无超参数神经混沌学习分类算法

Akhila Henry
Department of Mathematics
Amrita Vishwa Vidyapeetham
Amritapuri Campus
Kollam, 690525, Kerala, India.
akhilahenryu@am.amrita.edu

Nithin Nagaraj
Complex Systems Programme
National Institute of Advanced Studies
Indian Institute of Science Campus
Bengaluru, 560012, Karnataka, India.
nithin@nias.res.in

ABSTRACT

神经混沌学习（NL）是一种受大脑启发的分类框架，它利用混沌动力学从输入数据中提取特征，并在分类任务上实现了最先进的性能。然而，NL 需要调整多个超参数并为每个输入样本计算四个混沌特征。本文我们提出自动混沌网络——一种新颖且无超参数的 NL 算法变体，消除了训练和参数优化的需求。AutochaosNet 利用从 Champernowne 常数派生出的通用混沌序列，并使用输入刺激定义特征提取的时间界限。评估了两个简化的变体——TM AutochaosNet 和 TM-FR AutochaosNet，并与现有的 NL 架构 ChaosNet 进行比较。我们的结果显示，AutochaosNet 实现了具有竞争力或更优的分类性能，同时由于计算量减少而大大减少了训练时间。除了消除训练和超参数调整之外，AutochaosNet 还表现出优秀的泛化能力，使其成为现实世界分类任务中可扩展且高效的选项。未来工作将重点识别各种混沌映射下的通用轨道，并将其纳入 NL 框架以进一步提高性能。

***K*eywords** 神经混沌学习 · 轨迹均值 · 点火率 · 普适轨道 · 自动混沌网络

1 介绍

当前的人工智能研究轨迹旨在使机器能够自主学习。各种各样的机器学习算法正在以加速的速度开发中。其中，受大脑启发的学习算法最近受到了广泛关注。其中一个方法是*神经混沌学习（NL）*算法 [1]，它从人脑的动力学行为中汲取灵感，特别是神经活动中的混沌现象。

在 2023 中，研究者引入了两种不同的 NL 框架架构：ChaosNet 和 CFX（混沌特征提取）+ML [2]。这些架构通过使用一种称为斜顶帐篷映射的一维混沌映射来运用混沌。对于每个数据集，都会计算一个神经轨迹，该轨迹在达到预定义的噪声阈值时终止。

给定一个包含 m 个样本和 n 个特征的数据集，ChaosNet 和 CFX+ML 算法为每个输入特征提取四个混沌特征——放电时间、放电率、能量和熵——从而形成一个大小为 $4n$ 的转换特征集。这些特征随后被馈入经典机器学习分类器或直接使用余弦相似度进行分类。该算法涉及初始条件和混沌映射的偏斜值等超参数，以及噪声水平，所有这些都需要调优。这一过程增加了特征提取的计算成本和复杂性。

在这项研究中，我们介绍了一种新颖的神经混沌学习算法变体，该变体通过利用最近提出的通用混沌轨道 [3] 完全消除了超参数调整的需求，该轨道是使用十进制移位映射导出的。这种方法旨在简化特征提取过程，同时保持混沌表示在学习任务中的有效性。

2 使用通用轨道的神经混沌学习

神经混沌学习（NL）是一种跨学科方法，它结合了来自神经科学、混沌理论和机器学习的概念。该领域的最新进展包括引入了一类新的混沌轨道，即通用轨道，这些轨道具体是相对于十进制移位映射（DSM）[4] 和连分数映射 [3] 定义的。在这项研究中，我们使用 DSM 下的通用轨道进行混沌特征提取。

DSM 的轨道是通过依次移位初始点的数字生成的。在 DSM 下的普适轨道可以从不可数多个初始点生成。其中，我们关注从 Champernowne 常数导出的轨道，记为

$$C = 0.1234567891011\ldots498499,$$

，在对应数字 499 之后截断。此外，该常数已知是基数 10 下正规 [5] 的，确保所有有限位序列以相等频率出现 [3] 。以下引理提供了任意自然数在此常数中出现的理论基础。

引理 1. *令 $c = 0.1\,2\,3\,4\,5\,6\,7\,8\,9\,10\,11\ldots N(N+1)(N+2)\ldots$ 为 Champernowne 常数，$c_N = 0.1234567891011\ldots(N-2)(N-1)$ 为 c 的截断部分，其中 $N = a_1a_2\ldots a_d$，$d \geq 2$。然后小数点后 c_N 的位数由*

$$dN - (10 + 10^2 + \ldots + 10^{d-1}) - 1. \tag{1}$$

换句话说，数字 N 将会在 c 的小数展开中，在小数点后

$$dN - (10 + 10^2 + \ldots + 10^{d-1}) - 1$$

位出现。[3].

引理的详细证明可以在 [3] 中找到。

使用带 champernowne 常数作为初始点的 DSM 消除了调整参数如偏斜值和噪声阈值的需求，这些参数在传统的 NL 算法中通常是必需的。不再使用预定义的噪声水平来终止神经轨迹，而是当其前三位小数与输入刺激（特征值）相匹配时停止轨迹。本研究中的三位小数阈值是任意选择的，在未来的工作中可以根据经验性能进行调整。这相当于噪声阈值为 0.001。

假设刺激（特征值）是 $0.b_1b_2b_3b_4b_5$。那么根据引理 1，数字 $N' = b_1b_2b_3$ 将出现在 c（Champernowne 常数）的小数展开中，在小数点后 $\alpha = 3N' - (10 + 10^2) - 1$ 位之后（这里，我们考虑三位小数的阈值即 $d = 3$）。因此，α 将是放电时间的上限。

具体来说，引理 1 允许我们确定每个刺激的放电时间界限和放电率，这作为提出的神经混沌学习算法变体中的关键混沌特征。

3 提出的算法：AutochaosNet

在本节中，我们将使用十进制移位映射生成的通用轨道和 Champernowne 常数来开发一种无需调整超参数的神经混沌学习架构，即 AutochaosNet。AutochaosNet 有两个版本，基于提取的不同特征：

- 自动混沌网络（轨迹均值自动混沌网络）
- 轨迹均值与放电率自动混沌网络 (TM-FR AutochaosNet)

提出的架构如下：

(a) **输入**：假设输入数据的大小为 $m \times n$，其中 m 表示样本数量，n 表示原始输入特征（或属性）的数量。设

$$\{(x_{11}, x_{12}, \ldots, x_{1n}), (x_{21}, x_{22}, \ldots, x_{2n}), \ldots, (x_{m1}, x_{m2}, \ldots, x_{mn})\}$$

为整个数据集。经过适当的归一化后，每个数据实例可以依次进行特征变换。最小-最大规范化在这里被使用。规范化的数据可以表示为:

$$\{(z_{11}, z_{12}, \ldots, z_{1n}), (z_{21}, z_{22}, \ldots, z_{2n}), \ldots, (z_{m1}, z_{m2}, \ldots, z_{mn})\}$$

其中 $z_{ij} = \frac{x_{ij} - min(\{x_{ij}:1 \leq i \leq m\})}{max(\{x_{ij}:1 \leq i \leq m\}) - min(\{x_{ij}:1 \leq i \leq m\})}$ 对于 $1 \leq j \leq n$。

(b) **特征转换**: 数据实例的每个特征属性 x_{ij} 都必须转换为基于混沌的特征。
在 AutochaosNet 的两个版本中，每个 x_{ij} 对应的放电时间界值都使用引理 1 中定义的公式计算到小数点后三位。设 T 是对应于 $x_{(ij)}$ 和 Champernowne 常数（截断至 500 位）的放电时间界，

$$c = 0.1234567891011 \ldots 498499$$

为初始点。则对应于 x_{ij} 的神经轨迹为

$$\tau = \{c, f(c), f^{(2)}(c), \ldots, f^{(T-1)}(c)\}$$

(c) **特征提取**：在 TM AutochaosNet 中，对于每个 x_{ij}，计算神经轨迹 τ 的均值。假设如果 t_{ij} 是对应于 x_{ij} 的神经轨迹的均值。那么提取的特征集将是:

$$\{(t_{11}, t_{12}, \ldots, t_{1n}), (t_{21}, t_{22}, \ldots, t_{2n}), \ldots, (t_{m1}, t_{m2}, \ldots, t_{mn})\}$$

在 TM-FR AutochaosNet 中，对于每个 x_{ij}，计算神经轨迹 τ 的均值及其放电率。对应于每个 x_{ij} 的放电率定义为神经轨迹 τ 超过值 0.5 以识别刺激的时间分数。

(d) **分类**: 可以使用余弦相似度对转换后的特征数据集进行分类。首先，计算每个类别的平均表示向量。假设 m 数据实例属于 k 类别。设

$$\{(x_{l_1 1}, x_{l_1 2}, \ldots, x_{l_1 n}), (x_{l_2 1}, x_{l_2 2}, \ldots, x_{l_2 n}), \ldots, (x_{l_r 1}, x_{l_r 2}, \ldots, x_{l_r n})\}$$

为类别 l 中的 r 样本。归一化后，提取的数据为:

$$\{(f_{l_1 1}, f_{l_1 2}, \ldots, f_{l_1 n}), (f_{l_2 1}, f_{l_2 2}, \ldots, f_{l_2 n}), \ldots, (f_{l_r 1}, f_{l_r 2}, \ldots, f_{l_r p})\}$$

在 TM AutochaosNet 中，由于只有一个特征，p 的值将是 n 本身。但在 TM-FR AutochaosNet 中，索引 p 的值将是 $2n$。然后，对应类别 l 的平均表示向量可以定义为:

$$M^{(l)} = \left(\frac{\sum_{i=l_1}^{l_r} f_{i1}}{r}, \frac{\sum_{i=l_1}^{l_r} f_{i2}}{r}, \ldots, \frac{\sum_{i=l_1}^{l_r} f_{ip}}{r} \right)$$

为了对特定的数据实例

$$X_i = (x_{i1}, x_{i2}, \ldots, x_{in}),$$

进行分类，计算该数据实例 $F_i = (f_{i1}, f_{i2}, \ldots, f_{ip})$ 提取的特征向量与每个类别的平均表示向量 $\{M^{(1)}, M^{(2)}, \ldots, M^{(k)}\}$ 之间的余弦相似度。余弦相似度定义如下:

$$\cos\theta = \frac{M^{(j)} \cdot F_i}{||M^{(j)}|| \ ||F_i||},$$

其中 $j = 1, 2, \ldots, k$。数据实例将属于具有最小值的类别。

(e) **输出**：输出指的是每个测试数据实例被分配的类别。AutochaosNet 根据每个类别的平均表示向量给出输出。

提出的算法如图 1 所示。

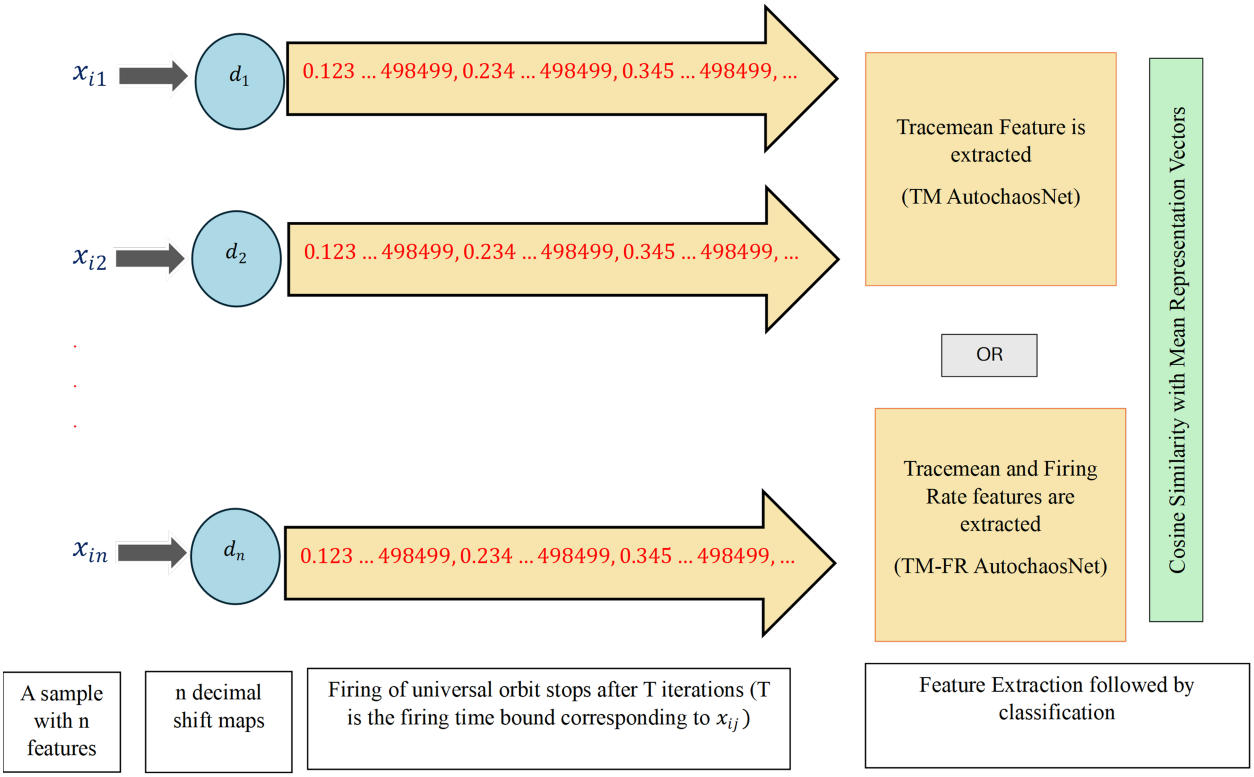

Figure 1: AutochaosNet 算法的特征提取与分类架构。

3.1 结果与分析

本节对提出的无超参数算法进行了比较分析：TM AutochaosNet 和 TM-FR AutochaosNet 对比 ChaosNet，使用 F1 分数作为评估指标。ChaosNet 算法利用四个混沌特征和三个超参数，而提出的算法最多使用两个混沌特征，并且不使用任何超参数。表 1 报告了 AutochaosNet 和 chaosNet 算法在十个不同数据集上的 F1 分数。具有最高值的 F1 分数用粗体突出显示。值得注意的是，尽管没有超参数，AutochaosNet 算法的表现仍与 ChaosNet 算法相当。

Table 1: **AutochaosNet 与 ChaosNet 的 F1 分数比较。**

Dataset	TM-FR AutochaosNet	TM AutochaosNet	ChaosNet
Iris	0.868	0.838	**1.000**
Haberman	0.537	0.516	**0.560**
Seeds	**0.878**	**0.878**	0.845
Statlog	**0.755**	**0.755**	0.738
Breast Cancer	0.784	0.717	**0.927**
Bank	0.821	0.806	**0.845**
Ionosphere	0.823	0.813	**0.860**
Wine	0.858	0.824	**0.976**
Sonar	**0.785**	0.760	0.643
Penguin	0.907	0.885	**0.964**

3.1.1 计算复杂性

除了达到可比较的 F1 分数外，AutochaosNet 算法的关键优势在于其能够比 chaosNet 更快地对数据集进行分类。为了比较这些算法的时间复杂度，选择了四个数据集——Iris、Seeds、Statlog 和 Sonar 进行评估。每个算法执行了 50 次迭代，并计算了 50 次迭代的平均耗时值。TM-AutochaosNet、TM-FR AutochaosNet 和 ChaosNet 在 50 次迭代中的平均耗时值如表 2 所示。

Table 2: **不同模型在 50 次迭代中的宏 F1 分数、平均值和标准差的比较**。

数据集	TM-FR 自动混沌网络			TM 自动混沌网络			混沌网（单次迭代）	
	F1 Score	Mean Time (s)	Standard Deviation	F1 Score	Mean Time (s)	Standard Deviation	F1 Score	Mean Time (s)
鸢尾花	0.868	0.741	0.123	0.838	0.596	0.018	1	52.15
种子	0.878	1.577	0.044	0.878	4.99	0.475	0.845	7310.19
统计日志	0.755	2.384	0.241	0.755	2.531	0.279	0.738	19750
声纳	0.785	14.506	0.858	0.76	14.702	0.584	0.643	10686.01

4 结论

在这项研究中，我们介绍了两种无超参数变体的神经混沌学习算法——TM AutochaosNet 和 TM-FR Autochaos-Net，它们利用了使用 Champernowne 常数（截断至 500 位）作为初始点生成的通用混沌轨道。通过消除对超参数调整的需求并依赖简化的一组混沌特征，这些模型显著降低了计算复杂性，同时保持分类准确性。

广泛的评估表明，所提出的 AutochaosNet 算法在多个基准数据集上的性能与 ChaosNet 相比具有竞争力。值得注意的是，AutochaosNet 在使用较少特征和无需超参数的情况下实现了这一点，提供了显著的效率提升。AutochaosNet 在不进行超参数调整的情况下跨数据集泛化的能力使其成为解决实际分类问题的可扩展且实用的解决方案。

在未来的工作中，我们旨在识别在各种混沌映射下生成通用轨道的点集。这些截断轨道将作为神经混沌学习框架中的起始触发器使用，目标是进一步提高分类准确性，同时保持算法的简单性和无参数特性。

致谢

NN 愿意感谢印度原子能部（政府）高等数学国家委员会（NBHM）的经济支持（资助编号 02011/8/2025/NBHM(R.P)/R&D II/1259）。

代码可用性

我们提出的方法的实现可以在以下位置获取：https://github.com/akhilahenry98/AutochaosNet.git

附录

A 数据集描述

本节概述了本研究中使用的十个数据集，详细列出了每个数据集的样本数量和类别数量。

Table 3: **每个本研究中使用的数据集的特征数量和样本数量。**

Dataset	Number of Features	Number of Classes	Number of Samples
鸢尾花	4	3	150
哈伯曼生存率	3	2	306
种子	7	3	210
心脏病统计日志	13	2	270
电离层	34	2	351
银行票据认证	4	2	1372
乳腺癌威斯康星	31	2	569
葡萄酒	13	3	178
企鹅	4	3	342
声纳	60	2	208

A.1 鸢尾花

鸢尾花 [6] 数据集包含关于三种鸢尾花品种的信息："红花 *Iris Setosa,* 变色鸢尾 *Iris Versicolor*，和花 *virginica*"。该数据集包含 150 个数据对象，每个对象具有四个属性："萼片长度、宽度，花瓣长度和宽度"。所有属性均以内容管理系统为单位测量。表 4 中的类别分布如下：山梅花属，变色紫罗兰，维吉尼亚斗红花属。

Table 4: **鸢尾花：每个类别中的样本数量。**

Class	Label	Number of Samples
Iris Setosa	0	50
Iris Versicolor	1	50
Iris Virginica	2	50

A.2 哈伯曼生存率

该数据集包括芝加哥大学比林斯医院在 1958 年至 1970 年期间进行的一项乳腺癌手术生存率研究的记录 [7]。它包含 306 个案例，每个案例都由三个数值特征来描述：发现的阳性腋窝淋巴结的数量、手术年份（表示为年份减去 1900）、以及患者在手术时的年龄。分类基于患者是否在手术后五年内死亡或至少存活了五年，如表 5 所示。

Table 5: **哈伯曼：每个类别的样本数量。**

Class	Label	Number of Samples
Less than 5 years	0	225
Greater than or equal to 5 years	1	81

A.3 种子

数据集种子 [8] 提供了三种不同小麦类型——卡玛，罗莎和加拿大的麦粒几何特征的测量值。七个实数值属性，即紧凑度、长度、宽度等，是使用软 X 射线方法和 GRAINS 软件包构建的。类别分布见表 6。

A.4 心脏统计日志

数据集统计日志 [8] 使用包括静息血压、胸痛类型、运动诱发的心绞痛等在内的 13 个属性来提供心脏病是否存在的情况。类别分布显示在表 7 中。

Table 6: **种子：每个类别的样本数量。**

Class	Label	Number of Samples
Kama	0	70
Rosa	1	70
Canadian	2	70

Table 7: **心脏：每个类别的样本数量。**

Class	Label	Number of Samples
Absence	0	150
Presence	1	120

A.5 电离层

电离层 [9] 数据集是一个维度为 34 的二分类数据集。类别表示从电离层接收到雷达信号的状态。本实验的目的是通过分析雷达数据来确定电离层的组成。类分布情况见表 8。

Table 8: **电离层：每个类中的样本数量。**

Class	Label	Number of Samples
Bad	0	126
Good	1	225

A.6 银行票据认证

银行票据认证 [10] 数据集用于区分真实和伪造货币。数据从真实和伪钞样本的照片中提取。该数据集包含四个属性，例如小波变换图像的方差（连续）、小波变换图像的偏度（连续）等。通过使用小波变换从图像中获得这些属性。类别分布见表 9。

Table 9: **银行：每个类中的样本数量。**

Class	Label	Number of Samples
Genuine	0	762
Forgery	1	610

A.7 乳腺癌威斯康星州

乳腺癌威斯康星州 [11] (诊断) 数据集预测癌症是良性还是恶性。每个细胞核的 31 特征，如半径、周长、纹理、光滑度等，是从乳腺肿瘤细针穿刺（FNA）的数字化图像中得出的。类别分布显示在表 10 中。

Table 10: **癌症：每个类别中的样本数量。**

Class	Label	Number of Samples
Malignant	0	212
Benign	1	357

A.8 葡萄酒

意大利同一地区生产的三种不同葡萄品种的葡萄酒的化学分析结果在葡萄酒数据集 [12] 中有提供。研究中对每个葡萄酒样本测量了十三种化学成分的浓度。带有标签“1, 2 和 3”的葡萄酒被分为三类。表 11 显示了这些类别的分布。

Table 11: **葡萄酒：每个类别中的样本数量**。

Class	Label	Number of Samples
1	0	59
2	1	71
3	2	48

A.9 帕尔默企鹅数据集

帕尔默群岛企鹅数据集 [13] 包含了三种企鹅的体型测量数据。数据来自南极洲帕默群岛的三个岛屿，包含了 344 只企鹅的信息。数据集中有三种企鹅：阿德利企鹅，金图企鹅和下颌 *Strap*。类别分布见表 12。

Table 12: **企鹅：每个类中的样本数量**。

Class	Label	Number of Samples
Adelie Penguin	0	151
Chinstrap Penguin	1	68
Gentoo Penguin	2	123

A.10 声纳

数据集声纳 [14] 包含来自类似条件下岩石的 97 个模式和从不同角度和条件反射金属圆柱体的声纳信号的 111 个模式。模式中的 60 数字范围从 0.0 到 1.0。每个记录的标签表示项目是地雷（用“M”表示）还是岩石（用“R”表示）。表 13 显示了类别的分布。

Table 13: **声纳：每个类别中的样本数量**。

Class	Label	Number of Samples
Mine	0	111
Rock	1	97

References

[1] Harikrishnan N B, Aditi Kathpalia, Snehanshu Saha, and Nithin Nagaraj. Chaosnet: A chaos based artificial neural network architecture for classification. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29:113125, 11 2019.

[2] Devashish Sethi and Nithin Nagaraj. Neurochaos feature transformation for machine learning. *Integration*, 90:157–162, 2023.

[3] Akhila Henry, Nithin Nagaraj, and Rajan Sundaravaradhan. Universal orbits: Unveiling the connection between chaotic dynamics, normal numbers, and neurochaos learning. *Chaos Theory and Applications*, 7(1):61–69, 2025.

[4] Steven H Strogatz. *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering*. CRC press, Taylor & Francis Group, 2018.

[5] D. Khoshnevisan. On the normality of normal numbers. Clay Mathematics Institute Annual Report, 2006.

[6] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188, 1936.

[7] Shelby J Haberman. The analysis of residuals in cross-classified tables. *Biometrics*, pages 205–220, 1973.

[8] Dheeru Dua, Casey Graff, et al. Uci machine learning repository. 2017.

[9] Vincent G Sigillito, Simon P Wing, Larrie V Hutton, and Kile B Baker. Classification of radar returns from the ionosphere using neural networks. *Johns Hopkins APL Technical Digest*, 10(3):262–266, 1989.

[10] Eugen Gillich and Volker Lohweg. Banknote authentication. *1. Jahreskolloquium Bild. Der Autom*, pages 1–8, 2010.

[11] W Nick Street, William H Wolberg, and Olvi L Mangasarian. Nuclear feature extraction for breast tumor diagnosis. In *Biomedical image processing and biomedical visualization*, volume 1905, pages 861–870. SPIE, 1993.

[12] Michele Forina, Riccardo Leardi, C Armanino, Sergio Lanteri, Paolo Conti, P Princi, et al. Parvus an extendable package of programs for data exploration, classification and correlation. 1988.

[13] Allison Marie Horst, Alison Presmanes Hill, and Kristen B Gorman. palmerpenguins: Palmer archipelago (antarctica) penguin data. *R package version 0.1. 0*, 2020.

[14] R Paul Gorman and Terrence J Sejnowski. Analysis of hidden units in a layered network trained to classify sonar targets. *Neural networks*, 1(1):75–89, 1988.