

基于LIME的改进机器学习可解释性方法

林志萍, 杨立洪

华南理工大学数学学院, 广东 广州

Email: 1406284563@qq.com

收稿日期: 2021年2月20日; 录用日期: 2021年3月23日; 发布日期: 2021年3月30日

摘要

随着机器学习技术不断取得突破性进展, 越来越多的决策交给复杂自动化的机器学习算法去做。但这些高性能的模型就像黑盒子, 缺乏决策逻辑的透明度和可解释性。LIME (Local Interpretable Model-agnostic Explanation)是由Marco Tulio Ribeiro等人提出的一种XAI (Explainable Artificial Intelligence)方法, 对于复杂的黑盒模型, LIME使用可解释性模型(线性模型)对黑盒模型进行局部近似, 局部解释复杂模型的决策行为。LIME中使用的线性模型(Ridge回归等)学习能力较弱, 不能很好地局部逼近复杂模型。对于复杂树模型(XGB、RF等)本文提出采用可解释性良好的广义加性树模型EBM去近似它们的局部行为, 而对于复杂神经网络模型, 本文提出利用广义加性神经网络模型GAMNET去局部逼近其局部行为。EBM (Explainable Boosting Machine)和GAMNET (广义加性神经网络模型)均具备可解释性并且拥有更强的学习能力, 能更好地逼近复杂机器学习模型。

关键词

机器学习可解释性, LIME, EBM, GAMNET

An Improved Method of Interpret Machine Learning Based on LIME

Zhiping Lin, Lihong Yang

School of Mathematics, South China University of Technology, Guangzhou Guangdong

Email: 1406284563@qq.com

Received: Feb. 20th, 2021; accepted: Mar. 23rd, 2021; published: Mar. 30th, 2021

Abstract

With the development of machine learning technology, more and more decisions are determined

by machine learning algorithms. However, those high-performance models are like black boxes, lacking transparency and interpretability of decision logic. LIME (Local Intepretable Model-agnostic Explanation) is a XAI (Explainable Artificial Intelligence) method proposed by Marco Tulio Ribeiro *et al.* For complex machine learning models, LIME use linear model to approximate and explain their decision behaviors locally. But the learning ability of linear models is weak, so they could not approximate complex models very well. In this paper, to improve LIME, for complex tree models (XGBoost, random forest, etc.), we propose use EBM to approximate (interpretable generalized additive tree model). For complex neural network models, we propose use GAMINET (generalized additive neural network model) to approximate. EBM and GAMINET are both generalized additive models. They have interpretability and stronger learning ability, which can better approximate complex models.

Keywords

Interpretable Machine Learning, LIME, EBM, GAMINET

Copyright © 2021 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

随着机器学习和深度学习的不断发展, 这些精度高、性能优异的复杂模型、应用场景贯穿我们生活的方方面面。数据导入到模型即可得到预测结果。我们无法知道模型是如何从数据当中捕获知识的。在一些涉及到生命健康、财产安全等较为重要的领域, 我们需要对模型的决策进行解释, 另外对错误结果的解释也可以反过来指导对模型的调整。因此除了模型的精度之外, 模型的可解释性也非常重要。对可解释机器学习的研究有着重要的意义。

近年来, 可解释机器学习在科研会议上成为关注热点, 可解释机器学习主要是研究如何使黑盒子模型决策更加透明、可信, 主要分为两大类: 一是内在可解释机器学习, 如逻辑回归; 二是事后解释方法, 如 LIME [1]、SHAP [2]等, 通过事后辅助的归因解析去对复杂模型进行解释。本文主要在 LIME 的基础上进行改进, 使用近年来新提出的可解释性加性模型去逼近复杂模型, 而不是 LIME 内部使用的简单线性模型, 进一步提升 LIME 解释复杂模型的能力。

2. 模型及原理介绍

2.1. 机器学习可解释性

机器学习可解释性是指人类对机器学习模型预测结果的理解程度大小。模型可解释性主要回答如何由输入数据得到预测结果的问题, 是对输入特征和预测结果之间关系的定性理解。基本的可解释性模型主要有线性回归、浅层决策树、朴素贝叶斯、以及 K 近邻等。这些模型具备较强的可解释性, 但令人觉得遗憾的是和复杂的机器学习模型如集成树模型以及神经网络模型相比, 它们的学习能力非常有限, 只能解决一些简单的问题。复杂的机器学习模型在许多目标任务上取得了良好的性能, 但绝大多数是黑盒子模型, 没法说明从输入到输出之间的因果关系。而在一些重要领域如金融、法律、医疗健康等, 模型的可解释性非常重要, 如风控借贷模型中, 对一个被拒绝贷款的客户, 我们有必要向其解释拒绝放贷的原因。再如机器学习模型在区分恶性肿瘤和不同类型的良性肿瘤方面是非常准确的, 但是我们依然需要

专家对诊断结果进行解释, 解释为什么一个机器学习模型将某个患者的肿瘤归类为良性或恶性将有助于医生信任和使用机器学习模型来支持他们工作。此外有些场景下, 模型的预测结果正确但原因却是错的, 故为模型赋予较强的可解释性非常有必要。本小节将利用来自 UCI 数据库[3]的糖尿病患者数据集 Diabetes 和 EBM [4]模型简单阐述一下模型可解性的必要性。

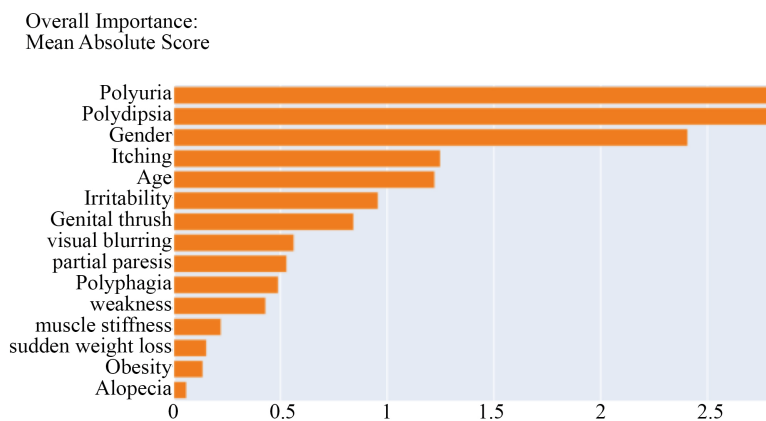


Figure 1. Feature importance of EBM
图 1. EBM 模型输出的全局特征重要性图

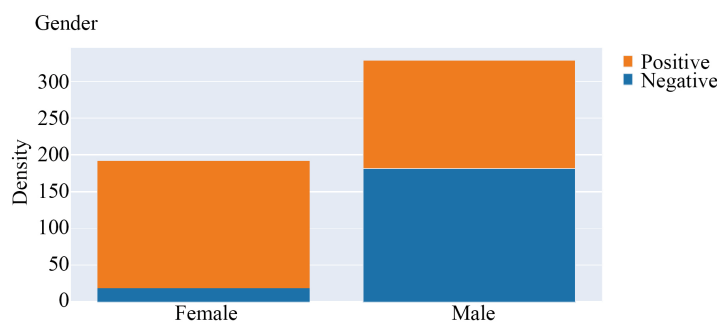


Figure 2. Feature age and label distribution
图 2. 性别特征和标签分布图

EBM 的预测结果达到 0.969, 但由图 1 可以看到, 性别特征(gender)对最终决策的影响程度排第三, 而实际上有关方面的调查显示在糖尿病患者中男女患病率之比接近 1:1, 因此模型预测结果正确度高但依赖的原因是不对的。对数据集的分布进一步查看发现, 性别为女性的样本当中结果为阳性的占绝大多数, 而性别为男性的样本中阳性阴性接近 1:1 (如图 2 所示)。故模型决策依据错误是输入数据不合理造成的。因此可解释性是非常有必要的, 它可以让我们知道模型做出相应决策的原因, 并及时发现模型存在的问题, 从而创造更加安全可靠的应用。

2.2. LIME 模型

LIME 由 Marco Ribeiro 等人[1]于 2016 年提出, 是帮助我们解释复杂黑盒模型如何做出决策的一个工具。它是一种模型无关的机器学习可解释性方法, 适用于解释所有模型包括神经网络、XGBoost、随机森林等。并且可应用于多种类型的数据, 包括表格数据、文本数据、图像数据等。LIME 主要有三个特点:

- 1) 只对模型进行局部解释, 而不提供整体解释, 及局部对每个样本进行解释;

2) 通过局部构建简单的可解释性模型进行预测去对重要特征进行解释;

3) 解释的是当前输入的特征和预测结果之间的关系, 而不包含复杂模型训练当中产生的抽象特征。

对于要解释的黑盒模型, 取要关注的样本点, 在该样本点进行扰动生成新的样本, 根据关注样本点的领域范围, 选择在领域范围内的样本并用黑盒模型进行预测得到预测值。于是得到一个新的数据集, 使用该数据集训练线性模型, 得到黑盒模型较好的局部近似, 利用可解释的模型可以知道黑盒模型局部的决策行为。

记 $x \in R^d$ 为要解释的样本, 首先选出较为重要的 d' 维特征, x 去掉不重要的分量之后变为 $x' \in R^{d'}$ 。

在 x' 附近扰动生成新的样本点 z' , 新样本点构成新的数据集 Z' 。样本点添加去掉的特征分量之后恢复成样本 $z \in R^d$ 。定义 $\pi_x(z)$ 为扰动前后样本的相似度, 其计算公式如下:

$$\pi_x(z) = \exp\left(-\frac{D(x, z)^2}{\sigma^2}\right) \quad (1)$$

其中 $D(x, z)$ 为距离公式, 不同的样本类型会有不同的定义, 当为图像数据时通常为 L2 范数距离, 当为文本数据时通常为余弦相似性。

记 f 为需要解释的复杂模型, g 为简单模型, 则衡量两个模型之间差异的目标函数如下所示:

$$\xi(x) = \sum_{z, z'} \pi_x(z) (f(z) - g(z'))^2 + \Omega(g) \quad (2)$$

其中 $\Omega(g)$ 为模型 g 的复杂度, 当 g 为线回归模型时, 模型复杂度为权重系数不为 0 的个数。

LIME 算法流程如下表 1 所示:

Table 1. Algorithm: LIME

表 1. LIME 算法流程

输入: 1) 复杂模型 f ; 2) 关注样本 x ; 3) 随机生成的样本数量 N 。

步骤:

1. 通过特征筛选, 初步获取较为重要的 d' 个特征, 得到关注样本的解释版本 x' ;
2. 在样本 x' 附近进行随机扰动生成新样本数据 z' , z' 恢复为和 x 同维度样本 z , 并用复杂模型进行预测获取标签;
3. 对新生成的数据集用线性模型进行拟合。

输出: 线性模型权重

2.3. EBM

EBM 于 2015 年被 Rich Caruana 等人[4]提出, EBM 被应用到医学领域并取得不错的效果。EBM 基于广义加性模型 GAMs (Generalized additive models) 而提出的模型。广义加性模型具有很好的可解释性。像逻辑回归这样的广义线性模型 GLMs (Generalized linear models) 实际上是广义加性模型的一种特殊形式记 x_j 为样本的第 j 个特征, 其中特征个数为 p , 则广义加性模型的标准形式为:

$$g(E[y]) = \beta_0 + \sum_j f_j(x_j) \quad (3)$$

其中 g 是连接函数, $E(f_j) = 0$ 。

EBM 在广义加性模型的基础上加入了两个特征的交互项, 一种自动搜索选取交叉项方式由 Yin Lou 等人提出[5], EBM 的公式如下所示:

$$g(E[y]) = \mu + \sum_j f_j(x_j) + \sum_{i,j} f_{i,j}(x_i, x_j) \quad (4)$$

其中 $f_j(x_j)$ 为岭函数(shape function), $f_{i,j}(x_i, x_j)$ 为交叉项岭函数, 对于回归问题 EBM 使用回归样条(Regression Splines)作为岭函数, 对于分类问题 EBM 采用二叉树或者集成二叉树作为岭函数。最后对每个特征以及筛选出的交叉项对应的岭函数进行线性组合得到预测结果, 具体细节以及迭代求解过程请参考文献[5]。值得注意的是本文将采用不含交叉项的 EBM 模型, 不含交叉项的 EBM 模型具有更强的可解释性。是一种传统的广义加性模型。

2.4. GAMINET

GAMINET 由香港大学教授张爱军于 2020 年提出[6], GAMINET 和 EBM 的原理架构非常相似, 最大的区别是 EBM 的岭函数采用二叉树或者集成二叉树生成, 而 GAMINET 则采用全连接神经网络, GAMINET 每个特征生成一个子网络或者两个特征交叉生成一个子网络, 每个子网络对应一个输出, 即为岭函数。最后对岭函数加上一层全连接层, 基于全连接层得到最终的输出结果。GAMINET 的模型表达式为:

$$g(E[y]) = \mu + \sum_{j \in S_1} f_j(x_j) + \sum_{i,j \in S_2} f_{i,j}(x_i, x_j) \quad (5)$$

其中 S_1 为主效应(main effect)构成的集合, 主效应即单个特征对应的岭函数, S_2 为交互效应构成的集合, 交互效应即为交叉项对应的岭函数。训练时, GAMINET 将每个特征输入子网络进行训练, 单个特征子网络模块训练完成后, 再将残差作为响应变量继续对交叉项子网络模块进行训练。同样, 为具备更强的可解释性, 本文将采用的不含交叉项的 GAMINET 模型。

3. LIME 模型的改进

LIME [1]算法当中采用线性模型对关注样本附近随机扰动生成的样本进行拟合, 局部逼近复杂模型在关注样本附近的决策行为。我们都知道, 线性模型解释性强, 但学习能力非常有限, 故本文将使用不含交叉项的 EBM [4]和 GAMINET [6]去近似复杂模型在关注样本附近的行为。不含交叉项的 EBM 模型和 GAMINET 模型属于广义加性模型, 具有良好的可解释性, 模型输出的特征重要性可与线性模型一般视为每个特征对最终决策的影响程度。由于 EBM、GAMINET 对单个特征通过集成树或者神经网络进行加工, 相比线性回归具备更强的特征学习能力。如当响应变量和特征分量之间的关系为 $y = x_1^2 + x_2^3 + \sin(x_3)$ 时, 对于线性回归模型是没有办法学习到这样的关系, 而 EBM、GAMINET 则可通过集成树或者神经网络将特征分量进行升华, 如:

$$\begin{aligned} x_1 &\rightarrow x_1^2 \\ x_2 &\rightarrow x_2^3 \\ x_3 &\rightarrow \sin(x_3) \end{aligned}$$

再通过最后的线性层将升华后的特征进行组合得到响应变量和特征分量之间的关系。由于最后一层结构是线性的, 故我们很容易计算每个特征分量对决策结果的贡献程度, 因此它们具有非常好的可解释性的同时精度也提高了。

3.1. 数据集介绍

本文将采用 UCI 数据库提供的 Adult 数据集[7]进行实验, Adult 数据集是从美国 1994 年人口普查数据库抽取而来, 用于预测居民收入是否超过 50 k, 标签为是否超过 50 k, 特征变量主要包含有性别、年龄、职业、教育程度、婚姻状况、种族等。一共有 48,842 条数据。

3.2. 模型对比

现选择 XGBoost 为复杂模型, LIME 模型算法内部利用的是线性模型(Ridge 回归、逻辑回归)去局部近似复杂模型的局部行为,本小节将在 LIME 模型算法的框架下,用 EBM 以及 GAMINET 去逼近 XGBoost 在关注样本点附近的行为,比较 EBM、GAMINET 和线性回归模型对复杂模型的逼近复杂模型的能力。使用 XGBoost 对 Adult 数据集进行拟合,在训练集上的准确率为 0.883, AUC 为 0.941,在测试集上准确率为 0.883, AUC 为 0.923。

对某个关注样本点,分别采用 10 个不同的随机数种子在该样本点附近扰动生成 10 个数据集,每次生成 5000 个样本,并分别用线性模型、EBM 以及 GAMINET 进行拟合,结果如下表 2 所示,其中括号左边是模型在生成的数据集上的 AUC 值,右边是准确率。

Table 2. Performance of three models on randomly generated datasets near certain sample (XGBoost)

表 2. 三种模型在关注样本附近随机生成的数据集上的拟合结果(XGBoost)

	线性模型	EBM	GAMINET
1	(0.967, 0.951)	(0.997, 0.992)	(0.983, 0.977)
2	(0.965, 0.950)	(0.999, 0.985)	(0.996, 0.968)
3	(0.964, 0.953)	(0.999, 0.986)	(0.995, 0.980)
4	(0.961, 0.951)	(0.999, 0.984)	(0.996, 0.979)
5	(0.957, 0.948)	(0.999, 0.984)	(0.997, 0.983)
6	(0.953, 0.951)	(0.999, 0.987)	(0.995, 0.980)
7	(0.965, 0.953)	(0.999, 0.986)	(0.996, 0.986)
8	(0.956, 0.950)	(0.999, 0.985)	(0.997, 0.981)
9	(0.967, 0.954)	(0.999, 0.987)	(0.994, 0.978)
10	(0.953, 0.953)	(0.999, 0.986)	(0.996, 0.988)

同样也是在 LIME 模型算法的框架下用不含交叉项的 GAMINET 去局部近似全连接神经网络在关注样本点附近的行为,并和 EBM 以及线性模型进行比较。注意到两个复杂模型的数据处理方式略不相同,故图中显示的特征值不一样,且复杂模型决策依据也有所出入。全连接神经网络在训练集上的 AUC 为 0.906,准确率为 0.848,在测试集上的 AUC 为 0.891,准确率为 0.840,三种模型在扰动生成的数据集上的结果如表 3 所示。

Table 3. Performance of three models on randomly generated datasets near certain sample (neural network)

表 3. 三种模型在关注样本附近随机生成的数据集上的拟合结果(全连接神经网络)

	线性模型	EBM	GAMINET
1	(0.955, 0.908)	(0.987, 0.856)	(0.990, 0.960)
2	(0.960, 0.919)	(0.989, 0.887)	(0.993, 0.972)
3	(0.948, 0.911)	(0.988, 0.882)	(0.992, 0.970)
4	(0.956, 0.914)	(0.987, 0.886)	(0.992, 0.973)
5	(0.939, 0.911)	(0.984, 0.877)	(0.989, 0.966)
6	(0.963, 0.951)	(0.989, 0.929)	(0.991, 0.982)
7	(0.969, 0.941)	(0.988, 0.909)	(0.993, 0.983)
8	(0.950, 0.910)	(0.985, 0.879)	(0.988, 0.964)
9	(0.955, 0.930)	(0.988, 0.898)	(0.992, 0.972)
10	(0.963, 0.933)	(0.988, 0.900)	(0.993, 0.976)

可以看到, 对两种复杂模型(XGBoost 和全连接神经网络), 整体上 EBM 和 GAMINET 的效果都比线性模型要好。对于 XGBoost, EBM 的拟合效果要比 GAMINET 要好一些, 而对于全连接网络, GAMINET 拟合的效果比 EBM 更佳。这可能是由于 EBM 是基于树的, 更加适合逼近复杂树模型, 而 GAMINET 是基于神经网络的, 和复杂神经网络的决策逻辑更相契合一些, 因此拟合神经网络的效果更好。

接下来将在 LIME 的算法框架下, 基于 EBM 拟合复杂模型 XGBoost 在关注样本点附近的行为并进行解释决策依据, 同样, 本文将基于 GAMINET 去拟合全连接网络在关注样本点附近的行为以及寻找复杂模型局部的决策原由。关注样本点的特征取值在局部解释图中有呈现。

3.3. 基于 EBM 的 LIME 模型改进算法

本小节用 XGBoost 在 Adult 数据集上训练一个预测模型, 并将选择一个关注样本点, 分别用原始的 LIME [1]模型以及基于 EBM [4]的改进 LIME 模型对 XGBoost 在该关注样本点附近的行为进行解释。复杂模型在训练集和测试集上的 AUC 分别为 0.941, 0.923, 准确率分别为 0.883, 0.873, 图 3 和图 4 分别为原 LIME

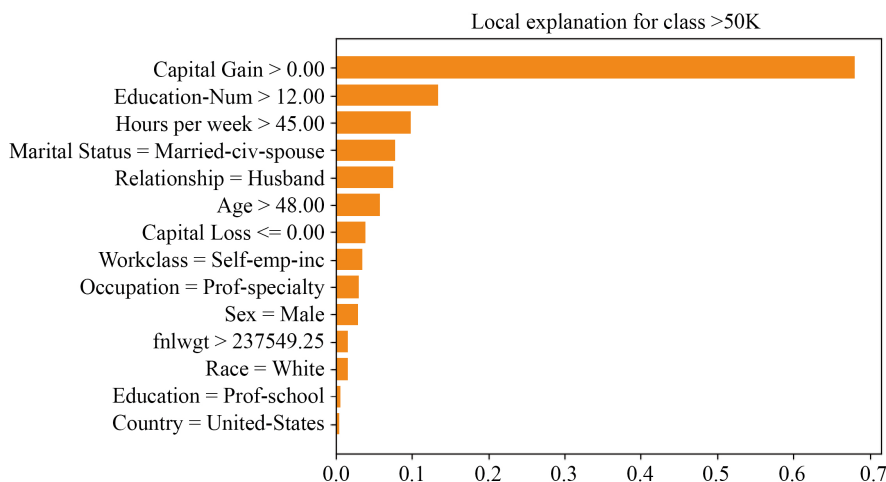


Figure 3. Feature importance of LIME globally
图 3. LIME 输出的全局特征重要性图

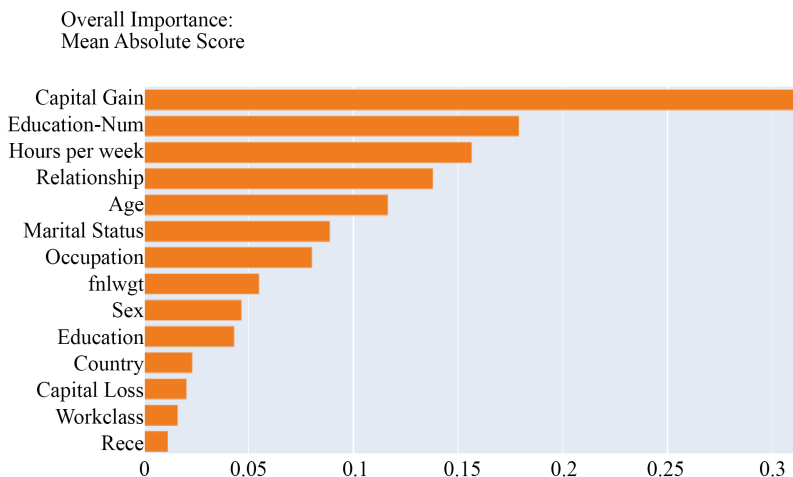


Figure 4. Feature importance of improved LIME based on EBM globally
图 4. 基于 EBM 的改进 LIME 模型输出的全局特征重要性图

模型和基于 EBM 改进后的 LIME 模型在关注样本点附近生成的数据集上整体特征重要性图(绝对值), 可以看到两个模型均认为 Capital Gain、Education-Num 以及 Hours per week 是最重要的特征, 并且均是和标签 >50 k 是正相关的, 关注样本点在这三个特征附近取值均较大, 因此被预测为 >50 k, 当我们取在这三个特征上取值相对较小的样本进行研究时, 绝大多数样本被复杂模型预测为 ≤ 50 k, 故不管是原始 LIME 模型还是基于 EBM 的改进 LIME 模型都可以让我们看到特征和复杂模型预测结果之间的关系, 探索其决策依据。

进一步可以看到 EBM 认为 Relationship 比 Marital Status 更加重要, 即前者比后者更加影响复杂模型的决策行为, 而 LIME 则反之。我们取 Relationship、Marital Status 这两个特征列联合标签放入一个逻辑回归模型中训练, 从图 5 可以看到逻辑回归输出的 Relationship 特征重要性绝对值要比 Marital Status 要大。基于 EBM 的 LIME 模型的解释性更加可靠, 这是因为 EBM 的学习能力比逻辑回归要强。

此外两个模型输出的局部特征重要性图提供了关注样本点的每个特征取值对标签的影响程度。如图 6 和图 7 所示。

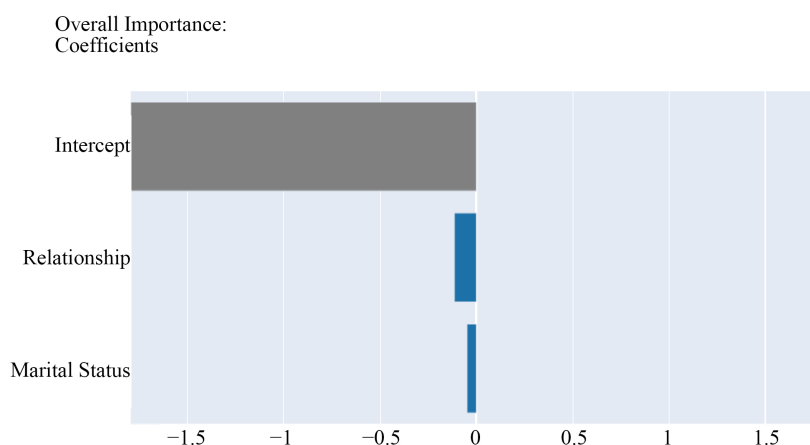


Figure 5. Feature comparison: Relationship VS. Marital Status

图 5. 特征 Relationship 和 Marital Status 比较图

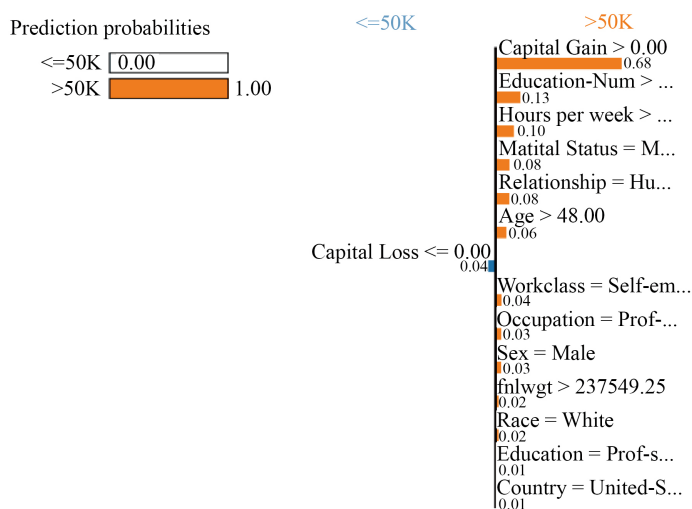


Figure 6. Feature importance of LIME locally

图 6. LIME 输出的局部特征重要性图

Predicted (1): 0.717 | Actual (1): 0.717

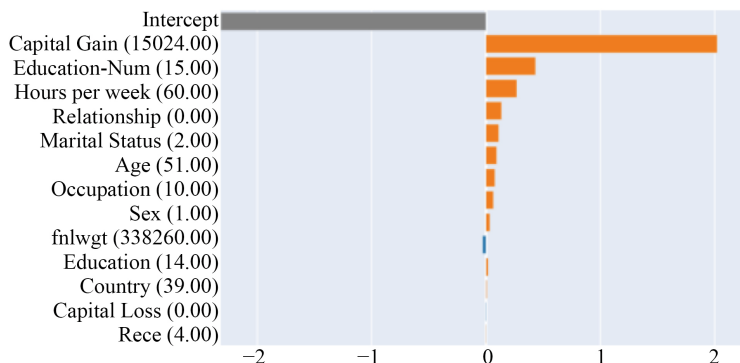


Figure 7. Feature importance of improved LIME based on EBM locally
图 7. 基于 EBM 的改进 LIME 模型输出的局部特征重要性图

3.4. 基于 GAMINET 的 LIME 模型改进算法

本小节用全连接神经网络在 Adult 数据集上训练一个预测模型, 选择和上一小节相同的关注样本点, 分别基于 LIME [1]模型以及基于 GAMINET [6]的改进 LIME 模型对复杂模型在关注样本点附近的行为进行解释, 寻找决策依据。复杂模型在训练集、测试集上的 AUC 分别为 0.907, 0.893, 准确率分别为 0.849, 0.843。从图 8 和图 9 可以看到两个模型特征重要性排前三的特征均为 capital_gain、education_num、marital_status, 原始 LIME 模型认为 capital_loss 比 hours_per_week 更加重要, 而基于 GAMINET 的改进模型则反之。同样我们取 capital_loss、hours_per_week 这两列特征以及标签列放入逻辑回归中训练, 得到如图 10 所示的逻辑回归特征重要性图。可以看到 hours_per_week 相对 capital_loss 来说特征重要性绝对值更大。基于 GAMINET 的 LIME 改进模型能够更加准确地逼近复杂模型的局部行为。图 11 为 GAMINET 每个子网络的输入到输出的函数曲线图, 当特征为连续特征时是一条光滑曲线, 若为离散特征则为柱形图。GAMINET 从原始输入到最后的输出都是公开透明的, 是结构上可解释性模型, 且它不是简单的线性模型, 具有更强的学习能力。

此外两个模型输出的局部特征重要性图提供了关注样本点的每个特征取值对标签的影响程度。如图 12 和图 13 所示。

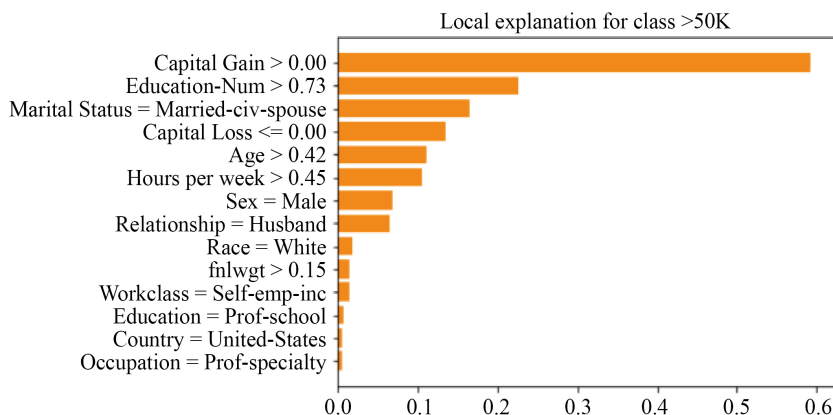


Figure 8. Feature importance of LIME globally
图 8. LIME 输出的全局特征重要性图

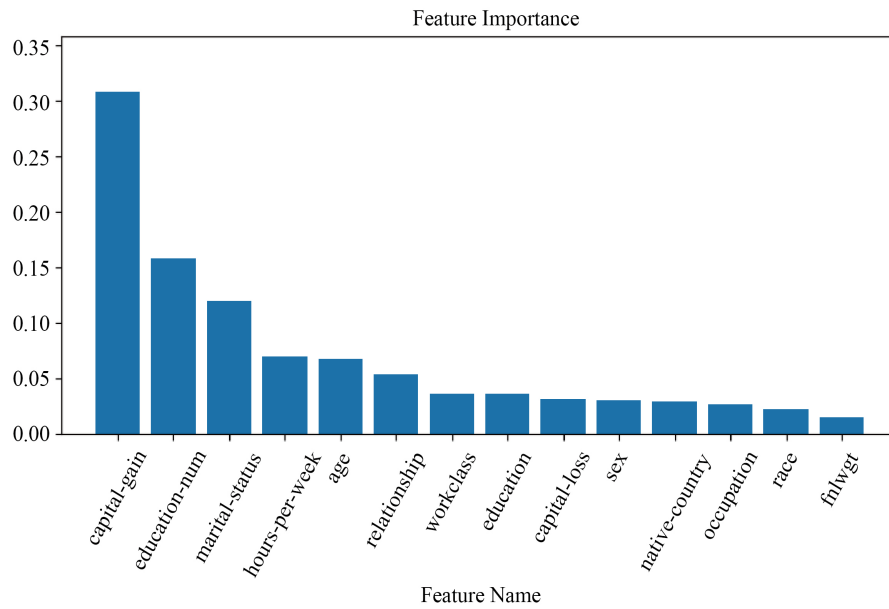


Figure 9. Feature importance of improved LIME based on GAMINET globally
图 9. 基于 GAMINET 的改进 LIME 模型输出的全局特征重要性图

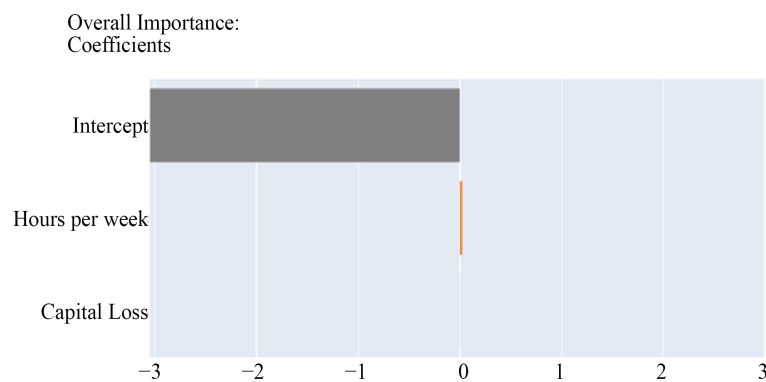
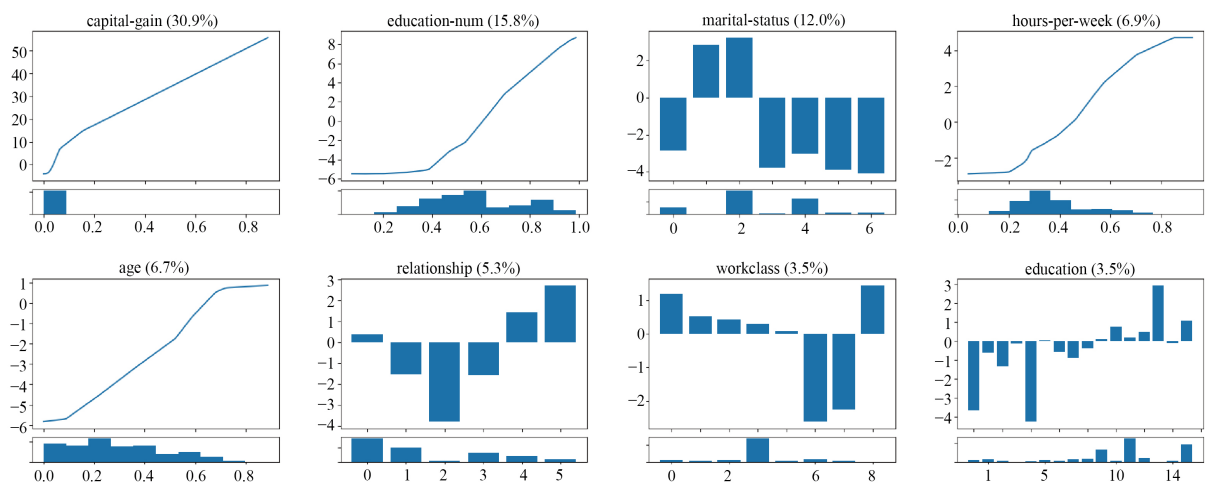


Figure 10. Feature comparison: Hours per week VS. Capital Loss
图 10. 特征 Hours per week 和 Capital Loss 比较图



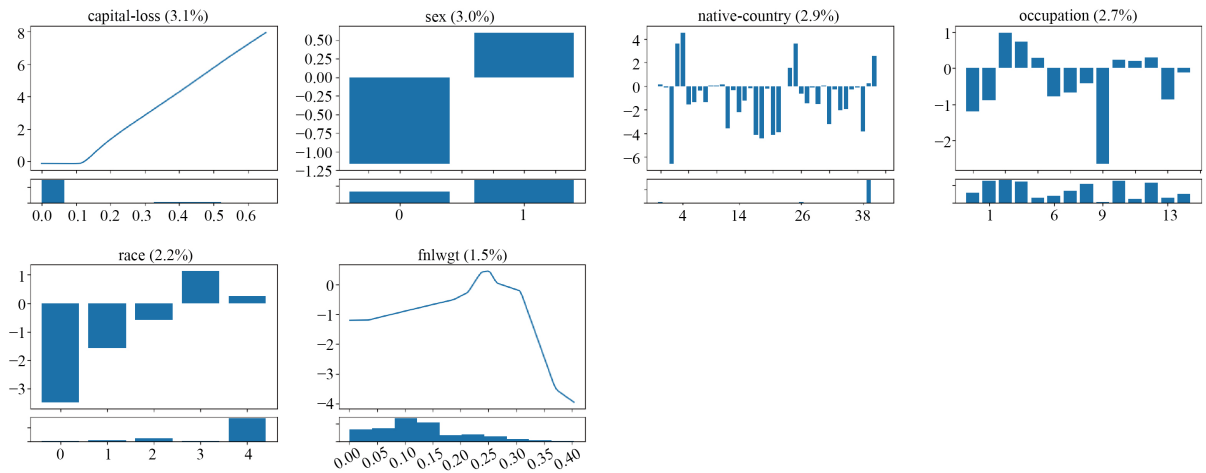


Figure 11. Shape function of each feature
图 11. 每个特征对应的岭函数图

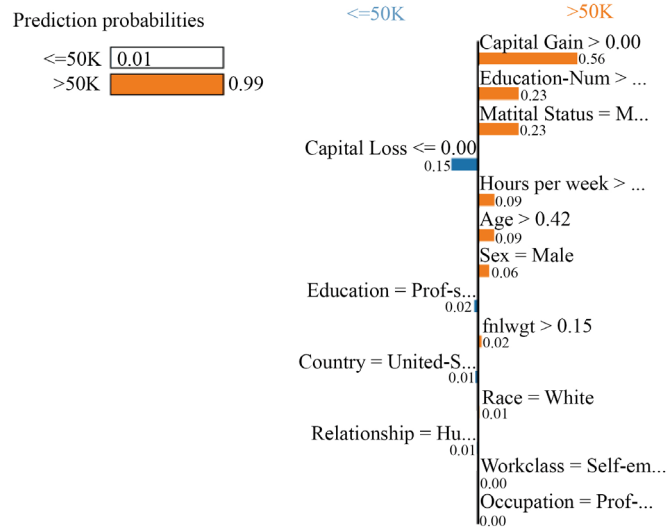


Figure 12. Feature importance of LIME locally
图 12. LIME 输出的局部特征重要性图

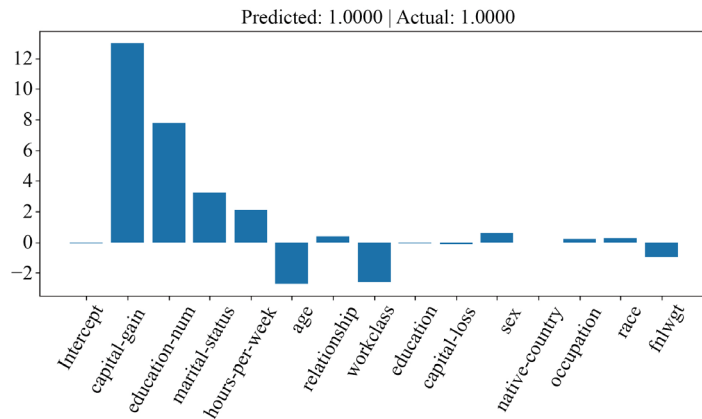


Figure 13. Feature importance of improved LIME based on GAMINET locally
图 13. 基于 GAMINET 的改进 LIME 模型输出的局部特征重要性图

4. 评价与展望

本文提出了基于机器可解释性方法 LIME [1]的改进算法, LIME 内部采用的是线性模型去逼近复杂模型的局部行为, 线性模型结构简单, 可解释性强但学习能力弱, 拟合效果一般。而本文采用的是较新提出的 EBM [4]模型和 GAMINET [6]模型, 两种模型均为结构上可解释性模型, 具备较强学习能力的同时也具备良好的可解释性, 进一步提升了 LIME 对复杂模型的解释能力。并且经实验发现, 基于 EBM 的改进 LIME 模型更适合去解释复杂树模型, 而基于 GAMINET 的改进 LIME 模型则更适合解释神经网络模型。LIME 原理简单, 适用范围广, 可以解释所有的黑盒模型, 但 LIME 在关注点附近进行扰动生成新样本时需要确定邻域范围, 不同的邻域范围得到的局部可解释性模型相差可能很大, 并且 LIME 在生成新样本点时, 新样本的每个特征分量是基于整个数据集的每列特征单独的分布生成的, 没有考虑到特征的交叉性, 导致可能生成不大可能出现的样本点, 因此还需进一步改进该方法。

参考文献

- [1] Ribeiro, M.T., Singh, S. and Guestrin, C. (2016) Why Should I Trust you?: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD*, 13, 1135-1144. <https://doi.org/10.1145/2939672.2939778>
- [2] Scott, M.L. and Su-in, L. (2016) A Unified Approach to Interpreting Model Predictions. *31st Annual Conference on Neural Information Processing Systems*, 4766-4775.
- [3] UCI Datasets. <http://archive.ics.uci.edu/ml/datasets/Diabetes>
- [4] Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M. and Elhadad, N. (2015) Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-Day Readmission. *Proceedings of the 22nd ACM SIGKDD*, August 2015, 1721-1730. <https://doi.org/10.1145/2783258.2788613>
- [5] Yin, L., Rich, C., Johannes, G. and Giles, H. (2013) Accurate Intelligible Models with Pairwise Interactions. *Proceedings of the 19th ACM SIGKDD*, 623-631.
- [6] Zebin, Y., Aijun, Z. and Agus, S. (2020) GAMI-Net: An Explainable Neural Network Based on Generalized Additive Models with Structured Interactions. <https://arxiv.org/>
- [7] UCI datasets. <http://archive.ics.uci.edu/ml/datasets/Adult>