Predicting Brain Hermorrhages

Uday Alla: 404428077
Chuyue(Camille) Zhang: 904461125
Department of Computer Science
University of California, Los Angeles

Background

- When treating ischemic stroke, recanalization therapy comes with certain risk of developing hemorrhagic transformation
- Our aim to predict the probability of hemorrhagic transformation with 618 derived features from MRI perfusion images.
- Pre-treatment DWI, source PWI and arterial input function (AIF) are used as input and the presence of HT observed in follow-up gradient recalled echo (GRE) images are used as groundtruth labels during training.
- Diffusion-weighted image uses the diffusion of water molecules to generate contrast in MR images.
- Perfusion-weighted image uses the passage of fluid through the lymphatic system or blood vessels to an organ or tissue.

Methods

- We imported a 50k entry csv file as our input, with each entry comprises 618 columns of input features and the last single column as the binary results.
- Both tensorflow and Scikit-learn models machine learning frameworks are used to implement the classfiiers.
- For tensorflow, we used logistic regression and multi-layer perception network models. We divided input sets into 80% training and 20% testing data.
- For Scikit-learn, we used SVM, Random Forest, SGD Classifiers, Nearest Neighbors, Nearest Centroids, Bagging Classifiers, and Gradient Classifer. We used cross validation of 3 fold, 5 fold and 7 fold respectively.

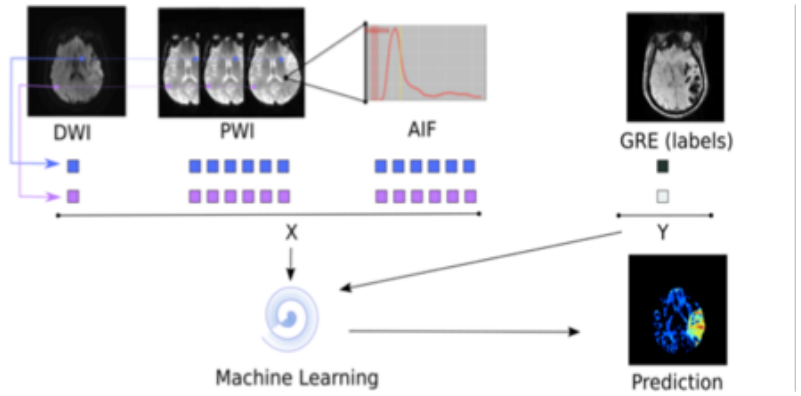

Fig1. Illustrative flowchart of our machine learning model.

Results

Tensorflow Classifiers

Classifier	Training/Testing Accuracy
Logistic Regressions	96.5% | 80.3%
MLP Neural Network	70.2% | 52.9%

Scikit-learn Classifiers

Classifier	#Folds	Testing Accuracy
NearestCentroid	3|5|7	69.3%|70.0%|70.0%
LogisticRegression	3|5|7	79.2%|79.2%|79.3%
MLPClassifier (activation=identity)	3|5|7	78.4%|80.0%|78.4%
SGDClassifer (loss=hinge)	3|5|7	65.5%|76.4%|77.5%
KNeighborsRegressor(n_neighbor=5)	3|5|7	70.0%|69.2%|68.7%
KNeighborsClassifier (n_neighbor=5)	3|5|7	70.0%|69.2%|68.7%
RadiusNeighborClassifier(n_neighbor=5)	3|5|7	46.4%|44.5%|44.2%
RandomForestClassifier(n-estimators=35)	3|5|7	64.6%|61.6%|65.1%
DecisionTreeClassifer	3|5|7	62.3%|60.0%|61.8%
SVC(kernel=rbf)	3|5|7	80.2%|79.8%|80.2%

Conclusions

- Based on testing accuracy, Logistic Regressions from Tensorflow and models such as LogisticRegression, MLPClassifer, SVC from Scikit-learn perform the best, with an average accuracy at around 80%.
- Even though traditionally more sophisticated models such as MLP neural network produces the best result, it is not the case for our prediction possibility due to the overfitting problem that comes with complex network.
- We weren't able to produce top-notch accuracy level due to lack of expertise in more sophisticated algorithm such as deep learning.

References

- Y.Yu, F. Scalzo, et al. Prediction of Hemorrhagic Transformation Severity in Acute Stroke from Source Perfusion MRI. 2013:31:961-969