

## 基于深度学习的医疗器械不良事件风险类型的研究\*

潘康宁<sup>1</sup>, 袁明辉<sup>2</sup>, 赵玉娟<sup>3</sup>, 玄怡<sup>3</sup>, 王茜<sup>2</sup>, 王洪杰<sup>1</sup>, 孙万晨<sup>4</sup>

(1. 威海市妇幼保健院 医疗设备科, 山东 威海 264200; 2. 威海市食品药品检验检测研究院, 山东 威海 264200; 3. 山东省药品不良反应监测中心, 山东 济南 250000; 4. 威海市胸科医院 医务科, 山东 威海 264200)

**摘要:** **目的** 探讨深度学习相关技术在医疗器械不良事件风险类型判别中的应用价值。**方法** 选取自 2023 年 4 月至 9 月国家医疗器械不良事件监测系统数据库中的 12 350 条数据为研究对象, 每条数据具有 25 个维度的信息, 其中包括企业、注册号、产品批号、不良事件描述等。采用 fastText 技术将文本特征转化为向量表示, 采用  $k$  均值聚类方法将具有相似主题或语义内容的文本数据分组到同一类别中, 搭建 BP 神经网络将不良事件的风险类型分为轻微危害、严重危害、死亡。**结果** 该研究选取的深度学习方法能够很好地处理文本数据且搭建的 BP 神经网络模型的准确度 (accuracy) 为 92.86%, 精密度 (precision) 为 93.65%, 召回率 (recall) 为 93.08%, F1-score 为 92.31%, 曲线下面积 (AUC) 为 0.98, 具有良好的准确性和泛化能力。**结论** 基于深度学习的医疗器械不良事件风险类型的研究可有效地为医疗器械不良事件监测工作提供帮助。

**关键词:** 深度学习; 医疗器械; 不良事件; 分类研究

**中图分类号:** TP18; R197.39

## Risk types of adverse events in medical devices based on deep learning\*

PAN Kangning<sup>1</sup>, YUAN Minghui<sup>2</sup>, ZHAO Yujuan<sup>3</sup>, XUAN Yi<sup>3</sup>, WANG Qian<sup>2</sup>, WANG Hongjie<sup>1</sup>, SUN Wanchen<sup>4</sup>

(1. Medical Equipment Department, Weihai Maternal and Child Health Hospital, Weihai, Shandong 264200, China;  
2. Weihai Food and Drug Inspection and Testing Research Institute, Weihai, Shandong 264200, China;  
3. Shandong Province Adverse Drug Reaction Monitoring Center, Jinan, Shandong 250000, China;  
4. Medical Department, Weihai Chest Hospital, Weihai, Shandong 264200, China)

**Abstract:** **[Objective]** To explore the value of deep learning related techniques in the risk type discrimination of medical device adverse events. **[Methods]** The study population was selected from April to September 2023, 12,350 data in the Medical Device Adverse Event Surveillance System database, each with 25 dimensions of information, including company, registration certificate number, product batch number, and description of adverse events. This study uses fastText technology to convert text features into vector representations, uses  $k$ -means clustering method to classify text data with similar themes or semantic content into the same category, and uses BP neural network algorithm to classify the severity of adverse events into mild injury, severe injury, and death. **[Results]** This method can handle text data well, and the accuracy of the BP neural network model is 92.86%, accuracy is 93.65%, recall is 93.08%, F1 score is 92.31%, AUC is 0.98, with good accuracy and generalization ability. **[Conclusion]** The research on the risk types of adverse events in medical devices based on deep learning can effectively provide assistance for the monitoring of adverse events in medical devices.

**Keywords:** deep learning; medical devices; adverse events; classification research

在现代医疗领域中，医疗器械的作用至关重要，然而，其伴随的安全问题已经引起了广泛的关注。医疗器械不良事件对患者的身心健康和生命安全构成潜在威胁，严重损害了医疗质量和医疗行业的信誉。如何有效地控制和降低医疗器械不良事件，是当前医疗器械安全管理中亟待解决的问题。因此，要建立医疗器械的安全监管体系，更新医疗器械不良事件监管的技术手段，预防和减少医疗器械不良事件的发生。目前，我国已经建立了一个覆盖全国的医疗器械不良事件监测系统，包含全国各省市级的药品不良反应监测中心、医疗器械厂商、销售公司以及医疗机构<sup>[1]</sup>。伴随着医疗机构、生产企业和公众对医疗器械不良事件监测工作的认知和参与度不断增加，每年上报的医疗器械不良事件也呈逐年上涨的趋势<sup>[2]</sup>。2020 年全年可疑医疗器械不良事件上报量突破 53 万件。但相对于报道数量的日益增多以及医疗器械的复杂性，如何及时地发现并判断不良事件危害的大小更富有挑战。传统人工分析方法呈现数据处理难度大，效率低，时间滞后，主观性等特点。这就给及时发现判断危害程度和检测可疑风险信号带来了严峻挑战，同时也决定了普通常规性方法不能充分满足当前的需要。因此，本研究针对当前医疗器械不良事件监测领域中评估危害程度和检测可疑风险信号时所遇到的问题，将深度学习技术引入到医疗器械不良事件风险程度的判定工作中，并结合国家医疗器械不良事件监测系统的报告数据进行深入分析，旨在为不良事件的风险信号程度判定提供一种创新方法，从而减轻医疗器械不良事件监测人员的工作负担。

## 1 资料与方法

### 1.1 一般资料

本研究数据选取自 2023 年 4 月至 9 月国家医疗器械不良事件监测系统数据库中的 12 350 条数据为研究对象，每条数据具有 25 个维度的信息，其中包括企业、注册证号、产品批号、不良事件描述等。

### 1.2 相关技术手段

1.2.1 fastText 技术 近年来，基于神经网络的自动特征抽取方法逐渐成为研究热点<sup>[3-4]</sup>。其中，一种名为 fastText 的基于词袋模型的词向量表示方法，以其在生成词向量和文本分类任务中展现出的卓越性能<sup>[5]</sup>，吸引了众多研究者的目光。fastText 模型

的独特之处在于它引入了子词 ( $n$ -grams) 的概念，这一设计使得模型能够更好地处理词汇外单词和多义词的语义多样性<sup>[6]</sup>。具体来说，fastText 将一个单词表示为其所有子词的向量的平均值，这样可以更好地捕捉到单词内部的语义信息和形态学特征。与传统的 Softmax 不同，分层 Softmax 使用了一棵哈夫曼树 (Huffman Tree) 将单词组织为一系列的二进制编码。这样，就可以通过树结构进行高效的概率计算，大大减少了计算量。这一创新性的设计，无疑为深度学习技术的进一步发展提供了强大的推动力。

fastText 模型是一种强大的文本分类工具，它主要由三个关键部分组成：输入层、隐藏层和输出层。在输入层，文本数据被接收并转换为单词和子词的向量表示。这种表示方式使得模型能够更好地理解和分析文本内容。隐藏层是模型的核心部分，它通过将输入层的向量表示进行加权求和得到。这个加权求和过程可以是平均池化、最大池化，或者是其他形式的池化操作。这种池化操作能够有效地提取文本数据中的关键信息，提高模型的预测精度。输出层是模型的最后部分，它使用分层 Softmax 方法计算每个类别的概率分布。这种概率分布能够告诉我们文本属于哪个类别，以及这个类别的可能性有多大。在本文中，我们通过最大化目标类别的条件概率来调整模型的参数，使得模型能够更好地预测文本的类别。这种调整方式能够使模型在训练过程中更加关注那些对分类结果影响最大的参数，从而提高模型的性能。

1.2.2  $k$  均值聚类 文本聚类是一种无监督学习的技术，用于将具有相似主题或语义内容的文本数据分组到同一类别中<sup>[7]</sup>。聚类算法的运行机制是通过深入分析文本之间的相似性和差异性来自动发现并呈现文本数据中的模式和结构，从而将文本数据划分为不同的聚类簇。 $k$  均值聚类 ( $k$ -means clustering) 是一种常用的无监督学习算法，其算法的核心思想是将数据集中的样本划分为  $k$  个簇，使得每个样本与所属簇的中心点 (质心) 之间的距离最小化<sup>[8]</sup>。该算法使用欧氏距离 (或其他距离度量) 来度量样本之间的相似性，并不停的迭代地更新质心和重新分配样本，直到满足收敛条件。这样，就能确保每个簇都包含相似的样本，而不同簇之间的样本则尽可能地不同。

1.2.3 BP 神经网络 神经网络在深度学习领域中

扮演者重要角色，BP (back propagation) 神经网络作为其中的一种技术，以其非线性映射能力强大和并行信息处理能力突出的特点，在实际工作中倍受青睐。该技术主要用来实现模式识别、数据分类、函数拟合等功能<sup>[9]</sup>。BP 神经网络是一种多层前馈神经网络，采用梯度下降法进行反向传播<sup>[10]</sup>。其网络结构具备非常强的非线性映射能力、自学习自适应能力、较强的收敛性和泛化能力，能够很好满足本研究的数据特点。

### 1.3 方法

1.3.1 数据处理 数据处理环节主要包括四个方面：①采用 fastText 算法将文本特征（如企业、不良事件描述等）转换为向量表示。②采用删除、填充平均数等方法处理缺失值、异常值和重复值。③标记不良事件的严重程度为类别标签，1 代表轻微危害，2 代表严重危害，3 代表死亡。④将数据分成训练集和测试集，采用 70% 的数据作为训练集，30% 的数据作为测试集。

1.3.2 聚类分析 本文采用 MATLAB 自带的 CH (contraction hierarchies) 指标评估聚类数，其 CH 值越大代表类自身越紧密，类与类之间越清晰，聚类效果更好。通过 MATLAB 中的 evalclusters 函数计算得出最佳聚类数为 3。该结果与风险程度大致分为轻度危害、严重危害、死亡相一致，具有一定的实际意义。

在实际操作中利用 MATLAB 调用 *k*-means 聚类算法。在计算距离时，选取了 sqEuclidean 函数，即采用欧几里得距离函数，此函数能够较为准确地反映数据之间的距离。为了提高聚类分析实验的可靠性，设置了代表重复聚类次数的 replicate 参数为 5，以此得到了更加稳定和准确的聚类结果。

1.3.3 搭建 BP 神经网络 在对 BP 神经网络模型进行优化设置时，本研究参考了数据量、数据输出情况以及数据特点等信息，对模型的层数、隐藏层神经元个数、激活函数和优化器等参数进行了精细调整。为了更加精确地判断不良事件的严重程度，笔者选择了一个包含三层隐含层的模型，以增强模型的复杂性和准确性。在神经元个数的选择上，本研究根据数据集的规模，确定了输入层神经元为 25 个，隐藏层神经元为 15 个。所有的激活函数都采用了 ReLU (rectified linear unit) 函数，输出层神经元个数为 3，意味着模型将输出三个预测数据。在模型优化过程中，本研究采用

了 RMSProp 模型优化器，并将学习率设定为 0.001。最大迭代次数设定为 3 000 次，损失函数选择均方误差 (mean-square error, MSE)，评测准则选择了 MSE 和平均绝对值误差 (mean absolute error, MAE)。这样的优化策略，旨在提高模型的精度和稳定性，使其能够更好地适应和处理实际问题。

1.3.4 模型训练与评价 本研究用预先划分的测试集数据来验证已经训练好的模型，并获取预测结果，并将预测结果与测试数据进行比对，通过综合运用准确度、精确度、召回率以及 F1 分数等评估指标全面评估模型的性能表现。

## 2 实验结果

本研究运用所建立的 BP 神经网络模型对不良事件危害程度进行了分类。选取注册证号、产品批号、生产日期、不良事件描述等 25 个特征作为输入参数，输出结果为三类不良事件的风险程度即轻微危害、严重危害、死亡。模型结果图见表 1，受试者工作特征 (receiver operating characteristic, ROC) 曲线见图 1。BP 神经网络模型的准确度 (accuracy) 为 92.86%，精密度 (precision) 为 93.65%，召回率 (recall) 为 93.08%，F1-score 为 92.31%，曲线下面积 (area under curve, AUC) 为 0.98，表明所建立的神经网络模型具有良好的准确性和泛化能力。

表 1 BP 神经网络在测试集上的表现

评价指标	数值
准确度	92.86%
精密度	93.65%
召回率	93.08%
F1-score	92.31%

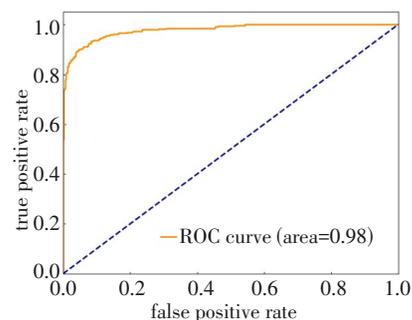


图 1 BP 神经网络模型 ROC 曲线

### 3 讨论

随着医疗质量的提高, 医疗器械不良事件的危害程度判定, 作为一项重要工作, 也越来越受到人们的重视。然而, 这一领域目前仍处于探索阶段, 许多方法和技术尚待进一步开发和完善。本研究展示了一种基于 BP 神经网络模型的医疗器械不良事件危害程度分类方法。经过实验验证, 该方法在风险程度判定方面表现优秀, 准确率高, 为医疗器械不良事件的监测工作提供了强有力的工具。然而, 本文提出的分类方法仍有提升空间。期望来者可以对聚类算法进行优化, 以提高分类的准确性和效率; 对数据进行更好的预处理, 以减少噪声和错误信息对分类结果的影响; 也可以尝试将不同的深度学习模型应用到危害程度的判定工作中, 并进行模型的对比, 以提高判定的准确性和效率。此外, 本研究也期望能够引发更多关于医疗器械不良事件风险程度检测方法的研究。深度学习作为一种新兴的技术, 其在医疗器械不良事件危害程度判定中的潜力是无可忽视的。我们期待未来能有更多的研究者在这个领域做出突破, 为我国的医疗器械不良事件监测工作提供更多的技术支持。

综上所述, 虽然医疗器械不良事件的危害程度判定工作目前仍处于探索阶段, 但是我们有理由相信, 随着人工智能技术的发展, 我们能够找到更好的方法来完成这项工作, 从而更好地保护公众的健康。

### 参考文献

- [1] 田燕, 肖桂荣, 徐斑. 我国医疗器械不良事件监测年度报告分析[J]. 中国药业, 2022, 31(6): 6-12.
- [2] 王东京, 吴维民. 我国医疗器械不良事件监测现状研究[J]. 中国医疗设备, 2022, 37(3): 127-130.
- [3] 伊恩·古德费洛, 约书亚·本吉奥, 亚伦·库维尔. 深度学习[M]. 北京: 人民邮电出版社, 2021.
- [4] SOCHER R, BAUER J, MANNING CD, et al. Parsing with compositional vector grammars[C]// 51st annual meeting of the Association for Computational Linguistics, August 4-9, 2013, Sofia, Bulgaria: Association for Computational Linguistics, 2013: 455-465.
- [5] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality[J]. arXiv E Prints, 2013: arXiv: 1310. 4546.
- [6] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[EB/OL]. 2013: arXiv: 1301. 3781. <http://arxiv.org/abs/1301.3781>.
- [7] 张孝飞, 黄河燕. 一种采用聚类技术改进的KNN文本分类方法[J]. 模式识别与人工智能, 2009, 22(6): 936-940.
- [8] 陶莹, 杨锋, 刘洋, 等. K均值聚类算法的研究与优化[J]. 计算机技术与发展, 2018, 28(6): 90-92.
- [9] 周梦, 吕志刚, 邸若海, 等. 基于小样本数据的BP神经网络建模[J]. 科学技术与工程, 2022, 22(7): 2754-2760.
- [10] 张明琪, 邓鑫. 基于集成学习后融合神经网络算法构建肝硬化代偿期中医智能辨证模型的探讨[J]. 广州中医药大学学报, 2023, 40(10): 2650-2660.

(张咏 编辑)