

《Exploratory Machine Learning with Unknown Unknowns》论文复现报告

论文作者：Peng Zhao, Yu-Jie Zhang, Zhi-Hua Zhou

复现者：郑凌霄

摘要

在传统的监督学习中，训练数据集中所有样本的真实标签都属于一个已知标签集，所以机器学习模型就会将未知标签的样本分类到已知的标签中。在实际情况下，当学习到的模型不能很好地工作时，研究者通常会将模型的失败归咎于学习算法的选择不充分或缺乏足够的训练样本。在该论文中，作者提出了一个会引起失败可能性，即训练数据集中的样本标签存在错误。它们在一开始就被视为已知类，但是它们其实真正属于隐藏的未知类。因此在传统的机器学习下，这些隐藏的未知类是无法被正确识别出来的，而且很难通过修改学习算法或者增加数据集中样本的数量来解决。所以，该论文提出了一种新的机器学习框架，探索性机器学习（ExML），来解决这个问题。这个框架可以检查是否有隐藏的未知类存在，只要发现在原始的训练模型下，机器学习模型的学习效果不好，就可以考虑数据集中有隐藏的未知类的存在。然后，对这些可能存在的隐藏的未知类的样本进行特征空间增强，并再次进行学习，从而使得它们可以被观察到。在合成数据集和公开数据集上进行的理论分析和实验研究，验证了该方法的有效性。同时，我们尝试使用 Python 代码，并修改部分参数，对该论文进行复现，得到了不错的结果。

关键词：机器学习；未知的未知；ExML

1 引言

机器学习在许多现实世界的应用中取得了巨大的成功。合适的学习算法和足够的监督训练数据是模型训练成功的关键。因此，面对模型训练的失败，研究者总是会怀疑算法的选择是否充分或者训练的数据是否足够。为了解决这个问题，常见的做法是尝试使用其他的算法或者积累更多的数据，当没有其他因素导致训练失败时，这种方法通常可以有效地解决问题。

在传统的监督学习中，训练数据集中所有样本的真实标签都属于一个已知标签集，所以机器学习模型就会将未知标签的样本分类到已知的标签中。在实际情况下，当学习到的模型不能很好地工作时，研究者通常会将模型的失败归咎于学习算法的选择不充分或缺乏足够的训练样本。但是，有一个一直以来都被忽视的原因：隐藏在训练数据集中的未知的未知类。也就是说，训练数据集中某些未知类的训练样本被错误地视为其他类，这些训练样本没有正确的标签，因为他们的正确标签不属于已知标签集，在给定的监督学习下，模型不知道它们的存在。当训练数据集中出现这些未知的未知数时，传统的监督学习（Supervised Learning, SL）无法获得一个令人满意的结果，于是该论文提出了探索性机器学习（Exploratory Machine Learning, ExML）来检测未知的未知数的存在并进行学习。

探索性机器学习（ExML）主要包括三个部分：拒绝模型、特征探索和模型级联。本次课程的论文复现工作拟使用 ExML，通过对更多的特征进行探索来检测和研究训练数据集，从而发现训练数据集中未知的未知类并进行识别。与传统的监督学习（SL）相比，ExML 可以检测训练集中未知的未知类的存在，并将它们识别出来。即使与带有拒绝模型的 SL（可以发现未知类）相比，ExML 在发现未知类和预测已知类等方面也更具有优势。

2 相关工作

在本节中，我们将讨论与探索性机器学习相关的一些工作。

2.1 开放分类学习

开放分类学习也被称为带有新类的学习，它专注于处理只出现在测试阶段的未知类^[1]。虽然这些研究也关注到了未知类的检测，但与本文的研究有很大的不同，不能适用于本文的情况。具体来说，这种学习方法不考虑特征质量的缺陷，而且在本文的设置中，训练数据集中存在隐藏的未知类，而他们的未知类只出现在测试阶段。

2.2 高置信度错误预测

高置信度错误预测的出现，是由于模型没有意识到存在这类错误，这样的样本和预测也被称为一种”未知的未知”^[2]。由于模型对这些错误的预测给出了很高的置信度，因此几乎不可能纠正这种错误。所以，现有的研究通常要求人类专家帮助识别高置信度的错误预测，然后在指导下重新训练模型。虽然这些工作也考虑了未知的未知，并诉诸于外部的人类知识，但它们的设置和方法与本论文不同：该论文的未知的未知是由于特征不足造成的，因此学习者需要增强特征而不是查询标签。

2.3 主动学习

主动学习的目的是通过要求人类专家对未标记的数据进行标记，从而实现用更少的标签以获得更高的准确性^[3]。此外，还有一些人试图通过尽可能少地查询的条目值（样本的特征）来改进特征学习的缺失。与它们的设置不同，在本文的工作中，增加了新的特征来帮助识别未知类，而不是查询给定特征的缺失值来提高已知类的分类能力。

2.4 带有拒绝模型的学习方法

带有拒绝模型的学习方法可以选择拒绝样本，但不是提供低置信度的预测。这为设计有效的算法和建立理论基础提供了大量的工作^[4]。如前所述，该方法不能直接应用于探索性机器学习中，因为它会导致对已知类样本的不准确拒绝，同时也不能像探索性机器学习一样增加新特性来进行学习。

3 本文方法

考虑如图 1 所示的医疗诊断任务。

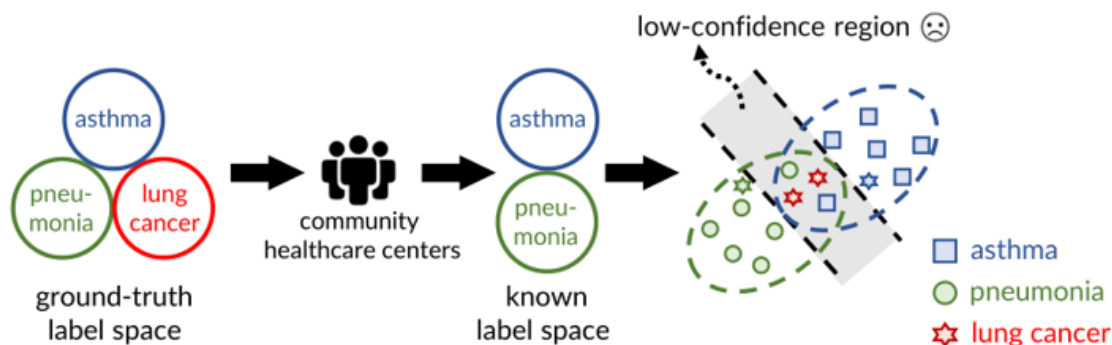


图 1: 医学诊断任务中的未知的未知数（图片来源于原论文^[5]）

根据当地医疗机构的记录训练模型，诊断出现哮喘或肺炎症状患者的病因。一共有三种类型的原因：两种常见类型（哮喘和肺炎）和一种罕见的病因（肺癌，需要昂贵的计算机断层扫描设备，如 CT）。

由于上述原因，当地医疗机构没有诊断出肺癌，因此在收集的数据集（已知标签空间）中没有肺

癌 的 类 别 。 因 此 ， 学 习 模 型 需 要 同 时 发 现 未 知 类 和 对 已 知 类 进 行 分 类 。

3.1 本文方法概述

在本节中，我们将介绍探索性机器学习（ExML）。由于缺乏特征，在机器学习中需要利用人类来处理未知的未知数，从而更深入地挖掘特征。研究者检查数据集并进行探索（发现新的特征信息），以发现隐藏的未知类并对现有类进行分类。具体包括以下内容：

- （1）当学习模型被赋予大量数据且性能较低时，学习者怀疑未知类的存在并进行特征强化。
- （2）如图 2 所示，由于学习者注意到了一些分类效果不好的类并进行适当的特征增强，从而发现存在一个隐藏的未知类。

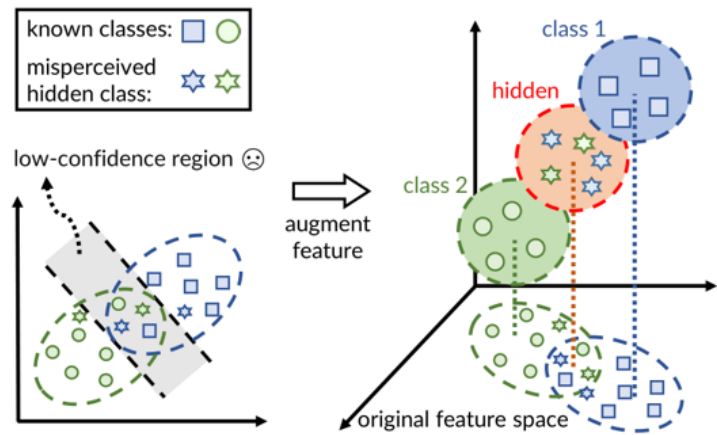


图 2: 信息特征可以显著提高低置信度样本的可分离性，并使隐藏类可区分（图片来源于原论文[5]）

为了处理由于特征不足而导致的未知的未知，我们在学习过程中依靠人与环境交互来增强数据收集，更具体点，就是积极地增加特征空间。其思想是，当一个学习过的模型即使有更多的训练数据，仍然表现不佳时，就会怀疑隐藏类的存在，然后寻找几个候选特征来增强。模型接收到一个数据集，并观察到有两个可分离性较差的类，导致明显的低置信度区域。在经过适当的特征增强后，就会意识到之前由于特征的不足，在训练数据中隐藏了一个额外的类。

传统的 SL 将训练数据集视为环境的可观察表示，并利用它训练一个模型来预测标签。相比之下，ExML 认为训练数据集是可操作的，模型可以通过探索更多的特征信息来检查和研究数据集，从而发现由于特征不足而导致的未知的未知。ExML 与传统的 SL 的对比如图 3 所示：

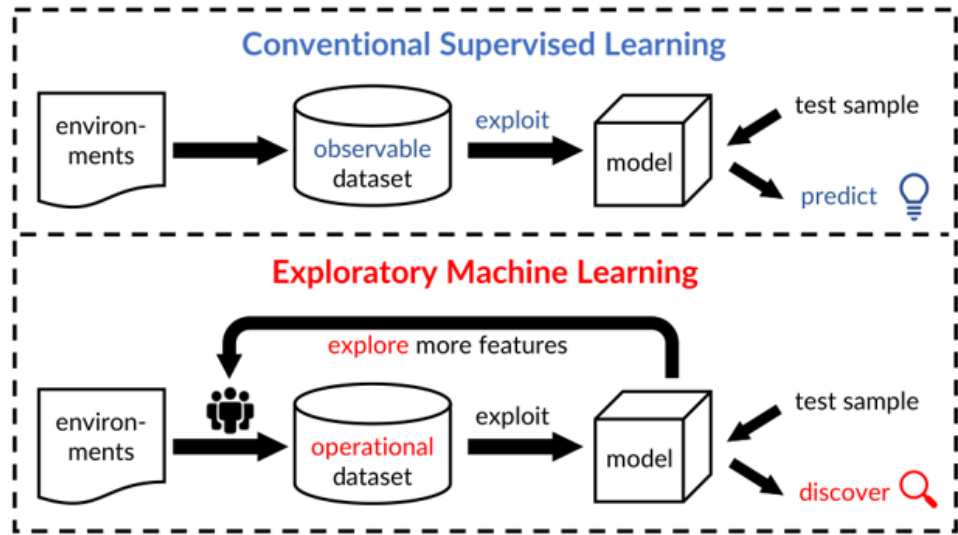


图 3: 两种机器学习框架的比较（图片来源于原论文[5]）

本论文所提出的 ExML 包含三个要素，如图 4 所示：

- (1) 拒绝模型 (rejection model)：鉴定哪些实例可能是未知未知类；
- (2) 特征探索 (feature exploration)：指导特征的探索，并基于增强特征空间重新训练模型；
- (3) 模型级联 (model cascade)：逐层处理以优化可疑实例的选择。

从初始模型 (蓝色部分) 开始，然后探索最佳候选特性 (绿色部分)。基于增强数据集对模型进行再训练，最后与初始模型级联，发现隐藏类 (红色部分)。

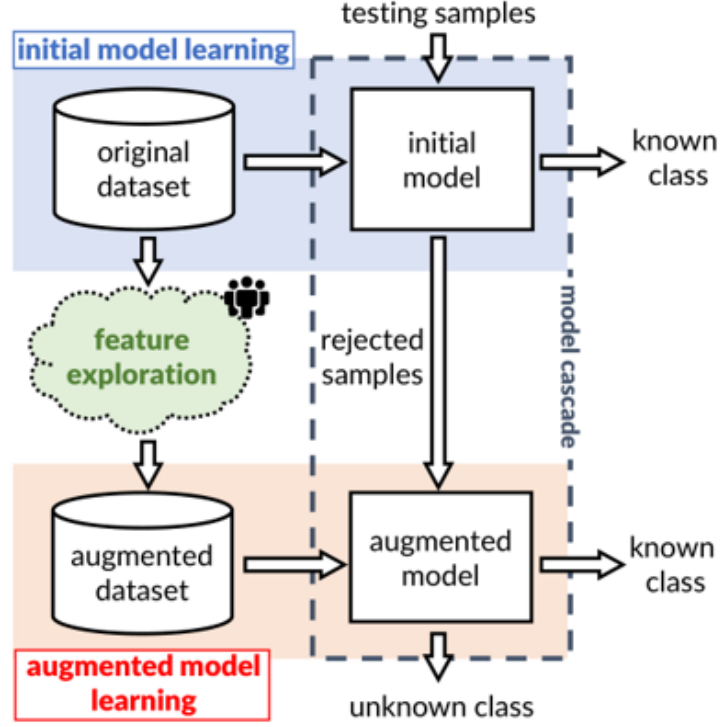


图 4: ExML 的总体过程 (图片来源于原论文^[5])

3.2 问题表述

(1) 训练集: $\hat{D}_t = \{(\hat{x}_i), \hat{y}_i\}_{i=1}^m$, 其中 $\hat{x}_i \in \hat{\mathcal{X}} \subseteq \mathbb{R}^d$ 是来自自己观测特征空间的实例，即数据集中的实例。 $\hat{y}_i \in \hat{\mathcal{Y}}$ 来自有 N 个已知类别的不完备标签空间。为了简便起见，只讨论二分类。

(2) 候选特征和代价预算: 令 $\mathcal{C} = \{c_1, \dots, c_k\}$ 表示候选特征的集合，其中的每个值在购买之前都是未知的。例如之前的诊断案例中，去看病的人花钱做了 CT。此外，购买任意的特征都需要付出代价。这和主动学习不同之处就是一个买样本一个买特征。对于学习器而言，需要在低于预算 B 的同时获取 K 个最有信息的特征。这里，每一次购买的代价都是 1，且设置 $K = 1$ 。

(3) 测试阶段: 假设学习器找到了最好的特征 c_i ，此特征将用于测试样本的增强: $\mathcal{X}_i = (\hat{\mathcal{X}} \cup \mathcal{X}^i) \subseteq \mathbb{R}^{d+1}$ ，其中 \mathcal{X}^i 是 c_i 对应的特征空间。简单说来，就是找一系列特征，加到原始的特征空间。此时，已有模型就需要预测增强测试样本的标签。

3.3 拒绝模型

首先在原始数据集上训练一个初始模型，具有识别低置信度样本的能力。这些低置信度样本可能来自自己知的或隐藏类 (不考虑高置信度的可疑类)，因此它们只被检测为可疑的，需要进一步分析。

论文通过拒绝技术来实现这一目标，在这种技术中，被学习的模型将避免预测最大条件概率低于给定值 $1-\theta$ 的样本。 θ 越小，被拒率越高，但对已知类的预测精度越高。

具体的，习得函数 $f = (h, g)$ ，其中 $h: \hat{\mathcal{X}} \mapsto \mathbb{R}$ 是一个已知类的预测函数， $g: \hat{\mathcal{X}} \mapsto \mathbb{R}$ 是一个是否

拒绝为未知的未知类的阀函数。如果 $g(\hat{x}) < 0$ 就接受，否则就预测为 $\text{sign}(h(\hat{x}))$ ，这种拒绝模型可以通过优化以下目标来训练：

$$\min_f \mathbb{E}_{(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \sim \hat{D}} [\ell_{0/1}(f, \hat{\mathbf{x}}, \hat{\mathbf{y}}; \theta)], \quad (1)$$

其中， $\ell_{0/1}(f, \hat{\mathbf{x}}, \hat{\mathbf{y}}; \theta) = \mathbb{I}_{\hat{\mathbf{y}} \cdot h(\hat{\mathbf{x}}) < 0} \cdot \mathbb{I}_{g(\hat{\mathbf{x}}) > 0} + \theta \cdot \mathbb{I}_{g(\hat{\mathbf{x}}) \leq 0}$ 是一个 0-1 损失函数， $\theta \in (0, 0.5)$ ， \hat{D} 数据和标签的混合 $\hat{\mathcal{X}} \times \hat{\mathcal{Y}}$ 。

θ 越小拒绝能力越强，其被预测为已知类的结果就越可信。不过这个损失函数是一个非凸优化，有一定的困难。Cortes, C 等人提出了一个替代的函数^[6]： $\ell_{\text{surr}} := \ell_{\text{surr}}(f, \hat{\mathbf{x}}, \hat{\mathbf{y}}; \theta)$ ，即：

$$\ell_{\text{surr}} = \max \left\{ 1 + \frac{1}{2}(g(\hat{\mathbf{x}}) - \hat{\mathbf{y}} \cdot h(\hat{\mathbf{x}})), \theta \left(1 - \frac{g(\hat{\mathbf{x}})}{1 - 2\theta} \right), 0 \right\} \quad (2)$$

其中， $h(\hat{\mathbf{x}})$ 是预测函数， $g(\hat{\mathbf{x}})$ 是阀函数。

在这里，由于分布未知，不能够直接计算风险，这里的风险可以理解是：将未知的未知类当作已知类的风险评估。自然，风险是越小越好，这是模型需要优化的，论文里的经验风险为：

$$\min_{f \in \mathbb{H} \times \mathbb{H}} \frac{1}{m} \sum_{i=1}^m \ell_{\text{surr}}(f, \hat{\mathbf{x}}_i, \hat{\mathbf{y}}_i; \theta) + C_h \|h\|_{\mathbb{H}}^2 + C_g \|g\|_{\mathbb{H}}^2 \quad (3)$$

其中， C_h 和 C_g 是正则参数， \mathbb{H} 是由核 $K : \hat{\mathcal{X}} \times \hat{\mathcal{X}} \mapsto \mathbb{R}$ 推出的再生核希尔伯特空间（Reproducing Kernel Hilbert Space, RKHS），简单的说就是核函数构成的空间，有线性空间的数乘、加减操作，也有内积操作等。

得益于 Schölkopf, B 等人的工作^[7]，这里 $h(\hat{\mathbf{x}}) = \sum_{i=1}^m u_i K(\hat{\mathbf{x}}, \hat{\mathbf{x}}_i)$ ， $g(\hat{\mathbf{x}}) = \sum_{i=1}^m w_i K(\hat{\mathbf{x}}, \hat{\mathbf{x}}_i)$ ，其中 u_i 和 w_i 是需要学习的参数。至此，式（3）可以使用二次编程来解决 (quadratic programming)。

3.4 特征探索

在该论文的设置中，学习者需要从 K 个候选中选择最好的特征来重训练，如图 5 所示：

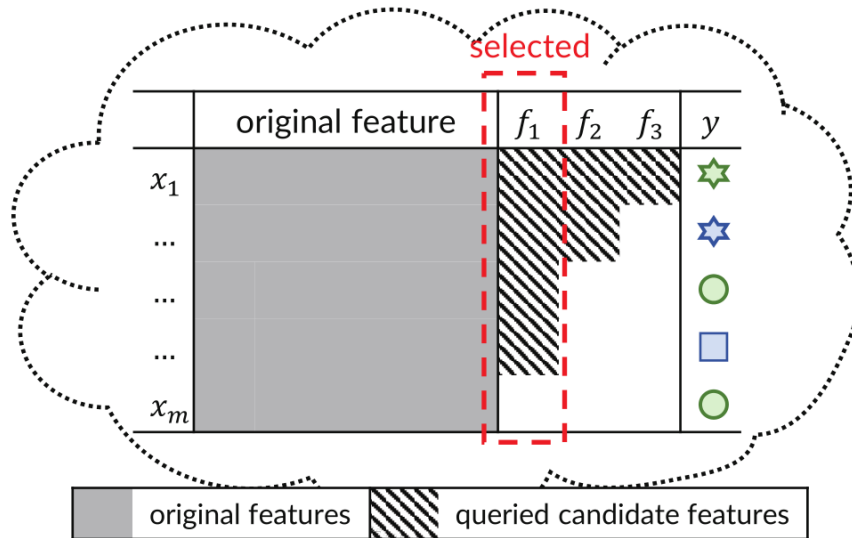


图 5: 特征探索（图片来源于原论文^[5]）

这里的重点是，传统的特征选择不适合该论文中问题的设置，因为候选特征在查询之前是不可知的。所以，该论文提出了特征探索，从而在无需提前提供特征值且低于预设代价的情况下找到最优信息的特征。这里存在两个问题：（1）怎样度量候选特征的质量？（2）怎么有效分配预算的同时又找到最好？

令 \mathcal{D}_i 表示覆盖 $\mathcal{Y}_i \times \hat{\mathcal{Y}}$ 的数据分布，其中 \mathcal{X}_i 是增强特征空间的第 i 个候选特征。该论文使用 \mathcal{D}_i 上的贝叶斯风险来作为特征质量度量：

$$R_i^* = R_i(f_i^*) = \min_f \mathbb{E}_{(x, \hat{y}) \sim \mathcal{D}_i} [\ell_{0/1}(f, x, \hat{y}; \theta)] \quad (4)$$

其中 $R_i(f)$ 是给定 f 下的预期 0/1 风险， f_i^* 是使得 $R_i(f)$ 最小化的函数。

其中贝叶斯风险的近似为：

$$\hat{R}_{D_i} = \hat{R}_i(\hat{f}_i) = \sum_{j=1}^{n_i} \ell_{0/1}(\hat{f}_i, \mathbf{x}_j, \hat{y}_j; \theta) \quad (5)$$

其中 $D_i = \{(x_j, \hat{y}_j)\}_{j=1}^{n_i} \sim \mathcal{D}_i$, $x_j \in \mathcal{X}_i$, $\hat{y}_j \in \hat{\mathcal{Y}}$, \hat{f}_i 是在增强数据集 D_i 上优化公式 (3) 得到的。

根据公式 (5)，该论文给出了两个预算分布策略来寻找最好的候选特征。其中，预算分布策略是指：在不失一般性的前提下，假设特征按照其价值排序，即 $R_1^* \leq \dots \leq R_K^*$ 。我们的目标就是在给定预算下找到最好的特征，并基于增强数据集训练好的更一般性的模型。以下给出具体策略：

(1) 均匀分配：对于任意的候选特征 $c_i, i \in [K]$ ，每一个增强数据集 D_i 的预算分配为 $\lfloor B/K \rfloor$ 。这个策略将选择最大概率的特征作为最好的特征。

(2) 中位消除 (Median Elimination)^[8]：对于每一次迭代，首先去除一半错误的候选特征，再从余下的里面选一个最好的。具体的，评估过程将进行 $T = \lceil \log_2 K \rceil$ 轮，每一轮每个特征分配的预算为 $\lfloor B/T \rfloor$ 。具体的伪代码如下：

Procedure 1 Median Elimination for Feature Exploration.

Input: Feature exploration budget B , original dataset $\hat{D}_{tr} = \{(\hat{\mathbf{x}}_i, \hat{\mathbf{y}}_i)\}_{i=1}^m$, candidate feature pool $\mathcal{C} \{c_1, \dots, c_K\}$, threshold $\theta \in (0, 1)$.

Output: Selected feature $c_{i_s} \in \mathcal{C}$ and corresponding augmented model \hat{f}_{i_s} .

Initialize: dataset $D_i = \emptyset$ for each feature $c_{i_s} \in \mathcal{C}$, set of active features $\mathcal{C}_1 = \mathcal{C}$, $T = \lceil \log_2 K \rceil$.

for $t = 1, \dots, T$ **do**

1. Randomly select $n_t = \lfloor B / (T |\mathcal{C}_t|) \rfloor$ samples from \hat{D}_{tr} and query active features $c_i \in \mathcal{C}_t$;
2. Update D_i with selected samples and train a model \hat{f}_i on D_i by ERM(2), for all $c_i \in \mathcal{C}_t$;
3. Compute \hat{R}_{D_i} according to (5), for all $c_i \in \mathcal{C}_t$;
4. Update \mathcal{C}_{t+1} as half of features in \mathcal{C}_t with lower \hat{R}_{D_i} ;

end

3.5 模型级联

在完成特征探索之后，学习模型将在增强的数据上重新进行训练。考虑到增强模型可能并不总是比初始模型要更好，特别是当预算不够或者候选特征的信息不充分时。该论文提出了模型级联机制，该机制将增强模型与初始模型进行级联。更具体的，初始模型接受高置信度的预测，其余低置信度的可疑样本传递到下一层增强模型进行特征挖掘，然后增强模型接受那些高置信度的增强样本，其余低置信度样本继续到下一层进行进一步增强。

该方法可以被视为一种用于识别隐藏类实例的逐层处理，并且可以随时停止该增强过程，直到人们发现仍然存在某些可疑的隐藏样本。为了简单起见，该论文只实现了一个两层的架构体系。

4 复现细节

4.1 与已有开源代码对比

该论文作者公开了他的代码，他的代码所实现的程序语言是 Matlab。本复现报告以该论文的算法为基础，使用 Python 语言中的 Numpy 等库，重新对 ExML 算法进行编程，并复现了该算法过程和运行结果。

4.2 实验细节

本次实验所使用的数据集是 Mfeat 数据集^[9]，这是一个多视图数据集 1，包含 2000 个样本和 6 个不同方法提取的特征视图。它们的语义信息和统计信息为：

- Fac: 剖面相关系数，216-维；
- Pix: 在 2×3 窗口的像素平均值，240 维；
- Kar: Karhunen-Love 系数，64-维；
- Zer: Zernike 时刻，47-维；
- Fou: 形状特征的傅里叶系数，76-维；
- Mor: 形态特征，6-维；

按照计算机领域的知识，可将以上特征按照质量大小排序为：Fac > Pix > Kar > Zer > Fou > Mor。

在训练阶段，随机抽取 600 个样本，形成有标记的训练数据集。重复执行 10 次，生成不同的数据集配置。由于 Mfeat 是一个多类数据集，所以需要随机采样其中 5 个配置，将其转换为二分类任务，其中每个已知类和隐藏类加起来共包含三个原始类。来自隐藏类的样本会被随机错误地标记为已知类之一。总共有 50 个随机配置用于训练。关于候选特征，六个特征中的每一个都会作为一次原始特征，然后其余的五个特征都在保留在候选特征集中进行准备。在训练之前，将所有特征数据归一化到 $[0, 1]$ 范围中。同时在包含 1200 个样本的测试数据集上进行消融实验，评估所有的不同方法，这些不同方法是指带有拒绝模型的传统 SL 算法、使用均匀分配算法进行特征探索同时带有模型级联的 ExML 算法、使用中位消除算法进行特征探索但不使用模型级联的 ExML 算法、使用中位消除算法进行特征探索同时带有模型级联的 ExML 算法。

分别用准确率和召回率评价实验结果。准确率是指在 50 个配置的测试数据集上的预测正确率的平均值和标准差，其中隐藏类的真实标签是可被观察到的。召回率是指在特征探索时，所找到的增强特征属于候选特征集中质量最高的两个特征，占所有的比例。

4.3 界面分析与使用说明

运行项目根目录下的 main.py 文件，即可运行程序。在进行机器学习的过程中，程序会输出每轮迭代后各个算法的分类准确率、召回率等评价结果，并保存在文件中。

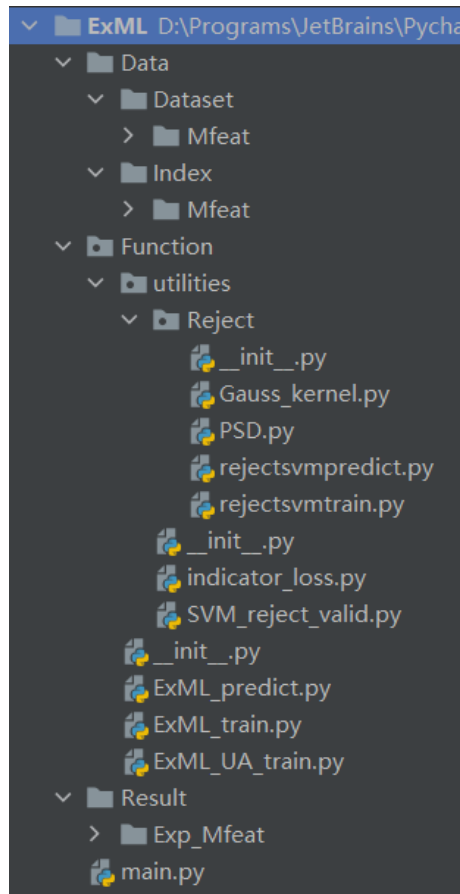


图 6: 项目文件结构

5 实验结果分析

表 1: 在 Mfeat 数据集上的实验结果

Feature	Budget	SL	EXML ^{ME} _{aug}	EXML ^{UA} _{csd}	EXML	Recall
Fac	10%	92.79 ± 1.42	73.42 ± 8.23	93.09 ± 2.59	92.49 ± 2.05	49%
	20%	92.79 ± 1.42	84.91 ± 7.31	91.55 ± 4.12	92.85 ± 3.12	47%
	30%	92.79 ± 1.42	87.39 ± 5.71	91.05 ± 3.93	92.14 ± 2.95	44%
Pix	10%	92.35 ± 2.52	72.23 ± 7.23	91.87 ± 5.07	91.55 ± 7.41	57%
	20%	92.35 ± 2.52	80.55 ± 6.13	91.84 ± 6.97	91.87 ± 6.87	54%
	30%	92.35 ± 2.52	87.57 ± 5.31	91.05 ± 4.72	91.74 ± 5.56	65%
Kar	10%	85.97 ± 3.31	73.67 ± 9.25	86.42 ± 4.48	87.85 ± 6.47	57%
	20%	85.97 ± 3.31	83.42 ± 6.44	85.28 ± 4.99	88.47 ± 8.71	58%
	30%	85.97 ± 3.31	87.00 ± 5.51	87.52 ± 4.85	84.41 ± 3.44	59%
Zer	10%	75.21 ± 8.63	69.91 ± 11.7	71.52 ± 10.2	79.74 ± 8.52	83%
	20%	75.21 ± 8.63	81.85 ± 4.01	78.96 ± 7.48	82.78 ± 6.48	81%
	30%	75.21 ± 8.63	87.01 ± 5.86	82.48 ± 6.74	87.48 ± 6.62	85%
Fou	10%	67.83 ± 8.98	70.52 ± 8.65	64.78 ± 11.6	77.25 ± 9.14	85%
	20%	67.83 ± 8.98	84.15 ± 9.47	75.84 ± 8.95	88.84 ± 7.84	87%
	30%	67.83 ± 8.98	88.53 ± 4.18	81.24 ± 6.95	89.95 ± 5.42	91%
Mor	10%	59.47 ± 15.1	70.01 ± 10.8	67.51 ± 14.1	75.48 ± 12.2	85%
	20%	59.47 ± 15.1	77.59 ± 10.0	76.74 ± 8.58	77.25 ± 8.45	84%
	30%	59.47 ± 15.1	88.55 ± 8.12	77.58 ± 8.78	87.85 ± 8.01	91%

根据实验结果可知，本次实验成功复现了 ExML 算法，得到了较好的结果。同时分析表格可知，当原有的特征具有较高的质量时，SL 可以获得较好的性能，无需探索新的特征。但是在提供信息不丰富的原始特征的情况下，SL 会严重退化，而 ExML 的效果会更加好。同时，使用中位消去方法去进行特征探索的 ExML 要略好使用均匀分配的方法的 ExML。还有，在低质量初始特征及有限预算的

条件下,不使用级联模型的 ExML 算法的效果比不上使用级联模型的 ExML 算法,有时甚至不能保证效果好于传统的 SL 算法。因此这些结果也验证了模型级联组件在整个模型中的有效性。

6 总结与展望

在该论文中,作者提出了除了学习算法的选择不充分和缺乏足够的训练样本外,未知的未知数也可能导致机器学习模型的训练失败。由于特征信息不足,训练数据集中的一些样本属于未知的隐藏类,但他们会被模型错误地视为已知类。

为了解决这个问题,本论文提出了探索性机器学习 (ExML),从而支持机器学习模型通过探索更多的特征来检查和强化训练数据集,以发现潜在的隐藏类。根据这一思想,本论文设计了一种包含三个步骤的方法:拒绝模型、特征探索和模型级联。同时利用消融实验,验证了该算法的合理性和有效性。

未来可以考虑提高特征探索的效率,从而在更低的代价下,探索到更高质量的特征,从而强化样本数据。同时,也可以考虑增加代价成本,提高模型级联体系的层数,从而可以更加精准地对未知的隐藏类数据进行分类。

参考文献

- [1] SCHEIRER W J, de REZENDE ROCHA A, SAPKOTA A, et al. Toward Open Set Recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 35(7): 1757-1772.
- [2] ATTENBERG J, IPEIROTIS P., PROVOST F. Beat the Machine: Challenging Humans to Find a Predictive Model's Unknown Unknowns[J]. ACM Journal of Data and Information Quality, 2015, 6(1): 1-17.
- [3] SETTLES B. Active Learning[J]. Synthesis lectures on artificial intelligence and machine learning, 2012, 6(1): 1-114.
- [4] YUAN M, WEGKAMP M H. Classification Methods with Reject Option Based on Convex Risk Minimization[J]. Synthesis lectures on artificial intelligence and machine learning, 2010, 11(1).
- [5] ZHAO P, ZHANG Y J, ZHOU Z H. Exploratory Machine Learning with Unknown Unknowns[J]. 10.48550/arXiv.2002.01605, 2020.
- [6] CORTES C, DESALVO G, MOHRI M. Learning with Rejection[J]., 2016: 67-82.
- [7] *Schölkopf B*, SMOLA A J. Learning with Kernels: support vector machines, regularization, optimization, and beyond[J]. Adaptive computation and machine learning series, 2018.
- [8] EVEN-DAR E, MANNOR S, MANSOUR Y. Action Elimination and Stopping Conditions for the Multi-Armed Bandit and Reinforcement Learning Problems[J]. Journal of Machine Learning Research, 2006, 7: 1079-1105.
- [9] Van BREUKELEN M, DUIN R P. W, TAX D M J, et al. Handwritten digit recognition by combined classifiers[J]. Kybernetika, 1998: 381-386.