Alzheimer’s PET and Clinical Risk Predictor

Intended Use:

This model is intended to classify different levels of cognitive impairment given metrics such as amyloid/tau levels, cognitive scores, and demographics.

This model is not for clinical use in any shape or form. This model is not always correct, and it is not clinically approved. For serious medical issues please consult a doctor instead of relying on this model.

Training Data:

This model was trained on the ADNI (Alzheimer's Disease Neuroimaging Initiative) dataset. In order to gain access to this dataset, the team had to apply and get approved first.

Multiple dataframes from this database were combined to make a final pandas dataframe with the core independent variables of amyloid levels, tau levels, cognitive scores, demographics, and diagnosis. When merging, they were merged so that all biomarkers and cognitive tests connected to the same clinical visit, avoiding temporal data leakage. The final dataframe the model trained on consisted of 944 patients.

This data was solely numerical data. There were six independent variables: amyloid level, tau level, MMSE score, CDR rating, gender, and years of education. The diagnosis (target variable) had three classes: CN (cognitively normal), MCI (mild cognitive impairment), and AD (Alzheimer’s disease).

Model Type:

This model is a Gradient-Boosted Decision Tree (XGBoost) Multi-Class Classifier.

Input: Tabular clinical and biomarker features

- AMY_BIOMARKER (amyloid PET biomarker)
- TAU_BIOMARKER (tau PET biomarker)
- MMSE_TOTAL (cognitive score on MMSE test)
- CDR_SB (Clinical Dementia Rating score)
- SEX (0=Male, 1=Female)
- EDUC (years of education)

Output: Classification based on three classes (CN/MCI/AD)

Objective: multi:softprob (predicts probabilities for multiple classifications)
Trees: Gradient Boosted Decision Trees, max depth 4, 5000 boosting rounds. Early stopping was implemented to prevent overfitting.

Performance:

The model performed with the following metrics:

- Test Accuracy: 90.85%
- Weighted Precision: 0.906
- Weighted Recall: 0.908
- Weighted F1: 0.906

These metrics show a strong overall performance for all classes.

The macro sensitivity for the model was 0.852, showing that the model identifies about 85% of true positive cases. The macro specificity of the model was 0.941, indicating that the model identifies around 94% of true negative cases, meaning low false positive rates.

All metrics were computed on a held-out test set not touched during training or validation.

Limitations:

This model does not take into account any MRI data, which is also an important part of diagnosing Alzheimer's. Also, instead of using actual PET scans, it uses biomarker data, which only captures one aspect of the scan. Overall, this model takes many variables, but they don't span the entirety of the causing factors of Alzheimer's.

Bias and Fairness:

The ADNI cohort used might not properly represent all the demographics of the global populations, with different biases towards certain ages, ethnicities, education levels, or health care access. Because of this, model performance may change on underrepresented groups, and these biases were not accounted for during training.

Ethical Concerns:

This model predicts 5.9% of inputted negative cases as false positives, which, if utilized wrong in a clinical environment, could lead to unnecessary actions and procedures. This is a core reason as to why this model should never be professionally used.

Any data inputted into this model is highly personal to the user, and they should take discretion when inputting in case of privacy. This is why no data such as name or location is taken in case of data leakage.