

玻璃文物特征分析及风化预测模型

——基于机器学习与统计分析

马毅, 魏小柯, 王逸鸣, 陈麟*

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2022年10月13日; 录用日期: 2022年11月4日; 发布日期: 2022年11月14日

摘要

本文基于2022年高教杯C题, 首先将数据集行one-hot编码, 对化学成分数据进行填补缺失值, 并使用Person相关系数对数据相关性进行分析。后基于奇异值分解的PCA算法对数据降维。通过随机森林、支持向量机、Xgboost、Logistic回归对降维后的数据进行分类并求得决策边界, 基于软分类器预测出了各个文物的风化程度, 得到在文物信息数据的验证集分类准确度达到了86.7%, 化学成分数据达到了94.1%。之后利用召回率、f1值、精度、ROC曲线等评价指标对模型进行了评价与选择。最终得到了各特征的相关性以及各文物的风化概率, 同时得出了Xgboost在该数据集中预测的优越性。

关键词

One-Hot编码, 奇异值分解, 机器学习, 风化概率

Characteristic Analysis and Weathering Prediction Model of Glass Cultural Relics

—Based on Machine Learning and Statistical Analysis

Yi Ma, Xiaoke Wei, Yiming Wang, Lin Chen*

School of Optoelectronic Information and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Oct. 13th, 2022; accepted: Nov. 4th, 2022; published: Nov. 14th, 2022

Abstract

Based on the 2022 Higher Education Cup A question, this paper first codes the data set line one-hot,

*通讯作者。

fills in the missing values for the chemical composition data, and analyzes the data correlation using the Person correlation coefficient. After the PCA algorithm based on singular value decomposition reduces the dimensionality of the data. Through random forest, support vector machine, Xgboost, logistic regression to classify the dimensionality reduction data and find the decision boundary, based on the soft classifier, the degree of weathering of each cultural relics is predicted, and the classification accuracy of the verification set of the cultural relics information data reaches 86.7%, and the chemical composition data reaches 94.1%. After that, the model was evaluated and selected using evaluation indicators such as recall rate, f1 value, accuracy, and ROC curve. Finally, the correlation of each feature and the weathering probability of each artifact are obtained, and the superiority of Xgboost's prediction in this data set is obtained.

Keywords

One-Hot Coding, Singular Value Decomposition, Machine Learning, Weathering Probability

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

古代玻璃制品是研究早期丝绸之路各国贸易往来的重要文物，具有非常宝贵的研究意义[1]。但由于各国制造玻璃制品的取材不同致其化学成分各不相同，且受掩埋环境影响大，容易风化，内部与环境元素进行大量交换，玻璃制品的成分比例发生变化，故影响对其真实类别的判断[2]，因此，如何对风化后的玻璃制品进行成分分析与鉴别，从而还原其研究价值是较有意义的研究课题。姜中宏、张勤远[3]利用同位素标记法分析铅钡玻璃的化学成分；赵凤燕、陈斌等人[4]利用 PXRF 分析研究文物成分；李鹏艳、谢承利[5]等人利用激光诱导击穿光谱法分析玻璃成分。本文提出使用 Person 相关系数对数据相关性进行分析，后基于奇异值分解的主成分分析算法对数据降维。通过随机森林、支持向量机、Xgboost、Logistic 回归对降维后的数据进行分类并求得决策边界，通过一系列指标得出最优模型，并基于软分类器预测出了各个文物的风化概率。

2. 数据处理

2.1. One-Hot 编码

One-Hot 编码，又称为一位有效编码，主要是采用 N 位状态寄存器来对 N 个状态进行编码，每个状态都由他独立的寄存器位，并且在任意时候只有一位有效，其数学的本质在于将文本或字符串映射到 n 维二进制空间上 $words \rightarrow R^n$ [6]。

One-Hot 编码是分类变量作为二进制向量的表示。这首先要将分类值映射到整数值。然后，每个整数值被表示为二进制向量，除了整数的索引之外，它都是零值，它被标记为 1。则：对于纹饰 A、B、C 有： $y \in \{(1,0,0), (0,1,0), (0,0,1)\}$ 其中 A、B、C 分别对应集合当中的三行向量。对于玻璃类型高钾类、铅钡类，可以把特征映射到以下集合： $y \in \{(1,0), (0,1)\}$ 其中铅钡玻璃对应，高钾玻璃对应。同理，对于玻璃颜色，一共有浅绿、浅蓝、深绿、深蓝、紫、绿、蓝绿、黑共 8 种颜色，可以将其映射到对应编码为：

其中 X_m 为 m 阶方阵, δ_n 为对角线为 $\delta_1, \delta_2, \dots, \delta_n$ 的对角矩阵, $(X_n)^T$ 为 n 阶方阵转置, 上述矩阵可化简为:

$$A = U\Sigma V^T \tag{3}$$

其中 U, Σ 是由 $A^T A$ 的特征值和特征向量得到, V 是由 AA^T 的特征向量得到。

2.3.2. 主成分分析降维模型

针对奇异值分解算法当中的 V^T 是一个 n 矩阵, 其中前 i 个列, 对应 δ 矩阵当中前 i 个奇异值, 也分别可以作为该数据集的主成分, 可以认为主成分分析的本质就是奇异值分解。一、二数据集的奇异值分解如图 1 所示。

利用 sklearn 中的 PCA 求解器进行计算, 发现在数据集一当中得到了六个主成分所包含整个数据集的信息量分别为 36.7%、26.2%、13.2%、6.51%、5.00%、3.71%, 共计能够包含 91.5% 的信息; 数据集二当中能够提取到三个主成分的信息量分别为 84.2%、9.38%、3.73%, 共计 97.3% 的信息。

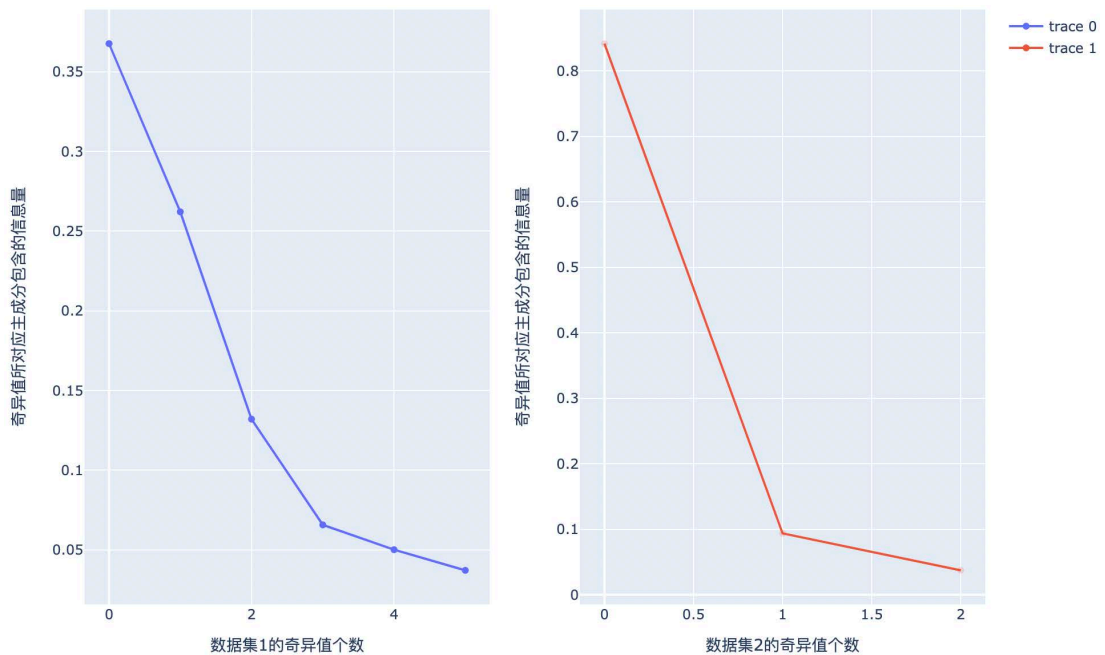


Figure 1. The amount of retention of the dimensionality reduction difference between the two datasets obtained by principal component analysis

图 1. 主成分分析所得到的两个数据集的降维后方差的保留量

显然, 对于数据集二的降维的效果更好, 原因是问题一当中数据分布较为稀疏, 且矩阵值为离散变量, 数据二当中的数值较为连续, 且较为密集, 因此可以更好地利用奇异值分解提取主成分。图 2 是数据集二的数据降维可视化, 可以看出数据二中 92.6% 特征降维到 PC1 和 PC2 上。

3. 基于不同机器学习算法的风化概率预测模型

机器学习因其优秀的数据分析能力以及模型的泛化能力在回归预测、分类等问题上表现出优越性。由于本文的研究是基于离散数据得出文物是否风化的二分类问题, 为了得到连续化数据, 本文将二分类问题转化为求解风化概率的连续性问题, 通过将是否风化映射到 $[0, 1]$ 并设置 0.5 作为分界线, 利用回归模型预测出一系列风化概率。

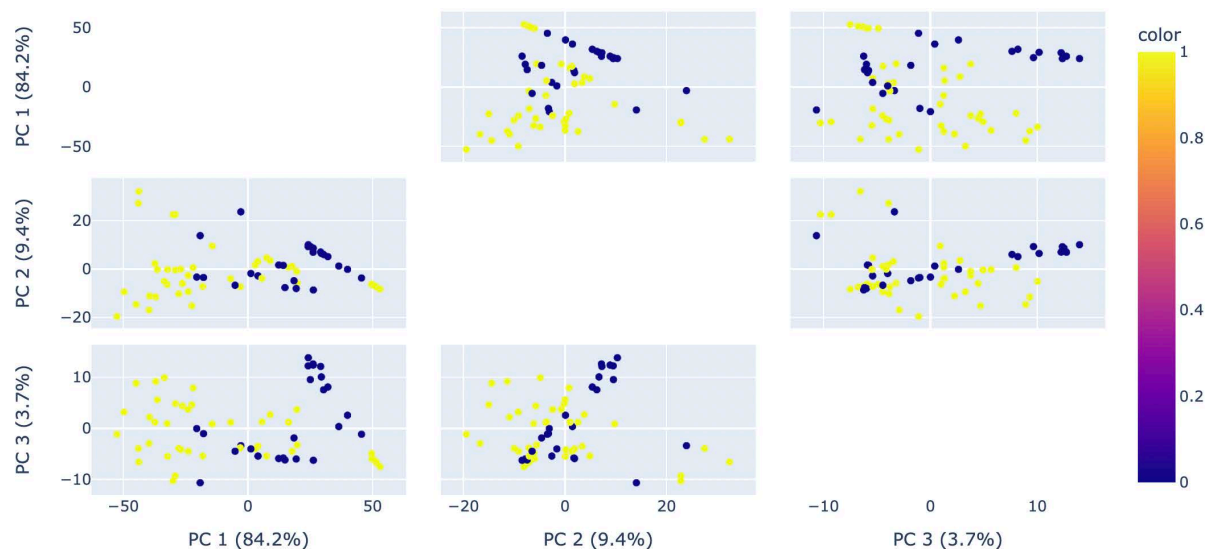


Figure 2. Visualization of PCA reduction results in dataset two

图 2. 数据集二降维可视化

针对文物风化的概率，本文分别使用 Logistic 回归、极致梯度提升(Xgboost)、随机森林、软间隔支持向量机模型对降维后的数据分类并求出决策边界，基于软分类器求出各文物风化的概率。然后利用召回率、F1 值、精度、ROC 曲线等评价指标选出最好的预测模型。机器学习使用经降维后的数据集，两个数据集分别提取到了 6、3 个主成分。对于数据集的划分，本文选择 75% 的数据作为训练集，25% 的数据作为测试集。

3.1. Logistic 回归模型

Logistic 回归是统计学中重要的经典分类方法，属于对数线性模型，所以也被称为对数几率回归。将对数概率函数取反后得到 sigmoid 曲线，首先利用极大似然函数拟合出参数的最佳取值，然后将降维后的特征输入分类模型得到各文物被风化的概率，由求出的参数拟合决策边界从而将连续的风化概率离散为是否风化两种情况。

设 P 为某件文物发生风化的概率，则 $1 - P$ 为未被风化的概率。显然，概率是非负的，为了将概率记为输入特征值的线性表达式，将概率取对数，得到：

$$\begin{aligned} \ln(y = 1 | x) &= w_0 x_0 + w_1 x_1 + \cdots + w_n x_n \\ &= \sum_{i=0}^n w_i x_i \\ &= W^T x \end{aligned} \quad (4)$$

上述公式表示当输入特征 x 后，文物被分为风化的概率，其中为决策边界系数矩阵。将对数函数取反，令 $Z = X$ ，可以得到：

$$f(z) = \frac{1}{1 + e^{-z}} \quad (5)$$

该函数就是 sigmoid 函数，为了实现 logistic 回归分类器，我们可以在每个特征上都乘以一个回归系数，然后把所有的结果值相加，将这个总和带入 sigmoid 函数中。进而得到一个范围在 0~1 之间的数值。最后设定一个阈值，在大于阈值时判定为 1，否则判定为 0。

关于参数 W 求解, 本文以极大似然函数为代价函数, 分别让代价函数对 W 中 j 个权重求偏导, 沿着梯度下降的方法, 找到使极大似然函数取最小值时的参数 W [9]。

极大似然函数 $L(w)$:

$$\begin{aligned} L(w) &= P(y | x; w) \\ &= \prod_{i=1}^n P(y^{(i)} | x^{(i)}; w) \\ &= \left(\phi(z^{(i)}) \right)^{y^{(i)}} \left(1 - \phi(z^{(i)}) \right)^{1-y^{(i)}} \end{aligned} \quad (6)$$

将极大似然函数取对数, 将连乘转化为累加求和的形式, 并对 j 个权重求偏导:

$$\frac{\partial}{\partial w_j} J(w) = \left(-y \frac{1}{\phi(z)} + (1-y) \frac{1}{1-\phi(z)} \right) \frac{\partial}{\partial w_j} \phi(z) \quad (7)$$

然后对 sigmoid 函数求偏导:

$$\begin{aligned} \frac{\partial}{\partial w_j} \phi(z) &= \frac{\partial}{\partial z} \frac{1}{1+e^{-z}} \\ &= \frac{1}{(1+e^{-z})^2} e^{-z} \\ &= \frac{1}{1+e^{-z}} \left(1 - \frac{1}{1+e^{-z}} \right) \\ &= \phi(z)(1-\phi(z)) \end{aligned} \quad (8)$$

将(8)式代入(7)式即得到了函数对 j 个权重的偏导, 并按照式(9)的规则不断更新权重, 最终求出参数 W 。

$$w := w - \eta \nabla J(w) \quad (9)$$

在 logistic 回归中, 决策边界方程定义为 $W^T x = 0$, 上文得到假设函数:

$$P(y = 1 | x; w) = g(w^T x) = \frac{1}{1+e^{-w^T x}} \quad (10)$$

通过计算得到数据集一、二的决策边界函数分别为:

$$h_{\theta_1}(x) = g(-0.44x_1 + 0.05x_2 - 0.48x_3 - 1.87x_4 - 1.27x_5 - 0.33x_6) \quad (11)$$

$$h_{\theta_2}(x) = g(-0.0319x_1 - 0.0447x_2 - 0.0743x_3) \quad (12)$$

决策边界也可以是一个软边界分类器, 即可以分析风化程度的概率, 因此在 logistic 回归当中, 把数据代入决策函数, 可以得到该行数据是否风化的概率, 由于数据集是离散的, 求概率的过程相当于利用提取到的特征对离散数据连续化处理, 因此可以得到每个文物的风化概率, 如果风化概率小于 0.5 则可以说明该文物没有风化。图 3 是由 logistic 回归得到的部分文物的风化概率。

3.2. 软间隔支持向量机模型(Support Vector Machines, SVM)

支持向量机是一类通过求解最大边距超平面进而解决一个二元分类问题, 而本数据集的非线性性质较高, 无法做到线性可分, 因此需要进行一个软间隔的支持向量机分类模型。

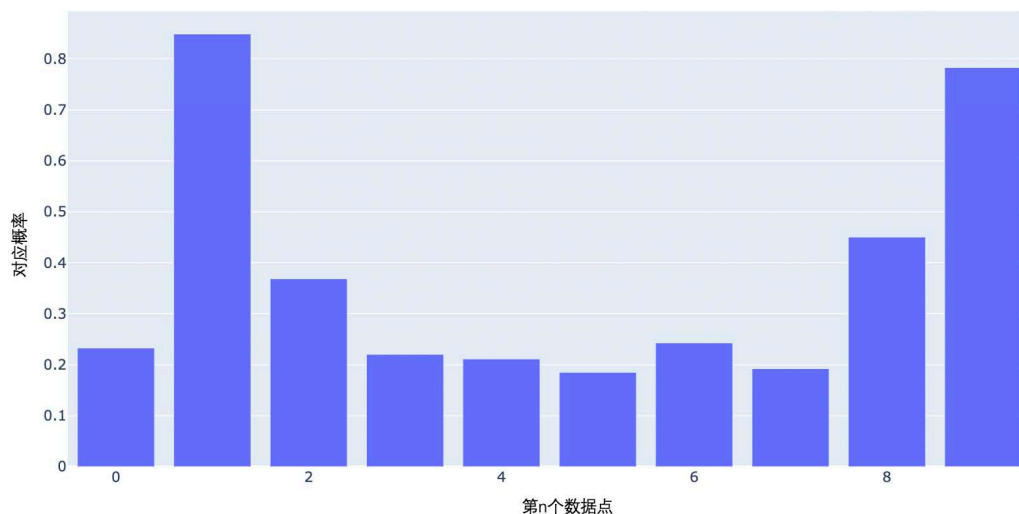


Figure 3. Based on Logistic regression of the weathering probability of some cultural relics
图 3. 基于 Logistic 回归部分文物风化概率

支持向量机本身是一个约束凸二次优化问题，其目标函数为 $\frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^N \varepsilon_i$ 。

其中 $\|\omega\|$ 是支持向量机模型中计算的超平面的 L2 范数； $C > 0$ 称之为惩罚参数，当 C 较大的时候，惩罚更大，避免过拟合效果更好，但过大会导致分类精度较低；而 C 较小的时候，容易过拟合，但分类精度大，因此需要在模型验证的时候选择合理的参数 C 。而 ε_i 是一个松弛变量。针对线性不可分的支持向量机模型可以写作一个凸二次规划问题[10]：

$$\begin{aligned} \max \quad & \sum_{i=1}^m \alpha_i - \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j (x_i - x_j) \\ \text{Subject to: } & \sum_{i=1}^m \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, i = 1, \dots, m \end{aligned} \quad (13)$$

3.3. 随机森林模型(Random Forest, RF)

随机森林是一个 bagging 法的集成学习方法，每次对数据集进行随机特征采样，再创建一个简单的决策树模型，对各个决策树的模型的预测结果进行投票，得到一个最终的结果。其思想如下所示：

Input: 分类数据集 X, Y

Output: 投票生成的类标

1: **for** $i=1$ to B **do**

2: 对数据集进行随机采样，在数据与类标 X, Y 当中采样 n 个，称为 X_b, Y_b

3: 训练弱模型 f_b

4: **end for**

5: 对于每个弱模型，其预测的算数平均数 $\hat{h} = \frac{1}{B} \sum_{b=1}^B f_b(x')$

3.4. 极致梯度提升模型(Xgboost)

Xgboost 是一种 boosting 算法，其目的在于利用迭代优化，对上一次迭代的预测结果进行加权，使得误差越来越小。而 Xgboost 算法是一种由 k 个基模型进行叠加的一个带有正则化项 ω 的目标函数，利用泰勒公式将目标函数展开后得到[11]：

$$\begin{aligned}
 Obj &= \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{i=1}^k \Omega(f_i) \\
 &= \sum_{i=1}^n l(y_i, \hat{y}_i^t) + \sum_{i=1}^t \omega(f_i) \\
 &= \sum_{i=1}^n l(y_i, \hat{y}_i^{t-1} + g_i f_t(x_i)) + \sum_{i=1}^t \Omega(f_i) \\
 &= \sum_{i=1}^n \left(l(y_i, \hat{y}_i^t) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right) + \sum_{i=1}^t \Omega(f_i)
 \end{aligned} \tag{14}$$

其中 g_i, f_i 分别是前一项损失函数的一阶导和二阶导。在决策树模型的 Xgboost 当中，由于决策树的复杂度可以通过叶子数 T 确定，因此目标函数的正则化项可以写为：

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \tag{15}$$

经简化，目标函数可写为：

$$Obj = \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \tag{16}$$

3.5. 模型评价及选择

本文选取了 logistic 回归、随机森林、支持向量机、Xgboost 模型分别对数据进行拟合分类，通过召回率 Recall、f1 值，support，精确度，AUC，ROC 曲线进行模型的评价，几项指标的定义如下[12] (其中 TP、FN、TN、FP 代表 True Positive、False Negative、True Negative、False Positive，前两者指的是预测正确的情况，后两者指的是预测反的情况)：(其中 AUC 指 ROC 曲线包围的面积)

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{17}$$

$$precision = recall = \frac{TP}{TP + FP} \tag{18}$$

$$F1 = \frac{2 * precision * recall}{precision + recall} \tag{19}$$

$$AUC = S(Roc) \tag{20}$$

Table 2. Dataset 1 Learning performance of each model

表 2. 数据集一各模型学习表现

模型	召回率	F1 值	Support	Accuracy	AUC	Precision
LR	1	0.90	9	0.867	0.8407	0.82
RF	1	0.88	9	0.867	0.8995	1.00
SVM	1	0.78	7	0.733	0.7708	0.73
Xgboost	1	0.89	7	0.867	0.9056	0.87

上表 2 展示了各模型对数据集一的预测情况，可以看出：线性模型和树模型对数据一的拟合效果都比较好，最高的拟合精度在验证集上能够达到 86.7%，对于机器学习任务来说，可以认为该数据集的拟合效果比较好，并且不存在过拟合。随机森林算法的精度能够达到 1.00，相当于预测正类全部预测正确。

Table 3. Dataset 2 Learning performance of each model
表 3. 数据集二各模型学习表现

模型	召回率	F1 值	Support	Accuracy	AUC	Precision
LR	0.90	0.86	10	0.823	0.7601	0.82
RF	0.91	0.95	11	0.941	0.9971	1.00
SVM	1.00	0.96	11	0.941	0.9610	0.92
Xgboost	0.91	0.95	11	0.941	1.0000	1.00

针对数据集二，利用上述指标评价各个模型得到如表 3 所示数据。仍然能够发现树的模型表现要比线性模型要好，并且无论是随机森林还是 Xgboost 和 SVM，表现均优于数据集一，但是发现线性模型的预测能力反而有所下降。因此可以得出结论：数据二数据的非线性性较强，由于随机森林可以叠加不同分类器，Xgboost 可以自动抽取非线性特征，而 SVM 可以通过内核进行无穷升维，进而对非线性性有一个非常好的拟合效果，Xgboost 的效果在非线性的数据集当中效果是最好的。

机器学习还有另外一大评价指标：ROC 曲线，也叫接收者操作特征曲线。在 ROC 曲线当中，曲线与横轴的面积即为 AUC 值，面积越大代表模型分类效果越好。根据图 4 的 ROC 曲线，可以看出分类效果随机森林最好，Xgboost 次之。支持向量机在分类当中的表现略弱。

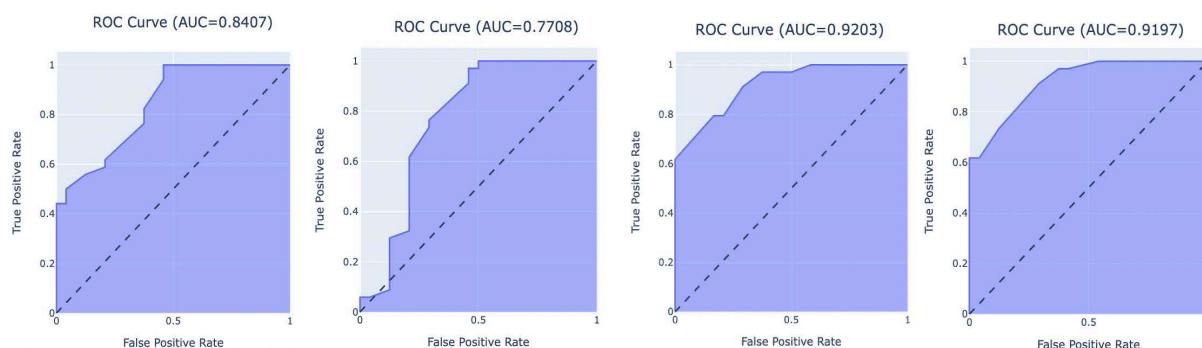


Figure 4. ROC curves for each model, from left to right, logistic regression, support vector machine, random forest, Xgboost
图 4. 各个模型的 ROC 曲线，从左到右依次为 logistic 回归、支持向量机、随机森林、Xgboost

对于各个模型的 AUC 曲线，把正类预测对的概率为红色曲线，预测错的曲线为蓝色曲线，评价指标是红色曲线和蓝色曲线之间的面积差，可以看出对于各个模型来说，红色曲线数值是蓝色曲线的一个上界，证明了机器学习算法的有效性。由以下结果可以看出，基于树的模型分类效果较好，线性模型表现优于支持向量机，Xgboost 表现是最好的。

从图 5 可以看出红色曲线与蓝色曲线之间的面积达到最大的时候，分类正确次数最多，分类错误次数最少，Xgboost 的表现最好。

针对数据集二，利用上述指标评价各个模型。任然能够发现树的模型表现要比线性模型要好，并且无论是随机森林还是 Xgboost 和 SVM，表现均优于数据集一，但是发现线性模型的预测能力反而有所下降。因此可以得出结论：数据二数据的非线性性较强，由于随机森林可以叠加不同分类器，Xgboost 可以自动抽取非线性特征，而 SVM 可以通过内核进行无穷升维，进而对非线性性有一个非常好的拟合效果，Xgboost 的效果在非线性的数据集当中效果是最好的。

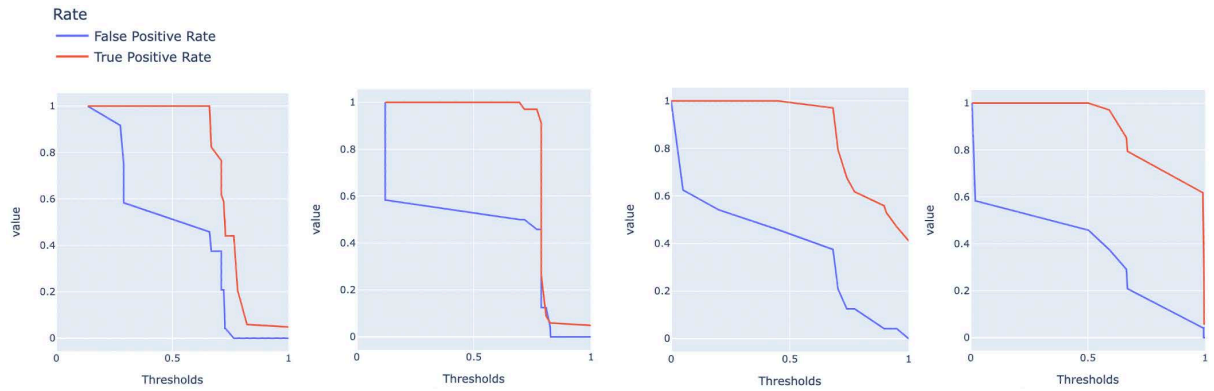


Figure 5. The AUC curves for each model, from left to right, logistic regression, support vector machines, random forests, Xgboost

图 5. 各个模型的 AUC 曲线，从左到右依次为 logistic 回归、支持向量机、随机森林、Xgboost

3.6. Xgboost 效果解释

由下图 6 可见，Xgboost 的 ROC 曲线下的面积为 1，并且 AUC 曲线当中红色曲线和蓝色曲线之间的面积差值非常大，根据概率预测和正负类划分的预测图表明，Xgboost 分类的时候仅仅会把极少数“模棱两可”的数据误分类。

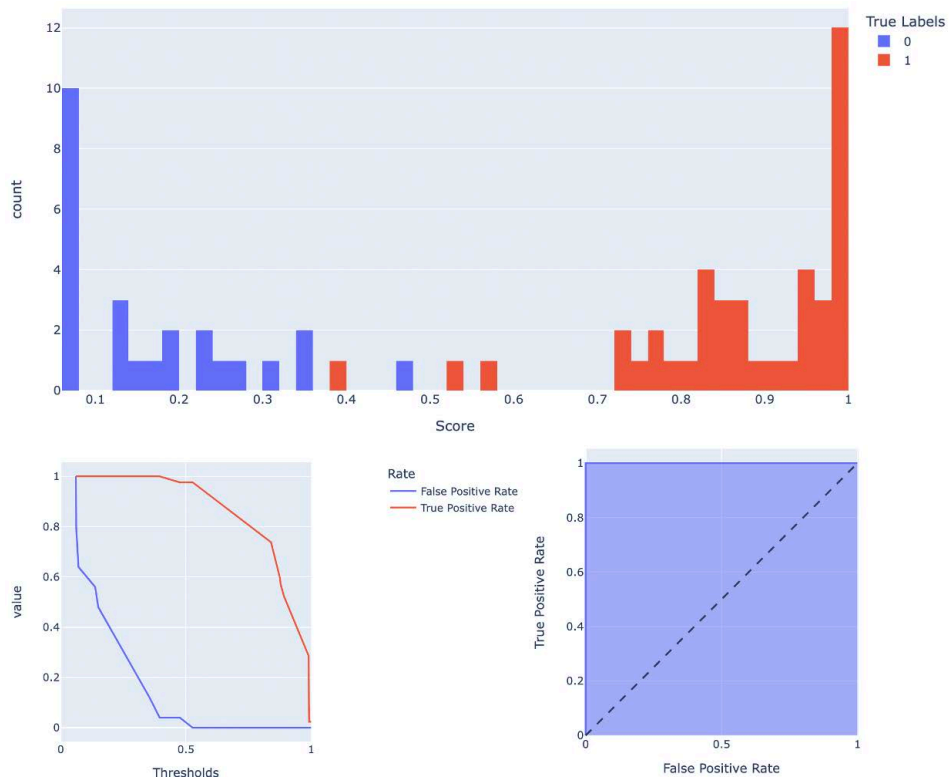


Figure 6. Xgboost's ROC curve, AUC curve, and classification positive and false label curves

图 6. Xgboost 的 ROC 曲线、AUC 曲线与分类正误标签曲线

对于 Xgboost 在数据集上表现良好的原因，可以总结为：

- 1) 鲁棒性强，对异常点不敏感，不需要归一化

- 2) Xgboost 的决策流形可以看做超平面的决策边界
- 3) 利用贪心算法、自动特征选择, 可以提取到更多的非线性特征
- 4) 数据量带来的增益不大, 适合小样本学习

通过上文四种模型对比得出 Xgboost 在预测模型中的优越性, 得到了如下表 4 所示的部分风化概率:

Table 4. Some artifact prediction data based on Xgboost

表 4. 基于 Xgboost 的部分文物预测数据

数据点 1	数据点 2	数据点 3	数据点 4	数据点 5	数据点 6	数据点 7
0.35040572	0.711119956	0.29035667	0.29035667	0.29035667	0.29035667	0.781010075
数据点 8	数据点 9	数据点 10	数据点 11	数据点 12	数据点 13	数据点 14
0.781010075	0.781010075	0.764448616	0.781010075	0.275775233	0.18600099	0.275775233

4. 结果分析

本文首先通过 person 相关系数矩阵得出各特征之间可能存在相关性并使用基于奇异值分解的 PCA 降维将两个数据集维数降低。将处理后的数据按照训练集与测试集 3:1 的比例划分训练 Logistic 回归、随机森林、软间隔支持向量机和 Xgboost 四种模型, 训练结果表明: 数据集一线性相关较强, 数据集二的非线性相关较强, 三随机森林、支持向量机和 Xgboost 三种非线性模型中 Xgboost 在数据集二中表现更好。

根据本文结论, 后续研究可以利用 Xgboost 模型对数据集二进一步分析并依据相关特征划分文物类型。然后根据划分出的类型, 通过灰色关联分析、线性模型和多层感知机模型分别拟合文物中各化学元素的关联性。

参考文献

- [1] 史美光, 何欧里, 吴宗道, 等. 一批中国古代铅玻璃的研究[J]. 硅酸盐学报, 1986(4): 17-23.
- [2] Chen, X., Qi, X.B. and Xu, Z.Y. (2020) Determination of Weathered Degree and Mechanical Properties of Stone Relics with Ultrasonic CT: A Case Study of an Ancient Stone Bridge in China. *Journal of Cultural Heritage*, **42**, 131-138. <https://doi.org/10.1016/j.culher.2019.08.007>
- [3] 姜中宏, 张勤远. 用铅同位素特征研究中国古代铅(钡)玻璃[J]. 硅酸盐学报, 1998(1): 112-116.
- [4] 赵凤燕, 陈斌, 柴怡, 董俊卿, 李青会. 西安出土若干玻璃器的 pXRF 分析及相关问题探讨[J]. 考古与文物, 2015(4): 111-119.
- [5] 李鹏艳, 谢承利, 陆继东, 林兆祥, 李捷, 杨立飞. 激光诱导击穿光谱法分析玻璃成分的实验研究[J]. 应用激光, 2009, 29(1): 21-25.
- [6] 数据预处理之 onehot 编码[EB/OL]. <https://www.jianshu.com/p/38f9f426e246>, 2022-09-15.
- [7] 彭海. 皮尔逊相关系数应用于医学信号相关度测量[J]. 电子世界, 2017(7): 163.
- [8] Ryo, N., Masaaki, O., Michiya, M., Kunimasa, Y. and Hideyuki, H. (2022) Investigation on Application of Singular Value Decomposition Filter in Element Domain for Extraction of Ultrasonic Echoes from Blood Cells in Jugular Veins. *Japanese Journal of Applied Physics*, **61**, SG1011. <https://doi.org/10.35848/1347-4065/ac4680>
- [9] 朱华锋. Logistic 模型的参数估计及其实证研究分析[J]. 科技信息, 2011(1): 169-170.
- [10] Marín, A., Martínez-Merino, L.I., Puerto, J. and Rodríguez-Chía, A.M. (2022) The Soft-Margin Support Vector Machine with Ordered Weighted Average. *Knowledge-Based Systems*, **237**, Article ID: 107705. <https://doi.org/10.1016/j.knosys.2021.107705>
- [11] Chen, T.Q. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [12] 机器学习模型评价指标整理[EB/OL]. https://blog.csdn.net/weixin_43199584/article/details/105722477, 2022-09-15.