FeaFiner: Biomarker Identification Through Feature Generalization and Selections

Supplemental Materials

Jiayu Zhou, Zhaosong Lu, Jimeng Sun, Lei Yuan, Fei Wang, Jieping Ye

Appendix I. Extension of FeaFiner to General Multi-Task Learning on Hilbert Space

The formulation of FeaFiner and the algorithm provided in this paper can be extended to a more general case: multi-task learning with shared feature groups. In the multi-task learning setting we are required to learn together a set of related tasks that has the same feature space. Note that in the appendix we use different notations as in the main part of the paper. We assume that for all tasks the features share the same group structure $\mathbf{G}$, but has different predicted models on the generalized features. Assume we have T tasks and for task i we have a predictive model $\mathbf{s}_i$, Let $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T) = ((\boldsymbol{x}_1, \mathbf{y}_1), \ldots (\boldsymbol{x}_T, \mathbf{y}_T)) = ((x_{ti}, y_{ti}) : 1 \le i \le n, 1 \le t \le T)$ be the dataset, where $x_{ti} \in H, y_{ti} \in \mathbb{R}$. For simplicity of the analysis we suppose that in every task there are n samples. The multi-task orthogonal MT-FeaFiner learning solves the following optimization problem:

$$\begin{aligned}
\min_{\mathbf{s},\mathbf{G}} \quad & \frac{1}{T}\sum_{i=1}^{T}\frac{1}{n}\ell\left(\langle \mathbf{X}_{(i)}, \mathbf{G}\mathbf{s}_i\rangle, \mathbf{y}_{(i)}\right) \\
\text{subject to:} \quad & \tau_{min} \le \|\mathbf{G}\|_1 \le \tau_{\max}^{\lambda_G}, \\
& \|\mathbf{s}_i\|_1 \le \alpha, \forall i = 1 \ldots T \\
& \mathbf{G}^T\mathbf{G} = \mathbf{I}, \mathbf{G} \ge 0
\end{aligned} \tag{13}$$

When $T = 1$ the multi-task FeaFiner reduces to the single task FeaFiner in Eq. (12). The augmented Lagrange method based Algorithm 3 can be used to obtain a local solution to this problem.

In the appendix we extend our discussion to a more general setting and let the feature space to be a finite or infinite dimensional Hilbert space and denote by H. We define two sets $\mathcal{G}_K = \{\mathbf{G} \in \mathbb{R}^{p\times k} : \tau_{min} \le \|\mathbf{G}\|_1 \le \tau_{\max}^{\lambda_G}, \mathbf{G}^T\mathbf{G} = \Delta, \mathbf{G} \ge 0\}$ and $\mathcal{S} = \{\mathbf{S} \in \mathbb{R}^k : \|\mathbf{s}\|_1 \le \alpha\}$. The problem in Eq. (13) can be written as

$$\min_{\mathbf{G}\in\mathcal{G},\mathbf{S}\in(\mathcal{S}_\alpha)^T} \frac{1}{T}\sum_{t=1}^{T}\frac{1}{n}\sum_{i=1}^{n}\ell\left(\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle, y_{ti}\right) = \min_{\mathbf{G}\in\mathcal{G}} \frac{1}{T}\sum_{t=1}^{T}\min_{\mathbf{s}\in\mathcal{S}_\alpha}\frac{1}{n}\sum_{i=1}^{n}\ell\left(\langle \mathbf{G}\mathbf{s}, x_{ti}\rangle, y_{ti}\right) \tag{14}$$

where ℓ is a normalized loss function, which has values in $[0, 1]$ and is Lipschitz continuous with Lipschitz constant L. Assume that the $i.i.d.$training sample $\boldsymbol{z}_t$ is drawn from $(\mu_t)^n$ and the data $\mathbf{Z} \sim \prod_{t=1}^{T}\mu_t^n$.

Definition 6.1. *Expected and empirical error, global solutions. Given any $\mathbf{G}$ and $\mathbf{S}$, we denote the expected risk as:*

$$\mathbb{E}(\mathbf{G}, \mathbf{S}) = \mathbb{E}_{(\boldsymbol{x},\mathbf{y})\sim\boldsymbol{\mu}}[\frac{1}{T}\sum_{t=1}^{T}\ell(\langle \mathbf{G}\mathbf{s}_t, x_t\rangle, y_t)] = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{(x,y)\sim\mu_t}[\ell(\langle \mathbf{G}\mathbf{s}_t, x\rangle, y)]$$

Also, given data $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$, the empirical risk is defined as:

$$\hat{\mathbb{E}}(\mathbf{G}, \mathbf{S}|\mathbf{Z}) = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{n}\sum_{i=1}^{n}[\ell(\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle, y_{ti})]$$

Let $\mathbf{G}^*$ and $\mathbf{S}^*$ be the global optimal solution of the expected risk, i.e.,:

$$(\mathbf{G}^*, \mathbf{S}^*) = \underset{\mathbf{G}\in\mathcal{G}_K,\mathbf{S}\in(\mathcal{S}_\alpha)^T}{\arg\min}\ \mathbb{E}(\mathbf{G}, \mathbf{S}) = \underset{\mathbf{G}\in\mathcal{G}_K,\mathbf{S}\in(\mathcal{S}_\alpha)^T}{\arg\min}\ \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{(x,y)\sim\mu_t}[\ell(\langle \mathbf{G}\mathbf{s}, x\rangle, y)]$$

and denote $\mathbf{G}^*_{(\mathbf{Z})}$ and $\mathbf{S}^*_{(\mathbf{Z})}$ be the optimal solution by minimizing the empirical risk, i.e.,:

$$\left(\mathbf{G}^*_{(\mathbf{Z})}, \mathbf{S}^*_{(\mathbf{Z})}\right) = \underset{\mathbf{G}\in\mathcal{G}_K,\mathbf{S}\in(\mathcal{S}_\alpha)^T}{\arg\min}\ \hat{\mathbb{E}}(\mathbf{G}, \mathbf{S}|\mathbf{Z}) = \underset{\mathbf{G}\in\mathcal{G}_K,\mathbf{S}\in(\mathcal{S}_\alpha)^T}{\arg\min}\ \frac{1}{T}\sum_{t=1}^{T}\frac{1}{n}\sum_{i=1}^{n}[\ell(\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle, y_{ti})]$$

The follow Theorem shows the asymptotic convergence behavior of the learning process of MT-FeaFiner in Eq. (14).

Theorem 6.2. *Let $\delta > 0$ and let $\mu_1, \ldots, \mu_T$ be probability measure on $H \times \mathbb{R}$. With probability of at least $1 - \delta$ in the draw of $\mathbf{Z} \sim \prod_{t=1}^{T}\mu_t^n$, we have:*

$$\begin{aligned}
\mathbb{E}(\mathbf{G}^*_{(\mathbf{Z})}, \mathbf{S}^*_{(\mathbf{Z})}) - \mathbb{E}(\mathbf{G}^*, \mathbf{S}^*) &= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{(x,y)\sim\mu_t}[\ell(\langle \mathbf{G}^*_{(\mathbf{Z})}\mathbf{s}^*_{t(\mathbf{Z})}, x\rangle, y)] - \inf_{\mathbf{G}\in\mathcal{G}_K}\frac{1}{T}\sum_{t=1}^{T}\inf_{\mathbf{s}\in\mathcal{S}_\alpha}\mathbb{E}_{(x,y)\sim\mu_t}[\ell(\langle \mathbf{G}\mathbf{s}, x\rangle, y)] \\
&\le L\alpha\sqrt{\frac{2\mathcal{C}_1(\mathbf{X})(K+12)}{nT}} + L\alpha\sqrt{\frac{8\mathcal{C}_\infty(\mathbf{X})\ln(2K)}{n}} + \sqrt{\frac{8\ln 4/\delta}{nT}}
\end{aligned} \tag{15}$$

where $\mathcal{C}_1(\mathbf{X}) = \frac{1}{T}\sum_{t=1}^{T} \|\hat{\Sigma}(\boldsymbol{x}_t)\|_1 := \frac{1}{T}\sum_{t=1}^{T} \operatorname{tr}\left(\hat{\Sigma}(\boldsymbol{x}_t)\right)$, *where* $\operatorname{tr}(.)$ *is the trace of a matrix,* $\mathcal{C}_\infty(\mathbf{X}) = \frac{1}{T}\sum_{t=1}^{T} \|\hat{\Sigma}(x_t)\|_\infty := \frac{1}{T}\sum_{t=1}^{T} \lambda_{\max}\left(\hat{\Sigma}(\boldsymbol{x}_t)\right)$ *where* $\lambda_{\max}$ *is the largest eigenvalue,* $\hat{\Sigma}(\boldsymbol{x}_t)$ *is the empirical covariance of the i-th task. If* $\boldsymbol{x}_t \in \mathbb{R}^{n\times d}$, *then* $\hat{\Sigma}(\boldsymbol{x}_t) = \frac{1}{n}\boldsymbol{x}_t^T\boldsymbol{x}_t \in \mathbb{R}^{d\times d}$.

Appendix II. Proof of Theorem 3.1 and Theorem 6.2

In this section we present the generalization bound for the learning process of FeaFiner in Eq. (13). When $T = 1$, the result is the same as in Theorem 3.1. The proof structure is standard and generally follows the multi-task dictionary learning problem in [21].

II.1 Proof Architecture

The main result in Theorem 6.2 upper bounds the difference between expected risks of the model estimated from a given dataset $\mathbf{Z}$ $\mathbb{E}(\mathbf{G}^*_{(\mathbf{Z})}, \mathbf{S}^*_{(\mathbf{Z})})$ and the global optimal solution: $\mathbb{E}(\mathbf{G}^*, \mathbf{S}^*)$. We can manipulate (as in the last proof) the terms to link the bound to the generalization error $\sup_{\mathbf{G},\mathbf{S}} |\mathbb{E}(\mathbf{G},\mathbf{s}) - \hat{\mathbb{E}}(\mathbf{G},\mathbf{s})|$ (Theorem 6.6). The key of bounding this generalization is to upper bound the Rademacher complexity of the risk function class (Proposition 6.5).

To bound the Rademacher complexity of the risk function class, we firstly use the Lipschiz property to remove the loss function from our analysis. We then transform the expectation into an integration, and the key is to get the bound on the integration. To do so the author develops two important lemma (Lemma 6.3 and Lemma 6.4) that can relate the integration to the two quantity $\mathcal{C}_\infty(\mathbf{X})$ and $\mathcal{C}_1(\mathbf{X})$. The proof of the two lemma uses the orthogonality properties of the groups $\mathbf{G}$ and the sparsity of the model $\mathbf{S}$.

II.2 Proof towards Theorem 6.2

Before proceeding to prove Theorem 6.2, we need some auxiliary results.

Fix $\mathbf{X} \in H^{nT}$ and for $\mathbf{S} = (\mathbf{s}_1, \ldots, \mathbf{s}_T) \in (\mathbb{R}^K)^T$, and define the random variable:

$$F_{\mathbf{S}} = F_{\mathbf{S}}(\boldsymbol{\sigma}) = \sup_{\mathbf{G}\in\mathcal{G}_K} \sum_{t,i} \sigma_{ti}\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle$$

The idea behind defining the auxiliary function/variable $F_{\mathbf{S}}(\boldsymbol{\sigma})$ is that after we 'remove' the loss function using Eq. (6.13), we need to bound the expectatio of $F_{\mathbf{S}}$.

Lemma 6.3. *If* $\mathbf{S} = (\mathbf{s}_1, \ldots, \mathbf{s}_T)$ *satisfies* $\|\mathbf{s}_t\|_1 \le \alpha, \forall t,\ \alpha > 0$*, then:*

$$\mathbb{E}_{\boldsymbol{\sigma}} F_{\mathbf{S}} \le \alpha\sqrt{nTK\mathcal{C}_1(\mathbf{X})}$$

Proof.

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\sigma}} F_{\mathbf{S}} &= \mathbb{E}_{\boldsymbol{\sigma}} \sup_{\mathbf{G}\in\mathcal{G}_K} \sum_{t,i} \sigma_{ti}\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle \quad \text{(by definition)} \\
&= \mathbb{E}_{\boldsymbol{\sigma}} \sup_{\mathbf{G}\in\mathcal{G}_K} \sum_{t,i} \sigma_{ti}\langle \sum_k \mathbf{g}_k \mathbf{s}_{tk}, x_{ti}\rangle = \mathbb{E}_{\boldsymbol{\sigma}} \sup_{\mathbf{G}\in\mathcal{G}_K} \sum_k \langle \mathbf{g}_k, \sum_{t,i} \sigma_{ti}\mathbf{s}_{tk}x_{ti}\rangle \\
&\le \mathbb{E}_{\boldsymbol{\sigma}} \sup_{\mathbf{G}\in\mathcal{G}_K} \sum_k \|\mathbf{g}_k\| \left\|\sum_{t,i} \sigma_{ti}\mathbf{s}_{tk}x_{ti}\right\| \quad \text{(Cauchy-Schwarz)} \\
&\le \mathbb{E}_{\boldsymbol{\sigma}} \sup_{\mathbf{G}\in\mathcal{G}_K} \left(\sum_k \|\mathbf{g}_k\|^2\right)^{1/2} \left(\sum_k \left\|\sum_{t,i} \sigma_{ti}\mathbf{s}_{tk}x_{ti}\right\|^2\right)^{1/2} \quad \text{(again Cauchy-Schwarz)} \\
&= \sup_{\mathbf{G}\in\mathcal{G}_K} \left(\sum_k \|\mathbf{g}_k\|^2\right)^{1/2} \mathbb{E}_{\boldsymbol{\sigma}} \left(\sum_k \left\|\sum_{t,i} \sigma_{ti}\mathbf{s}_{tk}x_{ti}\right\|^2\right)^{1/2} \quad \text{(Expectation on } \boldsymbol{\sigma}\text{, sign of } \mathbf{S} \text{ not matter)} \\
&= \sqrt{K} \left(\sum_k |\mathbf{s}_{tk}|^2 \mathbb{E}_{\boldsymbol{\sigma}} \left\|\sum_{t,i} \sigma_{ti}x_{ti}\right\|^2\right)^{1/2} \quad \text{(From the orthogonal constraints we have } \|\mathbf{g}_k\| = 1 \Rightarrow \textstyle\sum_k \|\mathbf{g}_k\|^2 = K) \\
&\le \sqrt{K} \left(\sum_k |\mathbf{s}_{tk}|^2 \mathbb{E}_{\boldsymbol{\sigma}} \sum_{t,i} \|\sigma_{ti}x_{ti}\|^2\right)^{1/2} \quad \text{(Triangle Inequality)} \\
&\le \sqrt{K}\alpha \left(\sum_{t,i} \|x_{ti}\|^2\right)^{1/2} \quad (\|\mathbf{s}_t\| = \sqrt{\textstyle\sum_k |\mathbf{s}_{tk}|^2} \le \|\mathbf{s}_t\|_1 \le \alpha,\ \|\sigma_{ti}x_{ti}\|^2 = \|x_{ti}\|^2) \\
&= \alpha\sqrt{nKT\mathcal{C}_1(\mathbf{X})} \quad (\mathcal{C}_1(\mathbf{X}) = \tfrac{1}{T}\operatorname{tr}\left(\hat{\Sigma}(\boldsymbol{x}_t)\right) = \tfrac{1}{nT}\textstyle\sum_{t,i} \|x_{ti}\|^2)
\end{aligned}$$

This completes the proof. □

LEMMA 6.4. *If $\mathbf{S}$ satisfies $\|\mathbf{s}_t\|_1 \leq \alpha, \forall t$, $\alpha > 0$, then for any $s \geq 0$*

$$\Pr\{F_{\mathbf{S}} \geq \mathbb{E}[F_{\mathbf{S}}] + s\} \leq \exp\left(\frac{-s^2}{\alpha^2 8nT\mathcal{C}_\infty(\mathbf{X})}\right)$$

PROOF. For any configuration $\boldsymbol{\sigma}$ of the Rademacher variables, let

$$\mathbf{G}(\boldsymbol{\sigma}) = \arg\max_{\mathbf{G}\in\mathcal{G}_K} F_{\mathbf{S}}(\boldsymbol{\sigma}) = \arg\max_{\mathbf{G}\in\mathcal{G}_K} \sum_{t,i} \sigma_{ti}\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle.$$

For any $s \in \{1, \ldots, T\}, j \in \{1, \ldots, n\}$ and any $\sigma' \in \{-1, 1\}$ to replace σ_{sj} we have:

$$\begin{aligned}
&F_{\mathbf{S}}(\boldsymbol{\sigma}) - F_{\mathbf{S}}(\boldsymbol{\sigma}_{(sj)\leftarrow\sigma'})\\
=&\sup_{\mathbf{G}\in\mathcal{G}_K}\sum_{t,i}\sigma_{ti}\langle\mathbf{G}\mathbf{s}_t, x_{ti}\rangle - \sup_{\mathbf{G}\in\mathcal{G}_K}\sum_{t,i}\sigma'_{ti}\langle\mathbf{G}\mathbf{s}_t, x_{ti}\rangle \quad (\sigma'_{ti} = \sigma_{ti}, \text{ except } \sigma'_{sj})\\
=&\sum_{t,i}\sigma_{ti}\langle\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_t, x_{ti}\rangle - \sup_{\mathbf{G}\in\mathcal{G}_K}\sum_{t,i}\sigma'_{ti}\langle\mathbf{G}\mathbf{s}_t, x_{ti}\rangle \quad \text{(By definition)}\\
\leq&\sum_{t,i}\sigma_{ti}\langle\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_t, x_{ti}\rangle - \sum_{t,i}\sigma'_{ti}\langle\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_t, x_{ti}\rangle \quad (\sup_{\mathbf{G}\in\mathcal{G}_K}\textstyle\sum_{t,i}\sigma'_{ti}\langle\mathbf{G}\mathbf{s}_t, x_{ti}\rangle \geq \sum_{t,i}\sigma'_{ti}\langle\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_t, x_{ti}\rangle)\\
=&\sigma_{sj}\langle\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_s, x_{sj}\rangle - \sigma'_{sj}\langle\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_s, x_{sj}\rangle \leq 2|\langle\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_s, x_{sj}\rangle| \quad (\sigma_{sj}, \sigma'_{sj} \in \{-1, 1\})
\end{aligned}$$

For any $\mathbf{G} \in \mathcal{G}_K$ We also have that

$$\begin{aligned}
\|\mathbf{G}\mathbf{s}_t\| &= \|\sum_k \mathbf{s}_{tk}\mathbf{g}_k\| \leq \sum_K \|\mathbf{s}_{tk}\mathbf{g}_k\| \quad \text{(Triangle Inequality)}\\
&= \sum_K |\mathbf{s}_{tk}|\|\mathbf{g}_k\| = \sum_k |\mathbf{s}_{tk}| \quad (\text{Again from the orthogonal constraint } \|\mathbf{g}_k\| = 1, \forall K)\\
&= \|\mathbf{s}_t\|_1 \leq \alpha
\end{aligned} \tag{16}$$

Therefore we have

$$\begin{aligned}
&\sum_{sj}\left(F_{\mathbf{S}}(\boldsymbol{\sigma}) - \inf_{\sigma'\in\{-1,1\}} F_{\mathbf{S}}(\boldsymbol{\sigma}_{(sj)\to\sigma'})\right)^2\\
\leq&4\sum_{sj}\langle\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_s, x_{sj}\rangle^2 = 4n\sum_{t=1}^T\frac{1}{n}\sum_{i=1}^n\langle\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_t, x_{ti}\rangle^2 = 4n\sum_{t=1}^T(\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_t)^T\hat{\Sigma}(\boldsymbol{x}_t)(\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_t)\\
\leq&4n\sum_{t=1}^T\lambda_{max}(\hat{\Sigma}(\boldsymbol{x}_t))\|\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_t\|^2 = 4n\sum_{t=1}^T\|\hat{\Sigma}(\boldsymbol{x}_t)\|_\infty\|\mathbf{G}(\boldsymbol{\sigma})\mathbf{s}_t\|^2 \quad \text{(by definition)}\\
\leq&4n\alpha^2\sum_{t=1}^T\|\hat{\Sigma}(\boldsymbol{x}_t)\|_\infty \quad \text{(Eq. (16))}\\
=&4nT\alpha^2\mathcal{C}_\infty(\mathbf{X})
\end{aligned}$$

Denote

$$B^2 = \sup\sum_{sj}\left(F_{\mathbf{s}}(\boldsymbol{\sigma}) - \inf_{\sigma'\in-1,1} F_{\mathbf{S}}(\boldsymbol{\sigma}_{(sj)\to\sigma'})\right)^2 = 4nT\alpha^2\mathcal{C}_\infty(\mathbf{X})$$

Applying Theorem 6.9, we have that

$$\Pr\{F_{\mathbf{S}} \geq \mathbb{E}[F_{\mathbf{s}}] + s\} \leq \exp\left(\frac{-s^2}{2B^2}\right) = \exp\left(\frac{-s^2}{8nT\alpha^2\mathcal{C}_\infty(\mathbf{X})}\right)$$

This thus completes the proof. □

Now we proceed to the key part that bounds the partial Rademacher complexity of the risk function $nT \cdot \hat{\mathcal{R}}(\mathbf{F}|\mathbf{Z})$, where function class $\mathbf{F} \sim \ell(\mathbf{G}, \mathbf{S})$:

PROPOSITION 6.5. *For every fixed $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) \in (H, \mathbb{R})^{nT}$ we have*

$$\mathbb{E}_\sigma \sup_{\mathbf{G}\in\mathcal{G}_k, \mathbf{S}\in(\mathcal{S})^T}\sum_{t,i}\sigma_{it}\ell(\langle\mathbf{G}\mathbf{s}_t, x_{ti}\rangle, y_{ti}) \leq L\alpha\sqrt{2nT\mathcal{C}_1(\mathbf{X})(K+12)} + L\alpha T\sqrt{8n\mathcal{C}_\infty(\mathbf{X})\ln(2K)}$$

PROOF. It is sufficient to prove the case $\alpha = 1$ and the general results can be obtained from rescaling. Because of the Lipschitz property of the loss function ℓ, according to Lemma 6.13, we have

$$\mathbb{E}_\sigma \sup_{\mathbf{G}\in\mathcal{G}_k,\mathbf{S}\in(\mathcal{S})^T} \sum_{t,i} \sigma_{it}\ell(\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle, y_{ti}) \leq L\mathbb{E}_\sigma \sup_{\mathbf{G}\in\mathcal{G}_k,\mathbf{S}\in(\mathcal{S})^T} \sum_{t,i} \sigma_{it}\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle$$
$$= L\mathbb{E}_\sigma \max_{\mathbf{S}\in(\mathcal{S})^T} \sup_{\mathbf{G}\in\mathcal{G}_k} \sum_{t,i} \sigma_{it}\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle = L\mathbb{E}_\sigma \max_{\mathbf{S}\in(\mathcal{S})^T} F_{\mathbf{S}}$$

And because that linear functions (Recall that we removed the non-linear loss function using Lemma 6.13) on a compact convex set attain their maxima at the *extreme points*, we denote $\text{ext}(\mathcal{S})$ the extreme points of the set $\mathcal{S}$. We thus have

$$\mathbb{E}_\sigma \max_{\mathbf{s}\in(\mathcal{S})^T} F_{\mathbf{S}} = \mathbb{E}_\sigma \max_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} F_{\mathbf{S}}.$$

This completes the proof. □

Now we need to upper bound the term $\mathbb{E}_\sigma \max_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} F_{\mathbf{S}}$. Because that $F_{\mathbf{s}} \geq 0$ (because the sup we will always manipulate the sign to make it positive. For any $\delta > 0$ we have:

$$\begin{aligned}
\mathbb{E}_\sigma \max_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} F_{\mathbf{S}} &= \int_0^\infty \Pr\left\{\max_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} F_{\mathbf{S}} > s\right\} ds \quad \text{(Lemma 6.7)}\\
&= \int_0^{\sqrt{nKT\mathcal{C}_1(\mathbf{X})}+\delta} \Pr\left\{\max_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} F_{\mathbf{S}} > s\right\} ds + \int_{\sqrt{nKT\mathcal{C}_1(\mathbf{X})}+\delta}^\infty \Pr\left\{\max_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} F_{\mathbf{S}} > s\right\} ds\\
&\leq \int_0^{\sqrt{nKT\mathcal{C}_1(\mathbf{X})}+\delta} 1 ds + \int_{\sqrt{nKT\mathcal{C}_1(\mathbf{X})}+\delta}^\infty \Pr\left\{\max_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} F_{\mathbf{S}} > s\right\} ds \quad (\Pr(*) \leq 1)\\
&= \sqrt{mKT\mathcal{C}_1(\mathbf{X})} + \delta + \sum_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} \int_{\sqrt{nKT\mathcal{C}_1(\mathbf{X})}+\delta}^\infty \Pr\{F_{\mathbf{S}} > s\} ds \quad \text{(union bound)}\\
&= \sqrt{nKT\mathcal{C}_1(\mathbf{X})} + \delta + \sum_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} \int_\delta^\infty \Pr\left\{F_{\mathbf{S}} > \sqrt{nKT\mathcal{C}_1(\mathbf{X})} + s\right\} ds\\
&\leq \sqrt{nKT\mathcal{C}_1(\mathbf{X})} + \delta + \sum_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} \int_\delta^\infty \Pr\{F_{\mathbf{S}} > \mathbb{E}F_{\mathbf{S}} + s\} ds \quad \text{(Lemma 6.3)}\\
&\leq \sqrt{nKT\mathcal{C}_1(\mathbf{X})} + \delta + \sum_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} \int_\delta^\infty \exp\left\{\frac{-s^2}{8nT\mathcal{C}_\infty(\mathbf{X})}\right\} ds \quad \text{(Lemma 6.4)}\\
&= \sqrt{mKT\mathcal{C}_1(\mathbf{X})} + \delta + (2K)^T \int_\delta^\infty \exp\left\{\frac{-s^2}{8nT\mathcal{C}_\infty(\mathbf{X})}\right\} ds \quad (\text{card}(\text{ext}(\mathcal{S}))\text{=2K})\\
&\leq \sqrt{nKT\mathcal{C}_1(\mathbf{X})} + \delta + \frac{4nT\mathcal{C}_\infty(\mathbf{X})(2K)^T}{\delta} \exp\left(\frac{-s^2}{8nT\mathcal{C}_\infty(\mathbf{X})}\right) \quad \text{(Gaussian variable estimate)}
\end{aligned}$$

Set $\delta = \sqrt{8nT\mathcal{C}_\infty(\mathbf{X})\ln(e(2K)^T)}$, we have that:

$$\mathbb{E}_\sigma \max_{\mathbf{s}\in\text{ext}(\mathcal{S})^T} F_{\mathbf{S}} \leq \sqrt{nKT\mathcal{C}_1(\mathbf{X})} + \delta + \frac{4nT\mathcal{C}_\infty(\mathbf{X})(2K)^T}{\delta} \exp\left(\frac{-s^2}{8nT\mathcal{C}_\infty(\mathbf{X})}\right) = \sqrt{2nT(K+12)\mathcal{C}_1(\mathbf{X})} + T\sqrt{8n\mathcal{C}_\infty(\mathbf{X})\ln(2K)}.$$

Therefore we have bounded:

$$\mathbb{E}_\sigma \sup_{\mathbf{G}\in\mathcal{G}_k,\mathbf{s}\in(\mathcal{S})^T} \sum_{t=1}^T\sum_{i=1}^m \sigma_{it}\ell(\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle, y_{ti}) \leq L\mathbb{E}_\sigma \max_{\mathbf{S}\in(\mathcal{S})^T} F_{\mathbf{S}} \leq L\sqrt{2nT(K+12)\mathcal{C}_1(\mathbf{X})} + LT\sqrt{8n\mathcal{C}_\infty(\mathbf{X})\ln(2K)}$$

THEOREM 6.6. *Let $\delta > 0$, fix K and let $\mu_1, \ldots, \mu_T$ be probability measures on $H \times \mathbb{R}$. With probability of at least $1-\delta$ in the draw of $\mathbf{Z} \sim \prod_{t=1}^T(\mu_t)$ we have $\forall \mathbf{G} \in \mathcal{G}_K$ and $\forall \mathbf{S} \in \mathcal{S}_\alpha^T$ that*

$$\begin{aligned}
\mathbb{E}(\mathbf{G},\mathbf{S}) - \hat{\mathbb{E}}(\mathbf{G},\mathbf{S}|\mathbf{Z}) &= \frac{1}{T}\sum_{t=1}^T \mathbb{E}_{(x,y)\sim\mu_t}[\ell(\langle \mathbf{G}\mathbf{s}_t, x\rangle, y)] - \frac{1}{T}\sum_{t=1}^T \frac{1}{n}\sum_{i=1}^n \ell(\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle, y_{ti})\\
&\leq 2L\alpha\sqrt{\frac{2(K+12)\mathcal{C}_1(\mathbf{X})}{nT}} + 2L\alpha\sqrt{\frac{8\mathcal{C}_\infty(\mathbf{X})\ln(2K)}{n}} + \sqrt{\frac{9\ln 2/\delta}{2nT}}
\end{aligned}$$

Proof. We have that:

$$\begin{aligned}
\mathbb{E}(\mathbf{G},\mathbf{S}) &= \frac{1}{T}\mathbb{E}_{(x,y)\sim\boldsymbol{\mu}}[\sum_{t=1}^{T}\ell(\langle \mathbf{G}\mathbf{s}_t, x\rangle, y)] = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{(x,y)\sim\mu_t}[\ell(\langle \mathbf{G}\mathbf{s}_t, x\rangle, y)] \\
&\leq \frac{1}{T}\sum_{t=1}^{T}\frac{1}{n}\sum_{i=1}^{n}\ell(\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle, y_{ti}) + \hat{\mathcal{R}}(\mathbf{F}|\mathbf{Z}) + \sqrt{\frac{9\ln 2/\delta}{2Tn}} \quad \text{(Theorem 6.12)} \\
&= \hat{\mathbb{E}}(\mathbf{G},\mathbf{S}|\mathbf{Z}) + \mathbb{E}_\sigma \sup_{\mathbf{G}\in\mathcal{G}_k, \mathbf{s}\in(\mathcal{S})^T}\frac{2}{Tn}\sum_{t=1}^{T}\sum_{i=1}^{n}\sigma_{it}\ell(\langle \mathbf{G}\mathbf{s}_t, x_{ti}\rangle, y_{ti}) + \sqrt{\frac{9\ln 2/\delta}{2Tn}} \\
&\leq \hat{\mathbb{E}}(\mathbf{G},\mathbf{S}|\mathbf{Z}) + 2L\alpha\sqrt{\frac{2T(K+12)\mathcal{C}_1(\mathbf{X})}{Tn}} + 2L\alpha\sqrt{\frac{8\mathcal{C}_\infty(\mathbf{X})\ln(2K)}{n}} + \sqrt{\frac{9\ln 2/\delta}{2Tn}} \quad \text{(Proposition 6.5)}
\end{aligned}$$

This is to say we have for any $\mathbf{Z}\sim\prod_{t=1}^{T}(\mu_t)$, $\mathbf{G}\in\mathcal{G}_K$ and $\mathbf{S}\in(\mathcal{S}_\alpha)^T$:

$$\mathbb{E}(\mathbf{G},\mathbf{S}) \leq \hat{\mathbb{E}}(\mathbf{G},\mathbf{S}|\mathbf{Z}) + 2L\alpha\sqrt{\frac{2(K+12)\mathcal{C}_1(\mathbf{X})}{Tn}} + 2L\alpha\sqrt{\frac{8\mathcal{C}_\infty(\mathbf{X})\ln(2K)}{n}} + \sqrt{\frac{9\ln 2/\delta}{2Tn}}$$

or equivalently for any $\mathbf{Z}\sim\prod_{t=1}^{T}(\mu_t)$ we have the generalization bound:

$$\sup_{\mathbf{G}\in\mathcal{G}_K, \mathbf{S}\in(\mathcal{S}_\alpha)^T}|\mathbb{E}(\mathbf{G},\mathbf{S}) - \hat{\mathbb{E}}(\mathbf{G},\mathbf{S}|\mathbf{Z})| \leq 2L\alpha\sqrt{\frac{2(K+12)\mathcal{C}_1(\mathbf{X})}{Tn}} + 2L\alpha\sqrt{\frac{8\mathcal{C}_\infty(\mathbf{X})\ln(2K)}{n}} + \sqrt{\frac{9\ln 2/\delta}{2Tn}}$$

This completes the proof. □

Now we are ready to prove Theorem 6.2 by using the above results.

Proof. By Definition 6.1, we have that

$$\hat{\mathbb{E}}(\mathbf{G}^*,\mathbf{S}^*|\mathbf{Z}) - \hat{\mathbb{E}}(\mathbf{G}^*_{(\mathbf{Z})},\mathbf{S}^*_{(\mathbf{Z})}|\mathbf{Z}) \geq 0$$

We thus have by manipulate the terms:

$$\begin{aligned}
\mathbb{E}(\mathbf{G}^*_{(\mathbf{Z})},\mathbf{S}^*_{(\mathbf{Z})}) &= \mathbb{E}(\mathbf{G}^*_{(\mathbf{Z})},\mathbf{S}^*_{(\mathbf{Z})}) - \mathbb{E}(\mathbf{G}^*,\mathbf{S}^*) + \mathbb{E}(\mathbf{G}^*,\mathbf{S}^*) \\
&\leq \hat{\mathbb{E}}(\mathbf{G}^*,\mathbf{S}^*|\mathbf{Z}) - \hat{\mathbb{E}}(\mathbf{G}^*_{(\mathbf{Z})},\mathbf{S}^*_{(\mathbf{Z})}|\mathbf{Z}) + \mathbb{E}(\mathbf{G}^*_{(\mathbf{Z})},\mathbf{S}^*_{(\mathbf{Z})}) - \mathbb{E}(\mathbf{G}^*,\mathbf{S}^*) + \mathbb{E}(\mathbf{G}^*,\mathbf{S}^*)
\end{aligned}$$

This is to say

$$\begin{aligned}
\mathbb{E}(\mathbf{G}^*_{(\mathbf{Z})},\mathbf{S}^*_{(\mathbf{Z})}) - \mathbb{E}(\mathbf{G}^*,\mathbf{S}^*) &\leq \hat{\mathbb{E}}(\mathbf{G}^*,\mathbf{S}^*|\mathbf{Z}) - \hat{\mathbb{E}}(\mathbf{G}^*_{(\mathbf{Z})},\mathbf{S}^*_{(\mathbf{Z})}|\mathbf{Z}) + \mathbb{E}(\mathbf{G}^*_{(\mathbf{Z})},\mathbf{S}^*_{(\mathbf{Z})}) - \mathbb{E}(\mathbf{G}^*,\mathbf{S}^*) \\
&\leq \sup_{\mathbf{G},\mathbf{S}}|\mathbb{E}(\mathbf{G},\mathbf{S}) - \hat{\mathbb{E}}(\mathbf{G},\mathbf{S}|\mathbf{Z})| + \hat{\mathbb{E}}(\mathbf{G}^*,\mathbf{S}^*|\mathbf{Z}) - \mathbb{E}(\mathbf{G}^*,\mathbf{S}^*)
\end{aligned}$$

The last two terms can be upper bounded using Hoeffding inequality (Theorem 6.14). With probability of at least $1-\delta$, we have that

$$\begin{aligned}
\mathbb{E}(\mathbf{G}^*_{(\mathbf{Z})},\mathbf{S}^*_{(\mathbf{Z})}) - \mathbb{E}(\mathbf{G}^*,\mathbf{S}^*) &\leq \sup_{\mathbf{G},\mathbf{S}}|\mathbb{E}(\mathbf{G},\mathbf{S}) - \hat{\mathbb{E}}(\mathbf{G},\mathbf{S}|\mathbf{Z})| + \sqrt{\frac{\log(2/\delta)}{2nT}} \\
&\leq 2L\alpha\sqrt{\frac{2(K+12)\mathcal{C}_1(\mathbf{X})}{Tn}} + 2L\alpha\sqrt{\frac{8\mathcal{C}_\infty(\mathbf{X})\ln(2K)}{n}} + \sqrt{\frac{8\ln 2/\delta}{Tn}} \quad \text{(Theorem 6.6)}
\end{aligned}$$

This completes the proof. □

Theorem 3.1 is an immediate consequence of the above theorem given $T=1$.

II.3 Background Materials

Below are some results for concentration inequality, Rademacher complexities we used in the proof. Many theorems below are standard in the learning theory field and thus we don't give detailed complete proof. The following Lemma links the expectation of a random variable to its density function

Lemma 6.7. *For any nonnegative random variable X with corresponding density function $f(x)$,*

$$\mathbb{E}[X] = \int_0^\infty P(X\geq t)dt = \int_0^\infty xf(x)dx$$

PROOF.

$$\begin{aligned}\int_0^\infty P(X \geq t)dt &= \int_0^\infty \int_t^\infty f(x)dxdt \quad \text{(by definition)} \\ &= \int_0^\infty \int_0^x f(x)dtdx \quad \text{(order of integration)} \\ &= \int_0^\infty xf(x)dx\end{aligned}$$

□

THEOREM 6.8. *(Concentration Inequality 1: Bounded difference inequality). Let $F : \mathcal{X}^n \to \mathbb{R}$ and define A by*

$$A^2 = \sup_{\boldsymbol{x}\in\mathcal{X}^n} \sum_{k=1}^n \sup_{y_1,y_2\in\mathcal{X}} (F(\boldsymbol{x}_{k\to y_1}) - F(\boldsymbol{x}_{k\to y_2}))^2$$

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a vector of independent random variables with values in $\mathcal{X}$, and let $\mathbf{X}'$ be i.i.dto $\mathbf{X}$. Then for any $s > 0$

$$\Pr\{F(\mathbf{X}) > \mathbb{E}F(\mathbf{X}') + s\} \leq e^{-2s^2/A^2}$$

PROOF. See McDiamid's inequality. □

THEOREM 6.9. *(Concentration Inequality 2). Let $F : \mathcal{X}^n \to \mathbb{R}$ and define B by*

$$B^2 = sup_{\boldsymbol{x}\in\mathcal{X}^n} \sum_{k=1}^n \left(F(\boldsymbol{x}) - \inf_{y\in\mathcal{X}} F(\boldsymbol{x}_{k\to y})\right)$$

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a vector of independent random variables with values in $\mathcal{X}$, and let $\mathbf{X}'$ be i.i.dto $\mathbf{X}$. Then for any $s > 0$

$$\Pr\{F(\mathbf{X}) > \mathbb{E}F(\mathbf{X}') + s\} \leq e^{-s^2/(2B^2)}.$$

DEFINITION 6.10. *(Rademacher variable and complexity). The Rademacher variable $\{\sigma_n\}_{n=1}^N$ is a set of random variables independently taking either value from $-1, 1$ with equal probability. For $A \subseteq \mathbb{R}^n$, the Rademacher average of A is defined by*

$$\mathcal{R}(A) = \mathbb{E}_\sigma \sup_{(x_1,\ldots,x_n)\in A} \frac{2}{n} \sum_{i=1}^n \sigma_i x_i$$

Let $\mathbf{F}$ be a class of real valued functions on a space $\mathbf{X}$ and $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbf{X}^n$, we write

$$\mathbf{F}(\boldsymbol{x}) = \mathbf{F}(x_1, \ldots, x_n) = (f(x_1), \ldots, f(x_n)) \subseteq \mathbb{R}^n.$$

The empirical Rademacher complexity of $\mathbf{F}$ on $\boldsymbol{x}$ is given by

$$\hat{\mathcal{R}}(\mathbf{F}|\boldsymbol{x}) := \mathbb{E}_{\boldsymbol{\sigma}} \sup_{f\in\mathbf{F}} \frac{2}{n} \sum_{i=1}^n \sigma_i f(x_i)$$

Let $\mu_1, \ldots, \mu_m$ be probability measures on $\mathbf{X}$ with product measure $\boldsymbol{\mu} = \prod_i \mu_i$ on $\mathbf{X}^m$, we define the (non-empirical) Rademacher complexity as

$$\mathcal{R}(\mathbf{F}) := \mathbb{E}_{\boldsymbol{x}\sim\boldsymbol{\mu}} \hat{\mathcal{R}}(\mathbf{F}|\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{x}\sim\boldsymbol{\mu}} \mathbb{E}_{\boldsymbol{\sigma}} \sup_{f\in\mathbf{F}} \frac{2}{n} \sum_{i=1}^n \sigma_i f(x_i).$$

Here we define the Rademacher complexity using the factor $2/n$ instead of $1/n$ in order to be more convenient in the following Theorem.

The following two Theorems leads to the generalization bound based on the above Rademacher complexity.

THEOREM 6.11. *Define random variable*

$$\varphi(\boldsymbol{x}) = \sup_{f\in\mathbf{F}} \mathbb{E}_{\boldsymbol{x}} f - \hat{\mathbb{E}} f = \sup_{f\in\mathbf{F}} \left(\frac{1}{m} \sum_{i=1}^m \mathbb{E}_{x\sim\mu_i}[f(x)] - \frac{1}{m} \sum_{i=1}^m f(x_i)\right) = \sup_{f\in\mathbf{F}} \frac{1}{m} \sum_{i=1}^m (\mathbb{E}_{x\sim\mu_i}[f(x)] - f(x_i))$$

Then we have $\mathbb{E}_{\boldsymbol{x}\sim\boldsymbol{\mu}} \varphi(\boldsymbol{x}) \leq \mathcal{R}(\mathbf{F}) = \mathbb{E}_{\boldsymbol{x}\sim\boldsymbol{\mu}} \hat{\mathcal{R}}(\mathbf{F}|\boldsymbol{x})$.

THEOREM 6.12. *(Generalization Error Bound based on Rademacher Complexity). Let* $\mathbf{F}$ *be a* $[0,1]$*-valued function class on a space* $\mathbf{X}$*, and* μ *as above, For* $\delta > 0$*, we have with probability of at least* $1-\delta$ *in sample* $\boldsymbol{x} \sim \boldsymbol{\mu}$ *and* $\forall f \in \mathbf{F}$

$$\mathbb{E}_{\boldsymbol{x}\sim\boldsymbol{\mu}}[f(x)] \leq \frac{1}{m}\sum_{i=1}^{m} f(x_i) + \mathcal{R}(\mathbf{F}(x)) + \sqrt{\frac{\ln(1/\delta)}{2m}}$$

$$\mathbb{E}_{\boldsymbol{x}\sim\boldsymbol{\mu}}[f(x)] \leq \frac{1}{m}\sum_{i=1}^{m} f(x_i) + \hat{\mathcal{R}}(\mathbf{F}(x)|\boldsymbol{x}) + \sqrt{\frac{9\ln(2/\delta)}{2m}}$$

where m *is the sample size.*

Denote $\mathbb{E}f := \mathbb{E}_{x\sim\boldsymbol{\mu}}[f(x)]$ and $\hat{\mathbb{E}}f := \frac{1}{m}\sum_{i=1}^{m} f(x_i)$, we can express above as

$$\mathbb{E}f \leq \hat{\mathbb{E}}f + \mathcal{R}(\mathbf{F}(x)) + \sqrt{\frac{\ln(1/\delta)}{2m}}$$

$$\mathbb{E}f \leq \hat{\mathbb{E}}f + \hat{\mathcal{R}}(\mathbf{F}(x)|\boldsymbol{x}) + \sqrt{\frac{9\ln(2/\delta)}{2m}}$$

And the theorem can be interpret as the generalization bound:

$$\sup_{f\in\mathbf{F}} |\mathbb{E}f - \hat{\mathbb{E}}f| \leq \mathcal{R}(\mathbf{F}(x)) + \sqrt{\frac{\ln(1/\delta)}{2m}}$$

$$\sup_{f\in\mathbf{F}} |\mathbb{E}f - \hat{\mathbb{E}}f| \leq \hat{\mathcal{R}}(\mathbf{F}(x)|\boldsymbol{x}) + \sqrt{\frac{9\ln(2/\delta)}{2m}}$$

The following Lemma is used to remove loss function from the analysis

LEMMA 6.13. *Let* $A \subseteq \mathbb{R}^n$*, and let* $\phi_1, \ldots, \phi_n$ *be real functions such that* $\phi_i(s) - \phi_i(t) \leq L|s-t|, \forall i$ *and* $s, t \in \mathbb{R}$*. Define* $\boldsymbol{\phi}(A) = \phi_1(x_1), \ldots, \phi_n(x_n) : (x_1, \ldots, x_n) \in A$*. Then*

$$\mathcal{R}(\boldsymbol{\phi}(A)) \leq L\mathcal{R}(A)$$

THEOREM 6.14. *(Concentration Inequality 3: Hoeffding). Let* $Z_1, \ldots Z_n$ *be* m *i.i.d random variables with* $f(Z) \in [a,b]$*. Then* $\forall \varepsilon > 0$ *we have:*

$$\Pr\left[\frac{1}{m}\sum_{i=1}^{m} f(Z_i) - \mathbb{E}[f(Z)] > \varepsilon\right] \leq \exp\left(-\frac{2n\varepsilon^2}{(b-a)^2}\right)$$

Set $\delta = \exp\left(-\frac{2n\varepsilon^2}{(b-a)^2}\right)$*, the consequence of the above inequality is: with probability at least* $1-\delta$*,*

$$\hat{\mathbb{E}}f - \mathbb{E}f \leq (b-a)\sqrt{\frac{\log(1/\delta)}{2m}}.$$

The corresponding version for two tails:

$$\Pr\left[\left|\frac{1}{m}\sum_{i=1}^{m} f(Z_i) - \mathbb{E}[f(Z)]\right| > \varepsilon\right] \leq 2\exp\left(-\frac{2n\varepsilon^2}{(b-a)^2}\right)$$

Equivalently, with probability of at least $1-\delta$*,*

$$|\hat{\mathbb{E}}f - \mathbb{E}f| \leq (b-a)\sqrt{\frac{\log(2/\delta)}{2m}}.$$

Appendix III. ADNI Data, Feature List and Feature Groups Learned by FeaFiner

III.1 The ADNI Data

Data used in the preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public- private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimers disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California San Francisco. ADNI is the result of efforts of many co- investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

III.2 Feature List

The patients involved in the study of FeaFiner are those who have baseline MRI scans and the scans have passed the quality controls. There are in total 306 features used in the study of FeaFiner, including the baseline MMSE score and 305 imaging features. The imaging features are extract by FreeSurfer from the MRI scans, and the names of the features are list below:

Cortical Thickness Average There are 68 bilateral symmetric cortical thickness average features: L. Bankssts, L. Cau.Ant.Cingulate, L. Cau.Mid.Frontal, L. Cuneus, L. Entorhinal, L. FrontalPole, L. Fusiform, L. Inf.Parietal, L. Inf.Temporal, L. Insula, L. IsthmusCingulate, L. Lat.Occipital, L. Lat.Orbitrontal, L. Lingual, L. Med.Orbitrontal, L. Mid.Temporal, L. Paracentral, L. Parahippocampal, L. ParsOpercularis, L. ParsOrbitalis, L. ParsTriangularis, L. Pericalcarine, L. Postcentral, L. Pos.Cingulate, L. Precentral, L. Precuneus, L. Ros. Ant. Cingulate, L. Ros.Mid.Frontal, L. Sup.Frontal, L. Sup.Parietal, L. Sup.Temporal, L. Supramarginal, L. TemporalPole, L. Tra.Temporal, R. Bankssts, R. Cau.Ant.Cingulate, R. Cau.Mid.Frontal, R. Cuneus, R. Entorhinal, R. FrontalPole, R. Fusiform, R. Inf.Parietal, R. Inf.Temporal, R. Insula, R. IsthmusCingulate, R. Lat.Occipital, R. Lat.Orbitrontal, R. Lingual, R. Med.Orbitrontal, R. Mid.Temporal, R. Paracentral, R. Parahippocampal, R. ParsOpercularis, R. ParsOrbitalis, R. ParsTriangularis, R. Pericalcarine, R. Postcentral, R. Pos.Cingulate, R. Precentral, R. Precuneus, R. Ros.Ant.Cingulate, R. Ros.Mid.Frontal, R. Sup.Frontal, R. Sup.Parietal, R. Sup.Temporal, R. Supramarginal, R. TemporalPole, R. Tra.Temporal,

Cortical Thickness Standard Deviation There are 68 bilateral symmetric cortical thickness average features: L. Bankssts, L. Cau.Ant.Cingulate, L. Cau.Mid.Frontal, L. Cuneus, L. Entorhinal, L. FrontalPole, L. Fusiform, L. Inf.Parietal, L. Inf.Temporal, L. Insula, L. IsthmusCingulate, L. Lat.Occipital, L. Lat.Orbitrontal, L. Lingual, L. Med.Orbitrontal, L. Mid.Temporal, L. Paracentral, L. Parahippocampal, L. ParsOpercularis, L. ParsOrbitalis, L. ParsTriangularis, L. Pericalcarine, L. Postcentral, L. Pos. Cingulate, L. Precentral, L. Precuneus, L. Ros. Ant. Cingulate, L. Ros. Mid. Frontal, L. Sup.Frontal, L. Sup.Parietal, L. Sup.Temporal, L. Supramarginal, L. TemporalPole, L. Tra.Temporal, R. Bankssts, R. Cau.Ant.Cingulate, R. Cau.Mid.Frontal, R. Cuneus, R. Entorhinal, R. FrontalPole, R. Fusiform, R. Inf.Parietal, R. Inf.Temporal, R. Insula, R. IsthmusCingulate, R. Lat.Occipital, R. Lat.Orbitrontal, R. Lingual, R. Med.Orbitrontal, R. Mid.Temporal, R. Paracentral, R. Parahippocampal, R. ParsOpercularis, R. ParsOrbitalis, R. ParsTriangularis, R. Pericalcarine, R. Postcentral, R. Pos.Cingulate, R. Precentral, R. Precuneus, R. Ros.Ant.Cingulate, R. Ros.Mid.Frontal, R. Sup.Frontal, R. Sup.Parietal, R. Sup.Temporal, R. Supramarginal, R. TemporalPole, R. Tra.Temporal,

Surface Area There are 70 bilateral symmetric surface area features: L. Bankssts, L. Cau.Ant.Cingulate, L. Cau.Mid.Frontal, L. Cuneus, L. Entorhinal, L. FrontalPole, L. Fusiform, L. Hemisphere, L. Inf.Parietal, L. Inf.Temporal, L. Insula, L. IsthmusCingulate, L. Lat.Occipital, L. Lat.Orbitrontal, L. Lingual, L. Med.Orbitrontal, L. Mid.Temporal, L. Paracentral, L. Parahippocampal, L. ParsOpercularis, L. ParsOrbitalis, L. ParsTriangularis, L. Pericalcarine, L. Postcentral, L. Pos.Cingulate, L. Precentral, L. Precuneus, L. Ros.Ant.Cingulate, L. Ros.Mid.Frontal, L. Sup.Frontal, L. Sup.Parietal, L. Sup.Temporal, L. Supramarginal, L. TemporalPole, L. Tra.Temporal, R. Bankssts, R. Cau.Ant.Cingulate, R. Cau.Mid.Frontal, R. Cuneus, R. Entorhinal, R. FrontalPole, R. Fusiform, R. Hemisphere, R. Inf.Parietal, R. Inf.Temporal, R. Insula, R. IsthmusCingulate, R. Lat.Occipital, R. Lat.Orbitrontal, R. Lingual, R. Med.Orbitrontal, R. Mid.Temporal, R. Paracentral, R. Parahippocampal, R. ParsOpercularis, R. ParsOrbitalis, R. ParsTriangularis, R. Pericalcarine, R. Postcentral, R. Pos.Cingulate, R. Precentral, R. Precuneus, R. Ros.Ant.Cingulate, R. Ros.Mid.Frontal, R. Sup.Frontal, R. Sup.Parietal, R. Sup.Temporal, R. Supramarginal, R. TemporalPole, R. Tra.Temporal.

Volume (Cortical Parcellation) There are 68 volume features from cortical parcellation: Icv, L. Bankssts, L. Cau.Ant.Cingulate, L. Cau.Mid.Frontal, L. Cuneus, L. Entorhinal, L. FrontalPole, L. Fusiform, L. Inf.Parietal, L. Inf.Temporal, L. Insula, L. IsthmusCingulate, L. Lat.Occipital, L. Lat.Orbitrontal, L. Lingual, L. Med.Orbitrontal, L. Mid.Temporal, L. Paracentral, L. Parahippocampal, L. ParsOpercularis, L. ParsOrbitalis, L. ParsTriangularis, L. Pericalcarine, L. Postcentral, L. Pos.Cingulate, L. Precentral, L. Precuneus, L. Ros.Ant.Cingulate, L. Ros.Mid.Frontal, L. Sup.Frontal, L. Sup.Parietal, L. Sup.Temporal, L. Supramarginal, L. TemporalPole, L. Tra.Temporal, R. Bankssts, R. Cau.Ant.Cingulate, R. Cau.Mid.Frontal, R. Cuneus, R. FrontalPole, R. Fusiform, R. Inf.Parietal, R. Inf.Temporal, R. Insula, R. IsthmusCingulate, R. Lat.Occipital, R. Lat.Orbitrontal, R. Lingual, R. Med.Orbitrontal, R. Mid.Temporal, R. Paracentral, R. Parahippocampal, R. ParsOpercularis, R. ParsOrbitalis, R. ParsTriangularis, R. Pericalcarine, R. Postcentral, R. Pos.Cingulate, R. Precentral, R. Precuneus, R. Ros.Ant.Cingulate, R. Ros.Mid.Frontal, R. Sup.Frontal, R. Sup.Parietal, R. Sup.Temporal, R. Supramarginal, R. TemporalPole, R. Tra.Temporal.

Volume (White Matter Parcellation) There are 31 volume features from white matter parcellation: Brainstem, CorpusCallosumCentral, CorpusCallosumMidAnt., Csf, FourthVentricle, L. Amygdala, L. Caudate, L. CerebellumCortex, L. CerebellumWM, L. CerebralCortex, L. CerebralWM, L. ChoroidPlexus, L. Hippocampus, L. Pallidum, L. Putamen, L. Thalamus, L. VentralDC, OpticChiasm, R. Amygdala, R. Caudate, R. CerebellumCortex, R. CerebellumWM, R. CerebralCortex, R. CerebralWM, R. ChoroidPlexus, R. Hippocampus, R. Pallidum, R. Putamen, R. Thalamus, R. VentralDC, ThirdVentricle.

Some abbreviations are used in the list: Superior (Sup), Inferior (Inf), Middle (Mid), Lateral (Lat), Posterior (Pos), Anterior (Ant), Transverse (Tra), Medial (Med), Rostral (Ros), Caudal (Cau).

III.3 High Level Concepts Learned by FeaFiner

In this section we present the high-level concept (feature groups) obtained from building predictive models on the ADNI data.

Table 2: Partial high level concepts (feature groups) obtained from MMSE M12 predictive modeling via FeaFiner.

CT Avg. R.ParsTriangularis, CT Avg. L.ParsTriangularis, CT Avg. R.ParsOrbitalis, CT Avg. L.ParsOpercularis, CT Avg. L.Ros.Mid.Frontal, CT Avg. R.Sup.Frontal, CT Avg. L.Sup.Frontal, CT Avg. R.Ros.Mid.Frontal, CT Avg. R.ParsOpercularis
MMSE, Vol.(WM) R.Hippocampus, Vol.(WM) L.Hippocampus, Vol.(WM) L.Amygdala, Vol.(WM) R.Amygdala
CT Avg. L.Inf.Temporal, CT Avg. L.Sup.Temporal, CT Avg. L.Fusiform, CT Avg. L.Bankssts, CT Avg. L.Mid.Temporal
Vol.(CO) R.Precentral, Vol.(CO) R.Paracentral, Vol.(CO) L.Precentral, Vol.(CO) R.Postcentral, Vol.(CO) L.Lat.Occipital, Vol.(CO) L.Postcentral
CT Std. L.ParsOpercularis, CT Std. R.Bankssts, CT Std. R.ParsTriangularis, CT Std. R.ParsOpercularis, CT Std. L.ParsTriangularis, CT Std. L.Bankssts
CT Std. R.Pos.Cingulate, CT Std. L.Pos.Cingulate, CT Std. R.Ros.Ant.Cingulate
CT Std. R.Cau.Ant.Cingulate, CT Std. L.Cau.Ant.Cingulate, CT Std. R.Cuneus, CT Std. L.Ros.Ant.Cingulate
CT Avg. R.Lat.Occipital, CT Avg. R.Bankssts, CT Avg. R.Fusiform, CT Avg. R.Mid.Temporal, CT Avg. R.Inf.Parietal, CT Avg. R.Supramarginal, CT Avg. R.Sup.Temporal, CT Avg. R.Inf.Temporal
CT Std. R.Sup.Frontal, CT Std. L.Sup.Frontal, CT Std. L.Paracentral, CT Std. R.Cau.Mid.Frontal, CT Std. R.Paracentral, CT Std. L.Precentral, CT Std. R.Precentral
CT Std. R.Insula, CT Std. L.Insula
CT Avg. R.Postcentral, CT Avg. R.Paracentral, CT Avg. L.Cau.Mid.Frontal, CT Avg. R.Precentral, CT Avg. L.Precentral, CT Avg. R.Sup.Parietal, CT Avg. L.Postcentral, CT Avg. L.Lat.Occipital, CT Avg. R.Cau.Mid.Frontal, CT Avg. L.Sup.Parietal, CT Avg. L.Paracentral, CT Avg. R.Precuneus, CT Avg. L.Supramarginal, CT Avg. L.Inf.Parietal, CT Avg. L.Precuneus
Suf. Area L.Sup.Parietal, Vol.(CO) L.Sup.Parietal, Suf. Area R.Sup.Parietal, Vol.(CO) R.Sup.Parietal
Suf. Area R.Bankssts, Vol.(CO) R.Bankssts
Suf. Area R.ParsOrbitalis, Vol.(CO) R.ParsOrbitalis, Vol.(CO) L.ParsOrbitalis, Suf. Area L.ParsOrbitalis
Vol.(WM) R.Pallidum, Vol.(WM) R.Putamen, Vol.(WM) R.Caudate, Vol.(WM) L.Putamen, Vol.(WM) L.Pallidum, Vol.(WM) L.Caudate
CT Std. L.Lat.Occipital, CT Std. L.Cuneus, CT Std. R.Lingual, CT Std. R.Lat.Occipital, CT Std. L.Lingual
CT Std. R.FrontalPole, Suf. Area R.FrontalPole, Vol.(CO) R.FrontalPole
Vol.(CO) R.TemporalPole, Vol.(CO) L.TemporalPole, CT Avg. L.Entorhinal, CT Avg. L.TemporalPole, Vol.(CO) L.Entorhinal, CT Avg. R.TemporalPole, CT Avg. R.Entorhinal
Vol.(CO) R.Parahippocampal, CT Avg. R.Parahippocampal, Vol.(CO) L.Parahippocampal, CT Avg. L.Parahippocampal
CT Std. L.Inf.Parietal, CT Std. L.Postcentral, CT Std. R.Sup.Parietal, CT Std. R.Supramarginal, CT Std. R.Postcentral, CT Std. L.Sup.Parietal, CT Std. L.Supramarginal, CT Std. L.Cau.Mid.Frontal, CT Std. R.Inf.Parietal
Suf. Area R.ParsOpercularis, Vol.(CO) R.ParsOpercularis
Vol.(CO) R.Ros.Ant.Cingulate, Suf. Area R.Ros.Ant.Cingulate
Suf. Area L.ParsOpercularis, Vol.(CO) L.ParsOpercularis
Vol.(CO) R.ParsTriangularis, Suf. Area L.ParsTriangularis, Vol.(CO) L.ParsTriangularis, Suf. Area R.ParsTriangularis

Table 3: Partial high level concepts (feature groups) obtained from MMSE M24 predictive modeling via FeaFiner.

Vol.(CO) L.Pericalcarine, Vol.(CO) L.Lingual, Suf. Area L.Lingual, Suf. Area L.Pericalcarine, Suf. Area R.Lingual, Suf. Area R.Pericalcarine, Vol.(CO) R.Lingual, Vol.(CO) R.Pericalcarine
Vol.(CO) L.Cuneus, CT Std. L.Cuneus, CT Avg. L.Cuneus
Vol.(WM) R.Putamen, Vol.(WM) R.Caudate, Vol.(WM) L.Putamen, Vol.(WM) L.Caudate
Vol.(CO) R.Ros.Mid.Frontal, Vol.(CO) R.Inf.Parietal, Vol.(CO) L.Supramarginal, Vol.(CO) L.Inf.Parietal, Vol.(CO) L.Sup.Temporal, Vol.(CO) R.Supramarginal, Vol.(CO) R.Fusiform, Vol.(CO) L.Precuneus, Vol.(CO) L.Ros.Mid.Frontal, Vol.(CO) L.Fusiform, Vol.(CO) L.Mid.Temporal, Vol.(WM) L.CerebralCortex, Vol.(CO) L.Insula, Vol.(CO) R.Pos.Cingulate, Vol.(CO) R.Precuneus, Vol.(WM) R.CerebralCortex, Vol.(CO) L.Ros.Ant.Cingulate, Vol.(CO) R.IsthmusCingulate, Vol.(CO) R.Mid.Temporal, Vol.(CO) R.Sup.Temporal, Vol.(CO) L.Precentral, Vol.(CO) L.Sup.Frontal, Vol.(CO) R.Insula, Vol.(CO) R.Sup.Frontal, Vol.(CO) R.Precentral
CT Std. R.Postcentral, CT Std. L.Inf.Parietal, CT Std. R.Precentral, CT Std. R.Inf.Parietal, CT Std. L.Sup.Parietal, CT Std. L.Postcentral, CT Std. L.Precentral, CT Std. L.Lat.Occipital, CT Std. R.Sup.Parietal, CT Std. R.Lat.Occipital, CT Std. R.Supramarginal
CT Std. L.Entorhinal, CT Std. L.Lingual, CT Std. R.Entorhinal, CT Std. R.Lingual
Suf. Area R.TemporalPole, Suf. Area L.Entorhinal, Suf. Area L.TemporalPole, Vol.(WM) OpticChiasm, Suf. Area R.Entorhinal
Vol.(CO) R.FrontalPole, Suf. Area L.FrontalPole, Vol.(CO) L.FrontalPole, Suf. Area R.FrontalPole
Vol.(WM) CorpusCallosumCentral, Vol.(WM) CorpusCallosumMidAnt.
CT Std. R.Cuneus, CT Avg. R.Cuneus
CT Avg. L.TemporalPole, Vol.(CO) L.TemporalPole, CT Avg. R.TemporalPole, Vol.(CO) R.TemporalPole
Vol.(WM) L.ChoroidPlexus, Vol.(WM) R.ChoroidPlexus
CT Std. L.Lat.Orbitrontal, CT Std. R.Lat.Orbitrontal, CT Std. R.ParsTriangularis, CT Std. R.Ros.Mid.Frontal, CT Std. L.Ros.Mid.Frontal, CT Std. L.ParsTriangularis, CT Std. R.FrontalPole, CT Std. L.ParsOrbitalis, CT Std. L.FrontalPole, CT Std. R.ParsOrbitalis
CT Avg. L.Lat.Orbitrontal, CT Avg. R.ParsOrbitalis, CT Avg. L.Med.Orbitrontal, CT Avg. R.FrontalPole, CT Avg. L.ParsOrbitalis, CT Avg. R.Lat.Orbitrontal, CT Avg. R.Med.Orbitrontal, CT Avg. L.FrontalPole
CT Std. L.Paracentral, Vol.(CO) L.Paracentral, CT Std. R.Paracentral, Vol.(CO) R.Paracentral
Vol.(CO) L.Parahippocampal, Vol.(CO) R.Parahippocampal, CT Avg. L.Parahippocampal, CT Std. L.Parahippocampal, CT Avg. R.Parahippocampal, CT Std. R.Parahippocampal
Suf. Area R.Cau.Mid.Frontal, Vol.(CO) R.Cau.Mid.Frontal
CT Avg. R.Inf.Parietal, CT Avg. R.IsthmusCingulate, CT Avg. R.Inf.Temporal, CT Avg. R.Mid.Temporal, CT Avg. R.Bankssts, CT Avg. R.Fusiform, CT Avg. R.Pos.Cingulate, CT Avg. R.Supramarginal, CT Avg. R.Sup.Temporal, CT Avg. R.Insula
Suf. Area L.Bankssts, Vol.(CO) L.Bankssts
Suf. Area R.ParsOpercularis, Vol.(CO) R.ParsOpercularis
Suf. Area R.Cuneus, Vol.(CO) R.Cuneus, Vol.(CO) L.Lat.Occipital, Vol.(CO) R.Lat.Occipital, Suf. Area L.Cuneus, Suf. Area L.Lat.Occipital, Suf. Area R.Lat.Occipital
CT Std. L.Pos.Cingulate, CT Std. L.IsthmusCingulate, CT Std. R.IsthmusCingulate
CT Avg. L.Pericalcarine, CT Avg. R.Pericalcarine, CT Std. R.Pericalcarine, CT Std. L.Pericalcarine
CT Std. R.Sup.Frontal, CT Std. R.Bankssts, CT Std. L.Supramarginal, CT Std. L.ParsOpercularis, CT Std. L.Sup.Frontal, CT Std. L.Bankssts, CT Std. R.ParsOpercularis, CT Std. R.Cau.Mid.Frontal, CT Std. L.Cau.Mid.Frontal
Suf. Area R.Med.Orbitrontal, Suf. Area L.Lat.Orbitrontal, Suf. Area L.Med.Orbitrontal, Suf. Area R.Lat.Orbitrontal, Suf. Area R.ParsOrbitalis, Suf. Area R.IsthmusCingulate

Table 4: Partial high level concepts (feature groups) obtained from MMSE M36 predictive modeling via FeaFiner.

CT Std. R.Lat.Orbitrontal, CT Std. L.ParsOrbitalis, CT Std. R.ParsTriangularis, CT Std. L.ParsTriangularis, CT Std. R.ParsOrbitalis, CT Std. L.Med.Orbitrontal, CT Std. L.Ros.Mid.Frontal, CT Std. R.Ros.Mid.Frontal, CT Std. R.Med.Orbitrontal
Vol.(CO) L.FrontalPole, Suf. Area L.FrontalPole
Vol.(CO) L.Lat.Orbitrontal, Vol.(CO) L.IsthmusCingulate, Vol.(CO) R.Pos.Cingulate, Vol.(CO) L.Med.Orbitrontal, Vol.(CO) L.Insula, Vol.(CO) R.Lat.Orbitrontal, Vol.(CO) R.Insula, Vol.(CO) R.Med.Orbitrontal, Vol.(CO) L.Ros.Ant.Cingulate
CT Std. L.Cuneus, CT Std. L.Lat.Occipital, CT Std. L.Lingual, CT Std. R.Lingual, CT Std. R.Lat.Occipital
Vol.(CO) R.Postcentral, Suf. Area R.Postcentral
Vol.(CO) L.Bankssts, Suf. Area L.Bankssts
CT Std. L.Paracentral, CT Std. R.Paracentral
Vol.(CO) L.ParsTriangularis, Suf. Area L.ParsTriangularis, Vol.(CO) R.ParsTriangularis, Suf. Area R.ParsTriangularis
Vol.(CO) R.Bankssts, Suf. Area R.Bankssts, Vol.(CO) R.Inf.Temporal, Vol.(CO) L.Sup.Temporal, Vol.(CO) L.Inf.Parietal, Vol.(CO) R.Mid.Temporal, Vol.(CO) R.Fusiform, Vol.(CO) L.Fusiform, Vol.(CO) R.Sup.Temporal, Vol.(CO) L.Mid.Temporal, Vol.(CO) L.Inf.Temporal, Vol.(CO) R.Inf.Parietal
Vol.(CO) R.IsthmusCingulate, Vol.(CO) R.Sup.Parietal, Vol.(CO) R.Sup.Frontal, Vol.(CO) L.Precuneus, Vol.(WM) R.CerebralCortex, Vol.(WM) L.CerebralCortex, Vol.(CO) L.Sup.Parietal, Vol.(CO) L.Ros.Mid.Frontal, Vol.(CO) L.ParsOrbitalis, Vol.(CO) R.Supramarginal, Vol.(CO) L.Sup.Frontal, Vol.(CO) R.Precentral, Vol.(CO) L.Postcentral, Vol.(CO) L.Precentral, Vol.(CO) R.Ros.Mid.Frontal, Vol.(CO) L.Supramarginal, Vol.(CO) R.Precuneus
Vol.(WM) R.CerebellumCortex, Vol.(WM) L.CerebellumCortex
CT Avg. R.Parahippocampal, Vol.(CO) R.Parahippocampal, CT Avg. L.Parahippocampal, CT Std. L.Parahippocampal, CT Std. R.Parahippocampal, Vol.(CO) L.Parahippocampal
CT Std. L.Postcentral, CT Std. L.Supramarginal, CT Std. L.Precentral, CT Std. R.Supramarginal, CT Std. R.Precentral, CT Std. R.Postcentral, CT Std. L.Sup.Parietal, CT Std. R.Inf.Parietal, CT Std. R.Sup.Parietal, CT Std. R.Cau.Mid.Frontal, CT Std. L.Inf.Parietal
Suf. Area L.ParsOpercularis, Vol.(CO) L.ParsOpercularis
Suf. Area R.ParsOpercularis, Vol.(CO) R.ParsOpercularis, Suf. Area L.Cau.Mid.Frontal, Vol.(CO) L.Cau.Mid.Frontal, Suf. Area R.Cau.Mid.Frontal, Vol.(CO) R.Cau.Mid.Frontal
MMSE, CT Avg. R.Inf.Parietal, CT Avg. R.Supramarginal, Suf. Area L.Supramarginal, CT Std. R.Sup.Temporal, Vol.(WM) L.Amygdala, CT Std. L.Cau.Mid.Frontal, Vol.(WM) R.Hippocampus, Vol.(WM) L.Hippocampus, Vol.(WM) R.Amygdala
CT Std. R.Mid.Temporal, CT Std. L.Sup.Frontal, CT Std. L.Lat.Orbitrontal, CT Std. R.Sup.Frontal, CT Std. R.Inf.Temporal, CT Std. L.Sup.Temporal, CT Std. L.Mid.Temporal, CT Std. L.Inf.Temporal, CT Std. R.ParsOpercularis
CT Std. L.Precuneus, CT Avg. R.Cau.Ant.Cingulate, CT Std. L.IsthmusCingulate, CT Std. R.Precuneus, CT Std. R.Fusiform, CT Std. R.IsthmusCingulate, CT Std. L.Fusiform
Suf. Area R.Lat.Occipital, Suf. Area L.Lat.Occipital, Vol.(CO) R.Lat.Occipital, Vol.(CO) L.Lat.Occipital
Suf. Area L.TemporalPole, Suf. Area R.TemporalPole, Suf. Area L.ParsOrbitalis
Vol.(CO) R.FrontalPole, Suf. Area R.FrontalPole
Vol.(WM) L.Caudate, Vol.(WM) L.Putamen, Vol.(WM) R.Caudate, Vol.(WM) R.Putamen
CT Std. R.Entorhinal, CT Std. R.TemporalPole, CT Std. L.Entorhinal, CT Std. L.TemporalPole
CT Std. R.FrontalPole, CT Std. R.Tra.Temporal, CT Std. L.FrontalPole
Vol.(CO) L.TemporalPole, Vol.(CO) R.TemporalPole

Table 5: Partial high level concepts (feature groups) obtained from ADAS-Cog M06 predictive modeling via FeaFiner.

Suf. Area L.ParsTriangularis, Vol.(CO) L.ParsTriangularis, Suf. Area R.ParsTriangularis, Vol.(CO) R.ParsTriangularis
CT Std. L.Tra.Temporal, CT Avg. R.Tra.Temporal, CT Std. R.Tra.Temporal, CT Avg. L.Tra.Temporal
CT Std. R.Ros.Ant.Cingulate, CT Std. L.Ros.Ant.Cingulate, CT Std. L.Cau.Ant.Cingulate, CT Std. R.Cau.Ant.Cingulate
Vol.(WM) L.CerebralCortex, Vol.(CO) L.Sup.Frontal, Vol.(CO) L.Fusiform, Vol.(CO) L.Supramarginal, Vol.(CO) L.Insula, Vol.(WM) R.CerebralCortex, Vol.(CO) R.Sup.Temporal, Vol.(CO) L.Inf.Temporal, Vol.(CO) L.Ros.Mid.Frontal, Vol.(CO) R.Fusiform, Vol.(CO) R.Supramarginal, Vol.(CO) R.Sup.Frontal, Vol.(CO) R.Ros.Mid.Frontal, Vol.(CO) R.Insula, Vol.(CO) R.Mid.Temporal, Vol.(CO) R.Pos.Cingulate, Vol.(CO) R.Inf.Temporal, Vol.(CO) L.Mid.Temporal, Vol.(CO) L.Sup.Temporal
CT Std. L.Lingual, CT Std. R.Lingual
CT Avg. L.Mid.Temporal, CT Avg. L.Sup.Temporal, CT Avg. L.Fusiform, CT Avg. L.Inf.Temporal, CT Avg. L.Entorhinal, CT Avg. L.Bankssts, CT Avg. L.Insula, CT Avg. L.TemporalPole
Vol.(WM) R.CerebellumWM, Vol.(WM) L.CerebralWM, Vol.(WM) R.CerebralWM, Vol.(WM) Corpus-CallosumMidAnt., Vol.(WM) L.CerebellumWM, Vol.(WM) CorpusCallosumCentral
MMSE, Vol.(CO) L.Entorhinal, Vol.(WM) L.Hippocampus, Vol.(WM) L.Putamen, Vol.(WM) R.Hippocampus, Vol.(WM) R.Putamen, Vol.(WM) R.Amygdala
Vol.(CO) L.Lat.Orbitrontal, Vol.(CO) L.ParsOpercularis, Vol.(CO) R.Lat.Orbitrontal, Vol.(CO) R.ParsOrbitalis, Vol.(CO) R.Tra.Temporal, Vol.(CO) L.ParsOrbitalis, Vol.(CO) L.Tra.Temporal
Vol.(CO) L.Parahippocampal, Suf. Area L.Parahippocampal, Suf. Area R.Entorhinal, Vol.(CO) R.Parahippocampal, Suf. Area R.Parahippocampal
Suf. Area L.Entorhinal, Suf. Area R.TemporalPole, Suf. Area L.TemporalPole
CT Std. L.Precuneus, CT Avg. R.IsthmusCingulate, CT Avg. R.Pos.Cingulate, CT Std. R.Precuneus, CT Avg. L.IsthmusCingulate, CT Avg. L.Pos.Cingulate
Vol.(CO) R.ParsOpercularis, Suf. Area R.ParsOpercularis
Vol.(CO) L.Lingual, Suf. Area L.Lingual
CT Std. L.Parahippocampal, CT Avg. R.Parahippocampal, CT Std. R.Parahippocampal, CT Avg. L.Parahippocampal
Suf. Area R.Tra.Temporal, Suf. Area L.Paracentral, Suf. Area R.ParsOrbitalis, Suf. Area L.Lat.Orbitrontal, Suf. Area R.Med.Orbitrontal, Suf. Area L.ParsOpercularis, Suf. Area R.Lat.Orbitrontal, Suf. Area R.Precentral, Suf. Area L.Med.Orbitrontal, Suf. Area L.Cau.Mid.Frontal, Suf. Area L.ParsOrbitalis, Suf. Area R.Paracentral, Suf. Area R.Cau.Mid.Frontal, Suf. Area L.Tra.Temporal
CT Avg. R.Lingual, CT Avg. L.Lingual, Vol.(WM) L.Amygdala
CT Std. L.Cuneus, CT Std. R.Lat.Occipital, CT Std. L.Lat.Occipital
CT Std. R.Pericalcarine, CT Avg. R.Pericalcarine
Vol.(CO) R.Pericalcarine, Suf. Area R.Pericalcarine, Suf. Area R.Lingual, Vol.(CO) R.Lingual, Vol.(CO) L.Pericalcarine, Suf. Area L.Pericalcarine
CT Std. R.Bankssts, CT Std. R.ParsTriangularis, CT Std. L.ParsTriangularis, CT Std. L.ParsOpercularis, CT Std. R.ParsOpercularis, CT Std. L.Bankssts
CT Std. R.Med.Orbitrontal, CT Std. L.Med.Orbitrontal
CT Std. L.Pericalcarine, CT Avg. L.Pericalcarine
Vol.(WM) L.Thalamus, Vol.(WM) R.VentralDC, Vol.(WM) L.VentralDC, Vol.(WM) R.Thalamus, Vol.(WM) L.Pallidum, Vol.(WM) R.Pallidum, Vol.(WM) Brainstem
Vol.(CO) R.Bankssts, Vol.(CO) L.Inf.Parietal, Vol.(CO) L.Precuneus, Vol.(CO) R.Inf.Parietal, Vol.(CO) R.Precuneus, Vol.(CO) R.IsthmusCingulate

Table 6: Partial high level concepts (feature groups) obtained from ADAS-Cog M12 predictive modeling via FeaFiner.

Suf. Area R.Lingual, Vol.(CO) L.Lingual, Vol.(CO) L.Pericalcarine, Vol.(CO) R.Pericalcarine, Suf. Area L.Lingual, Suf. Area R.Pericalcarine, Suf. Area L.Pericalcarine, Suf. Area L.Cuneus, Vol.(CO) R.Lingual, Vol.(CO) L.Cuneus
Vol.(CO) L.IsthmusCingulate, Suf. Area L.IsthmusCingulate
Vol.(CO) R.Cau.Ant.Cingulate, Suf. Area R.Ros.Ant.Cingulate, Vol.(CO) R.Ros.Ant.Cingulate, Suf. Area R.Cau.Ant.Cingulate
CT Std. R.Pericalcarine, CT Std. L.Pericalcarine, CT Avg. L.Pericalcarine, CT Avg. R.Pericalcarine
CT Std. R.Postcentral, CT Std. R.Supramarginal, CT Std. L.Supramarginal, CT Std. L.Postcentral, CT Std. R.Sup.Parietal, CT Std. R.Inf.Parietal, CT Std. L.Sup.Parietal, CT Std. R.Precentral, CT Std. R.Cau.Mid.Frontal, CT Std. L.Inf.Parietal, CT Std. L.Precentral, CT Std. L.Cau.Mid.Frontal
Vol.(WM) FourthVentricle, Vol.(WM) ThirdVentricle, Vol.(WM) Csf, Vol.(WM) OpticChiasm
CT Avg. L.ParsOpercularis, CT Avg. L.ParsTriangularis, CT Avg. R.ParsTriangularis, CT Avg. R.FrontalPole, CT Avg. R.Med.Orbitrontal, CT Avg. R.Lat.Orbitrontal, CT Avg. L.ParsOrbitalis, CT Avg. L.FrontalPole, CT Avg. L.Med.Orbitrontal, CT Avg. R.ParsOrbitalis, CT Avg. L.Ros.Mid.Frontal, CT Avg. L.Lat.Orbitrontal, CT Avg. R.Ros.Mid.Frontal
CT Std. L.Parahippocampal, CT Std. R.Parahippocampal
Suf. Area R.IsthmusCingulate, Vol.(CO) R.IsthmusCingulate
Suf. Area R.Cau.Mid.Frontal, Vol.(CO) R.Cau.Mid.Frontal, Suf. Area L.Cau.Mid.Frontal, Vol.(CO) L.Cau.Mid.Frontal
CT Avg. R.Ros.Ant.Cingulate, CT Avg. L.Ros.Ant.Cingulate, CT Avg. R.Cau.Ant.Cingulate, CT Avg. L.Cau.Ant.Cingulate
CT Std. R.Precuneus, CT Std. R.IsthmusCingulate, CT Std. L.Precuneus, CT Std. L.IsthmusCingulate
Vol.(CO) R.Ros.Mid.Frontal, Vol.(CO) L.Supramarginal, Vol.(CO) L.Fusiform, Vol.(CO) R.Fusiform, Vol.(CO) L.Mid.Temporal, Vol.(CO) L.Inf.Temporal, Vol.(WM) L.CerebralCortex, Vol.(WM) R.CerebralCortex, Vol.(CO) L.Sup.Temporal, Vol.(CO) R.Supramarginal, Vol.(CO) L.Ros.Mid.Frontal, Vol.(CO) R.Pos.Cingulate, Vol.(CO) R.Inf.Temporal, Vol.(CO) L.Pos.Cingulate, Vol.(CO) L.Sup.Frontal, Vol.(CO) R.Mid.Temporal, Vol.(CO) R.Sup.Temporal, Vol.(CO) R.Sup.Frontal
Vol.(WM) R.CerebellumCortex, Vol.(WM) L.CerebellumCortex, Suf. Area L.Parahippocampal, Suf. Area R.Parahippocampal
Vol.(CO) R.Paracentral, Suf. Area R.Paracentral
CT Std. L.ParsOpercularis, CT Std. L.Tra.Temporal, CT Std. R.ParsTriangularis, CT Std. R.Bankssts, CT Std. L.Bankssts, CT Std. R.Tra.Temporal, CT Std. R.ParsOpercularis
CT Std. L.Mid.Temporal, CT Std. R.Inf.Temporal, CT Std. L.Inf.Temporal, CT Std. L.Sup.Temporal, CT Std. R.Fusiform, CT Std. L.Fusiform, CT Std. R.Sup.Temporal, CT Std. R.Mid.Temporal
CT Avg. R.Sup.Frontal, CT Avg. R.Postcentral, CT Avg. R.Paracentral, CT Avg. L.Precentral, CT Avg. L.Postcentral, CT Avg. L.Cau.Mid.Frontal, CT Avg. R.Supramarginal, CT Avg. R.Precentral, CT Avg. L.Sup.Frontal, CT Avg. L.Paracentral, CT Avg. R.Cau.Mid.Frontal, CT Avg. R.ParsOpercularis
Suf. Area L.Entorhinal, Suf. Area R.TemporalPole, Suf. Area L.TemporalPole, Suf. Area R.Entorhinal
CT Std. L.Cau.Ant.Cingulate, CT Std. R.Ros.Ant.Cingulate, CT Std. L.Ros.Ant.Cingulate, CT Std. L.Insula, CT Std. R.Cau.Ant.Cingulate, CT Std. R.Insula, CT Std. L.Pos.Cingulate, CT Std. R.Pos.Cingulate
CT Avg. L.Lat.Occipital, CT Avg. L.Lingual, CT Avg. R.Lat.Occipital, CT Avg. R.Cuneus, CT Avg. L.Cuneus, CT Avg. R.Lingual
Vol.(WM) L.Caudate, Vol.(WM) L.Pallidum, Vol.(WM) R.Caudate, Vol.(WM) R.Putamen, Vol.(WM) L.Putamen, Vol.(WM) R.Pallidum
Suf. Area R.FrontalPole, Vol.(CO) R.FrontalPole
Vol.(CO) R.Precentral, Vol.(CO) L.Precentral, Vol.(CO) R.Postcentral, Vol.(CO) L.Postcentral
MMSE, CT Avg. R.Entorhinal, CT Avg. L.Entorhinal, Vol.(WM) L.Hippocampus, Vol.(WM) R.Amygdala, Vol.(CO) L.Entorhinal, Vol.(WM) R.Hippocampus

Table 7: Partial high level concepts (feature groups) obtained from ADAS-Cog M24 predictive modeling via FeaFiner.

Vol.(WM) L.CerebellumWM
CT Std. L.Pericalcarine, CT Avg. L.Pericalcarine, CT Avg. L.Cuneus, CT Avg. R.Pericalcarine, CT Std. R.Pericalcarine
CT Std. R.Sup.Frontal, CT Std. L.Sup.Frontal, CT Std. L.Precuneus, CT Std. L.Ros.Mid.Frontal, CT Std. R.Ros.Mid.Frontal, CT Std. R.Cau.Mid.Frontal, CT Std. R.Precuneus
Suf. Area R.Precentral, Suf. Area L.Ros.Mid.Frontal, Suf. Area R.Tra.Temporal, Suf. Area R.Supramarginal, Vol.(CO) Icv, Suf. Area R.Inf.Temporal, Suf. Area R.Sup.Temporal, Suf. Area R.Pos.Cingulate, Suf. Area R.Mid.Temporal, Suf. Area L.Ros.Ant.Cingulate, Suf. Area L.Tra.Temporal, Suf. Area L.Med.Orbitrontal, Suf. Area R.Postcentral, Suf. Area R.Hemisphere, Suf. Area R.Lat.Orbitrontal, Suf. Area L.Mid.Temporal, Suf. Area L.Postcentral, Suf. Area L.Lat.Orbitrontal, Suf. Area L.Hemisphere, Suf. Area L.Sup.Frontal, Suf. Area R.Sup.Frontal, Suf. Area L.Precentral, Suf. Area L.Insula, Suf. Area L.Inf.Temporal, Suf. Area L.Sup.Temporal, Suf. Area R.Insula
Vol.(WM) R.CerebralWM, Vol.(WM) Brainstem, Vol.(WM) R.VentralDC, Vol.(WM) L.CerebralWM, Vol.(WM) L.Thalamus, Vol.(WM) CorpusCallosumCentral, Vol.(WM) CorpusCallosumMidAnt., Vol.(WM) R.Thalamus, Vol.(WM) L.VentralDC
CT Std. L.Entorhinal, CT Std. R.Entorhinal
CT Std. R.ParsTriangularis, CT Std. L.Bankssts, CT Std. L.ParsTriangularis, CT Std. R.Supramarginal, CT Std. L.ParsOpercularis, CT Std. R.Bankssts, CT Std. R.ParsOpercularis
CT Std. L.Paracentral, CT Std. R.Paracentral
CT Std. R.Parahippocampal, CT Avg. L.Parahippocampal, CT Avg. R.Parahippocampal, Vol.(CO) R.Parahippocampal, CT Std. L.Parahippocampal, Vol.(CO) L.Parahippocampal
Vol.(CO) R.Cuneus, Suf. Area R.Cuneus, Vol.(CO) L.Cuneus, Suf. Area L.Cuneus
CT Std. L.Sup.Parietal, CT Std. R.Postcentral, CT Std. R.Sup.Parietal, CT Std. L.Postcentral
Suf. Area L.TemporalPole, Suf. Area R.TemporalPole, Suf. Area L.Entorhinal, Suf. Area R.Entorhinal
Vol.(WM) R.Pallidum, Vol.(WM) R.Putamen, Vol.(WM) R.Caudate, Vol.(WM) L.Putamen, Vol.(WM) L.Pallidum, Vol.(WM) L.Caudate
Vol.(CO) L.Fusiform, Vol.(CO) R.Inf.Parietal, Vol.(CO) R.Fusiform, Vol.(CO) L.Inf.Temporal, Vol.(CO) R.Ros.Mid.Frontal, Vol.(CO) L.Mid.Temporal, Vol.(CO) L.Inf.Parietal, Vol.(WM) L.CerebralCortex, Vol.(CO) R.IsthmusCingulate, Vol.(WM) R.CerebralCortex, Vol.(CO) L.IsthmusCingulate, Vol.(CO) L.Precuneus, Vol.(CO) R.Precuneus, Vol.(CO) R.Inf.Temporal, Vol.(CO) R.Sup.Parietal, Vol.(CO) R.Mid.Temporal
Suf. Area L.ParsOpercularis, Vol.(CO) L.ParsOpercularis
Vol.(CO) L.Pericalcarine, Suf. Area R.Pericalcarine, Vol.(CO) R.Pericalcarine, Suf. Area L.Pericalcarine
CT Std. L.Inf.Parietal, CT Std. L.Cau.Mid.Frontal, CT Std. R.Precentral, CT Std. L.Precentral, CT Std. R.Inf.Parietal, CT Std. L.Supramarginal
Suf. Area R.Lingual, Suf. Area L.Lingual, Vol.(CO) L.Lingual, Vol.(CO) R.Lingual
CT Std. R.Tra.Temporal, CT Std. L.Tra.Temporal, CT Std. L.Lat.Occipital, CT Std. L.Cuneus, CT Std. R.Lat.Occipital
Vol.(CO) R.Lat.Orbitrontal, Vol.(CO) L.Med.Orbitrontal, Vol.(CO) L.Lat.Orbitrontal, Vol.(CO) R.ParsOrbitalis, Vol.(CO) R.Med.Orbitrontal, Vol.(CO) L.ParsOrbitalis
Vol.(CO) R.Cau.Mid.Frontal, Suf. Area L.Cau.Mid.Frontal, Suf. Area R.Cau.Mid.Frontal, Vol.(CO) L.Cau.Mid.Frontal
Suf. Area R.FrontalPole, Vol.(CO) R.FrontalPole
CT Std. L.Mid.Temporal, CT Std. L.Fusiform, CT Std. R.Mid.Temporal, CT Std. L.Inf.Temporal, CT Std. L.Sup.Temporal, CT Std. R.Fusiform, CT Std. R.Inf.Temporal, CT Std. R.Sup.Temporal
CT Std. R.Cau.Ant.Cingulate, CT Std. R.Insula, CT Std. R.Pos.Cingulate, CT Std. L.Pos.Cingulate, CT Std. L.Cau.Ant.Cingulate, CT Std. L.Ros.Ant.Cingulate, CT Std. L.Insula
MMSE, CT Avg. L.Entorhinal, Vol.(WM) L.Hippocampus, Suf. Area L.Supramarginal, CT Avg. R.Inf.Parietal, Suf. Area R.Ros.Mid.Frontal, Vol.(CO) L.Entorhinal, Vol.(WM) L.Amygdala, CT Avg. R.TemporalPole

Table 8: Partial high level concepts (feature groups) obtained from ADAS-Cog M36 predictive modeling via FeaFiner.

CT Avg. L.Cuneus, Vol.(CO) L.Cuneus, CT Std. L.Cuneus
MMSE, Vol.(CO) R.Ros.Mid.Frontal, Suf. Area R.Fusiform, Vol.(CO) L.Ros.Mid.Frontal
CT Avg. R.Pericalcarine, CT Std. R.Pericalcarine, CT Std. R.Cuneus, CT Avg. R.Cuneus
Vol.(CO) L.Precentral, Vol.(CO) R.Postcentral, Vol.(CO) R.Precentral, Vol.(CO) L.Postcentral
CT Std. R.Sup.Parietal, CT Std. R.Postcentral, CT Std. L.Sup.Parietal, CT Std. L.Postcentral
Vol.(CO) L.Ros.Ant.Cingulate, Vol.(CO) L.ParsOpercularis, Vol.(CO) L.Lat.Orbitrontal, Vol.(CO) L.Med.Orbitrontal, Vol.(CO) R.Lat.Orbitrontal, Vol.(CO) L.ParsOrbitalis, Vol.(CO) R.ParsOrbitalis, Vol.(CO) R.Pos.Cingulate, CT Avg. R.Cau.Ant.Cingulate, Vol.(CO) L.Pos.Cingulate, CT Std. R.Lat.Occipital
Vol.(CO) R.Inf.Parietal, Suf. Area L.Inf.Parietal, Suf. Area R.Inf.Parietal, Vol.(CO) L.Inf.Parietal
CT Avg. L.Ros.Ant.Cingulate, CT Std. L.Precuneus, CT Avg. L.Pos.Cingulate, CT Avg. R.Pos.Cingulate, CT Avg. L.IsthmusCingulate, CT Std. R.Precuneus, CT Avg. R.IsthmusCingulate
Vol.(CO) L.TemporalPole, Vol.(CO) R.TemporalPole, CT Avg. R.TemporalPole, Vol.(CO) L.Entorhinal, CT Avg. R.Entorhinal, CT Avg. L.Entorhinal, CT Avg. L.TemporalPole
Vol.(WM) R.ChoroidPlexus, Vol.(WM) L.ChoroidPlexus
Suf. Area L.Sup.Parietal, Vol.(CO) R.Lat.Occipital, Suf. Area L.Med.Orbitrontal, Suf. Area L.Fusiform, Suf. Area R.Postcentral, Suf. Area L.Precuneus, Suf. Area R.Sup.Parietal, Suf. Area L.Postcentral, Suf. Area R.Precuneus, Suf. Area L.Cuneus, Suf. Area R.Lat.Occipital, Suf. Area L.IsthmusCingulate, Suf. Area R.IsthmusCingulate, Vol.(CO) L.Lat.Occipital, Suf. Area L.Lat.Occipital
Suf. Area R.Tra.Temporal, Vol.(WM) R.CerebralWM, Vol.(WM) L.CerebralWM, Suf. Area R.Cau.Mid.Frontal, Suf. Area L.Precentral, Vol.(CO) Icv, Suf. Area R.Paracentral, Suf. Area L.Pos.Cingulate, Suf. Area L.Sup.Temporal, Suf. Area R.Hemisphere, Suf. Area L.Insula, Suf. Area R.Lat.Orbitrontal, Suf. Area L.ParsOpercularis, Suf. Area R.ParsOrbitalis, Suf. Area R.Pos.Cingulate, Suf. Area R.Sup.Temporal, Suf. Area L.Ros.Ant.Cingulate, Suf. Area L.ParsOrbitalis, Suf. Area L.Tra.Temporal, Suf. Area R.Precentral, Suf. Area L.Cau.Mid.Frontal, Suf. Area L.Hemisphere, Suf. Area R.Supramarginal, Suf. Area L.Ros.Mid.Frontal, Suf. Area L.Lat.Orbitrontal, Suf. Area R.Sup.Frontal, Suf. Area L.Sup.Frontal, Suf. Area L.Parahippocampal, Suf. Area R.Ros.Mid.Frontal, Suf. Area R.Insula
Vol.(CO) L.Sup.Parietal, Vol.(CO) L.Paracentral, Suf. Area L.Paracentral, Vol.(CO) R.Paracentral, Vol.(CO) R.Sup.Parietal
CT Std. R.Tra.Temporal, CT Avg. R.Tra.Temporal, Vol.(CO) R.Tra.Temporal
CT Std. R.FrontalPole, CT Std. L.FrontalPole
CT Std. L.Insula, CT Std. R.Insula, CT Std. R.Sup.Temporal, CT Std. L.Sup.Temporal
CT Std. L.Lat.Occipital, CT Std. R.Lingual, CT Std. L.Lingual
CT Std. R.Cau.Ant.Cingulate, CT Std. R.Ros.Ant.Cingulate, CT Std. L.Cau.Ant.Cingulate, CT Std. R.Pos.Cingulate, CT Std. L.Pos.Cingulate, CT Std. L.Ros.Ant.Cingulate
CT Std. L.ParsTriangularis, CT Std. R.Lat.Orbitrontal, CT Std. R.Ros.Mid.Frontal, CT Std. R.Med.Orbitrontal, CT Std. R.ParsTriangularis, CT Std. L.IsthmusCingulate, CT Std. R.ParsOrbitalis, CT Std. L.Ros.Mid.Frontal, CT Std. R.IsthmusCingulate, CT Std. L.Lat.Orbitrontal, CT Std. L.Med.Orbitrontal, CT Std. L.ParsOrbitalis
CT Std. L.ParsOpercularis, CT Std. L.Bankssts, CT Std. R.Bankssts, CT Std. R.ParsOpercularis
CT Std. L.Parahippocampal, Vol.(CO) L.Parahippocampal, CT Avg. R.Parahippocampal, CT Avg. L.Parahippocampal, CT Std. R.Parahippocampal
CT Std. R.Fusiform, CT Std. R.Mid.Temporal, CT Std. L.Mid.Temporal, CT Std. R.Inf.Temporal, CT Std. L.Fusiform, CT Std. L.Inf.Temporal
Suf. Area R.Entorhinal, Suf. Area L.Entorhinal
Vol.(CO) R.Med.Orbitrontal, Suf. Area R.Med.Orbitrontal
CT Avg. L.ParsTriangularis, CT Avg. R.Sup.Frontal, CT Avg. L.ParsOrbitalis, CT Avg. R.ParsTriangularis, CT Avg. L.ParsOpercularis, CT Avg. R.ParsOrbitalis, CT Avg. R.Ros.Mid.Frontal, CT Avg. L.Sup.Frontal, CT Avg. L.Ros.Mid.Frontal, CT Avg. R.ParsOpercularis