面向DRG/DIP支付改革的医院病案首页智能质控与费用预测方法研究

张卓然[1],陈红波[2],羊　平[2],郝晨颖[1],杨慧兰[1]

(1.内蒙古医科大学,呼和浩特 010100;
2.昆明理工大学附属安宁市第一人民医院,昆明 650000)

摘　要: 在疾病诊断相关分组(DRG)与病种分值付费(DIP)支付方式改革背景下,医院病案首页数据质量和费用预测能力直接影响医保结算效率与医疗成本控制水平。针对当前病案首页存在的编码不规范、字段缺失及费用预测依赖经验等问题,构建了一种面向DRG/DIP支付改革的病案首页智能质控和费用预测方法。该方法首先基于规则引擎与机器学习融合方法实现病案首页多维度智能质控;其次构建融合临床特征和费用结构的多模型集成费用预测模型,对住院总费用进行预测。实验结果表明,该方法在数据完整率、编码一致性及费用预测精度方面均优于传统方法,可有效提升医院医保精细化管理水平。

关键词: DRG;DIP;病案首页;智能质控;费用预测;机器学习
中图分类号: TP399　　**文献标识码:** A

Research on intelligent quality control and cost prediction algorithm for hospital medical record front page oriented to DRG/DIP payment reform

ZHANG Zhuoran[1], CHEN Hongbo[2], YANG Ping[2], HAO Chenying[1], YANG Huilan[1]
(1.Inner Mongolia Medical University, Hohhot 010100, China;
2.Kunming University of Science and Technology Affiliated Anning First People's Hospital, Kunming 650000, China)

Abstract: Under the background of the payment reform of Diagnosis-Related Groups (DRG) and Diagnosis-Intervention Packet (DIP), the data quality of hospital medical record front pages and cost prediction capability directly affect the efficiency of medical insurance settlement and the level of medical cost control. Aiming at the problems existing in current medical record front pages, such as non-standard coding, missing fields, and cost prediction relying on experience, an integrated method of intelligent quality control and cost prediction for medical record front pages oriented to DRG/DIP payment reform is constructed. The method first achieves multi-dimensional intelligent quality control of medical record front pages based on a fusion of rule engines and machine learning. Secondly, a multi-model ensemble prediction algorithm integrating clinical features and cost structure is constructed to predict total hospitalization costs. Experimental results show that this method outperforms traditional methods in data completeness rate, coding consistency, and cost prediction accuracy, and can effectively improve the refined management level of hospital medical insurance.

Key words: DRG, DIP, medical record front page, intelligent quality control, cost prediction, machine learning

1　引言

随着我国医保支付方式由按项目付费向DRG/DIP付费改革转型,医院运营模式正从“规模驱动”向“质量与效率驱动”转变。在此过程中,病案首页作为医保结算和疾病分组的核心数据源,其质量直接影响DRG的准确性与DIP的核算结果。然而,在实际应用中,病案首页普遍存在主要诊断编码不规范或错填、手术操作编码遗漏或不一致、费用结构填报逻辑异常、人工质控效率低、滞后性强、缺乏基于数据驱动的费用预测模型等问题。因此,构建智能化病案首页质控和费用预测方法,对提升医保管理水平具有重要意义[1]。

基金项目: 内蒙古自治区高等学校科学研究项目:激励相容理论视角下呼和浩特市基层公立医疗机构医疗服务绩效研究(NJSZ22687);内蒙古医科大学校级面上项目:内蒙古牧区医疗卫生服务能力的时空分异及提升路径(YKD2024MS020)

2　相关技术基础

在DRG/DIP支付方式改革的背景下，医保结算机制正逐步由按项目付费转向基于病种和资源消耗的精细化付费模式。DRG依据疾病严重程度、治疗方式及资源消耗水平对病例进行分组，并按组内权重进行付费；DIP则通过病种组合与分值体系对医疗服务进行量化结算[2]。二者均高度依赖病案首页数据的规范性、完整性和准确性。作为医保结算与临床数据分析的核心载体，病案首页的数据结构主要涵盖患者基本信息、主要诊断及ICD编码、手术操作信息、住院费用明细，以及住院天数和转归情况等关键字段，上述信息共同构成DRG/DIP分组与费用核算的基础。在医疗费用预测方面，现有研究主要采用多元线性回归、随机森林、XGBoost及LSTM等机器学习和深度学习方法对住院费用进行建模与预测。然而，此类方法多仅依赖历史费用数据或基础统计特征，较少融合DRG分组特征和临床路径信息，致使其在复杂病例及跨病种场景中的泛化能力与可解释性仍存在不足[3]。

3　智能质控与费用预测方法设计

3.1　总体框架

本文构建了面向DRG/DIP支付改革的病案首页智能质控与费用预测方法，整体由病案首页智能质控模型与费用预测模型两个核心模型构成。其中，智能质控模型用于提升病案首页数据质量，费用预测模型用于实现住院费用的精细化预测，二者共同支撑DRG/DIP支付改革下的病案首页管理。该方法整体可表示为：

$$F(x)=f_3(f_2(f_1(x))) \tag{1}$$

其中，f_1表示数据预处理；f_2表示质控处理过程；f_3表示费用预测模型。

3.2　病案首页智能质控模型

病案首页智能质控模型采用“规则引擎+机器学习异常检测”双机制融合策略，以提升数据质量控制的准确性和覆盖率。传统单一规则质控难以应对复杂编码错误与隐性异常问题，而单纯机器学习方法又缺乏医学逻辑约束[4]。为此，本文通过结构化规则约束结合数据驱动异常识别，实现多维度质控优化。该模型首先依据医疗编码规范构建规则库，对关键字段进行逻辑校验；其次利用异常检测算法识别费用和编码异常病例；最后通过融合策略输出统一质控评分，以实现对病案首页数据质量的综合评估。

3.2.1　规则引擎

规则引擎主要依据国家病案首页填写规范及ICD编码标准构建结构化规则库，实现对关键字段的逻辑一致性和合法性校验。具体规则包括ICD编码合法性检查、主要诊断唯一性约束、手术操作时间顺序合理性约束及费用结构比例约束等。例如，当主要诊断数量超过1条或ICD编码不符合标准库时，系统将自动标记为违规数据。规则引擎具有高可解释性与强约束能力，能够快速过滤显性错误数据，提升数据基础质量[5]。规则判定函数定义为：

$$Q_{\text{rule}}(x)=\begin{cases}1, \text{符合规则}\\0, \text{违反规则}\end{cases} \tag{2}$$

3.2.2　机器学习异常检测模型

针对规则引擎难以覆盖的隐性异常识别问题，本文引入孤立森林算法建立异常检测模型。该模型通过构建多棵随机树计算样本路径长度，异常样本通常具有较短路径，据此实现异常识别。异常评分函数表示为：

$$S(x)=-\sum_{t=1}^{T}\frac{h_t(x)}{c(n)} \tag{3}$$

其中，$h_t(x)$表示第t棵树的路径长度；$c(n)$为标准化因子；T为树的数量。

3.2.3　质控融合策略

为综合规则引擎与机器学习异常检测模型的优势，本文设计了质控融合策略，将两类结果加权融合，形成统一质控评分体系。该策略既保留规则模型的医学逻辑约束能力，又增强机器学习模型对复杂异常的识别能力，从而提升整体质控效果和稳定性。在实际应用中，可通过调整权重参数实现不同医院或科室的个性化质控策略配置：

$$Q=\alpha Q_{\text{rule}}+\beta Q_{\text{ml}},\ \alpha+\beta=1 \tag{4}$$

其中，α控制规则约束权重；β控制数据驱动权重。

3.3　费用预测模型

费用预测模型基于集成学习方法构建，用于在DRG/DIP支付体系下实现费用精细化预测与风险控制。该模型综合考虑临床特征和费用结构信息，可通过多模型融合方式提升预测精度与泛化能力。相较于传统单一回归模型，该模型能够更好地捕捉非线性关系和复杂特征交互，提升对高维医疗数据的建模能力。

3.3.1　特征构建

费用预测模型的性能高度依赖特征构建质量。本文从病案首页中提取多维特征，包括DRG分组信息、住院天数、手术等级、患者年龄和性别、合并症数量及检查与耗材费用占比等。同时，引入结构化临床路径信息，以增强模型对医疗过程的表达能力。通过特征工程处理，对类别变量进行编码、连续变量进行标准化处理，从而提升模型训练稳定性和预测精度。特征向量表示为：

$$\boldsymbol{x}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n] \tag{5}$$

其中，每个$\boldsymbol{x}_i$对应一个标准化后的临床或费用特征，以实现对复杂医疗信息的统一表达。

3.3.2　模型结构

本文采用XGBoost与LightGBM的集成学习方法构建费用预测模型。XGBoost具有较强的正则化能力，适用于处理稠密特征；LightGBM则在大规模数据训练中具有更高效率。通过加权融合两者输出结果，可兼顾预测

精度和计算效率，提升模型稳定性与泛化能力。模型整体形式表示为：

$$\hat{y}=w_1 f_{XGB}(x)+w_2 f_{LGB}(x) \quad (6)$$

其中，f_{XGB}，f_{LGB}分别表示XGBoost与LightGBM模型输出；w_1，w_2为融合权重。

3.3.3 损失函数

在模型训练过程中，本文采用均方误差（MSE）作为损失函数，以衡量预测费用和真实费用之间的偏差。MSE能够有效反映连续型变量的预测误差，适用于住院费用这类数值型预测任务。同时，通过最小化损失函数，可不断优化模型参数，提升预测精度与稳定性。损失函数定义为：

$$L=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \quad (7)$$

其中，y_i为真实费用；$\hat{y}_i$为预测费用；n为样本数量。

4 实验设计与结果分析

4.1 数据来源

为验证本文提出的病案首页智能质控与费用预测方法（由病案首页智能质控模型与费用预测模型构成）的有效性，选取某三级综合医院作为实验数据来源。共收集2024—2025年期间的住院病例数据50 000例，涵盖内科、外科、妇产科及重症医学等临床科室，具有较强的代表性和泛化意义。

4.2 评价指标

为全面评估所提方法在病案首页智能质控与费用预测中的性能，本文从质控效果和预测精度两个方面设计评价指标。（1）在质控任务中，采用质控准确率、编码错误识别率与漏检率作为评价指标。其中，质控准确率用于衡量系统对病案首页数据整体判定的正确性，编码错误识别率用于评估模型对ICD编码及关键字段错误的检测能力，漏检率则用于反映模型未能识别出的错误数据占比。（2）在费用预测任务中，采用平均绝对误差（MAE）和均方根误差（RMSE）作为核心评价指标，用于衡量预测值与真实住院费用之间的偏差程度。MAE能够反映整体预测误差水平，RMSE则对较大误差具有更高敏感性，更适用于医疗费用这类波动较大的数据场景。为验证本文方法的有效性，质控任务选取人工质控与规则引擎作为对比基线，费用预测任务选取线性回归与XGBoost作为对比基线。

4.3 实验结果

4.3.1 质控效果对比分析

实验结果表明，不同方法在病案首页质控任务中的表现存在明显差异。人工质控方式依赖人工经验和抽样检查，整体质控准确率为82.6%，漏检率较高，达17.4%，难以满足DRG/DIP精细化管理需求。规则引擎方法通过引入结构化规则库，在一定程度上提升了质控能力，其准确率提升至90.3%，漏检率降至9.7%，但仍难以识别复杂隐性错误。相比之下，本文智能质控模型在规则约束与机器学习异常检测的协同作用下，质控准确率达96.8%，漏检率降至3.2%。结果表明，该模型不仅能够有效识别显性编码错误，还可捕捉费用结构异常及隐性数据偏差，从而显著提升病案首页数据质量，为后续DRG/DIP分组和医保结算提供更可靠的数据基础。

4.3.2 费用预测效果对比分析

在费用预测实验中，不同模型表现出不同程度的预测能力差异。传统线性回归模型由于表达能力有限，难以刻画复杂的非线性关系，其MAE为1 860，RMSE为2 400，预测误差较大，仅适用于简单场景，难以满足DRG/DIP精细化管理需求。XGBoost模型通过引入梯度提升机制，在非线性建模方面表现更优，其MAE降至1 320，RMSE为1 780，预测精度显著提升。本文模型进一步融合XGBoost与LightGBM的优势，在多特征协同建模基础上实现更高精度预测，其MAE降至980，RMSE降至1 320，整体误差显著低于对比方法。实验结果表明，该模型能够更好地捕捉临床特征和费用结构之间的复杂关系，在DRG/DIP支付体系下具有更高的应用价值。

5 结束语

本文提出一种面向DRG/DIP支付改革的病案首页智能质控和费用预测方法，通过规则引擎与机器学习融合实现高精度数据质控，并构建集成学习费用预测模型。实验结果表明，该方法能够有效提升病案首页数据质量和费用预测精度，对医院医保精细化管理具有重要应用价值。

参考文献：

[1] 杨林朋.某三甲医院病案质控统计分析[J].中国保健营养，2020，30（19）：14-15.

[2] 洪嘉铭，裘以冰，郭蕴青，等.广东省医院病案质控平台使用和测评报告[J].现代医院，2005，5（6）：102-103.

[3] 徐一楠，胡晓星.电子病案系统对医院病案质控工作的促进与挑战[J].中国病案，2025，26（1）：4-6.

[4] 朱建军，褚新奇.海军医院住院费用预测模型的建立与应用[J].解放军医院管理杂志，2002，9（5）：433-434.

[5] 杨蕾，刘硕，马先莹.某三甲医院肺癌手术患者住院费用的结构变动、影响因素及预测研究[J].中国医院统计，2025，32（1）：1-7+13.

作者简介：

张卓然（1997—），硕士，研究方向：民族地区卫生事业健康发展战略。

杨慧兰（1975—），博士，教授，研究方向：卫生管理、卫生政策，E-mail：1220247538@qq.com（通信作者）。