

浙江大学

工程管理硕士学位论文 选题报告



论文题目 基于集成学习的软件质量
缺陷预测与控制方法研究

学 号 22464300

姓 名 熊依琳

指导老师 刘兴高 教授

学科(专业) 工程管理

所在学院 工程师学院

日 期 2025 年 03 月 10 日

目 录

1 选题背景及来源，包括研究对象	1
1.1 拟研究的对象及来源	1
1.2 针对所研究对象，拟解决工程中什么样的管理问题或通过管理理论方法解决什么样的技术问题	1
1.3 研究的意义（应用性价值），结合文献阅读和实际需求进行分析	2
2 与自己目前从事的工作的相关性	3
2.1 所研究的对象及其问题与自己工作岗位任务的关系	3
2.2 已完成、正在进行或即将进行的相关研究工作	3
3 预期的研究成果及其成果的可应用性	4
4 拟采取的论文形式及其理由	5
5 查阅的中文文献及英文文献目录	6

1 选题背景及来源，包括研究对象

1.1 拟研究的对象及来源

随着软件规模日益扩大、功能复杂度不断提升以及开发交付周期持续压缩，软件质量缺陷的精准预测已经成为软件工程管理中的核心诉求之一。缺陷的频繁出现不仅会影响软件产品的稳定性与可靠性，更会直接导致开发资源的浪费、项目延期以及用户满意度的下降，从而给企业带来巨大的经济和声誉风险。

本研究以 C 公司金融数据平台（Data Service Platform, DSP）项目为研究对象。该平台主要面向大规模金融数据处理、高频交易和实时风险管理等场景，具有以下特点：快速迭代开发、高度并行的开发任务以及跨团队协同工作。近年来，DSP 项目的软件质量缺陷频繁出现，传统基于专家经验或简单统计模型的预测方法在应对代码耦合度高、多模块并行、团队成员流动等复杂因素时表现不佳，导致缺陷定位不准确、预测规模偏差大，从而对项目交付进度和整体质量产生负面影响。

为了从根本上提升软件质量管理水平并降低缺陷带来的成本和风险，本研究拟引入当前主流的机器学习模型（如随机森林、XGBoost、支持向量机 SVM 等）和深度学习方法（如卷积神经网络 CNN、长短期记忆网络 LSTM），以充分挖掘软件项目中的结构化与非结构化特征。同时，结合集成学习思想（包括 Bagging、Boosting 和 Stacking 等策略），通过多模型融合与动态加权机制，构建适用于大型企业项目环境的软件质量缺陷精准预测与控制的综合解决方案。

为了进一步提升跨项目的适应性与泛化能力，研究将引入迁移学习机制，并针对不同项目间数据分布与上下文差异，探索基于特征重构与模型微调的迁移优化策略，从而增强预测框架的可移植性与鲁棒性。

1.2 针对所研究对象，拟解决工程中什么样的管理问题或通过管理理论方法解决什么样的技术问题

本研究拟重点解决以下关键管理与技术难题：

首先，传统的缺陷预测方法无法有效捕捉引起软件质量缺陷的多种复杂非线性因素，例如代码耦合性、变更频率、技术选型、人员流动和经验差异等，导致

预测模型与实际情况严重脱节。

其次，单一预测模型通常泛化性不足，面对不同软件项目的业务特性、团队结构、开发方法和技术环境的差异时，难以实现跨项目的有效预测和迁移，导致跨项目的预测准确性显著降低。

第三，目前缺陷预测方法与实际软件开发及质量控制流程的结合度低，预测结果难以直接转化为有效的质量控制行动，缺乏实际应用价值，难以在企业落地。

本研究拟通过引入多种主流机器学习与深度学习模型，结合 **Boosting** 或 **Stacking** 等集成策略，并配合动态加权融合机制，构建具备自适应能力的多模型集成架构。同时针对跨项目环境设计迁移优化策略，在提升预测准确性的基础上提高模型适应性与通用性。本研究将采用如 **F1-score**、准确率、缺陷修复率、误报率、平均检测时延、预警响应时间、反馈闭环周期等多项指标，对模型的预测精度、响应效率和实际应用效果进行评估，确保所构建方法具备良好的性能表现与实际落地可行性。

1.3 研究的意义（应用性价值），结合文献阅读和实际需求进行分析

在企业应用层面，本研究将构建一套高精度的软件质量缺陷预测与控制体系，并特别注重模型在真实开发环境中的落地与易用性。通过在持续集成（CI/CD）流水线中嵌入预测钩子，每次代码提交后都能自动触发缺陷检测，并将检测结果无缝对接到现有项目管理平台。当高风险模块被及时识别并可视化展示后，管理者可迅速分配责任人并采取针对性措施，借助实时预警和缺陷跟踪功能，形成高效的闭环质量管控流程。这一机制不仅提高了缺陷预防与修复的精细化程度，也显著缩短了反馈周期和减少了相关资源投入，从而降低整体项目风险、保障交付进度，并最终为企业带来切实的经济与管理效益。

在行业推广层面，本研究的成果具有广泛的行业适用性，能够快速融入不同技术栈与业务场景的软件项目管理之中，实现质量缺陷预测结果的可视化监控与风险提示。对项目团队而言，这种通用性有助于在多个项目中复用预测模型及管理流程，大幅节省人力与时间成本；对行业而言，标准化的预测输出与管理机制不仅能够提高团队协作效率，还能进一步缩短整体开发周期，为软件质量管控提

供可落地、可复制的工程管理解决方案，从而推动更广泛的软件开发与运维水平升级。

在理论贡献层面，本研究创新性地融合集成学习与迁移学习理论，构建具有动态适应能力与跨项目迁移能力的多模型融合框架，突破单一模型泛化能力弱的问题，并结合系统性的性能评价指标，提供从建模到管理应用的完整闭环路径，推动软件工程质量管理的理论深化与方法创新。

2 与自己目前从事的工作的相关性

2.1 所研究的对象及其问题与自己工作岗位任务的关系

本人目前在 C 公司担任高级软件开发工程师，主要负责 DSP 项目的开发与质量管理工作，对软件质量缺陷的产生机制和成因有深入了解，也积累了大量真实的缺陷数据与一线实践经验。实际工作中，DSP 项目常因缺陷预测不足导致延期或返工，带来额外的人力成本和经济损失；再加上项目复杂度高、团队协作频繁，传统缺陷管理方法已难以应对这一系列挑战。

针对这些痛点，本研究立足企业一线实践，旨在将机器学习模型与工程管理理论相结合，构建更高效的软件质量缺陷预测与控制框架，以实现精准缺陷预测与品质提升的双重目标。通过为企业的质量管理决策提供实证化、可操作的支持方案，本研究不仅能有效缓解 DSP 项目的缺陷管理难题，也可为其他类似项目提供可借鉴的思路和方法。

2.2 已完成、正在进行或即将进行的相关研究工作

已完成的研究内容：

1) 缺陷数据统计与问题归纳：收集了 C 公司近几年在 DSP 项目中的质量缺陷数据，进行了基础统计与关联分析，明确了缺陷产生的主要特征与风险因素。

正在进行的研究内容：

1) 文献调研与方法体系梳理：系统回顾当前学术界和业界在缺陷预测、迁移学习与集成学习方向的最新成果，识别主流方法的适用条件、优势局限和工程落地难点。

2) 算法选型与模型集成策略研究：深入分析随机森林、XGBoost、支持向

量机（SVM）、卷积神经网络（CNN）和长短期记忆网络（LSTM）等在缺陷预测中的应用情况，择优选取 Bagging、Boosting、Stacking 等典型集成策略，构建适合复杂企业环境的软件质量缺陷预测框架。

3) 模型泛化与适应性优化研究：分析预测模型在不同项目环境下的泛化能力，设计适合跨项目迁移的集成与迁移学习策略，提升模型跨项目的适应性。

即将开展的研究内容：

1) 模型实证与优化研究：基于实际软件质量缺陷数据，开展大规模的实证研究，进行模型验证与性能优化，使用多项指标（如 F1-score、缺陷修复率、响应延迟等）评估模型性能，确定最优的预测与控制策略。

2) 与项目管理平台的深度集成：将预测结果与企业现有开发流程及质量控制机制对接，通过实时可视化和即时告警等方式，让预测结果快速转化为质量管控措施，以期实现品质提升并切实提升工程管理效率。

3 预期的研究成果及其成果的可应用性

本研究预期建立一套面向复杂软件开发环境的软件质量缺陷预测与控制方法体系，能够在实际工程流程中发挥高效的风险预警与质量保障作用。通过集成随机森林、XGBoost、CNN、LSTM 等多种机器学习与深度学习模型，并择优选取 Bagging、Boosting、Stacking 等策略形成多层次的融合架构，力求在预测精度、模型稳定性与跨项目适应性等方面取得突破。

在工程实践中，该方法将在企业 CI/CD 流水线中以自动化钩子的形式集成，实现代码提交后的实时缺陷预测与风险提示。预测结果可视化集成到管理平台，通过仪表盘呈现模块级别的风险趋势，并结合即时告警与责任分配机制，辅助管理人员精准定位问题区域，推动缺陷响应与修复的流程闭环。

研究成果可直接服务于企业质量管控的多个关键场景：项目早期的高风险模块识别、测试资源调度优化、代码审查方向聚焦、缺陷复发概率研判等。通过提前介入、精准定位、智能提示，有助于显著减少后期返工与质量事故，提升研发节奏与资源利用率。

此外，研究过程中形成的预测特征体系、模型融合机制与迁移适应策略，具

具备良好的可扩展性与通用性。在同类项目或其他软件工程场景中具备可移植推广价值，为企业构建跨系统、跨产品线的质量预测标准体系奠定基础。

综上，本研究不仅将为企业提供一个“可集成、可追踪、可优化”的缺陷预测与控制机制，也将以实证为基础反哺学术研究，丰富工程管理与软件质量缺陷预测交叉领域的理论成果与实践模型。

4 拟采取的论文形式及其理由

本研究计划采取“问题驱动—模型构建—实验验证”的专题研究形式，具体包括：

1) 问题驱动：基于 C 公司 DSP 项目的真实缺陷案例，深度挖掘缺陷预测不准确所带来的管理与品质问题，明确当下缺陷管理所面临的痛点。从工程管理角度分析缺陷预测不准造成的延误、成本与质量风险，提出需求。

2) 模型构建：运用集成学习与迁移学习的相关理论与方法，设计适合大型企业级项目的软件质量缺陷预测框架，结合工程管理实践的需求进行优化。针对多维特征与多算法集成的挑战，探索自适应的融合策略，确保在高复杂度环境下依然能产出高精度预测结果，并有效提高软件品质。

3) 实验验证：在真实缺陷数据集上进行大规模训练与测试，通过与企业内部的管理流程对接，监测并量化缺陷预测对工程管理效率以及项目品质的提升效果。比较传统方法与本研究方法的预测准确率、缺陷修复率、项目延期与返工次数等指标，以证明研究成果的可操作性与实际价值。

之所以选择这一研究形式，是因为：

1) 紧密契合实际工程管理需求：软件质量缺陷预测与质量管理是强实践导向的研究课题，只有通过真实项目的案例与数据分析，才能准确提炼出核心问题及其对项目品质与管理效率的影响。

2) 建模能有效揭示复杂因素的多维度作用：通过多模型融合及动态加权机制，能更全面地刻画影响缺陷产生和演化的各项因素，进而提升缺陷预测的准确性，实现对软件品质的有效提升。

3) 数据与管理实践相结合，确保成果落地：通过实证研究与管理平台的深

度集成,不仅可以进一步验证方法的可行性与有效性,还能切实提高企业的工程管理效果,将研究成果转化为企业实际的生产力。

5 查阅的中文文献及英文文献目录

- [1] Dong X, Wang J, Liang Y. A Novel Ensemble Classifier Selection Method for Software Defect Prediction[J]. Ieee Access, 2025, 13: 25578-25597.
- [2] Goyal S R. Current Trends in Class Imbalance Learning for Software Defect Prediction[J]. Ieee Access, 2025, 13: 16896-16917.
- [3] 喻皓, 张莹, 李倩, et al. 基于机器学习的软件缺陷预测研究[J]. 重庆大学学报, 2025, 48(02): 10-21.
- [4] 崔梦天, 吴克奇, S M M. 基于特征提取和 Stacking 集成学习的软件缺陷预测[J]. 计算机应用与软件, 2025, 42(01): 25-29+48.
- [5] 朱晓燕, 王文格, 王嘉寅, 张选平. 基于细粒度代码表示和特征融合的即时软件缺陷预测方法[J]. 计算机科学, 2025, 52(01): 242-249.
- [6] 赵晨阳, 刘磊, 江贺. 基于多目标优化的工作量感知即时软件缺陷预测特征构建方法[J]. 计算机科学, 2025, 52(01): 232-241.
- [7] 吴海涛, 马景悦, 高建华. 融合静态分析警告的软件缺陷预测模型及其应用研究[J]. 计算机科学与探索, 2025, 19(03): 818-834.
- [8] Licorish S A. Understanding the Effect of Agile Practice Quality on Software Product Quality[J]. Ieee Transactions on Software Engineering, 2025, 51(2): 650-662.
- [9] Mustaqeem M, Alam M, Mustajab S, et al. Comprehensive Bibliographic Survey and Forward-Looking Recommendations for Software Defect Prediction: Datasets, Validation Methodologies, Prediction Approaches, and Tools[J]. Ieee Access, 2025, 13: 866-903.
- [10] Dar A W, Farooq S U. An ensemble model for addressing class imbalance and class overlap in software defect prediction[J]. International Journal of System Assurance Engineering and Management, 2024, 15(12): 5584-5603.

-
- [11] Tang Y, Dai Q, Yang M Y, et al. Software defect prediction ensemble learning algorithm based on 2-step sparrow optimizing extreme learning machine[J]. Cluster Computing-the Journal of Networks Software Tools and Applications, 2024, 27(8): 11119-11148.
- [12] Tong H A, Zhang D L, Liu J Q, et al. MASTER: Multi-Source Transfer Weighted Ensemble Learning for Multiple Sources Cross-Project Defect Prediction[J]. Ieee Transactions on Software Engineering, 2024, 50(5): 1281-1305.
- [13] Chen J F, Xu J P, Cai S H, et al. Software Defect Prediction Approach Based on a Diversity Ensemble Combined With Neural Network[J]. Ieee Transactions on Reliability, 2024, 73(3): 1487-1501.
- [14] Al-Fraihat D, Sharrab Y, Al-Ghuwairi A R, et al. Hyperparameter Optimization for Software Bug Prediction Using Ensemble Learning[J]. Ieee Access, 2024, 12: 51869-51878.
- [15] Ali M, Mazhar T, Arif Y, et al. Software Defect Prediction Using an Intelligent Ensemble-Based Model[J]. Ieee Access, 2024, 12: 20376-20395.
- [16] Fan X, Mao J G, Lian L J, et al. Software Defect Prediction Method Based on Stable Learning[J]. Cmc-Computers Materials & Continua, 2024, 78(1): 65-84.
- [17] Khleel N a A, Nehéz K. Software defect prediction using a bidirectional LSTM network combined with oversampling techniques[J]. Cluster Computing-the Journal of Networks Software Tools and Applications, 2024, 27(3): 3615-3638.
- [18] Alqasrawi Y, Azzeh M, Elsheikh Y. Analyzing the Role of Class Rebalancing Techniques in Software Defect Prediction[J]. International Journal of Information Technology & Decision Making, 2024, 23(06): 2167-2207.
- [19] Kumar R, Chaturvedi A. Software Bug Prediction Using Reward-Based Weighted Majority Voting Ensemble Technique[J]. Ieee Transactions on Reliability, 2024, 73(1): 726-740.
- [20] Yang Z Y, Lu L, Zou Q Y. Ensemble Kernel-Mapping-Based Ranking Support

- Vector Machine for Software Defect Prediction[J]. Ieee Transactions on Reliability, 2024, 73(1): 664-679.
- [21] 王越, 李勇, 张文静. 面向可解释性的软件缺陷预测主动学习方法[J]. 现代电子技术, 2024, 47(20): 101-108.
- [22] 郭育晨, 朱晓燕. 工作量感知软件缺陷预测中偏斜分布的影响及测试评估方法[J]. 西安交通大学学报, 2024, 58(07): 203-213.
- [23] 石海鹤, 周世文, 钟林辉, 肖正兴. 改进的采样算法与无监督聚类相结合的软件缺陷预测模型[J]. 江西师范大学学报(自然科学版), 2024, 48(03): 301-310.
- [24] 谢华祥, 高建华, 黄子杰. 基于 SDL-LightGBM 集成学习的软件缺陷预测模型[J]. 计算机工程与设计, 2024, 45(03): 769-776.
- [25] 邱少健, 陆璐, 邹全义. 基于联合特征分布匹配的跨项目缺陷预测[J]. 计算机工程与设计, 2024, 45(01): 204-211.
- [26] Omondiagbe O P, Licorish S A, Macdonell S G. Improving transfer learning for software cross-project defect prediction[J]. Applied Intelligence, 2024, 54(7): 5593-5616.
- [27] Das H, Das S, Gourisaria M K, et al. Enhancing Software Fault Prediction Through Feature Selection With Spider Wasp Optimization Algorithm[J]. Ieee Access, 2024, 12: 105309-105325.
- [28] Haque R, Ali A, Mcclean S, et al. Heterogeneous Cross-Project Defect Prediction Using Encoder Networks and Transfer Learning[J]. Ieee Access, 2024, 12: 409-419.
- [29] Sotto-Mayor B, Kalech M. A Survey on Transfer Learning for Cross-Project Defect Prediction[J]. Ieee Access, 2024, 12: 93398-93425.
- [30] Dong X, Liang Y, Miyamoto S, Yamaguchi S. Ensemble learning based software defect prediction[J]. Journal of Engineering Research, 2023, 11(4): 377-391.
- [31] Khleel N a A, Nehéz K. A novel approach for software defect prediction using

- CNN and GRU based on SMOTE Tomek method[J]. Journal of Intelligent Information Systems, 2023, 60(3): 673-707.
- [32] Gong L A, Zhang H X, Zhang J X, et al. A Comprehensive Investigation of the Impact of Class Overlap on Software Defect Prediction[J]. Ieee Transactions on Software Engineering, 2023, 49(4): 2440-2458.
- [33] Tang Y, Dai Q, Yang M Y, et al. Software defect prediction ensemble learning algorithm based on adaptive variable sparrow search algorithm[J]. International Journal of Machine Learning and Cybernetics, 2023, 14(6): 1967-1987.
- [34] Abbas S, Aftab S, Khan M A, et al. Data and Ensemble Machine Learning Fusion Based Intelligent Software Defect Prediction System[J]. Cmc-Computers Materials & Continua, 2023, 75(3): 6083-6100.
- [35] Pornprasit C, Tantithamthavorn C. DeepLineDP: Towards a Deep Learning Approach for Line-Level Defect Prediction[J]. Ieee Transactions on Software Engineering, 2023, 49(1): 84-98.
- [36] Tabassum S, Minku L L, Feng D Y. Cross-Project Online Just-In-Time Software Defect Prediction[J]. Ieee Transactions on Software Engineering, 2023, 49(1): 268-287.
- [37] 徐金鹏, 郭新峰, 王瑞波, 李济洪. 基于 GAN 数据增强的软件缺陷预测聚合模型[J]. 计算机科学, 2023, 50(12): 24-31.
- [38] 常志鹏, 马生忠. 基于排序集成的软件并发缺陷动态预测算法[J]. 计算机仿真, 2023, 40(11): 423-427.
- [39] 张丽, 沈雅婷, 朱园园. 基于改进 SMOTE 的软件缺陷预测[J]. 计算机工程与设计, 2023, 44(10): 2965-2972.
- [40] 张艳梅, 植胜林, 姜淑娟, 袁冠. 类不平衡对软件缺陷预测模型稳定性和预测性能的影响分析方法[J]. 电子学报, 2023, 51(08): 2076-2087.
- [41] 李显伟, 潘伟丰, 王家乐, et al. 基于网络度量元的 Solidity 智能合约缺陷预测[J]. 计算机应用研究, 2023, 40(12): 3545-3550.
- [42] 卫梅特, 任洪敏. 基于特征优选的软件缺陷预测集成学习方法[J]. 计算机

- 仿真, 2023, 40(07): 331-336.
- [43] 田笑, 常继友, 张弛, et al. 开源软件缺陷预测方法综述[J]. 计算机研究与发展, 2023, 60(07): 1467-1488.
- [44] 唐宇, 代琪, 杨志伟, et al. 基于优化随机森林的软件缺陷预测算法研究[J]. 计算机工程与科学, 2023, 45(05): 830-839.
- [45] 李汇来, 杨斌, 于秀丽, 唐晓梅. 软件缺陷预测模型可解释性对比[J]. 计算机科学, 2023, 50(05): 21-30.
- [46] 刘旭同, 郭肇强, 刘释然, et al. 软件缺陷预测模型间的比较实验:问题、进展与挑战[J]. 软件学报, 2023, 34(02): 582-624.
- [47] Mishra A S, Rathore S S. Implicit and explicit mixture of experts models for software defect prediction[J]. Software Quality Journal, 2023, 31(4): 1331-1368.
- [48] Jiang W C, Qiu S J, Liang T C, Zhang F L. Cross-project clone consistent-defect prediction via transfer-learning method[J]. Information Sciences, 2023, 635: 138-150.
- [49] Bala Y Z, Samat P A, Sharif K Y, Manshor N. Improving Cross-Project Software Defect Prediction Method Through Transformation and Feature Selection Approach[J]. Ieee Access, 2023, 11: 2318-2326.
- [50] Zheng Z Q, Liu Y S, Zhang B, et al. Software defect prediction method based on the heterogeneous integration algorithm[J]. Journal of Intelligent & Fuzzy Systems, 2023, 45(3): 4807-4824.
- [51] Bhutapuram U S, Sadam R. With-in-project defect prediction using bootstrap aggregation based diverse ensemble learning technique[J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34(10): 8675-8691.
- [52] Ngo G, Beard R, Chandra R. Evolutionary bagging for ensemble learning[J]. Neurocomputing, 2022, 510: 1-14.
- [53] Chen J F, Wang X L, Cai S H, et al. A software defect prediction method with

- metric compensation based on feature selection and transfer learning[J]. *Frontiers of Information Technology & Electronic Engineering*, 2022, 23(5): 715-731.
- [54] Tang S Q, Huang S, Zheng C Y, et al. A Novel Cross-Project Software Defect Prediction Algorithm Based on Transfer Learning[J]. *Tsinghua Science and Technology*, 2022, 27(1): 41-57.
- [55] Goyal S, Bhatia P K. Heterogeneous stacked ensemble classifier for software defect prediction[J]. *Multimedia Tools and Applications*, 2022, 81(26): 37033-37055.
- [56] Goyal S. Handling Class-Imbalance with KNN (Neighbourhood) Under-Sampling for Software Defect Prediction[J]. *Artificial Intelligence Review*, 2022, 55(3): 2023-2064.
- [57] Mi W B, Li Y, Wen M, Chen Y R. Using active learning selection approach for cross-project software defect prediction[J]. *Connection Science*, 2022, 34(1): 1482-1499.
- [58] Jiang F, Yu X, Gong D W, Du J W. A random approximate reduct-based ensemble learning approach and its application in software defect prediction[J]. *Information Sciences*, 2022, 609: 1147-1168.
- [59] Daoud M S, Aftab S, Ahmad M, et al. Machine Learning Empowered Software Defect Prediction System[J]. *Intelligent Automation and Soft Computing*, 2022, 31(2): 1287-1300.
- [60] Pal S, Sillitti A. Cross-Project Defect Prediction: A Literature Review[J]. *Ieee Access*, 2022, 10: 118697-118717.
- [61] Bhat N A, Farooq S U. An Improved Method for Training Data Selection for Cross-Project Defect Prediction[J]. *Arabian Journal for Science and Engineering*, 2022, 47(2): 1939-1954.
- [62] 赵艳阳. 基于机器学习的跨版本软件缺陷预测关键技术研究[D]. 北京邮电大学博士学位论文, 2022.

-
- [63] Rathore S S, Kumar S. An empirical study of ensemble techniques for software fault prediction[J]. *Applied Intelligence*, 2021, 51(6): 3615-3644.
- [64] Matloob F, Ghazal T M, Taleb N, et al. Software Defect Prediction Using Ensemble Learning: A Systematic Literature Review[J]. *Ieee Access*, 2021, 9: 98754-98771.
- [65] Wickramasinghe I, Kalutarage H. Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation[J]. *Soft Computing*, 2021, 25(3): 2277-2293.
- [66] Kumar P S, Behera H S, Nayak J, Naik B. Bootstrap aggregation ensemble learning-based reliable approach for software defect prediction by using characterized code feature[J]. *Innovations in Systems and Software Engineering*, 2021, 17(4): 355-379.
- [67] Mehta S, Patnaik K S. Improved prediction of software defects using ensemble machine learning techniques[J]. *Neural Computing & Applications*, 2021, 33(16): 10551-10562.
- [68] González S, García S, Del Ser J, et al. A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives and opportunities[J]. *Information Fusion*, 2020, 64: 205-237.
- [69] Qiao L, Li X S, Umer Q, Guo P. Deep learning based software defect prediction[J]. *Neurocomputing*, 2020, 385: 100-110.
- [70] Dong X B, Yu Z W, Cao W M, et al. A survey on ensemble learning[J]. *Frontiers of Computer Science*, 2020, 14(2): 241-258.
- [71] Chen J Y, Hu K K, Yang Y T, et al. Collective transfer learning for defect prediction[J]. *Neurocomputing*, 2020, 416: 103-116.
- [72] Sheng L, Lu L, Lin J H. An Adversarial Discriminative Convolutional Neural Network for Cross-Project Defect Prediction[J]. *Ieee Access*, 2020, 8: 55241-55253.

-
- [73] Song Q B, Guo Y C, Shepperd M. A Comprehensive Investigation of the Role of Imbalanced Learning for Software Defect Prediction[J]. Ieee Transactions on Software Engineering, 2019, 45(12): 1253-1269.
- [74] Tantithamthavorn C, McIntosh S, Hassan A E, Matsumoto K. The Impact of Automated Parameter Optimization on Defect Prediction Models[J]. Ieee Transactions on Software Engineering, 2019, 45(7): 683-711.
- [75] Mori T, Uchihira N. Balancing the trade-off between accuracy and interpretability in software defect prediction[J]. Empirical Software Engineering, 2019, 24(2): 779-825.
- [76] Moustafa S, Elnainay M Y, El Makky N, Abougabal M S. Software bug prediction using weighted majority voting techniques[J]. Alexandria Engineering Journal, 2018, 57(4): 2763-2774.
- [77] Tong H N, Liu B, Wang S H. Software defect prediction using stacked denoising autoencoders and two-stage ensemble learning[J]. Information and Software Technology, 2018, 96: 94-111.
- [78] Chen X, Zhao Y Q, Wang Q P, Yuan Z D. MULTI: Multi-objective effort-aware just-in-time software defect prediction[J]. Information and Software Technology, 2018, 93: 1-13.
- [79] Yang X L, Lo D, Xia X, Sun J L. TLEL: A two-layer ensemble learning approach for just-in-time defect prediction[J]. Information and Software Technology, 2017, 87: 206-220.
- [80] Rathore S S, Kumar S. Linear and non-linear heterogeneous ensemble methods to predict the number of faults in software systems[J]. Knowledge-Based Systems, 2017, 119: 232-256.
- [81] Li Z Q, Jing X Y, Zhu X K, et al. Heterogeneous Defect Prediction through Multiple Kernel Learning and Ensemble Learning[C]. 33rd IEEE International Conference on Software Maintenance and Evolution (ICSME), 2017: 91-102.
- [82] Xia X, Lo D, Pan S J, et al. HYDRA: Massively Compositional Model for

Cross-Project Defect Prediction[J]. Ieee Transactions on Software Engineering, 2016, 42(10): 977-998.

- [83] Laradji I H, Alshayeb M, Ghouti L. Software defect prediction using ensemble learning on selected features[J]. Information and Software Technology, 2015, 58: 388-402.