

基于卷积神经网络的软件缺陷预测研究

张晴

(曲阜师范大学信息科学与工程学院, 山东 日照 276826)

摘要: 软件缺陷预测技术能够帮助软件维护人员提前检测到隐藏在应用软件中的各类缺陷并合理分配测试资源。提出了一种基于卷积神经网络的软件缺陷预测方法, 为了验证该方法的预测效果, 我们在 PROMISE 数据集的 10 个软件项目上进行了实证研究, 并利用 Recall、F1、mcc、pf、gm 和 AUC 6 个评价指标进行验证分析, 实验结果证明了该方法的优越性。

关键词: 软件缺陷预测; 软件质量保证; 卷积神经网络; 深度学习

DOI:10.16184/j.cnki.comprg.2020.09.005

1 概述

目前, 人工智能技术已经成为信息技术发展的优先战略目标和前沿方向, 各种应用软件在人工智能技术的应用中发挥着重要作用并不断推动着信息产业的迅速发展。由于软件规模的不断扩大和系统复杂性的持续增加, 其内部隐藏的各种缺陷也在急剧增加, 严重威胁了软件的质量和安全性。软件缺陷预测技术作为一项关键的软件质量保证技术, 通过挖掘软件缺陷跟踪系统或软件历史仓库, 首先设计出各种内在度量元, 使其能够衡量软件质量好坏, 其次利用数据挖掘、数理统计或机器学习等方法, 构建出相应的软件缺陷预测模型, 从而提前发现并预测潜藏在软件内部的缺陷^[1]。软件缺陷预测技术能够为软件开发或测试人员调配软件检测优先级, 合理分配软件测试资源, 节约大量检测时间和人力物力, 并极大地提高软件产品的质量^[2]。

先前的研究对软件缺陷的预测主要是利用传统的机器学习方法来进行的, 例如支持向量机 (Support Vector Machine)、朴素贝叶斯 (Naive Bayes)、逻辑回归 (Logistic Regression)、K-近邻 (k-Nearest Neighbor)。然而许多机器学习方法都有自身的局限性, 导致预测效果不够理想, 如支持向量机对参数和核函数的选择比较敏感; 样本不平衡的数据集经过 K-近邻的预测, 其效果较差, 且计算量较大; 逻辑回归一般准确度不高, 容易欠拟合, 且要求数据线性可分。Lessmann 等人^[3]在 NASA 数据集的软件项目上使用 6 类机器学习方法进行软件缺陷预测, 包括统计方法、支持向量机、K-近邻、决策树、传统的神经网络以及集成学习方法, 并使用 AUC 作为评价指标, 他们发现最优的大部分机器学习方法之间预测效果并不存在显著的差异。Yang 等人^[4]利用决策

树和集成学习技术构建了一种动态的软件缺陷预测模型, 该模型在内层利用基于决策树技术的 bagging 集成学习方法构造了一种随机深林模型, 并在外层运用随机抽样和 stacking 方式的方法来训练这些随机森林模型, 并取得了一定的预测效果。目前, 深度学习技术^[5]是人工智能领域研究的前沿方向, 并在很多领域得到较好的应用, 包括语音识别、目标检测、自然语言处理和图像识别。因此, 采用一种主流的深度学习技术—卷积神经网络 (Convolutional Neural Network, CNN) 来进行软件缺陷预测。

2 方法

卷积神经网络是一种具有卷积计算能力且包含多个网络层次的前馈神经网络, 它具有强大的特征提取能力, 能够将低级的原始缺陷特征映射成高级的抽象深度语义特征, 并能够按其阶层结构对输入的软件缺陷特征进行平移不变分类。

卷积神经网络中具有隐含层内卷积核参数共享和层间连接的稀疏性质, 其性质主要作用于网络使其能够以较小的计算量格点化特征。

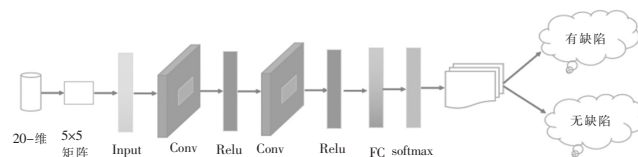


图1 用于软件缺陷预测的卷积神经网络结构图

由于文中使用的 PROMISE 中的 10 个软件项目均具有 20 维特征, 因此, 卷积神经网络的网络结构针对 20

作者简介: 张晴 (1999-), 女, 本科, 研究方向: 人工智能、软件工程。

维特征为例，包括输入层、卷积层、Relu 非线性激励、全连接层和输出层，网络结构如图 1 所示。

卷积层包含多个卷积核，能够对输入的缺陷特征进行特征提取，其内部的每个元素都对应一个权重系数和一个偏置量。矩阵元素乘法求和以及叠加偏差量运算是卷积核在感受野内对输入的缺陷特征进行的，卷积过程如公式（1）所示：

$$C^{l+1}(i,j)=[C^l \otimes W^{l+1}](i,j)+c, \quad (1)$$

其中 C^l 和 C^{l+1} 依次代表第 $l+1$ 层的卷积输入和输出，即特征图， W^{l+1} 表示网络权重， c 表示偏置量。

Relu 非线性激励函数：卷积层中的非线性激励函数可以对复杂的特征进行非线性转化，从而增强了特征的表达能力，其过程如公式（2）所示：

$$D_{i,j,k}^l = f(C_{i,j,k}^l), \quad (2)$$

其中 $f(\cdot)$ 表示非线性激励函数。

先将每个软件项目的 20 维特征映射成 25 维特征，并 reshape 为 5×5 的矩阵。然后进行连续的两轮卷积操作，卷积核大小分别为 3×3 、 5×5 ，并分别具有 32 和 16 个卷积核，每轮卷积操作中均具有 Relu 非线性激励函数。接下来全连接层则对提取的特征进行非线性组合，以得到 1 维特征的形式输出。最后通过 softmax 层对前面提取的缺陷特征进行分类，从而得到缺陷特征和无缺陷的特征。

采用交叉熵损失函数来训练卷积神经网络的网络结构，学习率为 0.001，批处理数量为 32，训练迭代次数为 1000。

3 结果与分析

3.1 数据集

在 PROMISE 数据集的 10 个软件项目上进行软件缺陷预测，这些软件项目均是广泛使用且开源的基线数据集，如表 1 所示。10 个项目的特征数量均为 20 维，实例数量最少为 293，最多为 803，缺陷率由 10.0% 到 48.2% 不等。由于文中的 10 个项目均存在类别不平衡的问题，在使用卷积神经网络进行缺陷预测之前，对这 10 个项目均进行了类别平衡处理。

表 1 PROMISE 数据集的 10 个软件项目的信息统计

数据集	特征数量	实例数量	有缺陷实例数量	有缺陷实例数量	缺陷率 (%)
ant-1.5	20	293	32	261	10.9
ant-1.6	20	351	92	259	26.2

数据集	特征数量	实例数量	有缺陷实例数量	有缺陷实例数量	缺陷率 (%)
camel-1.2	20	608	216	392	35.5
camel-1.4	20	872	145	727	16.6
poi-2.0	20	314	37	277	11.8
prop-6	20	660	66	594	10.0
xalan-2.4	20	723	110	613	15.2
xalan-2.5	20	803	387	416	48.2
ivy-2.0	20	352	40	312	11.4
jedit-4.1	20	312	79	233	25.3

3.2 评价指标

使用了 6 个评价指标对实验结果进行全面的统计和分析，分别是召回率（Recall）、F1、MCC（Matthews correlation coefficient）、误报率（pf）、gm（G-measure）和 AUC。除了误报率外，其他指标均是越大越好。

3.3 实验结果

使用卷积神经网络在 10 个软件项目上进行实验。Recall、F1、MCC、pf 和 gm 如表 2 所示，AUC 如图 2 所示。从表 2 中，可以看到，Recall、F1、MCC、pf 和 gm 5 个指标的平均值分别为 0.6742、0.7082、0.4374、0.2396 和 0.7113。重点关注了 F1 指标，除了在 camel-1.4 项目的值 0.5470 和 ant-1.6 项目的值 0.6800，在其余 8 个项目上的 F1 值均大于 0.7，在 10 个项目上的平均值 0.7082 也大于 0.7。从图 2 中，10 个项目的 AUC 值从 0.76 到 0.97 变化。显然，使用卷积神经网络获得的软件缺陷预测实验结果还是比较理想的。

表 2 卷积神经网络在 10 个项目上的 5 个评价指标统计结果

指标 数据集	Recall	F1	mcc	pf	gm
ant-1.5	0.6757	0.7246	0.4113	0.2593	0.7067
ant-1.6	0.6071	0.6800	0.4891	0.1389	0.7122
camel-1.2	0.5333	0.5470	0.1664	0.3676	0.5786
camel-1.4	0.6829	0.7241	0.5001	0.1880	0.7419
poi-2.0	0.7000	0.7368	0.3578	0.3333	0.6829
prop-6	0.6373	0.7263	0.5131	0.1333	0.7345
xalan-2.4	0.8021	0.7662	0.5127	0.2917	0.7523
xalan-2.5	0.6984	0.7154	0.4531	0.2462	0.7251
ivy-2.0	0.7273	0.7500	0.5015	0.2258	0.7500
jedit-4.1	0.6774	0.7119	0.4687	0.2121	0.7285
Average	0.6742	0.7082	0.4374	0.2396	0.7113



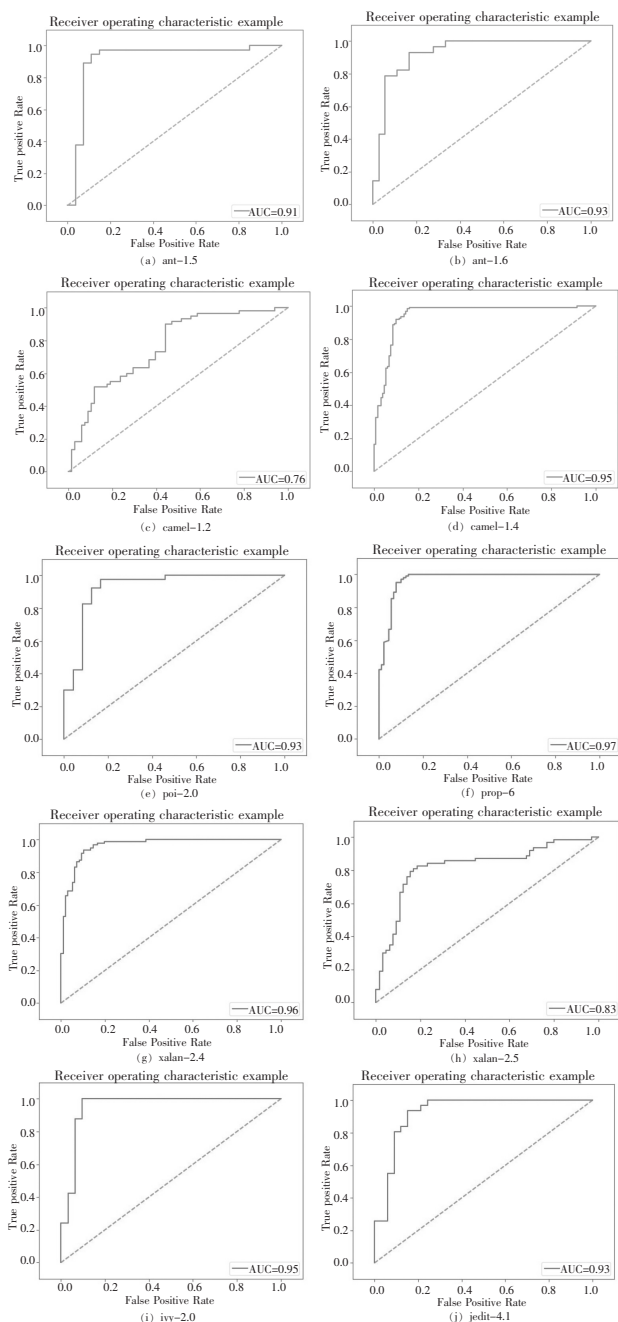


图2 卷积神经网络在10个项目上的AUC值和ROC (receiver operating characteristic curve) 曲线

(上接第11页)

4 结语

在网络技术飞速发展的当今社会中, 各类计算机编程语言层出不穷, 其中Python因具备简洁、易读、易维护等优点, 使其在计算机软件编程中具备独特的优势, 并成为一门受到人们欢迎的计算机程序设计语言, 被广泛应用于Web开发、视频游戏开发、图形用户界面等方面。

4 结语

使用卷积神经网络对软件项目进行缺陷预测, 并取得了较好的预测效果。这是由于卷积神经网络能够利用卷积层对隐藏在软件缺陷数据的特征进行深度挖掘和提取, 并将原始的缺陷特征转化为高级的抽象深度语义特征, 这样的深度语义特征对软件缺陷具有更强的判别能力, 而这种强大的判别能力则是传统的机器学习方法所不具备的。这也是深度学习技术在软件工程领域的成功应用。目前, 人类社会已经进入人工智能时代, 坚信像深度学习这样典型的人工智能技术将会更快更强地推动人类社会的发展。

参考文献

- [1] 李勇, 黄志球, 王勇, 等. 数据驱动的软件缺陷预测研究综述 [J]. 电子学报, 2017, 45 (4): 982-988.
- [2] 王青, 伍书剑, 李明树. 软件缺陷预测技术 [J]. 软件学报, 2008, 19 (7): 1565-1580.
- [3] Lessmann S, Baesens B, Mues C, et al. Benchmarking classification models for software defect prediction: A proposed framework and novel findings [J]. IEEE Transactions on Software Engineering, 2008, 34 (4): 485-496.
- [4] Yang X, Lo D, Xia X, et al. TLEL: A two-layer ensemble learning approach for just-in-time defect prediction [J]. Information and Software Technology, 2017, 87 (7): 206-220.
- [5] Zhong J, Yang B, Li Y, et al. Image Fusion and Super-Resolution with Convolutional Neural Network [C]. Singapore: Chinese Conference on Pattern Recognition, 2016, 78-88.

参考文献

- [1] 胡正雨, 刘文锐. Python的计算机软件应用技术研究 [J]. 计算机产品与流通, 2020, (07): 39.
- [2] 颜涛. 计算机软件技术在大数据时代的应用研究 [J]. 计算机产品与流通, 2020, (06): 25.
- [3] 王学庆. 基于Python的计算机软件应用技术研究 [J]. 信息系统工程, 2019, (04): 149.