Advances in Applied Mathematics 应用数学进展, 2024, 13(1), 199-207
Published Online January 2024 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2024.131023

Hans汉斯

基于宏基因组分析的机器学习疾病预测模型构建

张钰东

青岛大学数学与统计学院，山东 青岛

收稿日期：2023年12月17日；录用日期：2024年1月11日；发布日期：2024年1月17日

摘　要

随着高通量测序技术的发展，宏基因组数据库得到了极大的丰富，为利用其分析人类疾病与健康状况提供了可能，其中基于人类肠道微生物组分析的疾病预测成为了代表性研究方向之一。本文利用以门为单位的分类学肠道微生物数据，即操作分类单元数据，结合非负矩阵分解和变分自动编码器方法，提出了两类新的机器学习分类算法，这些算法旨在提取肠道微生物中的关键信息，以实现对疾病患者的预测。通过降维、数据生成以及引入惩罚约束项等技术手段，我们改善了预测效果、优化了模型的过拟合。在模拟数据、肝硬化数据和糖尿病数据上，我们的预测模型均表现出了较好的性能，AUC值分别达到了0.926、0.956和0.745。

关键词

操作分类单元，非负矩阵分解，变分自动编码器

Construction of Machine Learning Disease Prediction Model Based on Macro-Genomic Analysis

Yudong Zhang

School of Mathematics and Statistics, Qingdao University, Qingdao Shandong

Received: Dec. 17th, 2023; accepted: Jan. 11th, 2024; published: Jan. 17th, 2024

Abstract

With the advancements in high-throughput sequencing technologies, the macro-genomic databases have significantly expanded, offering possibilities for analyzing human health and diseases.

文章引用: 张钰东. 基于宏基因组分析的机器学习疾病预测模型构建[J]. 应用数学进展, 2024, 13(1): 199-207.
DOI: 10.12677/aam.2024.131023

Among these possibilities, disease prediction based on the analysis of the human gut microbiota has become a prominent research avenue. In this study, we utilized taxonomic gut microbiota data at the phylum level, known as Operational Taxonomic Units (OTU) data, and introduced two novel machine learning classification algorithms by combining non-negative matrix factorization and variational autoencoder methods. These algorithms are designed to extract critical information from the gut microbiota to predict diseases in patients. Through techniques such as dimensionality reduction, data generation, and the incorporation of penalty constraints in the models, we improve the prediction effect and optimize the overfitting of the model. Across simulated data, liver cirrhosis data, and diabetes data, our predictive models demonstrated significant performance, achieving AUC values of 0.926, 0.959, and 0.745, respectively.

Keywords

Operational Taxonomic Units, Non-Negative Matrix Factorization, Variational Auto Encoder

1. 引言

肠道微生物影响人体的营养代谢与免疫功能，且微生物丰度比例失衡会导致多种疾病的产生，例如肝硬化、二型糖尿病与肥胖等，因此肠道微生物逐渐成为了判断疾病状态重要标志[1] [2]。近年来随着高通量测序技术的发展，相关研究通常使用来自宏基因组测序分析的肠道微生物操作分类单元 Operational Taxonomic Units (OTU) [3]的丰度数据。目前已有一些机器学习模型应用于疾病预测，它能够学习微生物数据的关联和特征，并自动适应新的数据集特征，具有较好的泛化能力[4]，如支持向量机 Support Vector Machine (SVM)、随机森林 Random Forest (RF)、神经网络等分类方法已成功应用于结直肠癌、肝硬化、二型糖尿病等疾病的预测；Krizhevsky 等[5]提出了卷积神经网络 Convolution Neural Network (CNN)，Tsang 等[6]证明该方法可捕获到微生物数据的相互作用关系；Bokulich 等[7]针对微生物数据，使用机器学习方法对微生物数据进行有监督分类学习，并与回归模型分类效果进行比较；Lo 和 Marculescu 等[8]提出了用于分类宿主表型的神经网络模型 MetaNN，同时为了缓解过拟合问题，对微生物数据使用了数据扩充方法；Sharma 等[9]提出了一种基于分类学的预测方法 TaxoNN，本文将使用新提出的方法与该方法进行比较。我们针对二型糖尿病患者与肝硬化患者的微生物数据，构建了两个新的预测模型：DPNC、DPVC，考虑 OTU 数据高维且稀疏的特点，本研究使用了非负矩阵分解 Nonnegative Matrix Factorization (NMF) [10]和变分自动编码器 Variational Auto Encoder (VAE) [18]处理原始的 OTU 数据。其中 Karthik 等[11]证明 NMF 方法可以作为无监督学习方法，在生物计算学中执行多项任务，包括分子模式发现、类比和预测、跨平台和跨物种分析等。

2. 数据集与预处理

本文分别采用实证数据与模拟数据进行模型的训练与评估，本章将介绍数据集的来源、结构和建模的流程与架构。

数据集

实证数据包括来自两项研究的数据：二型糖尿病 Type 2 Diabetes (T2D)研究和肝硬化疾病

LiverCirrhosis (Cirr)研究。T2D 数据源自 Qin 等人的研究[12]，共包括 344 个样本，其中包含 174 个 T2D 患者样本和 170 个对照组样本。Cirr 数据取自 Qin 的研究[13]，这项研究探讨了肠道微生物组成对肝硬化疾病的影响，数据包含了 118 个患病样本和 114 个对照样本。模拟数据的生成基于克罗恩病实证研究数据，该研究由 Turpin 等人[14]提供，是 Genetic，Environmental，Microbial (GEM)项目的一部分。该项目旨在跟踪监测克罗恩病患者一级亲属的身体状况，以探究克罗恩病与潜在触发因素，包括遗传学、微生物组、环境和肠道屏障等之间的关系；克罗恩病数据集共包含 1796 个样本，其中包括 45 个阳性样本。

数据预处理

本文使用的模拟数据基于克罗恩病患者和对照组的原始数据，采用数据增强方法生成了近 10 万个样本的模拟数据集，这一过程旨在扩大样本规模，增加数据的多样性，从而提高模型的鲁棒性和泛化能力；通过复制和引入噪声，我们能够模拟真实生物群落中存在的变异性，有助于捕捉生物数据的特征和关联。

数据增强的步骤包括对原始数据集的复制和添加白噪声。对于每个样本，我们选择了 OTU 变量中的非零元素并对其引入噪声。我们假设噪声数据服从高斯分布，均值范围在 $\left[1\times10^{-6},2\times10^{-6}\right]$，标准差为 1×10^{-6}，此外，我们使用以下公式生成疾病状态：

$$p\left(y=1\right)=\frac{\exp\left(\alpha+\sum_{i=1}^{32}\beta_i\cdot V_i+\sum_{i=1}^{2}\sum_{j=i+1}^{3}\beta_{ij}V_i\cdot V_j\right)}{1+\exp\left(\alpha+\sum_{i=1}^{32}\beta_i\cdot V_i+\sum_{i=1}^{2}\sum_{j=i+1}^{3}\beta_{ij}V_i\cdot V_j\right)} \tag{1}$$

其中 V_i 表示第 i 个 OUT 特征或降为得到的第 i 个潜在特征，$y=1$ 表示疾病状态呈阳性，$y=0$ 表示疾病状态呈阴性，β_i 为特征回归系数，β_{ij} 为潜在特征成对相互作用系数，α 为疾病的基本流行率，本文取 $\alpha=25\%$，由此生成 16,626 个阳性样本，74,970 个对照样本。

3. 模型与方法

首先选取克罗恩病患者及其一级亲属的宏基因组测序数据为例展示建模流程，使用数据增强方法生成大量模拟数据样本，使用(1)式生成疾病状态，从模拟数据中抽样获取疾病组与对照组，对 OTU 数据降维处理并输入分类模型，整体流程如图 1 所示：

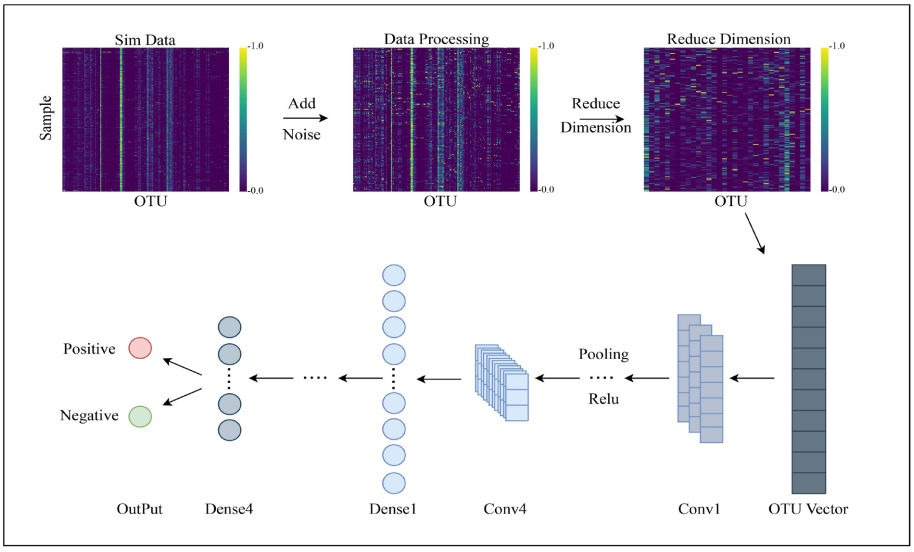

Figure 1. Model workflow diagram
图 1. 模型流程图

3.1. 非负矩阵分解(Nonnegative Matrix Factorization, NMF)

非负矩阵分解的基本思想是将一个非负矩阵 **X**，转换为两个非负矩阵的乘积：

$$\mathbf{X} \sim \mathbf{W} \times \mathbf{H} \tag{2}$$

为方便起见，我们称 **W** 为基矩阵，**H** 为系数矩阵，假设 **X** 是一个 $p \times n$ 的矩阵，则 **W** 是一个 $p \times k$ 的矩阵，**H** 为一个 $k \times n$ 的矩阵，其中 k 是矩阵分解所得潜在因子数量参数，矩阵 **W** 表示原始数据特征在潜在因子下的权重，矩阵 **H** 包含潜在因子在所有样本上的表达信息。

本文使用 sklearn.decomposition 库中的 NMF 函数实现矩阵分解，其中参数 components 决定潜在因子个数k，通常认为当且仅当 $k \times (p+n) \ll n \times p$，NMF 方法实现了对原始数据的降维，图 2 展示潜在因子个数为 8、16、32 时的特征模式热图：

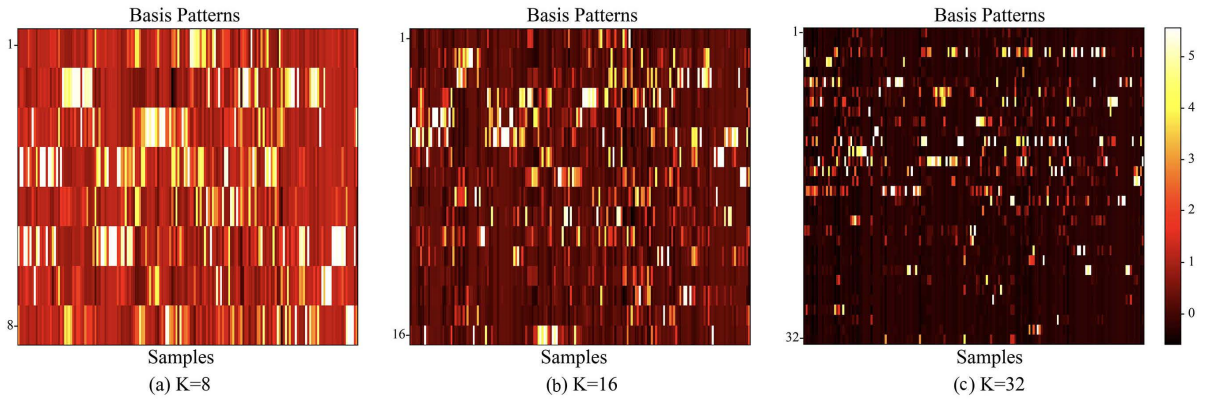

Figure 2. Heatmap of feature patterns across different latent factors
图 2. 不同潜在因子下的特征模式热图

Kong 等[15]在 NMF 中加入 L1、L2 正则化约束，以提高矩阵分解的鲁棒性，由此得到带有约束项的目标函数：

$$\min_{H} \left\| \mathbf{X} - \mathbf{W} \times \mathbf{H} \right\|_{2,1} + \lambda \left\| \mathbf{H} \right\|_{2,1}, \quad s.t. \mathbf{W} \geq 0, \mathbf{H} \geq 0 \tag{3}$$

其中 L1 约束表示通过赋予权重，仅保留部分特征发挥作用；L2 约束表示对 OTU 变量加以平滑约束，考虑了 OTU 之间的相互作用关系。Wang 等人的研究发现，在植物和肿瘤基因表达数据的分析中，使用基于 L1、L2 范数的 NMF 可以提取更多特征基因[16]；此外，为使 CNN 能够更充分地捕捉到微生物数据间的空间关系、相互作用关系，本文考虑聚类变量的两种排列方法，分别是基于相关关系的排序与基于欧式距离的排序。

3.1.1. $DPNC_{corr}$

该方法首先计算属性矩阵 **H** 中特征的 Spearman 相关系数矩阵记为矩阵 $\mathbf{K}_{k*k}$，其中k表示 NMF 的聚类数目，基于下式计算聚类变量的累积相关系数：

$$\rho_i = \sqrt[k]{\prod_{j=1}^{k} \left| \rho_{ij} \right|}, \quad i = 1, \cdots, k \tag{4}$$

其中 ρ_{ij} 表示 NMF 聚类得到的特征 i 与特征 j 的相关系数，式(4)计算的 ρ_i 则为第 i 个特征的累积相关系数，本文按照 ρ_i 值从大到小的顺序对特征进行排列，变量排序后的数据作为分类模型的输入，使得累积相关系数相近的特征能够被 CNN 的同一个滑动窗口捕捉到。

3.1.2. $DPNC_{dis}$

$DPNC_{dis}$ 需计算出 NMF 的聚类中心 N_{mid}，根据其余变量距聚类中心的距离对变量进行排序，聚类中心的计算方法如(5)式：

$$p(y=1)=\frac{\exp\left(\alpha+\sum_{i=1}^{32}\beta_i\cdot V_i+\sum_{i=1}^{2}\sum_{j=i+1}^{3}\beta_{ij}V_i\cdot V_j\right)}{1+\exp\left(\alpha+\sum_{i=1}^{32}\beta_i\cdot V_i+\sum_{i=1}^{2}\sum_{j=i+1}^{3}\beta_{ij}V_i\cdot V_j\right)} \tag{5}$$

其中 N_j 表示 NMF 聚类的特征 j，$d\left(n,N_j\right)$ 表示特征 n 与特征 j 的欧氏距离，$DPNC_{dis}$ 将特征按照距聚类中心由近到远的顺序排列，并采用与 $DPNC_{corr}$ 相同的方法，将排序后的变量作为分类模型的输入，使 CNN 的滑动窗口可以同时捕捉到。

3.2. 变分自动编码器(Variational Auto Encoder, VAE)

3.2.1. 自动编码器(Auto Encoder, AE)基本理论

自动编码器是用于处理无监督学习任务的机器学习方法，在降维和特征提取上的表现较好[17]，该方法分为编码层与解码层两部分，其流程框架如图 3 所示：

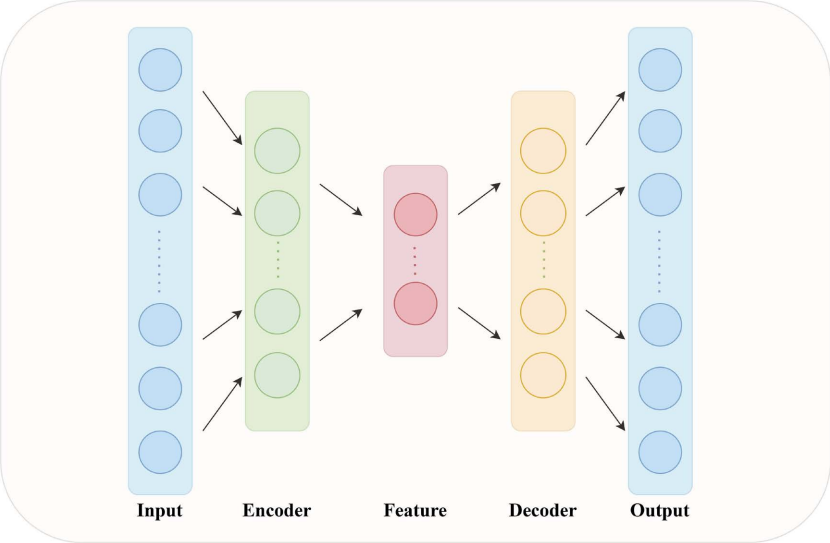

Figure 3. Auto Encoder workflow diagram
图 3. 自动编码器流程图

编码器部分负责将原始数据映射到编码空间，将初始特征转换为潜在特征；而解码器则负责从编码空间中的潜在特征重构数据，尽可能还原原始输入。假设输入为 $\mathbf{X}$，该方法的执行过程如下：

$$E^{(l)}=f_e\left(W_E^{(l)}\cdot E^{(l-1)}+B_E^{(l)}\right),\quad l=1,2,\cdots,L \tag{6}$$

式(6)中 l 指编码器的层数，$W_E^{(l)}$ 和 $B_E^{(l)}$ 表示编码器第 l 层的权重向量和偏置项向量，$E^{(0)}$ 即为输入 $\mathbf{X}$，$E^{(L)}$ 为低维深层特征；解码器传递公式如式(7)所示：

$$D^{(l)}=f_d\left(W_D^{(l)}\cdot D^{(l-1)}+B_D^{(l)}\right),\quad l=1,2,\cdots,L \tag{7}$$

同理，l 指解码器的层数，$W_D^{(l)}$ 和 $B_D^{(l)}$ 表示解码器第 l 层的权重项和偏置项，$D^{(0)}$ 等同于 $E^{(l)}$ 表示低维深层特征，$D^{(l)}$ 为重构数据 $\mathbf{X}'$，自编码器的目标是最小化重构误差，即原始数据与重构数据之间的差异：

$$\mathcal{L}(\mathbf{X},\mathbf{X}')=\left\|\mathbf{X}-\mathbf{X}'\right\|^2 \tag{8}$$

3.2.2. 变分自动编码器(VAE)基本理论

自动编码器具有映射分布无规律、空间不连续的问题，导致输入到潜在特征空间没有平滑的过度，其结果是降维得到的潜在特征矩阵稀疏，有效的潜在因子数量无法控制；实际上宏基因组变量数目庞大，相应的潜在变量空间也应连续，由 Kingma [18]等提出的 VAE 更适用于此背景下的数降维任务，其拟合后验分布并重采样生成数据的过程保证了潜在空间中变量连续性、生成样本多样性的要求，理论上可以优化模型过拟合情况。

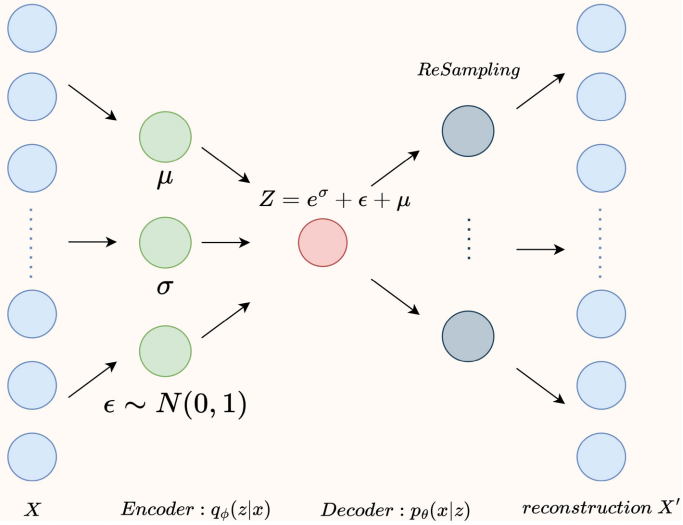

Figure 4. Variational auto encoder workflow diagram
图 4. 变分自动编码器流程图

变分自动编码器的整体流程如图 4 所示，模型的关键思想是将数据的生成过程建模为一个概率分布，学习该分布的参数以实现降维与新样本的生成，其中模型首先通过编码器 $q_\phi(z|x)$ 根据给定的输入 X 生成潜在变量 Z，再通过解码器 $p_\theta(x|z)$ 对潜在变量 Z 重抽样，生成新的样本。

VAE 的重点在于损失函数的构造，共由重构损失与 KL 散度正则化项两部分组成，损失函数是二者的线性组合，分别衡量了重构样本与输入样本的相似度、潜在变量分布与先验分布之间的差异度。重构损失(Reconstruction Loss)衡量解码器 $p_\theta(x|z)$ 生成重构数据 X' 与 X 的误差，通常使用交叉熵或均方误差：

$$\mathcal{L}_{recon}=-\mathbb{E}_{q_\phi(z|x)}\left[\log p_\theta(x|z)\right] \tag{9}$$

而 $\min \mathcal{L}_{\text{recon}}$ 的含义即在潜在变量空间中潜在变量分布 $q_\phi(z|x)$ 下，得到通过解码器生成重构数据分布 $p_\theta(x|z)$ 的对数似然最大化；另一部分则是通过 KL 散度衡量 $q_\phi(z|x)$ 与先验分布 $p(z)$ 之间的差异，即潜在分布与预先设定的理想分布之间的偏离程度，KL 散度的计算如(10)式所示：

$$\mathcal{L}_{KL}=KL\left[q_\phi(z|x)\middle\|p(z)\right] \tag{10}$$

实际上我们想要得到潜在变量的后验分布 $p(z|x)$，该计算涉及对潜在变量 Z 的计算，然而由于潜在空间维度较高的特性，使得潜在变量的后验分布的计算难以实现，因此我们使用 $q_\phi(z|x)$ 近似表示后验分布，$p(z)$ 相比较以得到我们想要的分布，考虑宏基因组数据特征，我们假定 $p(z)$ 服从高斯分布；则 VAE 完整的损失函数为：

$$\mathcal{L} = \mathcal{L}_{recon} + \mathcal{L}_{KL} \tag{11}$$

4. 实验结果分析

为实现分类预测效果，设置卷积神经网络的最后一层为带有 SoftMax 激活的两节点全连接层，其余激活函数均适用 Relu；卷积层的个数为 4，每个卷积层均带有一个最大池化层，卷积核个数分别为 32、64、128、256；随后共有 3 个全连接的隐藏层，其节点个数分别为 512、256、64；考虑特征数量大，可能带来训练模型过拟合的问题，我们在神经网络的全连接层中加入 L2 正则化处理；针对本研究的分类问题，使用预测结果绘制 ROC 曲线，计算对应的 AUC 值，比较本文提出的方法当下流行的宏基因组疾病预测方法。

4.1. 模拟数据预测结果

从模拟数据集的近 10 万份样本中，随机挑选 200 份疾病样本与 200 份对照样本，其中 80%用作训练集、20%用作测试集；我们将提出的方法与传统的分类方法进行比较，结果如图 5 所示：

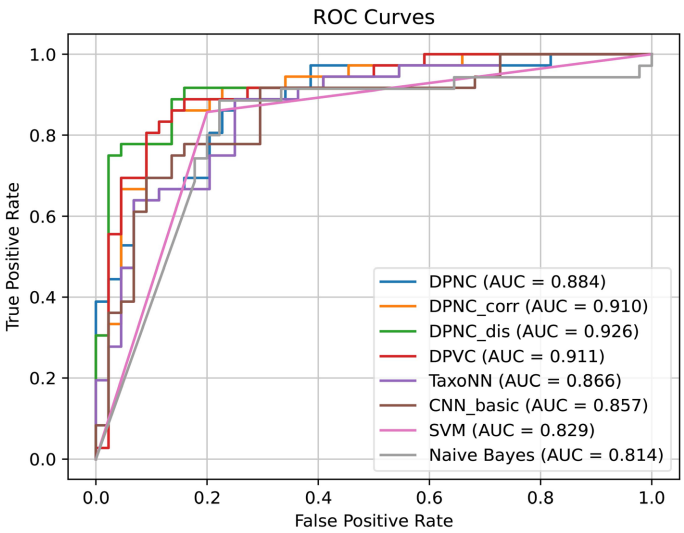

Figure 5. The ROC curves of 8 methods on the simulated dataset.
图 5. 8 种方法在模拟数据集上的 ROC 曲线

基于 NMF 降维的模型 DPNC 在模拟数据集上的预测的 AUC 值达到了 0.884，相较于 Sharma 提出的 TaxoNN 方法(AUC = 0.866)提升了 0.018；而 DPNC 的基于相关性与欧式距离排序的衍生模型，预测效果得到了更为显著的提升，AUC 分别为 0.910、0.926；本文提出的基于变分自动编码器的神经网络预测模型在模拟数据集上的表现较 TaxoNN 同样有显著的提升，AUC 达到了 0.911；而其余常见的机器学习方法如基础的卷积神经网络、支持向量、贝叶斯分类模型，预测结果所得的 AUC 分别为 0.857、0.829、0.814。

需要说明的是，加入了变分自动编码器的 DPVC 的 AUC 值虽不是最高，但由于其具有学习原始数据分布的能力，因此一定程度缓解了训练过拟合现象，可视化训练过程如图 6 所示，可以看出 DPVC 在训练集与测试集上均有较高的预测精度，表现出更好的模型性能。

4.2. 实证研究结果

我们采用了本文提出的四种方法，将其应用于实证数据，并与 TaxoNN 预测方法进行比较。评估分类效果时，我们使用 AUC 值作为评价指标，对肝硬化数据集与二型糖尿病数据集的预测结果进行了分析：

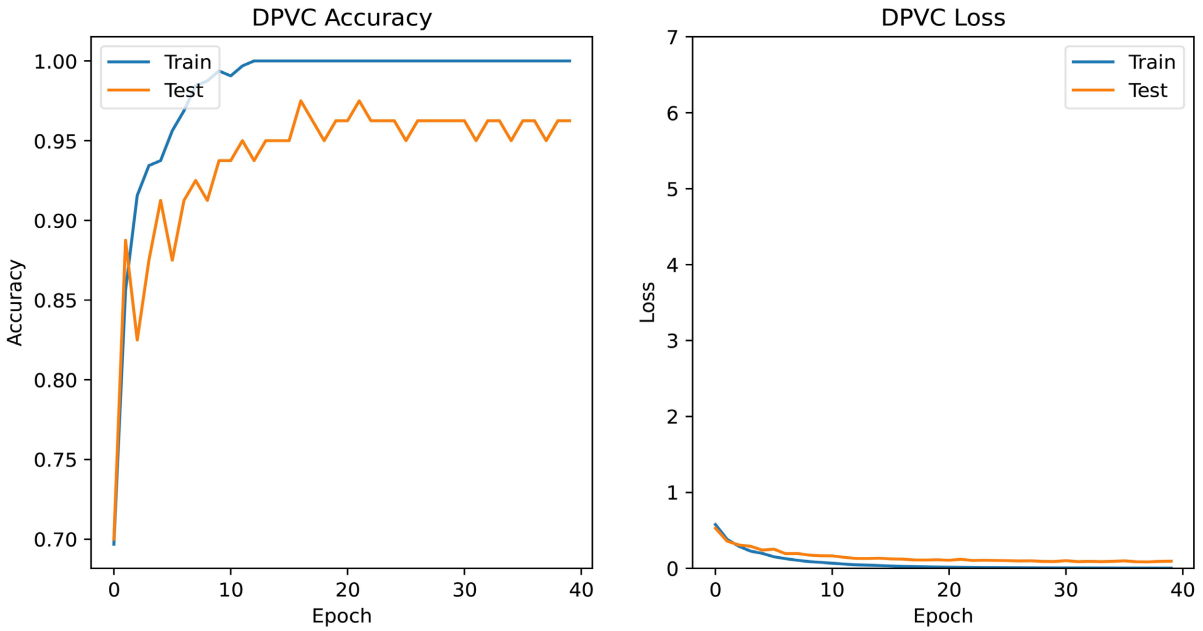

Figure 6. The performance of DPVC on the simulated dataset
图 6. DPVC 在模拟数据集上的表现

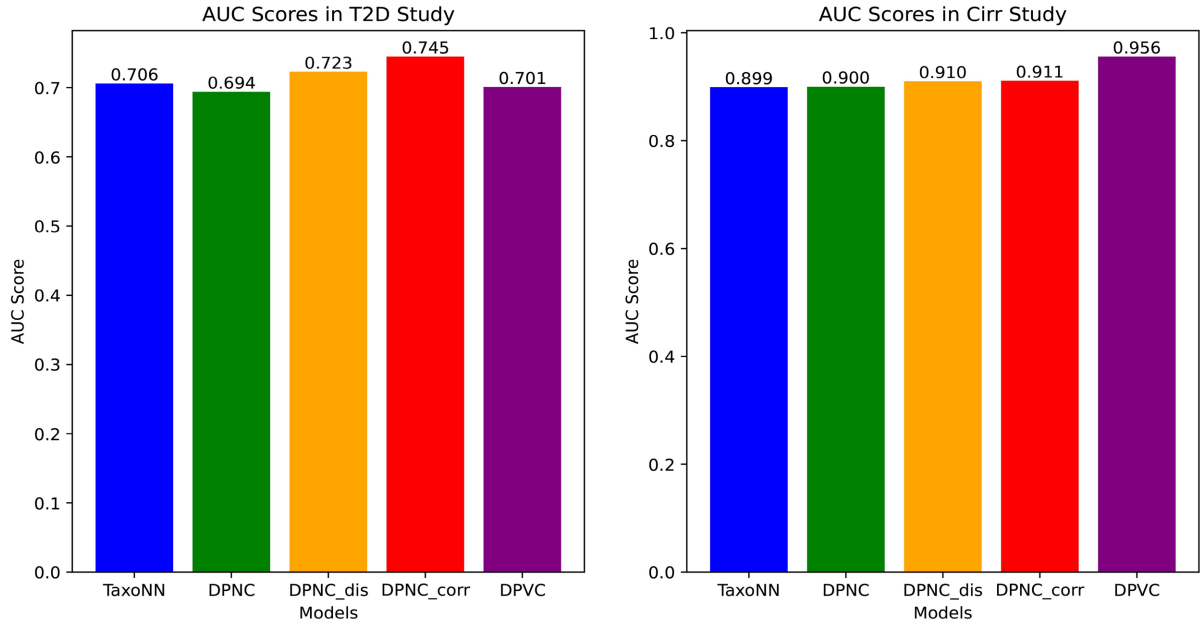

Figure 7. The performance of multiple new methods on the clinical dataset
图 7. 多种新方法在临床数据集上的性能表现

参考图 7，我们发现在肝硬化疾病患者预测中，加入了变分自动编码器的卷积模型 DPVC 表现突出，其 AUC 值达到了 0.956。对于二型糖尿病患者的预测，基于相关性排序的 $DPVC_{corr}$ 方法表现出色，其 AUC 值达到了 0.745，明显优于其他方法。值得注意的是，在宏基因组对糖尿病发病影响研究中，仅使用 VAE 或 NMF 的两种建模方法的预测效果未超越 Sharma 等提出的 TaxoNN，然而通过考虑对降维后特征的基于距离和相关性的重排方法，新模型的预测效果得到了显著提升，这表明该方法成功地捕捉到了微生物

间存在的空间距离关系与相互作用关系。

5. 总结

本文提出了两个基于宏基因组分析的疾病预测模型。这两个模型分别利用门级别的分类学数据作为原始特征，采用非负矩阵分解和变分自动编码器方法进行数据降维。进一步，针对前者，我们引入了基于特征空间距离和相关关系的变式新模型，然后建立了带有 L2 正则化约束全联接层的深度卷积模型作为分类模型。

新模型在模拟数据集、肝硬化患者数据集和二型糖尿病数据集上表现出色。与之前提出的方法相比较，三个数据集上的 AUC 指标分别提升了 0.06、0.039 和 0.057，显示出了显著的性能提升。

参考文献

[1] Sommer, F., Jacqueline, M., Richa, B., Jeroen, R. and Philip, R. (2017) The Resilience of the Intestinal Microbiota Influences Health and Disease. *Nature Reviews Microbiology*, **15**, 630-638. https://doi.org/10.1038/nrmicro.2017.58

[2] Jackson, A.M., Verdi, S., Maxan, M.E., Shin, C.M., Zierer, J., Bowyer, R., Martin, T., Williams, F., Menni, C., Bell, J., Spector, T. and Steves, C. (2018) Gut Microbiota Associations with Common Diseases and Prescription Medications in a Population-Based Cohort. *NatCommun*, **9**, Article No. 2655. https://doi.org/10.1038/s41467-018-05184-7

[3] Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R. and Abebe, E. (2005) Defining Operational Taxonomic Units Using DNA Barcode Data. *Philosophical Transactions of the Royal Society B*, **360**, 1935-1943. https://doi.org/10.1098/rstb.2005.1725

[4] Tsai, K., Lin, S., Liu, W. and Wang, D. (2015) Inferring Microbial Interaction Network from Microbiome Data Using RMN Algorithm. *BMC System Biology*, **9**, Article No. 54. https://doi.org/10.1186/s12918-015-0199-2

[5] Krizhevsky, A., Sutskever, I. and Hinton, G. (2012) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. https://doi.org/10.1145/3065386

[6] Tsang, M., Cheng, D. and Liu, Y. (2007) Detecting Statistical Interactions from Neural Network Weights. arXiv:1705.04977.

[7] Bokulich, N., Dillon, M., Bolyen, E., Kaehler, B. and Huttley, G. (2018) q2-Sample-Classifier: Machine-Learning Tools for Microbiome Classification and Regression. *Journal of Open Research Software*, **3**, Article 934. https://doi.org/10.21105/joss.00934

[8] Lo, C. and Marculescu, R. (2019) MetaNN: Accurate Classification of Host Phenotypes from Metagenomic Data Using Neural Networks. *BMC Bioinformatics*, **20**, Article No. 314. https://doi.org/10.1186/s12859-019-2833-2

[9] Sharma, D., Paterson, A., Xu, W. (2020) TaxoNN: Ensemble of Neural Networks on Stratified Microbiome Data for Disease Prediction. *Bioinformatics*, **36**, 4544-4550. https://doi.org/10.1093/bioinformatics/btaa542

[10] Lee, D. and Seung, H. (1999) Learning the Parts of Objects by Nonnegative Matrix Factorization. *Nature*, **401**, 788-791. https://doi.org/10.1038/44565

[11] Karthik, D. (2008) Nonnegative Matrix Factorization: An Analytical and Interpretive Tool in Computational Biology. *PLOS Computational Biology*, **4**, e1000029. https://doi.org/10.1371/journal.pcbi.1000029

[12] Qin, J., *et al*. (2012) A Metagenome-Wide Association Study of Gut Microbiota in Type 2 Diabetes. *Nature*, **490**, 55-60. https://doi.org/10.1038/nature11450

[13] Qin, N., *et al*. (2014) Alterations of the Human Gut Microbiome in Livercirrhosis. *Nature*, **513**, 59-64. https://doi.org/10.1038/nature13568

[14] Turpin, W., Sliverberg, M., Kevans, D., Smith, M., *et al*. (2016) Association of Host Genome with Intestinal Microbial Composition in a Large Healthy Cohort. *Nature Genetics*, **48**, 1413-1417. https://doi.org/10.1038/ng.3693

[15] Kong, D., Ding, C. and Huang, H. (2011) Robust Nonnegative Matrix Factorization Using L21-Norm. *Proceedings of the* 20*th ACM International Conference on Information and Knowledge Management*, 673-682. https://doi.org/10.1145/2063576.2063676

[16] Wang, D., Liu, J., Gao, Y., Zheng, C. and Xu, Y. (2016) An NMF-$l_{2,1}$-Norm Constraint Method for Characteristic Gene Secection. *PLOS ONE*, **11**, e0158494. https://doi.org/10.1371/journal.pone.0158494

[17] Wang, Y., Yao, H. and Zhao, S. (2016) Auto-Encoder Based Dimensionality Reduction. *Neurcomputing*, **184**, 232-242. https://doi.org/10.1016/j.neucom.2015.08.104

[18] Kingma, D.P. and Welling, M. (2014) Auto-Encoding Variationalbayes. arXiv:1312.6114.