homes.

The data set columns contain: `patient ID`, `age`, `sex`, `time interval since initial enrollment date`, `UPDRS_motor`, `UPDRS_total` and 16 voice biomedical measures. The variables that will be used are: `total_UPDRS` and 16 biomedical voice measurements. Each row corresponds to one of the 5,875 voice recordings of these individuals.

The main objective of the activity is **hyperparameter tuning** of a deep model for predicting the severity of Parkinson's from the 16 voice measures, where the severity of Parkinson's is based on a binary variable

that is defined from the binarization of the variable `total_UPDRS`, indicating severity when `total_UPDRS> 25` and not otherwise.

## Statement of the activity:

Implement a pipeline that performs the following items:

1. Load the data file `parkinsons_updrs.data`.
2. Create the binary variable of Parkinson's severity.
3. Normalize by means of the `min-max` transformation the variables of the 16 voice measurements.
4. Separate train and test.
5. Implement a dense deep neural network for severity prediction. Using `tfruns` package, determining which architecture among the following provides better accuracy:
   - with two hidden layers of 10 and 10 nodes each.
   - with three hidden layers of 20, 10, 5 nodes each.
   - adding 40% dropout between hidden layers.
   - adding l2 regularization on hidden layers.
6. For the selected model, include callback functions in the monitoring to save the best model.
7. Provide the predictions of the test dataset.

## Reference

Grover, Srishti, Saloni Bhartia, Akshama, Abhilasha Yadav, and Seeja K. R. 2018. "Predicting Severity of Parkinson's Disease Using Deep Learning." *Procedia Computer Science* 132: 1788–94. https://doi.org/https://doi.org/10.1016/j.procs.2018.05.154.