

编者按:食品是人类赖以生存的物质基础,而传统的食品研发、生产与管理方式正面临着诸多挑战,存在创新性不足的问题。近年来,人工智能(AI)技术发展迅猛,正深刻地影响着人类的思维、生产模式以及生活方式。AI技术是通过计算机模拟人类智能行为的一种技术,它包括机器学习、深度学习、自然语言处理、计算机视觉等多个领域。AI技术也赋能食品行业,其应用不断深化与拓展,已成为全行业创新发展的新质生产力与驱动力。目前,AI技术已被用于食品安全与监管、食品分析与检测、食品加工与生产、食品研发与创新、食品营养与健康等诸多领域。本刊现开设“人工智能+食品”特约专栏,连续刊载相关文章。

(本刊编辑部。客座编辑:江南大学人工智能与计算机学院吴小俊教授。)

基于时空 Transformer 模型的食品监督抽检分类预测研究

罗晓清^{1,2}, 郭林^{1,2}, 杨雨萌^{1,2}, 黄耐云^{1,2}, 吴小俊^{1,2}

¹江南大学人工智能与计算机学院 江苏无锡 214122

²江南大学 科技部中英人工智能联合实验室 江苏无锡 214122

摘要 目的:研究构建一种时空趋势与风险食品类别相关的多维分析模型,旨在实现对食品安全风险的高效预警,并为监管的智能化和精准化提供科学依据和决策支持。方法:依托江苏省市场监督管理局2019年至2024年间的食品抽检数据,设计一种基于时空Transformer的多分类预测模型。该模型深度挖掘食品安全事件在时间与空间维度的演变规律,并建立与风险类别的内在关联,实现对全时空点风险食品类别的预测。为解决多类别数据量差异导致的长尾分布问题,引入加权均方误差损失函数以优化训练,增强模型对尾部类别数据的敏感性。结果:采用4种最高风险类别分类指标与3种全类别风险分类指标综合评估模型的准确性,不仅在已发生风险区域的预测精度上与统计方法相当,而且在全时空范围内实现了全类别风险的高效预测。结论:时空Transformer多分类预测模型为食品安全监管部门提供了一种新颖且有效的工具,能够优化随机抽检策略并提升监管效率。

关键词 时空Transformer; 加权均方误差; 食品抽检; 风险预警; 多分类预测; 食品安全

文章编号 1009-7848(2024)11-0001-09 **DOI:** 10.16429/j.1009-7848.2024.11.001

食品安全直接关系到公共健康和经济发展,已经成为极为关键的民生问题^[1-2]。近年来,频发的食品安全事件给公众的生命健康带来了严重的威胁,也对社会造成了难以估量的损失^[3-5]。建立完善的食品安全监管机制已成为一项紧迫的任务。食品安全监督抽检作为保障体系中的重要防线^[6-7],在维护公众健康和市场秩序方面发挥着关键作用^[8-10]。传统的食品抽检方案因效率低、覆盖范围有限,故难以满足现代监管的需求^[11-12]。为此,各级市场监管部门积极践行“双随机,一公开”的抽检模式^[13-14],以提升抽检工作的有效性和覆盖面。然

而,面对庞大的食品市场和有限的监管资源,如何精准高效地确定抽检对象仍是一个亟待解决的难题。特别是在人力、物力和财力等资源有限的情况下,提升抽检靶向性成为优化抽检策略的关键^[15-17]。

历史抽检数据作为宝贵的信息资源,已被用于指导未来食品安全的健康发展。基于逆向传播(BP)神经网络的数据挖掘方法以抽检工作数据为基础,训练模型以预判抽检食品项目的合格与否,从而提高检测工作的效率^[18]。另外,统计分析抽检数据的分布情况,有力推动了不合格食品的靶向抽检^[19]。虽然这些方法在抽检工作中取得了一定的成效,但是它们往往妥协于时间和空间之间的条件独立性,难以全面考虑食品安全事件受时间、空间、类别等多维因素的关联影响。在优化抽检策略的研究中,识别和利用时空趋势与风险食品类别之间的潜在关联,可以为提升抽检工作的精准

收稿日期:2024-10-29

基金项目:国家重点研发计划项目(SQ2023YFF1100111);
国家自然科学基金项目(61772237)

第一作者:罗晓清,女,博士,教授

E-mail: xqluo@jiangnan.edu.cn

度提供新的思路。

随着人工智能技术的迅猛发展,其在食品安全领域的影响力日益增强^[20-21]。深度学习凭借其强大的特征提取和复杂的模式识别能力,在食品安全风险管理中展现出巨大潜力。利用深度学习技术构建食品安全风险知识图谱,能显著提升对海量、复杂数据的理解和处理能力^[22];同时,基于深度学习的食品安全事件综合展示平台,为提供精准的数据分析开辟了新途径^[23];此外,受益于先进的时间序列预测领域的方法,协同强化学习实现了各类别食品召回风险的预测和及时召回^[24]。

本文结合人工智能技术的优势,深入研究风险食品类别的预测问题,以应对食品安全监管的迫切需求。时空数据与食品类别数据之间的显著差异,增加了数据分布的复杂性和数据分析的难度。传统的统计分析方法在处理这些复杂数据时,往往面临预测准确性不足和对数据特征深层次规律捕捉不充分等问题。然而,深度学习中的Transformer模型通过多头自注意力机制,能够在全局范围内捕捉数据间的复杂交互和非线性特征关系。本文提出一种基于时空Transformer的食品监督抽检分类预测模型,以实现时空点与风险类别之间的协同学习,为提升监管靶向性与制定抽检工作计划提供科学依据。

1 研究方法

1.1 数据集

在风险食品类别预测的研究中,基础数据来源于江苏省市场监督管理局在政府信息公开平台发布的食品抽检不合格情况的通告,数据涵盖2019年6月至2024年6月,共5年的时间。首先对不合格批次数据进行去重、缺失值填补、异常值处理等数据清洗步骤。其次,根据研究目的,将不合格批次数据的通告公布时间视为食品安全事件 S 的发生时间,而被抽取单位的地址作为事件所发生的空间位置。即,“被抽样单位地址”字段用作空间信息 L ，“发布时间”字段用作时间信息 T ，“分类”字段代表不合格食品类别信息 C 。在做数据预处理时,空间文本转换成经纬度数值,并以区县级为单位精确到小数点后一位。考虑到食品的季节性特点,时间信息精确到“月”。类别信息涵盖食

用农产品、调味品、肉制品等28($N_c=28$)个食品大类。最终,收集的数据总量为2340批次,具体类别分布如图1所示。为了构建深度学习模型,划分1896批次作为训练数据集,444批次作为验证数据集。

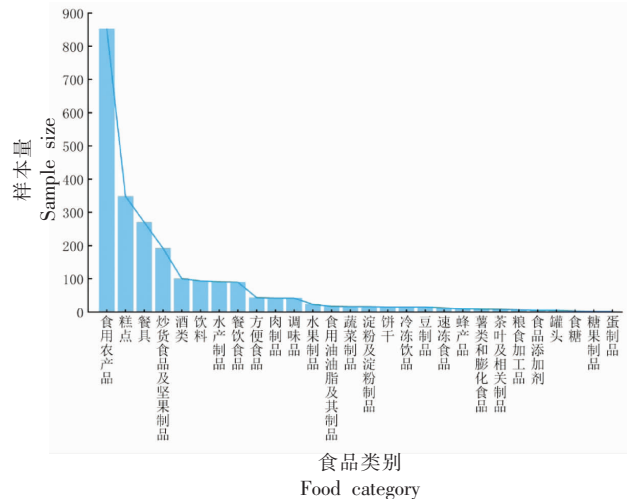


图1 食品类别分布柱状图

Fig.1 Bar chart of food category distribution

1.2 模型构建

基于时空Transformer的分类预测模型如图2所示。它由时空编码器、时空Transformer、多分类解码器和加权均方误差损失函数4个部分组成。时空编码器用于提取输入数据的有效表征,生成时空特征组向量。随后,通过加法组合策略融合空间和时间特征的线性关系,为模型提供融合时空线索的联合表征。融合后的特征接着被送入时空Transformer,以建模特征之间的高阶关系和依赖性。最终,多分类解码器输出与风险类别相关的预测概率值,并通过与真实概率值进行损失计算,以完成模型的训练。推理阶段,将江苏省内任意经纬度与时间输入该模型后,能够获得该时空点对应的全类别的预测概率值。

1.2.1 时间与空间编码器 为了对时空点进行动态建模,分别对时间数据和空间数据使用独立的编码器获得唯一表征。编码器的输入由事件 $S=(T,L,C)$ 组成。对于时间 T ,采用类似于位置编码(Positional encoding)^[25]的策略,将时间信息视为序列数据并编码为高维特征表示。具体来说,时间 $t(t \in T)$ 通过频率分量进行缩放,以便在不同的频

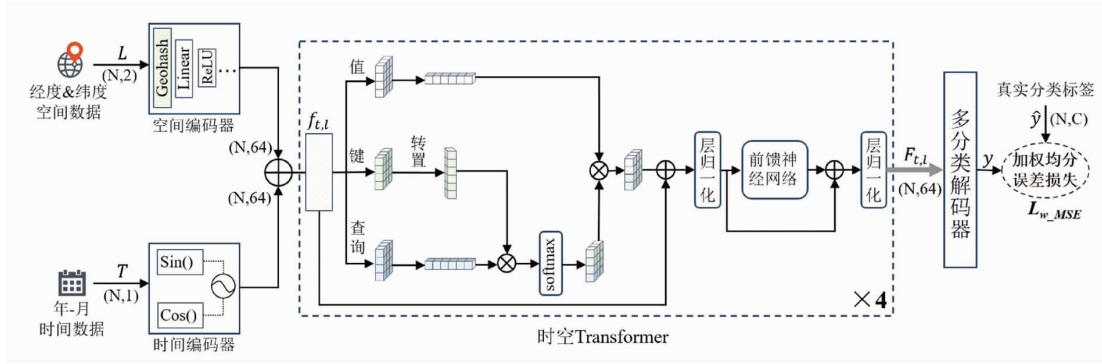


图 2 时空 Transformer 分类预测模型结构图

Fig.2 Spatio-temporal Transformer classification prediction model architecture diagram

率分量上被表达。然后,偶数索引的维度应用正弦函数,而奇数索引的维度应用余弦函数。时间编码器表示如下:

$$\begin{cases} \text{Te}(t, 2i) = \sin(t/10000^{\frac{2i}{m}}) \\ \text{Te}(t, 2i+1) = \cos(t/10000^{\frac{2i}{m}}) \end{cases} \quad (1)$$

式中:Te 为时间编码; i 为特征维度的索引; m 为特征的总维度。

编码交替的正、余弦变换使得时间信息可在不同尺度上被表达,从而捕捉到时间序列中的周期性和趋势性特征。

在空间编码器中,首先使用 Geohash 编码^[26]将经、纬度坐标值转换成字符串,以增强地理位置的表征能力。然后,采用 4 层全连接层和非线性激活函数(ReLU)来构建位置嵌入模块。给定低维地理位置坐标 $l=(lon, lat) \in L$,通过以下映射转换到高维特征空间中:

$$\text{Le}(l, 0) = \text{Geohash}(lon, lat) \quad (2)$$

$$\text{Le}(l, j) = \text{ReLU}[W^{(j)}\text{Le}(l, j-1)+b^{(j)}] \quad (3)$$

式中:Le 为空间编码; $j \in \{1, 2, 3, 4\}$ 为全连接层的层数; (lon, lat) 为经度和纬度; $W^{(j)}$ 和 $b^{(j)}$ 分别为第 j 层的权重矩阵和偏置向量,含有可学习的参数。

多层次的特征提取模块捕捉地理位置的复杂关系。随后,通过将时间编码和空间编码相加获得时空特征 $f_{t,l}$ 。

1.2.2 时空 Transformer 时空 Transformer 模块被应用于时空特征 $f_{t,l}$, 以生成联合时空特征 $F_{t,l}$ 。该模块主要由 2 个部分组成:多头自注意力机

制 (Multi-head self-attention) 和前馈神经网络 (Feedforward neural network)。具体来说,首先将时空特征 $f_{t,l}$ 映射到查询 (Query)、键 (Key) 和值 (Value) 空间,然后,采用缩放点积注意 (Scaled dot-product attention)^[27] 进行加权求和,从而聚合来自不同时空的信息,以捕捉全局时空依赖关系。计算表达如下:

$$Q, K, V = f_{t,l}W^Q, f_{t,l}W^K, f_{t,l}W^V \quad (4)$$

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right) \times V \quad (5)$$

式中: W^Q, W^K, W^V 为用于线性映射的权重矩阵; T^* 为转置; \sqrt{d} 为缩放因子,其中 d 为键向量的维度;Softmax 为激活函数。

在多头自注意力机制的输出和输入之间应用残差连接,之后进行层归一化,这将有助于稳定训练并加速模型收敛。前馈神经网络由两层线性变换和 GELU 激活函数组成,其中增加 Dropout 操作以防止过拟合。而后引入残差连接和层归一化以加速训练和缓解梯度消失问题。该部分计算表达如下:

$$f'_{t,l} = \text{LN}[\text{Attention}(Q, K, V)+f_{t,l}] \quad (6)$$

$$F_{t,l} = \text{LN}[\text{FFN}(f'_{t,l})+f'_{t,l}] \quad (7)$$

式中: $f'_{t,l}$ 为多头自注意力部分的输出特征;LN 为层归一化;FFN 为前馈神经网络。

1.2.3 多分类解码器 多分类解码器将融合特征 $F_{t,l}$ 解码为预测概率 y 。它由多个线性层 (Linear layer)、批归一化 (Batch normalization) 和 ReLU 激活函数组成,逐步将特征维度从 $m(m=64)$ 降低到 N_c ,最终输出一个 N_c 维向量 y ,代表了 N_c 个类别

的预测概率值。对预测向量 y 应用 Softmax 函数以归一化概率分布。真实类别标签 \hat{y} 是一个长度为 N_c 的向量,其中每个元素值代表对应的风险食品类别在相应的时空下出现的次数。由于多分类任务的分布差距,对真实标签 \hat{y} 中出现次数为 0 的类别视为无效数据,将其值设置为负无穷大。在做 Softmax 归一化时,无效类别的概率贡献将变为 0,从而确保模型只关注有效类别。

1.2.4 加权均方误差损失函数 均方误差(MSE)损失用于评估模型预测值与真实值之间的差异,它通过计算预测误差的平方平均值来衡量模型的准确性。其计算公式为:

$$L_{\text{MSE}}(\hat{y}, y) = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (8)$$

式中: N 为样本数量; y_i 为第 i 个样本的预测值; \hat{y}_i 为第 i 个样本的真实值。

均方误差损失对出现频率较高的类别更有偏好,从而忽略出现频率较低的类别。然而,实际数据分布差异较大(如图 1 所示),其中“食用农产品”的数量高达 800 余次,而“罐头”的数量低至 10 余次。类别数量分布呈现出显著的长尾分布模式,这将导致模型往往在尾部数据表现不佳。为了解决训练偏差问题,提出一种新的加权均方误差损失函数。该损失函数通过动态计算类别权重并应用平滑技术来调整每个类别对损失的贡献,从而缓解数据头尾不平衡所构成的障碍。其计算公式如下:

$$L_{\text{w_MSE}}(\hat{y}, y) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{N_c} w_j \times (\hat{y}_{i,j} - y_{i,j})^2 \quad (9)$$

式中: w_j 为类别权重。

对于给定的真实值 \hat{y} , 首先计算每个类别的出现次数 n_j , 从而得到类别频率 p_j 。为了增加频率较低类别在损失计算中的影响,采用类别频率的倒数作为初始权重 w_j' , 并进行归一化,使所有权重和等于类别数 N_c 。随后使用平滑因子 α 对初始权重进行平滑处理,以避免权重过度偏向于某些类别。类别权重 w_j 的计算公式为:

$$p_j = \frac{n_j}{N}, w_j' = \frac{1}{p_j + \epsilon}, w_j = N_c \times \frac{w_j'}{\sum_{k=1}^{N_c} w_k'} \quad (10)$$

$$w_j = (1 - \alpha) \times w_j' + \alpha \times w_j \quad (11)$$

式中: ϵ 为小常数,避免分母为零; α 为超参数,值为 0.6。

时空 Transformer 分类预测模型训练的伪代码如图 3 所示。

Algorithm 1 模型训练伪代码

```

input: Original training dataset  $S = (L, T, C)$ 
output: Prediction probabilities  $y$  for each category at spatio-temporal points
initialize ST_Transformer()
for epoch = 1 to epochs do
  for each time point  $t (t \in T)$  in  $T$  do
     $Time_{enc} = Te(t)$ 
  end for
  for each location  $l = (lon, lat) \in L$  in  $L$  do
     $Local_{enc} = Le(l)$ 
  end for
   $F_{t,l} = \text{ST\_Transformer}(Time_{enc}, Local_{enc})$ 
   $y = \text{Multi\_dec}(F_{t,l})$ 
  for  $j = 1$  to  $N_c$  do
     $\hat{y}[j] = \begin{cases} -Infinity, & \text{if value} = 0 \\ \hat{y}[j]_{N_c \times 1}, & \text{otherwise} \end{cases}$ 
  end for
  Use  $L_{\text{w\_MSE}}(\hat{y}, y)$  to backward
  Update optimizer
end for

```

图 3 时空 Transformer 分类预测模型训练的伪代码

Fig.3 Pseudocode for training the spatio-temporal Transformer classification prediction model

1.3 分析方法

研究采用单分类和多分类 2 种任务指标以全面评估模型的性能。单分类指标通过计算最高风险类别的概率来分析模型在重大风险类别上的决策质量,包括准确率(Accuracy)、召回率(Recall)、精确率(Precision)和 F_1 值(F_1 _Score)4 项指标。准确率评估模型在所有时空点上正确预测风险类别的整体能力。召回率反映模型探查特定风险类别的能力。精确率评估模型在预测为特定类别时的准确性。 F_1 值评估模型在特定风险类别上的综合表现,是精确率和召回率的调和平均数。计算公式如下:

$$\text{准确率} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (12)$$

$$\text{召回率} = \text{TP} / (\text{TP} + \text{FN}) \quad (13)$$

$$\text{精确率} = \text{TP} / (\text{TP} + \text{FP}) \quad (14)$$

$$F_1 \text{ 值} = (2 \times \text{精确率} \times \text{召回率}) / (\text{精确率} + \text{召回率}) \quad (15)$$

式中:TP 为模型正确预测为正类的样本数;

TN 为模型正确预测为负类的样本数;FP 为模型错误预测为正类的样本数;FN 为模型错误预测为负类的样本数。

多分类指标评估模型是在同时处理多种风险食品类别时的综合能力,包括余弦相似度(Cosine similarity, CosSim)、结构相似性(SSIM)和归一化皮尔逊相关系数(Normalized pearson correlation coefficient, NPCC)3项指标。余弦相似度评估模型预测概率分布与真实概率分布之间的相似性。结构相似性衡量模型预测与真实风险分布在时空特征结构上的一致性。归一化皮尔逊相关系数衡量模型预测与真实风险分布之间的线性相关性。计算公式如下:

$$\text{CosSim}(y, \hat{y}) = \frac{y \cdot \hat{y}}{\|y\| \|\hat{y}\|} \quad (16)$$

$$\text{SSIM}(y, \hat{y}) = \frac{(2\mu_y \mu_{\hat{y}} + C_1)(2\sigma_y \sigma_{\hat{y}} + C_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + C_1) + (\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2)} \quad (17)$$

$$\text{NPCC}(y, \hat{y}) = \frac{1}{2} \left(\frac{\sum (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum (y_i - \bar{y}_i)^2 \sum (\hat{y}_i - \bar{\hat{y}}_i)^2}} + 1 \right) \quad (18)$$

式中: $\|\cdot\|$ 为欧几里得范数; μ, σ, σ^2 分别为均值、协方差、方差; C_1 和 C_2 为小常数,避免分母为零, $C_1=0.001, C_2=0.001$ 。

2 结果与分析

表 1 单分类高风险食品类别预测性能

Table 1 Performance of high-risk food category prediction for single-class classification

	准确率	召回率	精确率	F_1 值
2019—2022	0.825	0.722	0.770	0.729
2019—2023	0.880	0.809	0.869	0.820
2019—2024($L_{w,MSE}$)	0.924	0.865	0.961	0.894
2019—2024(L_{MSE})	0.876	0.751	0.831	0.769

2.2 风险食品多类别分布

食品风险是多样化的,通常某个时空下会同时发生多种不合格食品的风险事件,因此模型需要能够准确预测多个类别的风险,以提供更全面的安全保障。采用3种多分类指标分析模型预测结果与真实风险分布之间的相似性和一致性。如表2所示,余弦相似度为96.6%,意味着模型能够有效地捕捉到不同风险食品类别的相对重要

2.1 最高风险食品类别预测

通常最高风险食品类别代表着对公众健康和安全的最大威胁,同时也是在特定时空点上发生概率最高的风险类别。准确预测这些类别对于食品安全管理至关重要。为了验证模型在不同时空点的预测能力和误报情况,采用4种单分类指标全面评估模型在重大风险食品类别上的决策质量,模型的预测性能如表1所示。准确率为92.4%,意味着模型在识别高风险食品类别时具有较高的可靠性,有助于食品安全管理者在资源有限的情况下优先处理最严重的风险事件。召回率为86.5%,表明模型能够有效捕捉到大多数的高风险食品事件。如果漏报一个高风险事件就可能导致较为严重的后果,如食物中毒或大规模的健康危机。精确率达到96.1%,表明模型在预测高风险食品类别时误报率较低,从而减少不必要的干预和资源浪费。综合指标 F_1 值为89.4%,说明模型在平衡识别能力和误报控制方面的综合能力较强。这确保了在复杂多变的现实环境中,模型依然能够提供稳定和可信的预测结果。最高风险食品类别预测的指标分析,充分验证了所研究的时空 Transformer 分类预测模型通过精准识别高风险食品类别,能够帮助监管部门有效应对当前时空点最严重的风险事件,减少食品安全事件的发生,降低不必要的经济损失。

表 2 多分类风险食品类别分布预测性能

Table 2 Performance of multi-class risk food category distribution prediction

	CosSim	SSIM	NPCC
2019—2022	0.859	0.795	0.927
2019—2023	0.920	0.875	0.959
2019—2024($L_{w,MSE}$)	0.966	0.923	0.982
2019—2024(L_{MSE})	0.923	0.840	0.960

性和发生概率,有助于在资源分配和风险优先级管理中做出更准确的决策。结构相似性为92.3%,表示模型在时空特征结构上与真实风险分布保持一致,这确保了模型不仅在整体概率分布上准确,而且在特定的时间和空间维度上能准确反映风险的变化和分布,从而支持更精细化的管理和干预措施。归一化皮尔逊相关系数为较高的98.2%,表明该模型能有效识别和反映风险事件的趋势和变化,预测潜在的风险高峰期,为提前制定应对措施提供了依据。综合3种多分类指标的分析,时空Transformer分类预测模型为决策者提供了可靠的数据支持。更为重要的是,表1和表2还展示2019年至2024年间,随着数据量的增加,各项指标呈现增长的趋势。这表明随着数据规模的扩大,模型的预测能力得到增强。随着后续食品抽检工作的持续开展,将实现更为精准的风险类别预测。

2.3 统计与模型预测的对比分析

统计方法在分析食品风险类别时,主要通过计算历史数据的规律来识别风险模式和趋势。为了检验模型预测的有效性,采用统计的方法计算验证数据集中各个时空点的全类别真实概率值。同时,模型通过训练权重得到验证集的全类别预测概率值。表3展示2种方法所计算的绝对误差率对比。结果显示,大部分数据的误差值均低于0.1,特别是有19.15%的数据的误差率是低于1%的。这说明模型在大多数情况下能够提供符合统计分析规律的预测概率值。仅有4.51%的数据的误差率大于等于100%,出现了与统计方法完全不同的预测结果。图4展示以12月份为例的统计方法与模型预测的分类地图。图中每个方格代表一个空间点(经、纬度精确到小数点后一位)。有颜色的方格表示该空间点曾发生食品安全事件,若方格中出现多种颜色,则表示该区域内存在多种概率相同的食品类别。星型标记表示模型预测的类别与统计方法所得类别一致。从图中可以看出,在已发生风险事件的区域,模型预测与统计方法的结果

基本一致。此时,在未发生风险事件的区域,统计方法无法预知风险类别的发生概率。

统计方法虽简单直接,但在处理大规模数据时效率将逐渐降低。更重要的是,它完全局限于历史数据的时空范围,难以有效预测未发生食品安全事件的时空点的风险类别及趋势。然而,在食品安全领域,对这种“空白”地区的预测同样重要。相比之下,时空Transformer分类预测模型不仅学习历史时空点的数据规律,而且能通过分析全局信息,预测“空白”时空点的风险类别。此外,当前风险食品类别的数据分布是不均衡的,尤其是在样本量有限的情况下,样本无法代表总体特征,此时统计分析更容易引入偏差。时空Transformer分类预测模型通过建立全局时空与类别的内关联,在处理动态变化的数据时,具有更强的适应性和预测能力,从而能够更有效应对数据不平衡带来的挑战。

2.4 加权均方误差损失分析

在分类预测模型的训练过程中,损失的训练趋势是评估模型收敛性和稳定性的关键度量。图5展示模型在验证集上的两种损失曲线。在前100个训练轮次内,2种损失值相近,并均呈现快速下降的趋势。然而,随着训练的深入,损失曲线趋于平稳, L_{w_MSE} 显示出更优越的收敛性。这表明 L_{w_MSE} 能够在数据分布不平衡的情况下,更有效地捕捉重要特征。相比之下, L_{MSE} 可能因少数类别样本不足而影响模型在某些类别上的表现。由表1和表2所示,基于 L_{w_MSE} 所获得的各项指标显著高于 L_{MSE} ,这进一步验证了加权均方误差损失在不平衡数据集上的显著优势。

综合上述试验,设计的时空Transformer分类预测模型充分捕捉了时空与风险类别之间的复杂关联,能够基于历史数据预测到全时空点的风险类别。这一重大突破将在食品安全监管中发挥关键作用,为监管部门提供坚实的理论指导和决策支持,为公众健康筑起一道更加坚固的防线。

表3 基于验证集的统计方法与模型预测概率值的绝对误差率对比

绝对误差率	[0%~1%)	[1%~10%)	[10%~20%)	[20%~100%)	≥100%
数量	68	179	43	49	16
百分比/%	19.15	50.42	12.11	13.80	4.51

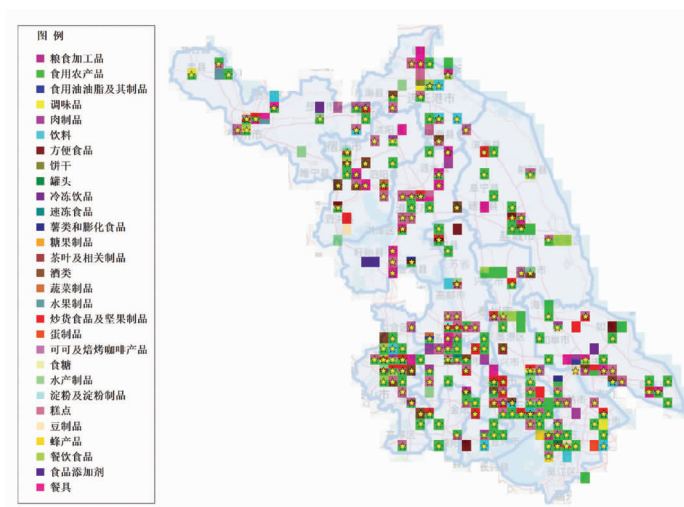


图 4 基于完整历史数据的统计与模型预测的 12 月份风险食品分类地图

Fig.4 Classification map of risk foods in december based on statistical and model predictions from complete historical data

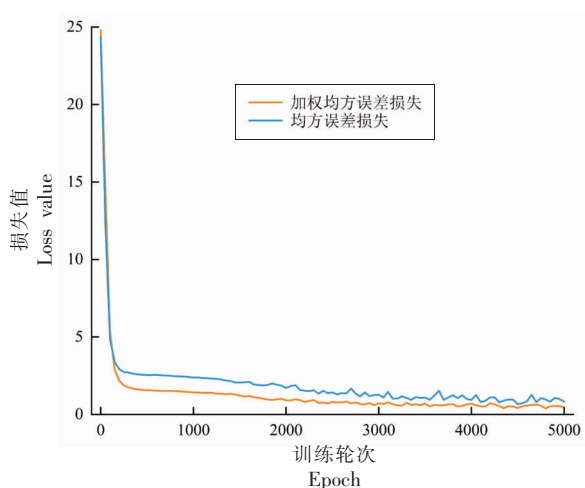


图 5 损失对比曲线

Fig.5 Loss comparison curve

3 结论

本文研究了时空 Transformer 分类预测模型在食品安全监管领域的创新应用, 依托历史抽检数据深度挖掘时空演变规律与风险类别之间的内在关联, 实现了对食品安全风险的精准预判和决策辅助。面对实际场景中普遍存在的数据长尾分布挑战, 模型克服了统计方法在趋势预测中的局限性。随着后续抽检数据量的持续增加, 可持续优化模型的预测性能, 不断提升抽检工作的靶向精准度和监管效率, 为抽检策略的科学制定和优化提供了强有力的技术支撑。未来, 模型将融合气象

监测、检验项目等多源外部数据, 进一步提升其鲁棒性和实用性, 为推动食品安全管理的智能化和精准化发展构建更加安全、高效的食品安全监管体系。

参 考 文 献

- [1] 吴林海, 陈宇环, 尹世久. 中国食品安全战略: 科学内涵、战略目标与实施路径[J]. 江西社会科学, 2022, 42(2): 112-123, 207.
WU L H, CHEN Y H, YIN S J. China's food safety strategy: Scientific connotation, strategic objectives and implementation path[J]. Jiangxi Social Sciences, 2022, 42(2): 112-123, 207.
- [2] 杨竞. 食品安全执法中的监管模式研究进展[J]. 食品科学, 2024, 45(18): 361-366.
YANG J. A review of the regulatory model of food safety enforcement in China[J]. Food Science, 2024, 45(18): 361-366.
- [3] DING J, QIAO P, WANG J X, et al. Impact of food safety supervision efficiency on preventing and controlling mass public crisis[J]. Frontiers in Public Health, 2022, 10: 1052273.
- [4] 陈默, 张景祥, 胡恩华, 等. 基于结构化分析和语义相似度的食品安全事件领域数据挖掘模型[J]. 食品科学, 2021, 42(7): 35-44.
CHEN M, ZHANG J X, HU E H, et al. Data

- mining model for food safety incidents based on structural analysis and semantic similarity[J]. *Food Science*, 2021, 42(7): 35–44.
- [5] 李佳敏. 食品安全事件修复策略对消费者报复行为的影响机制研究[D]. 无锡: 江南大学, 2022.
- LI J M. Research on the influence mechanism of food safety incident repair strategy on consumer's retaliation behavior[D]. Wuxi: Jiangnan University, 2022.
- [6] 刘世娟. 食品安全抽检数据质量问题分析及提升建议[J]. *现代食品*, 2023, 29(17): 169–172.
- LIU S J. Analysis of quality issues in food safety sampling data and suggestions for improvement[J]. *Modern Food*, 2023, 29(17): 169–172.
- [7] 曹慧, 王瑾. 基层食品安全监管存在的问题及对策分析[J]. *食品安全导刊*, 2024(25): 31–33.
- CAO H, WANG J. Analysis on the problems and countermeasures of food safety supervision at the grassroots level[J]. *China Food Safety Magazine*, 2024(25): 31–33.
- [8] 褚红英. 论食品监督抽检对食品安全监管的重大意义[J]. *现代食品*, 2024, 30(10): 145–147.
- CHU H Y. The significance of food supervision and sampling for food safety supervision[J]. *Modern Food*, 2024, 30(10): 145–147.
- [9] LEI Z. Research on food safety supervision[C]. *Proceedings of the E3S Web of Conferences*. Paris: EDP Sciences, 2020, 185: 04066.
- [10] YANG J, LIN Y. Study on evolution of food safety status and supervision policy — A system based on quantity, quality, and development safety[J]. *Sustainability*, 2019, 11(23): 6656.
- [11] 曹静. 食品抽检中基于大数据分析的食品安全评估研究[J]. *食品安全导刊*, 2023(10): 173–176.
- CAO J. Study on food safety evaluation based on big data analysis in food sampling inspection[J]. *China Food Safety Magazine*, 2023(10): 173–176.
- [12] 曹慧, 王瑾. 基层食品安全监管存在的问题及对策分析[J]. *食品安全导刊*, 2024(25): 31–33.
- CAO H, WANG J. Analysis on the problems and countermeasures of food safety supervision at the grassroots level[J]. *China Food Safety Magazine*, 2024(25): 31–33.
- [13] 张国. 食品安全监督抽检中的常见问题及对策研究[J]. *中国食品*, 2024(12): 70–72.
- ZHANG G. Study on common problems and countermeasures in food safety supervision and sampling inspection[J]. *China Food*, 2024(12): 70–72.
- [14] 张聪. 市场监管总局印发《关于全面深化“双随机、一公开”监管规范涉企行政检查服务高质量发展的意见》[J]. *食品安全导刊*, 2024(21): 5.
- ZHANG C. The State Administration For Market Regulation issued the opinions on comprehensively deepening the 'double random, one open' regulation to standardize enterprise-related administrative inspections and serve high-quality development[J]. *China Food Safety Magazine*, 2024(21): 5.
- [15] 毛磊. 我国基层食品抽检工作存在的问题及解决方法探究[J]. *食品安全导刊*, 2023(25): 1–3.
- MAO L. The existing problems and solutions of basic food sampling inspection in China[J]. *China Food Safety Magazine*, 2023(25): 1–3.
- [16] 高超, 吴遥. 食品抽检靶向性问题探讨[J]. *现代食品*, 2021(1): 16–18.
- GAO C, WU Y. Discussion on the targeting of food sampling[J]. *Modern Food*, 2021(1): 16–18.
- [17] 杨勇. 基于靶向目标优化的某市食品监管抽检计划管理系统[D]. 广州: 广东工业大学, 2019.
- YANG Y. A management system of optimizing a city's food supervision and sampling plan base on targeted goals[D]. Guangzhou: Guangdong University of Technology, 2019.
- [18] 王星云, 左敏, 肖克晶, 等. 基于BP神经网络的食品安全抽检数据挖掘[J]. *食品科学技术学报*, 2016, 34(6): 85–90.
- WANG X Y, ZUO M, XIAO K J, et al. Data mining on food safety sampling inspection data based on BP neural network[J]. *Journal of Food Science and Technology*, 2016, 34(6): 85–90.
- [19] 柳亚飞, 查怀忠. 2022—2023年江苏省Y市食品安全监督抽检不合格情况分析和建议[J]. *食品安全导刊*, 2024(16): 43–45, 58.
- LIU Y F, ZHA H Z. Unqualified situation analysis and suggestions of food safety supervision and sampling inspection in Y City, Jiangsu Province from 2022 to 2023[J]. *China Food Safety Magazine*, 2024(16): 43–45, 58.
- [20] CHEN T C, YU S Y. Research on food safety sampling inspection system based on deep learning[J]. *Food Science and Technology*, 2021, 42: e29121.
- [21] MAKRIDIS G, MAVREPIS P, KYRIAZIS D. A

- deep learning approach using natural language processing and time-series forecasting towards enhanced food safety[J]. *Machine Learning*, 2023, 112(4): 1287–1313.
- [22] 袁刚, 郭爽, 唐琦, 等. 基于深度学习的食品安全风险知识图谱构建方法[J]. *质量安全与检验检测*, 2023, 33(5): 60–65.
- YUAN G, GUO S, TANG Q, et al. Construction method of food safety risk knowledge graph based on deep learning[J]. *Quality Safety Inspection and Testing*, 2023, 33(5): 60–65.
- [23] 胡昊天, 吉晋锋, 王东波, 等. 基于深度学习的食品安全事件实体一体化呈现平台构建[J]. *数据分析与知识发现*, 2021, 5(3): 12–24.
- HU H T, JI J F, WANG D B, et al. An integrated platform for food safety incident entities based on deep learning[J]. *Data Analysis and Knowledge Discovery*, 2021, 5(3): 12–24.
- [24] MAKRIDIS G, MAVREPIS P, KYRIAZIS D, et al. Enhanced food safety through deep learning for food recalls prediction [C]. *Discovery Science: 23rd International Conference, DS 2020, Thessaloniki, Greece, Proceedings 23*. Berlin: Springer International Publishing, 2020: 566–580.
- [25] ZUO S M, JIANG H M, LI Z C, et al. Transformer hawkes process[C]. *Proceedings of the International Conference on Machine Learning*. New York: PMLR, 2020: 11692–11702.
- [26] NIEMEYER G. Tips & tricks about geohash[J]. *Int Res J Eng Technol*, 2008, 10: 95–102.
- [27] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]. *Proceedings of the Advances in Neural Information Processing Systems*. New York: Curran Associates, Inc., 2017: 5998–6008.

Studies on Classification Prediction of Food Supervision Sampling Inspection Based on Spatio-temporal Transformer Model

Luo Xiaoqing^{1,2}, Guo Lin^{1,2}, Yang Yumeng^{1,2}, Huang Naiyun^{1,2}, Wu Xiaojun^{1,2}

(¹*School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, Jiangsu*

²*Sino-UK Joint Laboratory on Artificial Intelligence of Ministry of Science and Technology, Jiangnan University, Wuxi 214122, Jiangsu*)

Abstract Objectives: This study aimed to develop a multidimensional analytical model for spatiotemporal trends and risk food categories, with the objective of achieving efficient early warning of food safety risks and providing a scientific basis and decision support for the intelligent and precise management of supervision. Methods: Based on the food sampling data from Jiangsu Market Supervision and Administration Bureau from 2019 to 2024, a spatio-temporal Transformer based multi-class prediction model was proposed. This model deeply explored the evolution patterns of food safety incidents in both temporal and spatial dimensions and established intrinsic associations with risk categories. To mitigate the long-tail distribution problem caused by disparities in data volume across multiple categories, a weighted mean squared error loss function was introduced to optimize model training, thereby enhancing the sensitivity of the model to tail categories. Results: The accuracy of the model was comprehensively evaluated using four high-risk category classification metrics and three all-category risk classification metrics. It not only achieved prediction accuracy comparable to statistical methods in regions where risks had occurred, but also efficiently predicted all-category risks across the entire spatiotemporal scope. Conclusions: The spatio-temporal Transformer multi-class prediction model provides a novel and effective tool for food safety regulatory authorities, enabling the optimization of random inspection strategies and enhancing regulatory efficiency. **Keywords** spatio-temporal Transformer; weighted mean squared error; food sampling inspection; risk early warning; multi-class prediction; food safety