

福建省海洋经济发展综合评估与分析

蔡晓倩, 林 翊

(福建师范大学 经济学院, 福州 350108)

摘要:福建省海洋经济发展综合指数评价过程中往往存在指标的多重共线性以及回归的非线性问题,这都可能导致模型评价结果无法客观反映实际经济发展规律。将海洋 GDP 作为被解释变量来替换传统综合指数,运用随机森林算法来减少建模过程中指标的共线性影响,建立了 4 种多层感知神经网络模型来验证回归任务中的非线性关系。结果表明:福建省海洋经济发展水平同构建的子系统指标之间存在显著的非线性关系;随机森林算法同多层感知神经网络模型结合后构建的回归模型能够很好地拟合福建省海洋经济发展状况,且福建省海洋经济在科技创新和生态环境方面存在明显的短板。

关键词:福建省海洋经济;随机森林;多层感知机;评价体系

中图分类号:P74 **文献标志码:**A **文章编号:**1671-1807(2023)16-0096-06

对福建省海洋经济发展进行综合评估分析,探究海洋经济发展的总体情况和内部各子系统之间的关系,并进一步探索海洋经济发展的主要驱动因素,将有助于准确把握海洋经济的可持续发展状态。

目前学术界对海洋经济发展水平的研究通常采用主客观赋权评价法来对构建的综合指数进行度量与分析,如鲁亚运等^[1]的信息熵权法、Song 等^[2]的德尔非法、朱坚真崔曦文^[3]的 TOPSIS 熵权法、赵晖等^[4]的层次分析法以及程曼曼等^[5]的全局熵权法,然而海洋经济发展综合指数编制过程中指标之间往往存在多重共线性问题,且海洋经济发展同评价指标之间通常存在复杂的非线性关系,简单地使用线性模型去构建发展指数可能无法客观地反映实际经济规律。值得关注的是,随着大数据时代的到来,机器学习中很多算法能够很好地克服回归过程中的非线性问题^[6],但已有的文献仅限于在回归模型中寻找指标与构造的线性指数的拟合关系^[7-9],根源上还是没有消除综合指数的共线性问题。

基于此,从四个方面进行探索:①为了避免综合指数编制过程中存在的共线性问题,区别于以往普遍将融合的综合指数作为被解释变量进行分析,直接将海洋 GDP 作为被解释变量来反映福建省海洋经济发展水平。②借鉴《福建省“十四五”海洋强

省建设专项规划》中的子系统设置,从 6 个维度来评估福建省海洋经济发展水平,并借助随机森林算法来规避指标间的多重共线性问题。此外,通过该算法进一步测算出评价体系中各指标的贡献度,以此来探索福建省海洋经济发展的驱动因素。③区别于常用的线性回归模型,借助 4 种多层感知机模型来验证并拟合不同贡献度阈值下筛选出来的评价指标与海洋经济发展水平之间的非线性关系,由此可以得到更为科学合理的福建省海洋经济发展评估体系。④借助实证结果对福建省海洋经济发展驱动因素进行更深层次的分析,对发展的薄弱环节提出对策建议,以促进福建省海洋经济综合实力的进一步提升。

1 指标选择与数据处理

1.1 指标选择

指标选择将遵循综合指标体系构建原则^[10],基于《福建省“十四五”海洋强省建设专项规划》中的子系统设置,依据目的性、完备性、可操作性、显著性与动态性的设计原则,进一步从海洋科技创新、海洋对外开放、海洋产业体系、海洋生态环境、海洋基础设施以及海洋社会民生 6 个层面(系统层)来评估福建省海洋经济发展水平。构建的初始海洋经济发展评价指标体系如表 1 所示。

收稿日期:2023-04-14

基金项目:福建省高校人文社科研究基地互联网创新研究中心基金(IIRC20190101);致公党福州市委会有关问题调研课题(B20ESH20136A06)。

作者简介:蔡晓倩(1995—),女,福建泉州人,福建师范大学经济学院,硕士研究生,研究方向为海洋产业经济;林翊(1973—),男,福建福州人,福建师范大学经济学院,副教授,硕士研究生导师,博士,研究方向为海洋产业经济。

表1 福建省海洋经济发展评级指标体系

系统层	指标层		
科技创新系统	投入	X ₁ 科技创新投入	
		X ₂ 海洋科研机构数	
		X ₃ 海洋科研机构课题数	
	产出	X ₄ 海洋科技著作发表数	
		X ₅ 海洋专利申请受理	
	人才资源	X ₆ 海洋科技研究人员	
		X ₇ 海洋相关毕业生人数(专科及以上)	
		X ₈ 科研机构中研究生以上学历人员占比	
生态环境系统	资源	X ₉ 海洋捕捞量	
		X ₁₀ 海盐产量	
		X ₁₁ 海洋矿业产量	
		X ₁₂ 海水养殖产量	
	压力	X ₁₃ 废水排放	
		X ₁₄ 固体废弃物	
		X ₁₅ 海洋碳排放	
		X ₁₆ 单位岸线灾害损失	
	响应	X ₁₇ 海洋自然保护区面积比重	
		X ₁₈ 海洋自然保护区个数	
		X ₁₉ 环保投入总额	
		X ₂₀ 固体废弃物综合利用量	
	产业体系系统	经济效率	X ₂₁ 海洋生产总值增长率
			X ₂₂ 海洋产业增长弹性系数
		经济结构	X ₂₃ 海洋产业结构转型升级规模
			X ₂₄ 海洋产业结构优化规模
X ₂₅ 第三产业增产率			
对外开放系统	对外贸易	X ₂₆ 国际旅游外汇收入	
		X ₂₇ 进出口总额	
		X ₂₈ 沿海地区旅行社数量	
		X ₂₉ 实际利用外商投资金额	
	客货流动	X ₃₀ 接待入境游客人数外国人	
		X ₃₁ 海洋货物远洋运量	
基础设施系统	交通基础设施	X ₃₂ 铁路运营密度	
		X ₃₃ 公路运营密度	
	港口基础设施	X ₃₄ 港口生产用码头泊位	
		X ₃₅ 港口货物吞吐量	
基础设施系统 社会民生系统	数字基础设施	X ₃₆ 互联网普及率	
		X ₃₇ 电话普及率	
		X ₃₈ 长途光缆线路密度	
	就业水平	X ₃₉ 移动电话基站	
X ₄₀ 沿海地区就业水平			
社会民生系统	生活水平	X ₄₁ 渔民人均纯收入	
		X ₄₂ 人均海水产品产量	

1.2 数据处理

1.2.1 数据来源

数据主要来源于2006—2019年《中国海洋统计年鉴》《中国能源统计年鉴》《福建省统计年鉴》《福建省国民经济和社会发展统计公报》《中国港口统计年鉴》以及福建省文化和旅游厅等相关资料。其中,缺失数据通过K近邻算法处理,即将数据集中每一个记录进行分类,最后通过K近邻算法训练模型预估缺失值。

1.2.2 训练集与测试集划分

基于矢量均值相似性的样本分割法,提出了一种适用于回归任务的样本集分解方法,可以在划分训练集和测试集时使得两者在统计意义上保持一致性。

第一步:假设采集到的样本集为 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, 其中 y_i 为第 i 个样本的被解释变量(因变量 GDP), m 为样本数。

第二步:进一步求出样本集 D 中样本示例 $x_i = (x_{i1}, x_{i2}, \dots, x_{ik})$ 的均值向量 $\mu = (\mu_1, \mu_2, \dots, \mu_k)$, 其中下标 k 为指标个数;计算出每个样本示例 x_i 与均值向量 μ 两者的余弦相似度 $\cos\theta_i = \mu^T x_i / (\sqrt{\mu^T \mu} \sqrt{x_i^T x_i})$, $\cos\theta_i$ 越接近于1,样本示例 x_i 与 μ 越相似。

第三步:按照样本余弦相似度对所有的样本从小到大进行排序;之后,将 D 中排序好的所有样本分成 $\frac{m}{3}$ 段,每段包含3个样本,抽取每段的第2个样本组成测试集 T ,剩下的样本组成单标签训练集 S 。

1.2.3 数据标准化处理

为消除数据中不同代谢指标的量纲和值域差异,使用Z-score标准化方法处理,使指标数据在0与1之间取值。

$$\hat{x} = \frac{x - \mu}{\sigma} \quad (1)$$

2 研究方法

2.1 随机森林算法

随机森林(random forest, RF)是由Breiman于2001年提出的一种监督学习算法^[11]。实际应用中,RF结合了多个决策树基学习器,借助于集成学习方法来为复杂的分类与回归问题提供解决方案,且通常具有如下优点:在许多主流公共数据集上,相对其他算法RF显示了较好的性能;RF并不需要通过降维方式来处理高维数据的多重共线性问题,训练完成后,可以输出每个特征的重要度;RF对泛化误差(generalization error)使用的是无偏估计,模型泛化能力强,训练速度快,较易实现并行化处理方法。随机森林回归流程如图1所示。

2.2 多层感知机模型

多层感知机(multilayer perceptrons, MLP)主要由输入层,隐藏层和输出层构成,且不同层之间是全连接的。MLP网络结合梯度下降算法,通过多轮的前向和反向传播过程实现参数(权值、阈值与偏置项)的更新,且对任何连续非线性结构函数都具有较好的拟合效果。采用Kolmogorov定理^[12]来

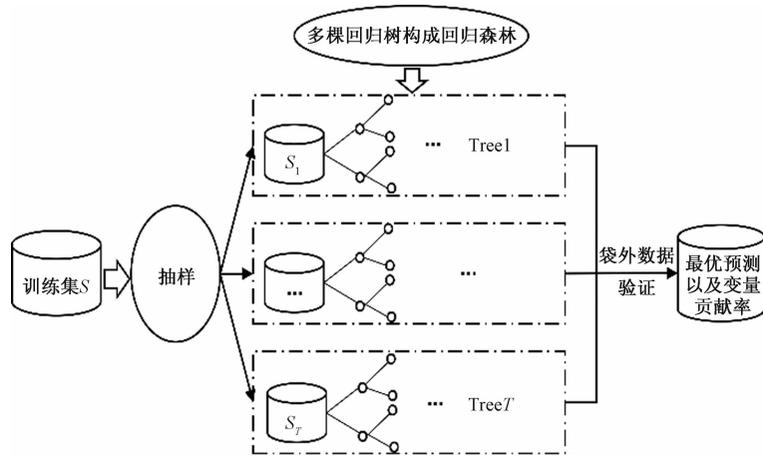


图 1 随机森林特征变量贡献率计算流程

确定网络结构。该定理指出给定任一连续函数 $F: U^d \rightarrow H^r$, 即 $F(X) = Y$, 其中 U 为闭区间 $[0, 1]$, 那么 F 可以精确地用一个三层前向网络实现。此网络的输入层有 q 个神经元, 隐层有 $2q + 1$ 个神经元, 输出层有 1 个神经元, 如图 2 所示。

多层感知机建模流程如下:

首先, 将输入解释变量提供给输入层神经元, 然后将信号逐层向前传, 直到输出层产生预测结果。

其次, 计算输出层的均方误差, 再将误差逆向传播至隐层神经元, 然后根据隐层神经元的误差来对连接权和阈值进行优化(学习率设置为 $\eta=0.1$)。

最后, 该迭代过程循环进行, 直到累计均方误差 $E < \epsilon (\epsilon = 0.1)$ 为止, 最终产生的最优化连接权与阈值构成了非线性回归函数的系数。

3 实证结果与分析

3.1 运用随机森林筛选指标重要度

使用福建省海洋 GDP 作为随机森林算法中的目

标值, 并运用随机森林算法对所有数据进行训练, 随机森林的一个重要功能是测算特征变量的重要度, 度量单个特征的重要度主要是通过计算变更该特征排序后整个森林袋外数据^[13] (out of bag, OOB) 均方误差均值 $[OOBMSE(X_j, \text{permuted})]$ 与不变更该特征排序整个森林 OOB 均方误差均值 (OOBMSE) 的差值来衡量的, 即 $OOBMSE(X_j, \text{permuted}) - OOBMSE$, 详细表达式如下:

$$OOBMSE(X_j, \text{permuted}) = \frac{1}{T} \sum_{i=1}^T \frac{1}{n_{OOB_i}}$$

$$\sum_{i=1}^n \sum_{i \in OOB_i} [y_i - \hat{y}_{i,t}(X_j, \text{permuted})]^2 \quad (2)$$

$$OOBMSE = \frac{1}{T} \sum_{i=1}^T \frac{1}{n_{OOB_i}} \sum_{i \in OOB_i} (y_i - \hat{y}_{i,t})^2 \quad (3)$$

当特定变量 X_j 的重要性越大, 则其求出来的 $OOBMSE_i(X_j, \text{permuted}) - OOBMSE_i$ 的差就越大。通过 Python 仿真可知最优(袋外误差最小)决策树数量为 39, 如图 3 所示。

在最优决策树数量下测算出各指标的贡献度, 如表 2 所示。

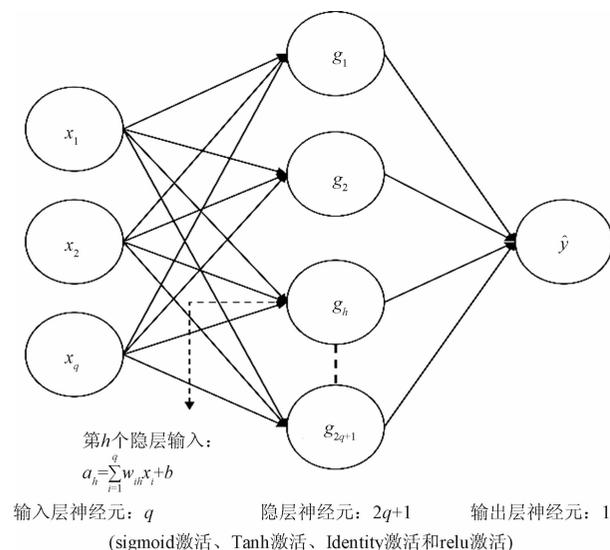


图 2 多层感知器回归结构

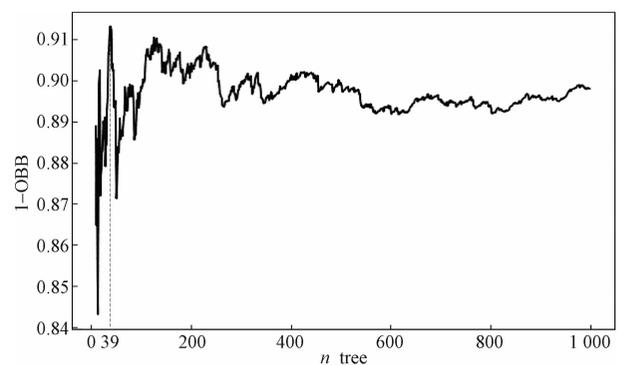


图 3 OOB 袋外数据错误率估计

表 2 指标贡献度

指标	贡献度/%
X ₂₄ 海洋产业结构优化规模	10.692
X ₂₆ 国际旅游外汇收入	8.909
X ₂₃ 海洋产业结构转型升级规模	6.280
X ₄₂ 人均海水产品产量	6.107
X ₃₀ 接待入境游客人数外国人	6.106
X ₈ 科研机构中研究生以上学历人员占比	5.113
X ₇ 海洋相关毕业生人数(专科及以上)	4.874
X ₃₉ 移动电话基站	4.782
X ₃₂ 铁路运营密度	4.684
X ₄₁ 渔民人均纯收入	4.442
X ₃₄ 港口生产用码头泊位	4.202
X ₁₇ 海洋自然保护区面积比重	3.158
X ₃₃ 公路运营密度	2.918
X ₅ 海洋专利申请受理	2.835
X ₁₂ 海水养殖产量	2.811
X ₃₈ 长途光缆线路密度	2.789
X ₄₀ 沿海地区就业水平	2.577
X ₃₅ 港口货物吞吐量	2.452
X ₂₉ 实际利用外商投资金额	2.277
X ₂ 海洋科研机构数	2.276
X ₃₇ 电话普及率	1.175
X ₂₇ 进出口总额	1.142
X ₁₈ 海洋自然保护区个数	1.056
X ₁₅ 海洋碳排放	0.748
X ₉ 海洋捕捞量	0.694
X ₃₇ 互联网普及率	0.641
X ₃₁ 海洋货物远洋运量	0.572
X ₁₃ 废水排放	0.496
X ₂₈ 沿海地区旅行社数量	0.463
X ₂₅ 第三产业增产量	0.441
X ₁₁ 海洋矿业产量	0.433
X ₁₄ 固体废弃物	0.427
X ₃ 海洋科技机构课题数	0.367
X ₁₉ 环保投入总额	0.297
X ₂₁ 海洋生产总值增长率	0.268
X ₁ 海洋科研机构经费投入	0.131
X ₁₀ 海盐产量	0.094
X ₂₀ 固体废弃物综合利用量	0.092
X ₁₆ 单位岸线灾害损失	0.071
X ₄ 海洋科技著作发表数	0.067
X ₆ 海洋科技研究人员	0.035
X ₂₂ 海洋产业增长弹性系数	0.007

从表 2 可以看出,由随机森林算法训练出来的指标贡献度有较大差别。从单个指标贡献度排序来看,排名前 5 的分别是海洋产业结构优化规模(X₂₄)、国际旅游外汇收入(X₂₆)、海洋产业结构转型升级规模(X₂₃)、人均海水产品产量(X₄₂)和接待入境游客人数外国人(X₃₀),其中指标 X₂₄和 X₂₃属于海洋产业体系系统,总贡献率为 16.97%,指标 X₂₆和 X₃₀属于对外开放系统,总贡献率为 15.02%,指标 X₄₂属于社会民生系统,贡献率为 6.11%。由此可以看出,海洋产业结构、海洋对外开放和社会民生对福建省海洋经济发展发挥着重要推动作用。

从指标体系系统层来看,当阈值设定为 0.01(贡献度大于 1%)时筛选出来的指标体系为科技创新系统(X₂、X₅、X₇、X₈合计贡献 15.10%)(X₁₂、X₁₇、X₁₈合计贡献 7.03%)、产业体系系统(X₂₃、X₂₄合计贡献 16.97%)、对外开放系统(X₂₆、X₂₇、X₂₉、X₃₀合计贡献 18.43%)、基础设施系统(X₃₂、X₃₃、X₃₄、X₃₅、X₃₇、X₃₈、X₃₉合计贡献 23.00%)和社会民生系统(X₄₀、X₄₁、X₄₂合计贡献 13.13%);当阈值设定为 0.02(贡献度大于 2%)时筛选出来的指标体系为科技创新系统(X₂、X₅、X₇、X₈合计贡献 15.10%)、生态环境系统(X₁₂、X₁₇合计贡献 5.97%)、产业体系系统(X₂₃、X₂₄合计贡献 16.97%)、对外开放系统(X₂₆、X₂₉、X₃₀合计贡献 17.29%)、基础设施系统(X₃₂、X₃₃、X₃₄、X₃₅、X₃₈、X₃₉合计贡献 21.83%)和社会民生系统(X₄₀、X₄₁、X₄₂合计贡献 13.13%);阈值为 0.03(贡献度大于 3%)时筛选出来的指标体系为科技创新系统(X₇、X₈合计贡献 9.99%)、生态环境系统(X₁₇合计贡献 3.16%)、产业体系系统(X₂₃、X₂₄合计贡献 16.97%)、对外开放系统(X₂₆、X₃₀合计贡献 15.01%)、基础设施系统(X₃₂、X₃₄、X₃₉合计贡献 13.67%)和社会民生系统(X₄₁、X₄₂合计贡献 10.55%)。

3.2 多层感知神经网络回归验证

考虑到回归问题中可能存在的非线性问题,将采用 4 种(identity 激活、relu 激活函数、sigmoid 激活和 tanh 激活)多层感知神经网络来进一步验证筛选出来的指标体系合理性。其中,随机森林算法的阈值分别设定为 0.01、0.02 和 0.03。此外,为了更直观地看出训练过程是否存在过拟合问题,将训练集以及测试集的衡量指标进行了分开度量。验证结果表 3 所示。

由表 3 可知:①不同阈值下 identity 激活函数的拟合性能基本都能达到最优,拟合优度最高能达到 0.99,这验证了筛选出来的评价指标与实际海洋经济发展存在非线性关系,且表明了非线性感知神经网络具备良好的拟合效果;②比较不同激活函数在贡献度大于 1%、2% 和 3% 指标体系中测试集的均方误差(MSE)、平均绝对误差(MAE)和拟合优度(R²)可知,贡献度大于 3% 的指标体系拟合结果最好,说明随机森林算法按照阈值为 0.03 时选出来的指标体系对福建省海洋经济发展水平具有较强的解释能力;③相同阈值相同激活函数下训练集与测试集在同一度量指标的差距并不大,进一步验证了感知神经网络模型没有出现拟合问题。

表 3 多层感知机网络验证回归结果比较

3 种不同贡献度 阈值下的指标体系	神经网络激活函数	测试集预测			训练集预测		
		MSE ↓	MAE ↓	R ² ↑	MSE ↓	MAE ↓	R ² ↑
贡献度大于 3% 指标体系	identity	0.012	0.084	0.992	0.017	0.112	0.978
	sigmoid	1.425	1.045	-0.046	0.767	0.728	0.011
	tanh	0.181	0.358	0.867	0.081	0.240	0.895
	relu	0.048	0.174	0.965	0.023	0.115	0.970
贡献度大于 2% 指标体系	identity	0.194	0.401	0.776	0.014	0.089	0.987
	sigmoid	0.818	0.750	0.055	0.941	0.833	0.108
	tanh	0.198	0.382	0.771	0.039	0.123	0.963
	relu	0.150	0.305	0.826	0.031	0.129	0.971
贡献度大于 1% 指标体系	identity	0.062	0.168	0.942	0.015	0.105	0.984
	sigmoid	1.210	1.032	-0.146	0.832	0.779	0.061
	tanh	0.187	0.306	0.823	0.028	0.144	0.968
	relu	0.125	0.290	0.882	0.016	0.105	0.982

注: MSE 为均方误差, MAE 为平均绝对误差, R² 为拟合优度, ↓ 代表越小越好, ↑ 代表越大越好。

3.3 福建省海洋经济发展驱动因素分析

由表 2 与表 3 可知, 当设定阈值为 0.03 时, 筛选出来的指标体系中科技创新系统贡献率为 9.987 3%、生态环境系统贡献率为 3.157 9%、产业体系系统贡献率为 16.972 4%、对外开放系统贡献率为 15.014 7%、基础设施系统贡献率为 13.668 2% 以及社会民生系统贡献率为 10.548 7%。从各子系统的贡献度看, 产业体系系统、对外开放系统和基础设施系统对福建省海洋经济发展水平具有较大的贡献, 其次是社会民生系统和科技创新系统, 最后是生态环境系统, 这也间接表明了福建省海洋经济的发展在科技创新与生态环保上存在明显的短板效应。从各子系统的核心驱动因素看, 指标 X₂₄ (海洋产业结构优化规模) 在产业体系系统中发挥着主要驱动作用, 指标 X₂₆ (国际旅游外汇收入) 是对外开放系统主要驱动因素, 指标 X₃₂、X₃₄ 和 X₃₉ (分别是铁路运营密度、港口生产用码头泊位和移动电话基站) 均对基础设施系统产生重要的推动作用, 指标 X₄₂ (人均海水产品产量) 是社会民生系统的驱动因素, 指标 X₇、X₈ (海洋相关毕业生人数(专科及以上) 和科研机构中研究生以上学历人员占比) 是科技创新系统的主要动力因素和指标 X₁₇ (海洋自然保护区面积比重) 是生态环境中最重要的影响因素, 在实际经济活动与政策制定中, 可以着重观察这些指标的动态变化情况。

4 结论与建议

为分析福建省海洋经济发展水平, 寻找福建省海洋经济发展的驱动因素, 依据评价指标的构建原则, 在借鉴前人的研究的基础上, 构建了 6 个子系统 42 个指标的福建省海洋经济发展水平初始评价模型。首先, 通过袋外数据误差率寻找到最优决策树

数量, 并运用随机森林算法对初始评价体系进行测算, 规避了指标之间存在的共线性问题, 统计计算后进一步得到了福建省海洋经济发展在各个子系统的主要驱动因素以及各子系统的贡献度。接着, 依据不同的贡献度阈值(0.01、0.02 和 0.03) 筛选出了 3 种不同的海洋经济评价指标体系。随后, 借助 4 种多层感知机网络对筛选出来的评价体系进行回归拟合, 验证了指标与经济发展水平之间的非线性关系, 且得出了贡献度阈值为 0.03 时的指标评价体系能够更好地诠释福建省海洋经济的发展状况。

上述实证结果表明, 福建省海洋科技创新系统和海洋生态环境系统存在明显的短板效应, 由此提出如下的建议。

1) 针对福建省海洋科技创新水平较低问题。一是打造良好的科技创新环境。从省、企、校三方同时推进, 加快建设海洋领域省创新实验室, 鼓励与支持涉海企业与“一带一路”沿岸国家或地区企业共建联合共建科技创新平台, 推进厦门大学、福州大学和福建师范大学等高校涉海实验室建设。二是解决海洋人才引进和培养问题。鼓励和引导涉海企业及高校等平台引进海洋领域的高层次人才, 集中力量支持福建省现有涉海专业的发展, 支持各高校做强涉海学科。三是发挥政府的主导作用。完善并出台人才引进、科研项目资金支出等相关政策措施, 提升科研成果的转化率。如山东省对海洋领域拥有关键核心技术、全职引进的海外顶尖人才团队, 省级财政给予最高 5 000 万元综合资助或 6 000 万元直投股权投资支持。

2) 针对福建省海洋生态环境水平较低问题。一是鼓励渔业从捕捞和养殖并举向养殖为主转型。

渔业的过度捕捞,使得近海的渔业资源已几乎枯竭,渔民作业的范围逐渐向远洋靠拢,对渔业资源的生态链条产生了巨大的影响。二是加快推进新型海洋环境监测设施的完善。利用当前的大数据、5G、云计算等先进数字技术,推动环保数字产业化,建立安全、高效的海洋环境监测体系,进一步加强对沿海地区的固废处理、环境修复和环境监测,进而达到保护海洋生态环境的目的。三是加强政府的法治监管力度。围绕海洋保护地建立、日常管护、科研监测等工作,加快出台海洋保护地管理的相关法律法规和制度规章,使海洋保护地的管理有法可依、有章可循,预防出现“建而不管,管而不力”的情形。

参考文献

- [1] 鲁亚运,原峰,李杏筠.我国海洋经济高质量发展评价指标体系构建及应用研究——基于五大发展理念的视角[J].企业经济,2019,38(12):122-130.
- [2] SONG W, GUO J, WANG Y, et al. An evaluation of the development of the ocean economy in china using an ocean economic development index [J]. Marine Policy, 2021, 132:104691.
- [3] 朱坚真,崔曦文.基于熵权 TOPSIS 模型的中国海洋综合实力评价[J].科技和产业,2022,22(9):283-289.
- [4] 赵晖,张文亮,张靖苓.天津海洋经济高质量发展内涵与指标体系研究[J].中国国土资源经济,2020,33(6):10-15.
- [5] 程曼曼,陈伟,杨蕊.我国海洋经济高质量发展指标体系构建及时空分析——基于海洋强国战略背景[J].资源开发与市场,2022,38(1):8-15.
- [6] 周亮,蒋练.机器学习驱动的大类资产因子配置研究[J].金融发展研究,2022(8):55-63.
- [7] 蹇令香,苏宇凌,曹珊珊.数字经济驱动沿海地区海洋产业高质量发展研究[J].统计与信息论坛,2021,36(11):28-40.
- [8] 张贺,范梦莹.基于 Lasso 回归模型的青岛海洋经济和海洋产业分析[J].海洋开发与管理,2022,39(8):22-28.
- [9] GAO S, SUN H, WANG J, et al. Evaluation and countermeasures of high-quality development of china's marine economy based on pso-svm [J]. Sustainability, 2022, 14(17):10749.
- [10] 彭张林,张爱萍,王素凤,等.综合评价指标体系的设计原则与构建流程[J].科研管理,2017,38(S1):209-215.
- [11] GAO W, XU F, ZHOU Z H. Towards convergence rate analysis of random forests for classification [J]. Artificial Intelligence, 2022, 313:103788.
- [12] SCHMIDT-HIEBER J. The Kolmogorov-Arnold representation theorem revisited [J]. Neural Networks, 2021, 137:119-126.
- [13] 周志华.机器学习[M].北京:清华大学出版社,2016.

Comprehensive Assessment and Analysis of Marine Economic Development in Fujian Province

CAI Xiaoqian, LIN Yi

(School of Economics, Fujian Normal University, Fuzhou 350108, China)

Abstract: In the evaluation process of the comprehensive index of marine economic development in Fujian Province, there are often multicollinearity of indicators and nonlinear problems of regression, which may lead to the model evaluation results not being able to objectively reflect the actual economic development law. The traditional composite index with ocean GDP was replaced as an explanatory variable, the random forest algorithm was used to reduce the collinear influence of the index in the modeling process, and four multilayer perceptron network models to verify the nonlinear relationship in the regression task was established. The results show that there is a significant nonlinear relationship between the development level of marine economy in Fujian Province and the subsystem indicators constructed. The regression model constructed by combining the random forest algorithm with the multilayer perceptron can fit the development of the marine economy in Fujian Province, and there are obvious shortcomings in scientific and technological innovation and ecological environment of the marine economy of Fujian Province.

Keywords: marine economy of Fujian Province; random forest; multilayer perceptron; evaluation system