
FeAture Explorer (FAE) 使用指南

(V0.3.6)

2020年9月03日

一、 简述.

FeAture Explorer (FAE) 意在提高机器学习建模过程中的超参数选择效率。针对不同的实际问题, 数据分布和超参数选择也不尽相同, 科研上往往需要科研人员花费大量时间进行超参数探索。FAE 意在提供一个超参数探索平台, 提供多种数据处理方法, 模型建立方法, 使得科研人员能够方便得横向比较不同方法在处理特定数据时的效果。

FAE 的建立受放射组学 (Radimocs) 的启发。在 Radiomics 问题中, 样本数量在 100 这个数量级, 而特征数往往会在 1000 的量级中, 核心问题是从众多特征中找到与临床诊断相关的特征并建立有效模型。对基于使用 pyradiomics 的引用。

FAE 包括三大模块: 特征提取 (Feature Extraction), 数据准备 (Feature, Preprocessing), 模型探索 (Model Exploration) 和可视化 (Visualization)。四大模块相互独立。

FAE 视频演示请参见: <https://www.bilibili.com/video/BV1yt4y1S79S/>

二、 声明

FAE 由华东师范大学上海市磁共振重点实验室和上海夏雨岛智能医疗科技有限公司联合开发, 该项目仅用于科研, 项目代码开源在 <https://github.com/salan668/FAE>。欢迎一切意见和建议, 如有任何合作意向和需求, 请联系:

宋阳, songyangmri@gmail.com

张敬, zhangjingmri@gmail.com

杨光, gyang@phy.ecnu.edu.cn

地点: 上海市普陀区中山北路 3663 号, 华东师范大学老图书馆一楼, 200062。

电话: +86-021-62233873

三、 模块介绍

特征提取（Feature Extraction）

(请参见视频)

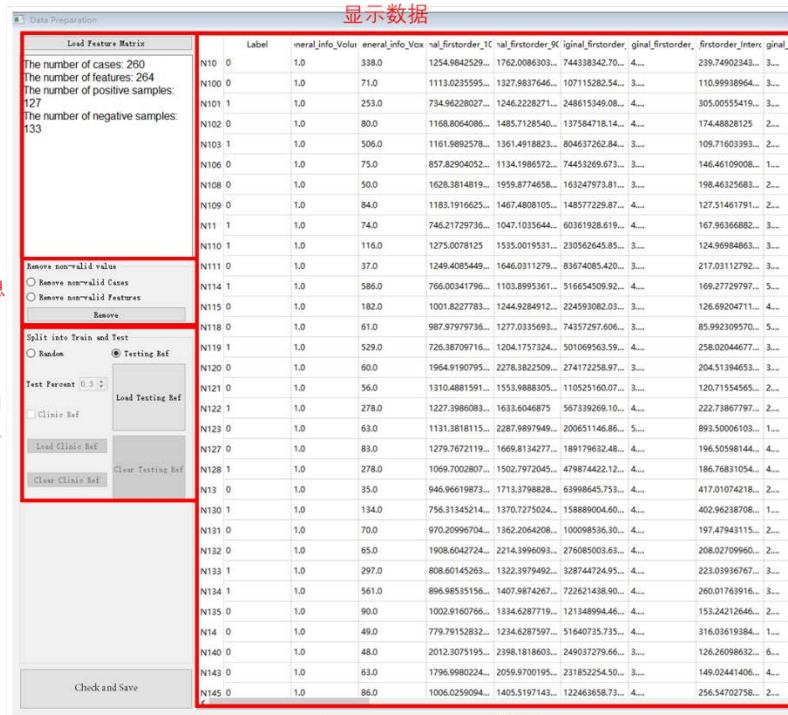
特征预处理（Data Preparation）

该模块用以处理已提取好的特征文件(csv 格式), 要求行方向为不同的 case。列方向为不同的 feature, 第一行为特征特征名称。其中需要一列标记为“label”, 如未提供, 则无法进行处理。同时目前 FAE 只能处理二分类问题, 即 label 标记为 0 和 1 (推荐字符串形式)。用 Excel 打开可以显示为下图:

	A	B	C	D	E	F	G
Row: Cases	1	ID	label	ADC_ge	ADC_ge	ADC_orig	ADC_orig
Column: Features	2	N10	0	1	338	1255	1762.009
	3	N100	0	1	71	1113	1327.984
	4	N101	1	1	253	734.96	1246.223
	5	N102	0	1	80	1168.8	1485.713
Case ID	6	N103	1	1	506	1162	1361.492
Label/Target	7	N106	0	1	75	857.83	1134.199
	8	N108	0	1	50	1628.4	1959.877
Feature Name	9	N109	0	1	84	1183.2	1467.481
	10	N11	1	1	74	746.22	1047.104
Feature Matrix	11	N110	1	1	116	1275	1535.002
	12	N111	0	1	37	1249.4	1646.031
	13	N114	1	1	586	766	1103.9

图 1, 特征矩阵准备形式。

该模块对数据进行检查, 确保数据矩阵中不存在字符数据(例如病人姓名), 不存在无效数据(例如某些病人的特征未进行提取)。同时该模块提供数据拆分功能, 可将数据拆分为训练集和测试集。界面如下:



显示数据统计信息

去除无效信息

数据集拆分和临床信息导入

图 2，数据准备界面示意图

界面左上角“Load Feature Matrix”按钮可以从本地导入特征数据，只支持 csv 格式，路径中不能包含中文。数据导入之后，可以对数据进行浏览。右侧框图进行数据检视，特征名称显示在 header 栏，case id 显示在左边 index 栏。“Load Feature Matrix”按钮下方显示数据基本信息，包括样本总数、特征总数、正样本数目（label=1）、负样本数据（label=0）。

由于特征数据中可能存在无效数据，例如文本信息（病人姓名），无效数字（标记为 nan）和空数据，含有这种无效数据的特征文件会使得后续数据处理无法进行。FAE 提供了简单的去除无效数据的操作，允许使用者选择一次性去除含有无效数据的 case 或者去除无效数据的 feature。如有更加复杂的需求，目前需要使用者自行处理。点选需要的处理方法，单击“Remove”按钮，即可完成。

在模型建立时，数据往往分为训练集（training data）和测试集（testing data），其中训练集有时又会被拆分为训练集和验证集（validation data）。Validation 用作确定模型超参数，例如选用特征数目。测试集为独立数据进行最终的模型评估。预处理模块可以进行数据拆分，以一定比例分为训练集和测试集。这里所选的参数为测试集所占总样本的比例。若需要拆分相同的 case 做对比实验，可以通过‘Load Testing Ref’来获取相同的训练集和测试集 case。通过“Clinic Ref”的选项可以导入临床特征。

界面左下角的“Check and Save”按钮可以用来进行数据保存。

模型探索（Model Exploration）

该模块用以对有效数据特征（经过预处理过的 csv 数据）进行建模，我们将数据到模型的处理过程称为一条流水线（pipeline），一条典型的流水线包括：数据归一化、数据降维、特征选择方法、选择的特征数目和分类器方法五个主要部分。该流水线基于 Scikit-learn（v 0.22.2）构建，界面简述如下：

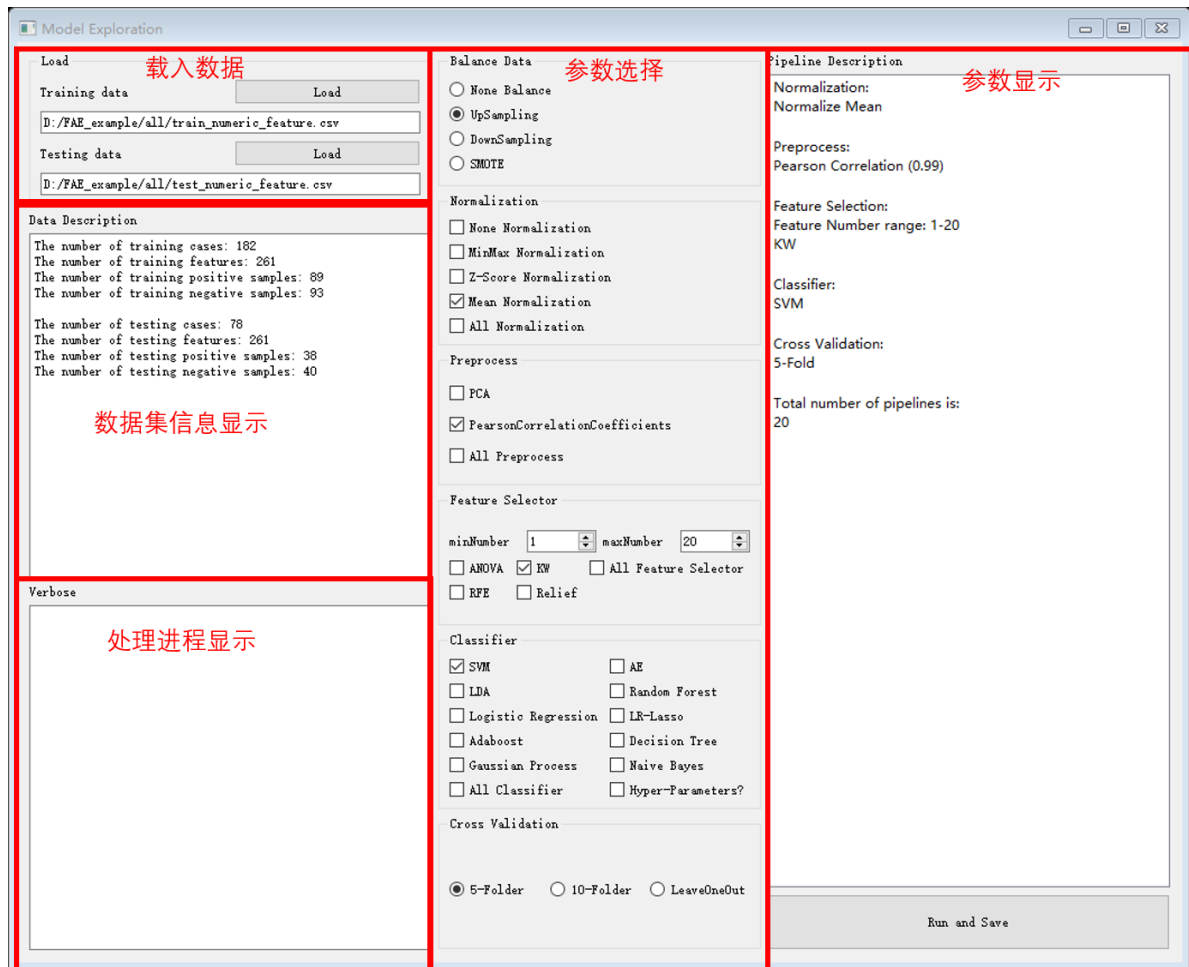


图 3，数据处理界面示意图。

界面左上角是数据导入部分，可以分别导入训练数据和测试数据。其中训练数据必须导入，测试数据为可选操作；如若不导入测试数据，则模型会进行交叉验证，输出训练集见过和验证集结果，所有数据必须为 csv 文件。载入数据后下边的 Data Description 中会显示 train 和 test 数据中的样本信息。

特征处理流水线（Pipeline）

由于每条 Pipeline 中的任何一个模块均有多种方法可供选择，在对特定问题进行分析时，我们需要对不同方法的组合进行探索。每个模块可供选择如下：

1. 数据平衡 (Balance Data)

有些时候，数据集中的样本比例会不均匀，需要通过特定策略进行数据平衡。FAE 提供随机上采样、随机下采样和 SMOTE 方法来实现数据平衡（这里进行的数据平衡操作只针对训练集进行，测试集数据不进行处理，同时原始训练集数据也会进行保存）。

2. 归一化处理

该模块是指对特征矩阵进行归一化处理。由于不同特征计算出的数值差别较大，为探索不同特征与 label 之间的相关性，需将不同特征直接的数值缩放到同一数量级，即对每个特征对应的所有 case 的数值进行处理。目前提供一下三种归一化方法：

1) Normalize to unit

$$\vec{x}_n = \frac{\vec{x}_n}{\sqrt{x_{1n}^2 + x_{2n}^2 + \dots + x_{mn}^2}}$$

其中 \vec{x}_n 表示第 n 个特征，其中 m 表示样本数。

2) Normalize to 0-center

$$\vec{x}_n = \frac{\vec{x}_n - \bar{x}_n}{std(\vec{x}_n)}$$

其中 \bar{x}_n 表示所有样本特征值的均值， $std(x_n)$ 表示特征 n 的方差。

3) Normalize to unit with 0-center

$$\vec{x}_n = \frac{\vec{x}_n - \bar{x}_n}{\sqrt{x_{1n}^2 + x_{2n}^2 + \dots + x_{mn}^2}}$$

该方法结合上述两种方法，做处理。

3. 数据降维处理

由于某些情况下特征数量远大于样本数量，为了后续进行高效的特征选择和模型建立，该模块可以对数据进行预处理，降低特征个数。目前该模块支持主成分分析（Principal component analysis, PCA）降维和利用皮尔森相似度（Pearson Correlation Coefficients, PCC）降维两种方式。

1) 主成分分析 (Principal Component Analysis, PCA)

PCA 可以把可能具有相关性的高维特征映射为线性无关的低维特征，映射后的低维数

据每个特征都线性无关。具体可参考 sklearn 文档：

<http://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>

注意：PCA 会使得特征数据中的每一个特征都是有原始数据中所有特征线性叠加组成，因此 PCA 之后，每个特征的 header 标记为“PCA feature”，因此后续再进行特征解释时，需要注意。

2) 皮尔森相关系数 (Pearson Correlation Coefficients, PCC)

遍历所有特征，两两计算皮尔森相关系数，当系数大于某个阈值时（目前默认为 0.90），随机去除其中一个，该方法可以使得降维后的特征不具有高相似度。

$$\cos \theta_{12} = \frac{\vec{x}_1 \cdot \vec{x}_2}{|\vec{x}_1| |\vec{x}_2|}$$

4. 特征数目选择

选择最小特征与最大特征，FAE 会遍历之间所有特征数，进行模型建立。从数据可解释性角度来讲，模型最大特征数不宜过高（例如 20），当进行遍历的特征数大于特征矩阵的特征数时，会使用所有特征进行建模。

5. 特征选择

该模块的功能为根据 label 数据，从中所有特征之中选取与 label 相关性最好的特征。目前支持的方法有 ANOVA、RFE、Relief。

ANOVA (Analysis of Variance)

多变量方差分析，通过对多变量进行统计分析，计算每个特征对 label 的 F 值权重，从大到小排序后，决定出与 label 最相关的特征。具体文档可参照：

http://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.f_classif.html

KW (Kruskal-Wallis)

Kruskal-Wallis (KW) 测试是一个关于三组或更多数据的非参数性测试。它是用来检测总体函数分布的一致性原假设和其替代假设，关于至少两个样本之间存在差异的假设

<https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kruskal.html>

RFE (Recursive Feature Elimination)

递归特征消除的主要思想是反复的构建模型（如 SVM 或者回归模型）然后选出最好的（或者最差的）的特征（可以根据系数来选），把选出来的特征放到一边，然后在剩余的特征上重复这个过程，直到所有特征都遍历了。这个过程中特征被消除的次序就是特征的排序。

http://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html

Relief

Relief 算法中特征和类别的相关性是基于特征对近距离样本的区分能力。算法从训练集 D 中随机选择一个样本 R ，然后从和 R 同类的样本中寻找最近邻样本 H ，称为 Near Hit，从和 R 不同类的样本中寻找最近邻样本 M ，称为 Near Miss，然后根据以下规则更新每个特征的权重：如果 R 和 Near Hit 在某个特征上的距离小于 R 和 Near Miss 上的距离，则说明该特征对区分同类和不同类的最近邻是有益的，则增加该特征的权重；反之，如果 R 和 Near Hit 在某个特征的距离大于 R 和 Near Miss 上的距离，说明该特征对区分同类和不同类的最近邻起负面作用，则降低该特征的权重。以上过程重复 m 次，最后得到各特征的平均权重。特征的权重越大，表示该特征的分类能力越强，反之，表示该特征分类能力越弱。

6. 分类器

该模块决定最终模型分类方法，使用选择出的特征进行建模。

SVM (Support Vector Machine)

SVM 是一种高效稳定的分类器，其思想是建立一个最优决策超平面，使得该平面两侧距离该平面最近的两类样本之间的距离最大化，从而对分类问题提供良好的泛化能力。对于一个多维的样本集，系统随机产生一个超平面并不断移动，对样本进行分类，直到训练样本中属于不同类别的样本点正好位于该超平面的两侧，满足该条件的超平面可能有很多个，SVM 正式在保证分类精度的同时，寻找到这样一个超平面，使得超平面两侧的空白区域最大化，从而实现线性可分样本的最优分类。

<http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>

AE (Auto-Encoder)

AE 是利用神经网络 (Neural Network) 感知机进行数据分类，经过对特征进行线性组合并配上激活函数，可以从理论上模拟任何复杂的函数，但其过多的参数往往导致模型产生过拟合。

http://scikit-learn.org/dev/modules/generated/sklearn.neural_network.MLPClassifier.html

LDA (Linear Discriminant Analysis)

LDA 是一种分类模型，它通过在 k 维空间选择一个投影超平面，使得不同类别在该超平面上的投影之间的距离尽可能近，同时不同类别的投影之间的距离尽可能远。此外，由于 LDA 会将 k 维数据投影到 $k-1$ 维的超平面，因此也具有降维的作用。不同于 PCA 会选择数据变化最大的方向，LDA 会主要以类别为思考因素，使得投影后的样本尽可能可分。

<http://scikit-learn.org/0.16/modules/generated/sklearn.lda.LDA.html>

Random Forest

随机森林是一种有监督学习算法。就像你所看到的它的名字一样，它创建了一个森林，并使它拥有某种方式随机性。所构建的“森林”是决策树的集成，大部分时候都是用“bagging”方法训练的。bagging 方法，即 bootstrap aggregating，采用的是随机有放回的选择训练数据然后构造分类器，最后组合学习到的模型来增加整体的效果。

<http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

Logistic Regression

回归是一种极易理解的模型，就相当于 $y=f(x)$ ，表明自变量 x 与因变量 y 的关系。最常见问题有如医生治病时的望、闻、问、切，之后判定病人是否生病或生了什么病，其中的望闻问切就是获取自变量 x ，即特征数据，判断是否生病就相当于获取因变量 y ，即预测分类。

http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html

LR-Lasso(Logistic Regression via Lasso)

在逻辑回归的基础上引入 Lasso 正则项以获得更好模型，lasso 回归可以适应的情况是：样本量比较小，但是指标非常多，即小 N 大 P 问题。适用于高维统计，传统的方法无法应对这样的数据。并且 lasso 可以进行特征选择。

http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso.html

Ada-boost

Adaboost 是一种常见的继承学习方法，它属于 boosting 系列算法中的，也就是说每个学习器之间存在强依赖关系。Adaboost 既可以用于分类也可以用于回归。两个核心步骤：

1、权值调整：Adaboost 算法提高那些被前一轮基分类器错误分类样本的权值，而降低那些被正确分类样本的权值。从而使得那些没有得到正确分类的样本，由于权值的加大而受到后一轮基分类器的更大关注。

2、基分类器组合：Adaboost 采用加权多数表决的方法：

- (1) 加大分类误差率较小的弱分类器的权值，使得它在表决中起较大的作用。
- (2) 减小分类误差率较大的弱分类器的权值，使得它在表决中起较小的作用。

<http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.AdaBoostClassifier.html>

Decision Tree

决策树是一种常见机器学习方法中的一种分类器。它通过训练数据构建一种类似于流程图的树结构，其中每个内部结点表示在一个属性上的测试对未知数据进行分类，每个分支代表一个属性输出，每个树叶结点代表类或类分布。

<http://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html>

Gaussian Process

针对机器学习的高斯过程(Gaussian Processes for Machine Learning, 即 GPML) 是一个通用的监督学习方法，主要被设计用来解决回归问题。它也可以扩展为概率分类(probabilistic classification)

http://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessClassifier.html

Native Bayes

贝叶斯分类器的分类原理是通过某对象的先验概率，利用贝叶斯公式计算出其后验概率，即该对象属于某一类的概率，选择具有最大后验概率的类作为该对象所属的类。

http://scikit-learn.org/stable/modules/naive_bayes.html

超参数选择

目前我们提供以上分类模型的超参数选择，其中超参数的配置文件在 FAE 的路径下的 HyperParameters 文件夹下，可以通过编辑其中的 json 文件实现超参数的设置。

勾选 ‘Hyper-Parameters’ 后，分类模型会遍历预先设置好的超参数进行建模，并保存验证集中 AUC 最高的结果。

交叉验证

该模块对训练数据进行拆分，分为训练集和验证集，来寻求已上策略合适的组合。为寻

求稳定的超参数组合，这里选择进行交叉验证，目前支持 5-folder 和 10-folder。其中 5-folder 和 10-folder 属于 K-folder，初始采样分割成 K 个子样本集，每个单独的子样本集被保留作为验证模型的数据，其他 K-1 个子样本集用来训练。交叉验证重复 K 次，每个子样本验证一次，将所有结果记录，与对应 label 进行统计分析，用以评价模型。

在对 Pipeline 每个模块进行选择之后，界面中央区域显示 Pipeline 配置情况。例如：

```
Pipeline Description
Normalization:
Normalize Mean

Preprocess:
Pearson Correlation (0.99)

Feature Selection:
Feature Number range: 1-20
KW

Classifier:
SVM

Cross Validation:
5-Fold

Total number of pipelines is:
20
```

中央界面显示如下，总流水线个数为 $1*1*20*1=240$ 个。

点击左下角的“Run and Save”按钮，选择结果的保存路径（此路径只能为空的文件夹），程序进行运行，右下角 Verbose 界面对程序运行进行监控，当 Pipelines 数目很多时，程序会需要大量运行时间，建议运行之后不对电脑进行操作，可以从右下角的界面中对过程进行监控。当 Verbose 出现“Merging CV Results DONE”时，表示程序运行完毕。

```
Verbose
Model Developing:
Done.

Cross Validation:
Done.

Merging CV Results:
Done.
```

可视化（Visualization）

该模块包含不同模型的横向比较和针对特定模型描述文件生成。

横向比较显示 FAE 对模型的探索结果，包括对 Pipeline 中每一部分不同方法的比较。界面分布如下：

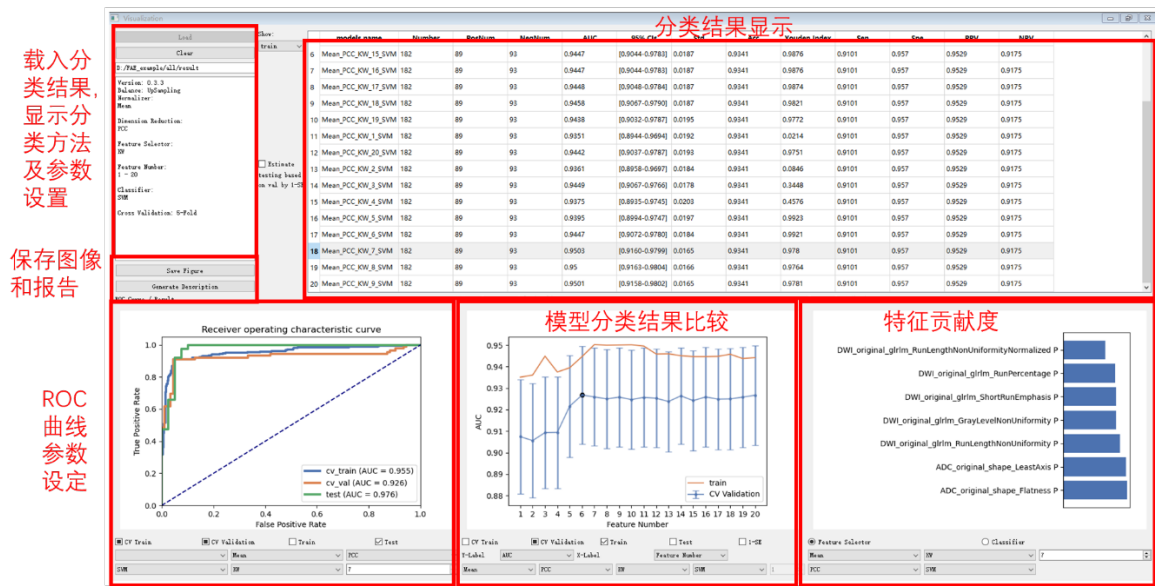


图 4，可视化模块界面

点击 Load 载入存储处理结果数据的文件夹，注意要选择上一次所保存的文件夹而不是该文件夹下的子目录，可以看到所有模型的简易描述。右面是每种超参数组合建模的统计结果，可以选择 Train、Validation 或 testing 来显示不同数据集下的结果。Train 是模型在训练集的表现，Validation 是模型在交叉验证集上的表现，testing 是每个模型在独立测试集中结果。严格来讲，testing 作为独立测试集，只在最终模型上进行评估，但对于部分课题，不需要独立 testing 数据，只根据独立 validation 进行分析时，此时的 testing 数据相当于独立 validation。勾选“Estimate testing based on validation by 1-SE”可以在比较不同模型在验证集 AUC，按照 1-SE 选择出候选模型。所有模型均记录样本总数、正样本数目、副样本数目、AUC 值、95% 置信区间下 AUC 值、准确率、Youden Index、敏感性、特异性、真阳性和假阳性。

ROC Curve/ Result (下方左侧)

绘制 ROC 曲线，该曲线经常用于评价模型表现。根据分类结果选择超参数设置，显示不同流水线模型下的 CV Train、CV validation、Train、Test 的 ROC 曲线。其中，CV Train 代表 K-fold cross validation 中训练集的 K-1 folder 训练数据集的平均结果，CV validation 代表 K-fold cross validation 中训练集的 1-folder 数据集的平均结果，Train 代表按全部训练集的结果，Test 代表测试集的结果。

Hyper Parameters Relationship (下方中央)

可横向比较不同流水线，Y 轴可以选择 AUC 值和准确率，X 轴为不同的 Pipeline 中的任一部分，如不同特征数比较或不同特征选择方法比较。允许固定其他超参数只比较指定参

数：也允许比较特定参数，其他参数选取对应 Y 的最大值（勾选 **Maximum**）。

在 AUC 和特征数的比较上，我们生成 AUC 值的方差，此使可以通过比较验证集的 AUC，来判断模型选择。选择方法可以使用 a)验证集最高 AUC 对应的模型或 b)验证集最高 AUC 再通过 1-Standard Error 方法选择到的更为鲁棒的模型。可以通过 1-SE 的选择图标进行修改。

Feature Contribution (下方右侧)

展示不同特征的贡献度特征贡献度（勾选 **Show**）。由于特征权重可以从特征选择角度获得，也可以从模型建立角度获得，可以通过按钮进行选择。**注意，某个特征可能与 label 相似度最高，被特征选择方法选出，但在与其他特征一起构建模型时，不一定所占最大权重。**

Pipeline Description 下方的“Save”按钮可将所有展示的图片进行保存，保存路径为文件夹，所有图片会保存为压缩过的 JPG 形式和高分辨率的 EPS 形式，其中 JPG 可以用 PPT 中展示，EPS 用于印刷使用。

Generation Description

用于生成对某一模型的详细描述文档，包括方法的介绍和模型权重，ROC 等结果展示

Materials and Methods

We selected 182 cases as the training data set (89/93 = positive/negative). We also selected another 78 cases as the independent testing data set (38/40 = positive/negative).

To Remove the unbalance of the training data set, we up-samples by repeating random cases to make positive/negative samples balance. We applied the normalization on the feature matrix. Each feature vector was subtracted by the mean value of the vector and was divided by the length of it. Since the dimension of feature space was high, we compared the similarity of each feature pair. If the PCC value of the feature pair was larger than 0.99, we removed one of them. After this process, the dimension of the feature space was reduced and each feature was independent to each other. Before build the model, we used Kruskal Wallis to select features. KruskalWallis was a common method to explore the significant features corresponding to the labels. F-value was calculated to evaluate the relationship between features and the label. We sorted features according to the corresponding F-value and selected top N features according to validation performance. We used support vector machine (SVM) as the classifier. SVM was an effective and robust classifier to build the model. The kernel function has the ability to map the features into a higher dimension to search the hyper-plane for separating the cases with different labels. Here we used the linear kernel function because it was easier to explain the coefficients of the features for the final model. To determine the hyper-parameter (e.g. the number of features) of model, we applied cross validation with 5-fold on the training data set. The hyper-parameters were set according to the model performance on the validation data set.

The performance of the model was evaluated using receiver operating characteristic (ROC) curve analysis. The area under the ROC curve (AUC) was calculated for quantification. The accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were also calculated at a cutoff value that maximized the value of the Youden index. We also estimated the 95% confidence interval by bootstrape with 1000 samples. All above processes were implemented with FeAture Explorer Pro (FAEPro, V 0.3.3) on Python (3.7.6).

Result

We found that the model based on 7 features can get the highest AUC on the validation data set. The AUC and the accuracy could achieve 0.926 and 0.934, respectively. In this point, The AUC and the accuracy of the model achieve 0.976 and 0.949 on testing data set. The clinical statistics in the diagnosis and the selected features were shown in Table 1 and Table 2. The ROC curve was shown in Figure 1.

Table 1. Clinical statistics in the diagnosis

Statistics	Value
Accuracy	0.9487
AUC	0.9757
AUC 95% CIs	[0.9375-1.0000]
NPV	1.0000
PPV	0.9048
Sensitivity	1.0000
Specificity	0.9000

Materials and Methods

We selected 203 cases as the training data set. 133 of them were marked as positive and the left 70 were marked as negative. We also selected another 87 cases as the independent testing data set (57/30 = positive/negative).

We applied the normalization on the feature matrix. Each feature vector was subtracted by the mean value of the vector and divided by the length of it. Since the dimension of feature space was high, we compared the similarity of each feature pair. If the cosine value of the feature pair was larger than 0.86, we removed one of them. After this process, the dimension of the feature space was reduced and each feature was independent to each other. Before build the model, we used analysis of variance (ANOVA) to select features. ANOVA was a common method to explore the significant features corresponding to the labels. F-value was calculated to evaluate the relationship between features and the label. We sorted features according to the corresponding F-value and selected specific number of features to build the model. We used logistic regression as the classifier. Logistic regression is a linear classifier that combines all the features. A hyper-plane was searched in the high dimension to separate the samples. To prove the performance of the model, we applied cross validation with leave-one-out on the data set.

The performance of the model was evaluated using receiver operating characteristic (ROC) curve analysis. The area under the ROC curve (AUC) was calculated for quantification. The accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were also calculated at a cutoff value that maximum the value of the Yorden index. We also boosted estimation 1000 times and applied paired t-test to give the 95% confidence interval. All above processes were implemented with FeAture Explorer (FAE, v0.1.1, <https://github.com/salen668/FAE>) on Python (3.5.4, <https://www.python.org/>).

Result

We found that the model based on 2 features can get the highest AUC on the validation data set. The AUC and the accuracy could achieve 0.751 and 0.704, respectively. In this point, The AUC and the accuracy of the model achieve 0.798 and 0.770 on testing data set. The clinical statistics in the diagnosis and the selected features were shown in Table 1 and Table 2. The ROC curve was shown in Figure 1.

Statistics	Value
Accuracy	0.7701
AUC	0.7982
AUC 95% CIs	[0.6967-0.8817]
NPV	0.6190
PPV	0.9111
Sensitivity	0.7193
Specificity	0.8667

Features	Coef in model
CT_original_firstorder_Uniformity	-1.852
CT_original_gldm_DependenceEntropy	2.145

图 5, FAE 生成的报告范文。

打开 report.pdf 后会有关于该模型所用数据、分类方法、统计方法、分类结果、ROC 曲线等细节的英文介绍,用以跟其他科研工作者进行交流。ROC 曲线单独保存为.jpg 格式,以备使用。