arXiv:2301.06815v1 [cs.SI] 17 Jan 2023

关注我们，成为名人！Instagram 互动机制的洞察与指南

Pier Paolo Tricomi*
帕多瓦大学，意大利
tricomi@math.unipd.it

Marco Chilese*
marco.chilese@trust.tu-darmstadt.de
达姆施塔特工业大学，德国

Mauro Conti 帕多瓦大学，意大利
conti@math.unipd.it

Ahmad-Reza Sadeghi 达姆施塔特工业大学，德国
ahmad.sadeghi@trust.tu-darmstadt.de

摘要

Instagram（Instagram）拥有 13 亿用户，也已成为一种商业工具。预计 2022年Instagram网红营销将创造 332.5 亿美元的收入，并推动公司和网红创作热门内容。目前已有多种方法用于预测帖子的受欢迎程度，也就是它能带来多少互动（例如点赞）。然而，这些方法存在局限：首先，它们着重预测点赞数，却忽视了评论数量，而评论在 2021年变得至关重要。其次，研究往往使用存在偏差或规模有限的数据。第三，研究人员专注于深度学习模型来提升预测性能，但这类模型难以解释。因此，终端用户只能在帖子发布后估算互动情况，既低效又昂贵。更好的做法是根据人们和 Instagram 的喜好来创作帖子，例如遵循相关指南。

在这项研究中，我们揭示了驱动 Instagram 互动的部分潜在机制。为实现这一目标，我们采用统计分析和可解释模型，而非深度学习（黑箱）方法。我们使用一个全球数据集开展了广泛实验，该数据集包含 10 万名来自九个不同类别的全球网红发布的 1,000 万条帖子。凭借简单而强大的算法，我们能够以高达 94% 的 F1分数预测互动，表现可与基于深度学习的方法相媲美，甚至更胜一筹。此外，我们提出了一种新颖的无监督算法，用于发现 Instagram 上互动度极高的话题。借助可解释的方法，我们最后概述了创作成功帖子的指南。

CCS 概念

• **信息系统** → **社交网络**；• **计算方法学** →**机器学习**。

关键词

Instagram、互动、人气、可解释人工智能、社交网络、帖子人气

1 引言

人们在 Instagram（IG）上发布照片有许多目的，例如传达个人身份、与他人建立联系，或推广有价值的内容 [41]。获得他人的认可会带来极大的满足感，以至于互动指标（例如点赞、评论、浏览量）已经变得令人上瘾，尤其是对于自尊较低的人而言 [26]。有些人每天只使用几分钟 IG，但对另一些人来说，例如网红，IG 已经成为一种生活方式。简而言之，网红是能够影响社会的人。由于他们拥有触达受众的能力，公司一直利用他们来营销产品 [27], ，以至于据估计，网红营销将产生 2022 [16, 17]亿美元的收入。无论出于何种原因，每个人都努力让自己的帖子获得尽可能多的互动，即使不惜花钱购买互动 [50, 52]。对于网红而言，策划受欢迎的帖子是一项耗时且成本高昂的活动，而且无法保证成功。在这方面，如果有一种工具能够提前预测帖子的受欢迎程度，将会非常有价值，尤其是在赞助帖子报酬丰厚的情况下（例如，克里斯蒂亚诺·罗纳尔多发布一条帖子就能获得约 100 万美元的报酬 [5]）。

研究人员提出了用于预测帖子受欢迎程度的算法，但这些算法远非完美（§2.1）。第一个局限是，它们仅以点赞来衡量互动，没有纳入更强形式的互动，也未考虑 Instagram 所偏好的指标，即评论数量 [48]。其次，缺乏用于此类预测任务的通用数据集，导致结果建立在有限或有偏的数据之上。此外，这些模型通常使用深度学习（DL）模型，而这类模型可能难以（甚至不可能）解释 [14, 44]。因此，终端用户必须先经历昂贵且耗时的流程来创建帖子，然后使用这类黑箱模型评估帖子的受欢迎程度。

贡献。我们的目标是理解并解释驱动 Instagram 互动的底层机制。我们利用深度学习模型广为人知的能力，提取与领域相关的特征，但将预测任务交由可解释机器学习（ML）算法完成 [31], ，从而得以提出指南。根据 Instagram 推荐算法，我们将点赞和评论都视为互动指标。我们针对一个近期数据集开展了大量实验，该数据集包含来自 3.4 万名网红的 1000 万条帖子。通过统计分析，实验表明，网红层级（即其受众范围）和类别（即他们主要涉及的主题）是预测帖子受欢迎程度的关键因素。图1 展示了吸引人的帖子和不吸引人的帖子，这印证了这样一种直觉：决定互动的特征因类别而异。



图1：各类别中吸引人和不吸引人的帖子。

*两位作者对本研究的贡献相同。

最后，我们提出一种新颖的无监督方法，用于检测每个类别中的热门话题（即能够引发高互动的话题），从而克服了需要领域知识才能提取有意义特征的限制。我们的贡献总结如下：

- 我们基于一个包含 1000 万条帖子的 数据集（划分为九个类别和五个网红层级），运用统计分析和可解释的机器学习算法，分析 Instagram 互动（点赞和评论）背后的机制；
- 我们提出一种可解释模型，用于预测帖子的互动，并利用最先进的深度学习模型提取的多种特征，制定实用指南；
- 我们提出一种新颖的无监督方法，用于发现各层级和类别中能够引发高互动的话题，同时考虑视觉内容和文本内容；
- 我们发布了经过扩充的数据集[1]，可作为未来研究的潜在基线。

论文结构。 §2 介绍相关研究。§3 描述数据集和初步评估；互动预测与解释见§4。热门话题检测见§5，最终的指南见§6。§7 总结全文。

透明度。 为提升透明度和可复现性，我们创建了一个代码仓库[2]，其中包含有关本研究的详尽信息、源代码和数据集。

2 相关研究

Instagram 帖子的受欢迎程度主要通过预测其获得的点赞数来评估，通常将其除以发帖用户的粉丝数，或进行对数尺度变换。Mazloom 等人 [29] 通过定义对营销至关重要的互动参数，并采用支持向量回归，预测与品牌相关帖子的受欢迎程度。作者进一步将研究 [28] 扩展到活动、风景、人物和动物等不同类别。De 等人 [6] 使用帖子元数据（例如创建日期、被标记的用户、话题标签）训练深度神经网络（DNN），以预测印度某生活方式杂志未来帖子的受欢迎程度。类似地，Zhourian 等人 [58] 将受欢迎程度预测视为回归和分类任务，重点研究伊朗商业 Instagram 账号发布的帖子。Zhang 等人 [56]没有预测总体受欢迎程度，而是采用双重注意力机制，预测用户特定的帖子受欢迎程度。Ding等人 [8] 尝试通过 DNN 预测图像固有受欢迎度，从而分离视觉内容的影响。Gayberi 等人 [10] 使用基于 Microsoft COCO 数据集 [24] 预训练的模型提取概念和物体特征，并采用多种机器学习算法预测帖子的点赞数。Riis 等人 [43] 通过迁移学习提取概念、场景和物体等视觉语义，并尝试为基于总体的受欢迎程度预测建立可解释基线。Carta 等人 [3] 提出了一种基于梯度提升的方法，并对用户和帖子元数据进行特征工程，以分类方式预测受欢迎程度。最后，Purba 等人 [39] 尝试创建一个全球数据集，其中包含来自 1.6 万名用户的约 2 万条帖子，并利用从话题标签、图像分析和用户历史中提取的特征，采用支持向量回归（SVR）预测点赞数与粉丝数之比。

2.1 现有文献的局限性

本节将简要说明该领域以往文献为何不够完善，以及我们的工作如何弥补这些不足。

不完善的热度指标。 以往研究仅 关注点赞数来衡量帖子热度，这种做法已过时，也与当前的Instagram算法不一致。Instagram算法在 2021 [48] 中进行了更改，旨在根据用户的兴趣而非仅仅依据其社交图谱向用户展示内容。转向这种推荐媒体 改变了帖子走红的方式。内容必须 获得互动，主要是通过点赞和评论，这样Instagram才会将其推送到更多用户的动态流中，之后才可能走红。因此，将评论数作为互动指标至关重要，因为评论比点赞 [1], 更能体现用户的表达 [48]，因而与Instagram推荐算法的相关性更高。据我们所知，我们是首个将评论纳入互动指标的研究。

数据集有限或存在偏差。 由于 Meta 的 API[3] 受到限制，因此没有可用作基线的公开数据集。大多数先前研究都自行收集数据集，且只关注总体中的有限部分 [6, 58]。此外，除了 Mazloom 等人 [28], 之外，其他研究都没有考虑创作者所属的不同类别和层级；然而，例如，狗的照片和顶级模特的照片走红的原因可能截然不同。此前也没有考虑网红层级。不过，拥有数百万粉丝的网红与只有几千名粉丝的网红，其互动率水平各不相同 [11],，仅对指标进行归一化并不足够。在 §3.3中，我们证明网红的类别和层级会显著影响互动指标（p-值<0.001），因此需要分别处理这些因素，才能得到准确的预测。

结果解释能力差。 随着深度学习算法和集成机器学习算法的性能不断提升，近期研究大多依赖端到端黑箱模-型 [8, 43, 56]，而不是提取特定特征来训练简单的回归器或分类器 [29, 58]。虽然模型的准确率更高，却难以（甚至无法）理解模型学到了什么 [14, 44, 45]。正如 Rudin 在《自然》上的里程碑式论文中充分论证的那样 [44],，当可解释性至关重要时，必须 优先选择可解释模型，而非（复杂的）黑箱模型。通常，如果问题涉及结构化数据和有意义的特征，那么更复杂的分类器（即深度神经网络、集成方法）与更简单的分类器在性能上并无显著差异。此外，在我们的场景中，使用黑箱模型预测帖子热度意味着用户必须先创建帖子，这可能代价极高 [5]。因此，我们使用可解释模型（即决策树），为希望提升帖子热度的用户提供可在发帖之前 遵循的指南。

3 数据集与初步评估

在下文中，我们将介绍数据集（§3.1）、互动指标（§3.2）、按类别和内部层级剖析数据的重要性（§3.3），以及我们在研究中考量并提取的特征（§3.4）。

[1]该内容将在发表后发布。 [2]该内容将在发表后公开。

[3]https://developers.facebook.com/docs/instagram-api/

3.1 数据集描述

在我们的研究中，我们使用了Kim等人提出的数据集。 [19]该数据集共包含来自33,935位网红的10,180,500条帖子，这些数据是在92天内收集的。这些网红根据其内容类型分为九个类别，分别是美容、家庭、时尚、健身、美食、室内、宠物、旅行和其他。此外，我们根据网红的关注者数量 [50]，将每位网红划分为五个广为人知的层级：纳米级 [1K，10K）、微型 [10K，50K）、中型 [50K，500K）、大型 [500K，1M）、超大型 [1M，+∞]。

每篇帖子由图像、配文、元数据（例如发布时间、位置）和互动指标（即点赞数和评论数）组成[4]。与以往研究类似 [39, 43, 58],，我们将目标特征（点赞数和评论数）除以帖子创作者的粉丝数量进行归一化，从而能够公平地比较不同用户发布的帖子[5]。鉴于创作者的粉丝数量仅在数据收集结束时统计，我们剔除了发布超过三十天的帖子；在这段时间内，粉丝数量的增长基本保持稳定 [34]。此外，由于 Instagram 帖子的互动增长通常持续一至三天 [11],，我们剔除了发布未满五天的帖子。最终，我们的数据集包含 33,935 位网红发布的 650,118 篇帖子。表 1 展示了各层级和类别的帖子数量（以及网红数量）。某些类别（例如美食、室内、宠物）在极受欢迎的网红中占比较低，这与 Instagram 实际的类别分布一致 [15]。

表 1：各类别和层级的帖子数（网红数）。

	Nano	微型	Mid	大型	Mega
美容	8449 (546)	7879 (537)	6998 (387)	745 (35)	835 (37)
家庭	29744 (1887)	23432 (1330)	12740 (674)	1267 (77)	2145 (102)
时尚	49622 (3154)	82895 (4841)	68737 (3238)	8833 (325)	8987 (355)
健身	5060 (301)	6194 (424)	6256 (342)	352 (27)	748 (39)
Food	27697 (1511)	28191 (1440)	14805 (583)	936 (25)	305 (6)
室内	6461 (373)	9606 (541)	5525 (261)	413 (13)	404 (7)
Pet	3416 (164)	4073 (260)	2929 (153)	87 (6)	115 (4)
旅行	24445 (1774)	19630 (1522)	13098 (838)	816 (49)	540 (27)
其他	73213 (2976)	38967 (1454)	31874 (1004)	4255 (120)	6399 (166)

3.2 互动指标：点赞与评论

先前的研究（§2）仅将点赞数作为人气指标。然而，自2021年起，评论已成为让帖子走红的重要互动指标 [48]。图2 展示了各类别和层级的点赞数与评论数箱线图。虽然存在一些共同趋势，但发表评论的频率低于点赞。这种差异可由二者所代表的公开表达程度不同来解释 [1]。评论需要付出更多精力，也更能暴露用户的观点；而点赞几乎是即时且本能的。因此，获得大量点赞的帖子未必会收到很多评论。为进一步说明这两个指标彼此独立，我们计算了点赞数和评论数分布之间的斯皮尔曼相关系数（rho）。结果（$rho = 0.58$, p-值 < 0.001）显示点赞数与评论数之间存在中等程度的相关性，说明二者并非高度依赖，应作为两种现象分别分析。因此，我们将$\#Likes\#Followers$ 和$\#Comments\#Followers$ 视为互动指标。

[4]我们没有对这些指标进行进一步处理，例如删除垃圾评论，因为 Instagram算法考虑的是数量，而非质量 [48]。[5]为方便起见，我们将归一化后的数字简称为点赞和评论。

3.3 层级与类别的重要性

北极光带来的互动会比可爱的小狗更多吗？那不勒斯的菠萝披萨又如何？由于这些概念无法相提并论，事先很难回答这些问题。同样，如果离婚的是一位名人与一位普通人物，人们的反应会一样吗？很可能不会。这些只是支持我们假设的几个例子：网红的层级和类别会对互动指标产生显著影响。为验证这一假设，我们进行多元方差分析（MANOVA） [25],，将类别和层级作为自变量，将点赞数和评论数作为因变量。通过这项统计检验，我们可以判断九个类别和五个层级之间的互动平均分是否存在差异。在进行多元方差分析之前，我们按照§3.2所述，对点赞数和评论数的分布进行了标准化。在多元方差分析的结果中，我们采用了 Pillai迹检验；当多元方差分析的假设不成立时，该检验具有较强的稳健性 [49]。Pillai迹检验对类别和层级的检验统计量分别为 0.0942 和 0.2646，二者的 p-值均为 < 0.0001。由于 p-值小于显著性水平 $\alpha = .0001$，我们拒绝多元方差分析的零假设，并得出结论：解释变量（层级和类别）会显著影响响应变量（点赞数和评论数）的取值。具体而言，层级的影响大于类别。

3.4 特征提取

从 §3.1 中经过筛选的帖子出发，我们利用各类数据源（例如元数据、图像和文本）中的特征扩充了数据集，现简要介绍如下。[6] 在此过程中，我们还采用了九种最先进的深度学习算法。

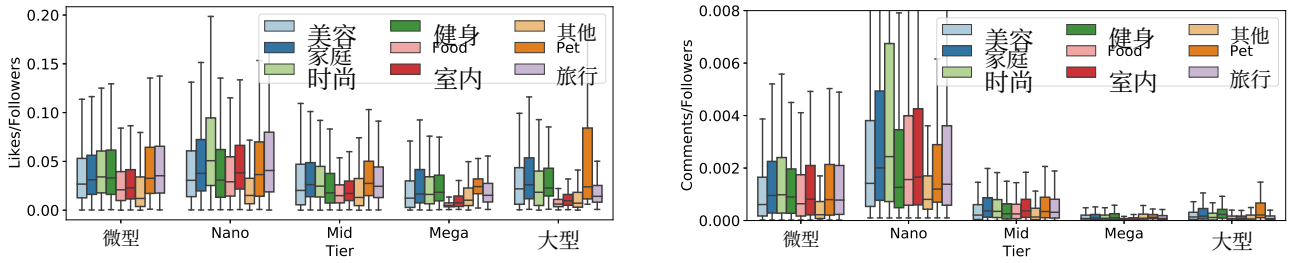
3.4.1 元数据特征。 帖子的元数据提供了有关其“可发现性”的信息。这个术语指的是能提高帖子可见度的特征，例如话题标签和提及。话题标签用于标记帖子的内容，而提及则可在帖子中标记某人，使其粉丝能够找到来源个人资料。因此，我们设置了两个计数器，分别记录话题标签的数量和被标记用户的数量。此外，我们还记录帖子是否为视频或赞助内容、是否包含位置以及与时间相关的信息，共计 10 个特征。

3.4.2 图像特征。 我们从多个层面提取图像特征，以全面描述图像内容，包括场景特征、人物特征和审美特征。

场景特征。 为了描述图片所处的环境，我们采用 Places365 深度学习模型 [57]。该模型最多可识别 365 种不同场所，并将其划分为 3 个大型类别（室内、自然户外、人工户外）和 16 个微型类别（例如购物/餐饮、交通）。此外，我们使用在 MS COCO 数据集 [24], 上训练的 Faster R-CNN MobileNetV3 [37]），对映射至 12 个类别的 80 种不同类别进行物体检测，并统计属于各类别的物体数量。

人物特征。 使用 RetinaFace [7] 进行人脸边界检测，随后估算检测到的人物的年龄和性别 [32] 。对于每个帖子，我们都会记录女性和男性的人数，以及人物年龄的最小值、最大值、平均值和标准差。此外，鉴于众所周知的广告中的裸露内容所产生的影响 [46],，我们使用 NudeNet [36]进行裸露内容检测

[6]完整的特征列表可在我们的代码仓库中查看。



(a) 点赞。 **(b) 评论。**

图2：不同类别和层级的点赞箱线图 和评论箱线图。请注意，y 轴采用了两种不同的刻度，es, 以便显示较低的评论数 一般评论。

适用于美容 和时尚 类别，这些类别的主体是人体。该模型会判断人体的 16 个部位（例如乳房、腹部、脚、臀部）是否裸露。

审美特征。受 Guntuku 等人启发， [4], 我们提取了图像的审美特征。具体而言，我们首先提取红、绿、蓝通道的占比。然后，根据 HSV（色相、饱和度、明度）表示，获取亮度、暖色和冷色的占比，以及愉悦度、唤醒度和支配度得分 [30, 51]。此外，我们利用 Kong 等人的 [20] 模型提取了 11 项高层次审美特征（例如，内容的色彩和谐、运动模糊和对称性）。最后，我们通过 Campos 等人提出的模型提取图像所传达的情感得分。 [2]。

其他特征。对于宠物类别，我们使用可爱动物检测器 [18] 计算宠物的可爱度得分。总共获得了 80 项视觉特征。

3.4.3 文本特征。我们从帖子图注中提取了标题长度、表情符号数量及其相对情感等特征 [21]。此外，我们利用 Google 认知服务（GCP） [12], 获取整段文本的情感，并以得分表示（情感得分 $\in [-1, 1]$，其中 -1 表示负面，0 表示中性， $+1$ 表示正面）和强度（情感强度 $\in [0, +\infty)$ ，以反映情感的强弱。我们使用 GCP 翻译非英语文本，总共获得了五项文本特征。

4 预测并解读互动

通过相关性分析，我们找出与互动相关的特征。然后，我们使用可解释模型预测互动，并制定制作高互动内容的指南。

4.1 相关性分析

为了确定哪些特征对提升互动的贡献最大，我们将这些特征与两个互动指标（点赞和评论）进行相关性分析。为此，我们采用斯皮尔曼秩相关系数 r_s [59]。该方法具有以下优点：能够生成特征排名、对离群值不敏感，并且不要求对数据进行任何特定的归一化。斯皮尔曼相关系数基于皮尔逊相关系数 [33] ，并且它定义如下。对于 n 个观测值， n 得分 X_i、 $Y_i X_i$、 Y_i被转换为排名，即 $R(X_i)$ 和 $R(Y_i)$，并计算 r_s 如下：

$$r_s = \rho_{R(X),R(Y)} = \frac{\mathrm{cov}(R(X), R(Y))}{\sigma_{R(X)}\sigma_{R(Y)}} \quad (1)$$

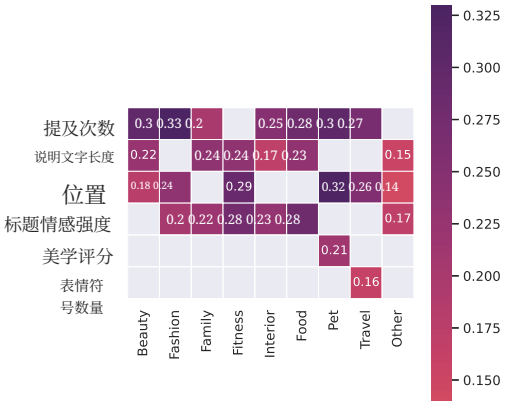
其中， ρ表示应用于秩变量的培生相关系数，$\mathrm{cov}(R(X)\ R(Y))$ 表示秩变量的协方差，$\sigma_{R(X)}$ 和$\sigma_{R(Y)}$ 表示秩变量的标准差。与培生相关系数一样，斯皮尔曼相关值的范围为 $r_s \in [-1, 1]$ ，同时还带有 ρ值，用以表示其显著性；该值越小，显著性越高。

对于每种网红类别（即美容、时尚等）以及每个层级（即纳米、微型等），我们都会分析特征与互动指标之间的相关性[7]。首先，我们立即注意到，如图3所示，当层级较低（例如纳米和微型）时，不同类别中最相关的特征非常相似。相反，随着层级规模增大（例如大型和超大型），行为会因类别而异；这一规律在点赞和评论方面都成立。正如我们所见，小型网红，或有志成为网红的用户，在各个类别中都会采用相似的策略，包括大量提及、较长的配文以及位置标签。

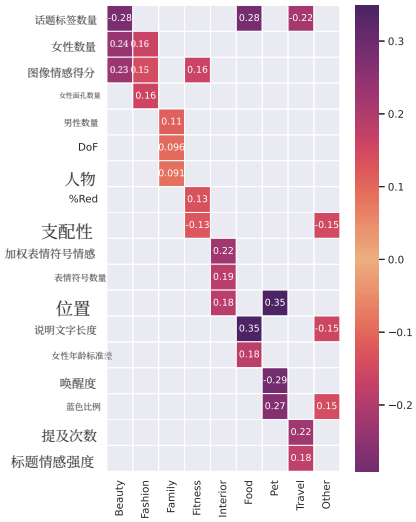
点赞互动。通过考察特征与点赞之间的相关性，可以观察到不同类型网红的互动机制有何差异。总体而言，我们可以发现，影响最大的特征与图像及其内容有关，而非文本（即配文）；评论的情况几乎恰恰相反。提及数量通常会对点赞产生正面影响，但其重要性会随着层级升高而降低。话题标签数量也呈现类似规律，会产生负面影响，且这一影响往往会随着层级升高而加剧。位置是否可用在微型级别及以下都非常重要，超过该级别后则取决于类别。

评论互动。与点赞互动类似，我们观察到提及次数整体呈正相关，尽管其相关性会随着层级提高而降低。即使在这种情况下，话题标签的数量也会对评论互动产生负面影响，而

[7]我们已在代码仓库中公开所有结果。



(a) 纳米级。



(b) 超大型。

图3：按相关值绝对值排序的前3项特征。评论互动。

帖子中的位置字段通常会有所帮助。这类互动会受益于文本特定特征，例如标题长度、情感强度和表情符号使用。

> 要点：点赞互动与评论互动的不同之处在于，前者面向图像，后者则面向图注。此外，低层级网红往往采用相同的增长策略，而高层级网红则表现出更多类别特定特征。

4.2 互动预测与指南方法论

除了说明 Instagram 帖子的哪些特征能够促进互动，我们还旨在制定指南，帮助创作出理想的高互动帖子。对于网红而言，遵循这类指南可以持续节省时间和金钱，因为制作高互动帖子的流程已有明确定义。为此，我们采用可解释模型，即便这可能会降低整体准确率。

深度学习模型以能够解决复杂任务而闻名，但从定义上来说，它们是黑箱，我们无法真正解释 [44]。因此，我们决定采用决策树（DT）[40]。通过训练决策树分类器来预测低互动或高互动（分别对应最低 0.75 分位数和最高 0.25 分位数），我们可以通过简单地遵循二元分类树，解释如何创作出互动率最高的帖子。通往最高 0.25 分位数叶节点的路径，代表了创作高互动帖子的指南。

4.2.1 实现。 由于网红会根据其类别和层级表现出不同的行为，并且他们希望触达不同的受众，因此我们的分析针对每个类别和层级分别进行。为构建准确的估算器，我们针对每个数据集，通过网格搜索（cv=5）对决策树分类器进行微调，评估超过 2 万种参数拟合组合，以尽可能获得最佳的 F1分数（宏平均加权）。为了进一步减少数据集随机划分带来的偏差，我们在三个不同的数据分区上重复进行了三次评估。随后，每个表现最佳的分类器都使用数据集的 75% 进行训练，并在剩余的 25% 上进行测试。根据 0.75 和 0.25 分位数划分低互动和高互动，会导致类别严重不平衡，使学习过程更加困难。因此，我们还将采用广为人知的欠采样和过采样技术作为调优参数，即 SMOTE 和 Tomek链接 [23, 54]。

4.2.2 结果。 测试集上的结果汇报于表 2中。所有结果均优于虚拟分类器，表明我们的方法能够有效预测帖子互动。此外，标准差相当低，说明模型稳定。在点赞方面，对于大型和超大型网红，预测通常更准确，F1分数约为 60%–80%（比虚拟分类器高 20%–40%）。原因可能是，正如相关性分析中发现的那样，这些高层级网红的内容往往更加多样化，因此某些特征更为有效。相应地，我们的分类器难以预测时尚类别中层级较低的网红，因为他们往往发布相似内容；对于其他类别也存在困难，因为其内容极为多样化。平均而言，我们在宠物、室内和美容类别上取得了最佳表现。在评论方面，我们发现其表现模式与点赞类似，但时尚和家庭类别的最佳表现出现在纳米网红中。一个可能的原因是，许多纳米网红可能不了解如何创作能够引发互动的图注，而正如相关性分析所示，这类图注与评论互动密切相关。预测效果最好的类别是宠物、美食和旅行。最后，在宠物类别的大型网红帖子上，我们取得了最佳点赞和评论预测分数（分别为 94% 和 87%）。一个[8] 采用决策树结构的指南示例见图 4。沿着节点条件（即帖子特征）逐层判断，到达叶节点时（即底部 0.75 或顶部 0.25 分位数）便会分配标签。我们将在 §6中展示更多指南示例。

4.2.3 基线比较。 正如 Mazloom 等人 [28], Gayberiand Oguducu [10], 以及其他类似研究所提到的，与该领域的其他研究进行比较并非完全可行。主要原因在于使用了私有算法和数据，以及问题的

[8]所有结果都可以在我们的代码仓库中找到。

表 2：决策树（DT）与虚拟分类器（Dum.）的性能。每个类别中点赞和评论的最佳分数以粗体显示。报告的数值为 F1 分数，按大型加权 *mean±std.*

	Nano				微型				Mid				大型				Mega			
	Like		评论		Like		评论		Like		评论		Like		评论		Like		评论	
	DT	Dum.	DT	Dum.	DT	Dum.	DT	Dum.	DT	Dum.	DT	Dum.	DT	Dum.	DT	Dum.	DT	Dum.	DT	Dum.
美容	0.61±0.002	0.46±0.004	0.65±0.003	0.46±0.004	0.60±0.003	0.47±0.012	0.62±0.006	0.46±0.017	0.61±0.004	0.47±0.011	0.61±0.002	0.47±0.016	0.69±0.012	0.45±0.054	0.67±0.011	0.47±0.046	**0.74±0.006**	0.48±0.029	**0.68±0.009**	0.49±0.025
时尚	0.57±0.006	0.47±0.008	**0.62±0.003**	0.47±0.001	0.53±0.002	0.46±0.002	0.57±0.004	0.46±0.004	0.53±0.004	0.47±0.004	0.56±0.007	0.47±0.002	**0.61±0.001**	0.47±0.009	0.60±0.011	0.48±0.015	0.57±0.013	0.46±0.011	0.56±0.002	0.47±0.005
家庭	0.57±0.005	0.47±0.010	**0.59±0.005**	0.47±0.007	0.53±0.002	0.46±0.005	0.56±0.002	0.47±0.008	0.55±0.007	0.48±0.005	0.55±0.002	0.47±0.003	**0.60±0.014**	0.47±0.031	**0.59±0.013**	0.47±0.017	0.59±0.006	0.47±0.026	0.57±0.001	0.46±0.007
健身	0.62±0.004	0.48±0.005	0.63±0.004	0.47±0.009	0.60±0.006	0.47±0.009	0.59±0.012	0.47±0.006	0.61±0.010	0.47±0.009	0.61±0.003	0.47±0.010	**0.72±0.010**	0.45±0.045	**0.69±0.032**	0.46±0.023	0.65±0.013	0.49±0.016	0.61±0.033	0.48±0.010
Food	0.59±0.003	0.46±0.003	0.62±0.001	0.47±0.010	0.57±0.001	0.47±0.007	0.61±0.001	0.47±0.006	0.57±0.005	0.46±0.010	0.63±0.002	0.46±0.011	**0.75±0.008**	0.46±0.020	0.71±0.024	0.47±0.061	0.71±0.023	0.57±0.091	**0.72±0.029**	0.53±0.098
室内	0.59±0.002	0.48±0.012	0.62±0.003	0.46±0.01	0.59±0.003	0.47±0.002	0.63±0.001	0.47±0.006	0.58±0.010	0.45±0.012	0.63±0.003	0.47±0.011	0.76±0.033	0.48±0.021	**0.83±0.020**	0.49±0.015	**0.77±0.011**	0.44±0.083	0.74±0.027	0.45±0.074
其他	0.58±0.002	0.47±0.001	0.57±0.0001	0.47±0.004	0.55±0.001	0.47±0.0021	0.56±0.005	0.47±0.002	0.59±0.001	0.47±0.006	0.58±0.001	0.46±0.005	**0.63±0.003**	0.46±0.013	**0.59±0.008**	0.47±0.023	0.60±0.003	0.46±0.005	0.58±0.008	0.46±0.012
Pet	0.69±0.004	0.47±0.019	0.72±0.006	0.49±0.009	0.60±0.008	0.45±0.018	0.62±0.005	0.46±0.007	0.61±0.018	0.46±0.006	0.64±0.003	0.47±0.024	**0.94±0.057**	0.42±0.042	**0.87±0.048**	0.42±0.042	0.78±0.045	0.51±0.032	0.77±0.012	0.5±0.065
旅行	0.60±0.001	0.47±^{0042}	0.61±0.001	0.47±0.006	0.59±0.0012	0.47±0.003	0.63±0.002	0.47±0.007	0.58±0.004	0.47±0.006	0.64±0.009	0.48±0.010	**0.72±0.008**	0.45±0.024	0.67±0.050	0.47±0.012	**0.72±0.010**	0.46±0.031	**0.71±0.017**	0.42±0.032

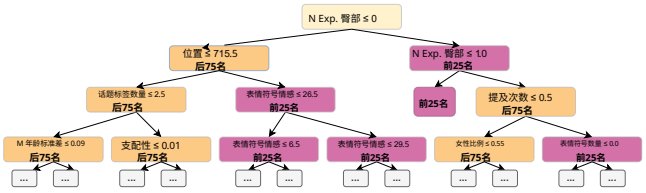


图4：决策树为美容类别、纳米层级、点赞互动生成的指南示例。该表示的最大深度限制为 3。

表述方式。遗憾的是，IG 的政策[9] 从未允许自动收集和发布普通用户的帖子，迫使先前的研究每次都创建新的（私有）数据集 [3, 6, 18, 28, 29, 39, 58]。

此外，由于缺乏可供研究使用的通用数据集，一些研究关注回归问题 [10, 39, 58]，，另一些则关注分类问题 [3, 6, 58]，，并采用了不同的指标，例如点赞数对数归一化值 [8, 10] 或点赞数除以关注者数量 [39, 43, 58]。因此，尽管存在上述限制，[10] 我们仍采用了四种模型来建立基线：(i) Ding等人的 I^{2}PA [8]（唯一公开发布的模型）；(ii) 决策神经网络（Dec-NN），用于那些先提取通用视觉或文本特征、随后训练不可解释分类器的先前研究（类似于 [10, 28, 43]）；(iii) 端到端深度神经网络（EE-DNN），用于那些依赖端到端黑盒深度学习模型、同时输入帖子图像和配文的先前研究（类似于 [56]）；(iv) 分层虚拟分类器，[11]，根据训练集分布预测目标。Dec-NN 和 EE-DNN 都会提取帖子的图像嵌入和配文嵌入（分别通过 ResNet50 [13] 和 Sentence-Bert [42]）；不过，EE-DNN 会在融合前对其进行微调，而 Dec-NN 则以早期融合后的结果作为输入。决策通过三个 ReLU 前馈层作出（层大小为 = 2048+768 → 256 → 128 → 2）。两种神经网络均使用 Adam 优化，并训练 50 轮，采用提前停止策略（耐心值为 = 5）。

表3 的结果表明，除时尚和其他类别外，我们的方法在各个类别中均优于基线模型；在这两个类别中，我们的表现与之相当。这证明了简单的决策树优于深度学习模型。对于以下类别美容、健身、美食、室内、宠物和旅行，我们的结果在统计上显著高于排名第二的模型（通过非配对 t 检验计算，双尾 p 值为 < 0.05）。尤其值得注意的是，与 I^{2}PA 和 EE-DNN 的对比结果；这两者代表了当前最先进的端到端深度学习模型。EE-DNN 的表现尤其不佳，可能是因为微调特征提取模块导致了过拟合。另一方面，与我们策略更相似的 Dec-NN 没有调整图像和文本的通用表征，因此泛化能力更强。我们可能主要是凭借提取的类别相关特征超越了这些基线模型，这再次说明开发跨类别互动预测器可能并不可行。因此，在其他类别中，我们可能无法击败 Dec-NN，因为缺少类别相关特征。

尽管出于上述原因，与先前研究的比较并不完全公平，但我们的结果 [6, 56]与之相当 [3, 39, 43]或优于其在各自数据上报告的结果。无论如何，我们提醒读者，我们的目标是解释互动，而非一定要超越现有不可解释模型的预测表现。最后，我们通过 Instagram API 收集了来自商业账户的数据集，因此该数据集可以共享。我们认为，该数据集可作为未来研究的基线。

表 3：我们的模型（决策树）与基线模型在预测点赞方面的平均 F1 分数比较。

类别	决策树（我们的模型）	I^{2}PA	Dec-NN	EE-DNN	虚拟模型
美容	**0.65±0.055**	0.587±0.026	0.582±0.097	0.362±0.037	0.466±0.024
时尚	0.563±0.030	**0.581±0.019**	0.572±0.043	0.327±0.013	0.464±0.008
家庭	**0.568±0.026**	0.567±0.019	0.507±0.071	0.347±0.042	0.476±0.017
健身	**0.640±0.043**	0.545±0.026	0.511±0.069	0.377±0.062	0.478±0.018
Food	**0.638±0.077**	0.550±0.030	0.518±0.053	0.464±0.269	0.48±0.031
室内	**0.660±0.087**	0.534±0.056	0.461±0.047	0.309±0.031	0.468±0.022
其他	0.590±0.026	0.540±0.022	**0.602±0.021**	0.318±0.013	0.463±0.004
Pet	**0.724±0.126**	0.564±0.046	0.630±0.127	0.342±0.0727	0.461±0.015
旅行	**0.642±0.064**	0.570±0.012	0.473±0.044	0.342±0.069	0.457±0.016

带下划线的结果在统计上显著高于第二高的结果（双尾 p-值 < 0.05）。

4.2.4 特征重要性。根据决策树分类器生成的树得出创建吸引人的帖子的指南。此外，与相关性分析类似，内容创作者可以查看模型的特征重要性，以确定哪些特征会影响互动预测。因此，我们研究了模型使用的特征，并检查它们是否与

[9] https://help.instagram.com/581066165581870，访问时间：2022 年 9 月。[10] 请注意，先前研究中使用的某些特征在我们的数据集中不可用，因此限制了比较。[11] https://scikit-learn.org/stable/modules/generated/sklearn.dummy.DummyClassifier.html

表4：时尚类别、微型层级的特征重要性。

#	特征	Imp.
1	提及次数	1.0
2	平均年龄	0.80
3	支配性	0.73
4	臀部暴露次数	0.46
5	户外自然环境	0.19

(a) 点赞。

#	特征	Imp.
1	N Exp. 臀部	1.0
2	提及次数	0.64
3	说明文字长度	0.18
4	表情符号数量	0.16
5	标题情感强度	0.13

（b）评论。

相关性分析结果一致。表[12]中展示了此项分析的一个典型示例4，结果表明，这与第§4.1节所述因素高度吻合。例如，小层级中存在共同特征，而随着层级规模增大，则出现类别特定的特征。与相关性分析一样，提及次数以及是否设置了位置，其重要性与层级规模成反比。

要点：如果利用领域知识特征，简单且易于解释的决策树可以胜过深度学习算法。预测结果和特征重要性分析证实了基于特征相关性得出的判断，展示了相似与不同层级和类别的表现。

5 发现 Instagram 热门话题

我们的特征能够较准确地预测帖子的互动情况，但仍有改进空间。在我们的可解释方法中，特征必须预先提取，而不是由深度学习模型"自动"学习。因此，我们的特征受限于我们对哪些内容可能引发互动的合理推测，以及现有最先进深度学习模型所能识别的概念。例如，如果有现成的模型，我们会使用爱情或婚礼场景检测器，因为这类场景很可能带来高互动。虽然可以通过经典方法实现这类检测器（例如微调一个图像识别神经网络，如 ResNet [13]），但我们选择定义一种无监督策略来检测通用热门话题。具体来说，我们旨在找出（如果存在的话）这样一些话题或概念：只要它们出现在帖子中，无论发布者是谁，都能带来高互动。在此语境下，无监督意味着我们不对哪些话题能够引发互动作任何假设（不同于我们在§4中提取类别相关特征时所做的假设），而是探索用户兴趣 [53]。

从§4我们了解到，点赞和评论主要分别由图像和配文驱动。因此，在接下来的实验中，我们着重通过视觉特征寻找与点赞相关的热门话题，并通过文本特征寻找与评论相关的热门话题。现在，我们将介绍方法论和研究发现。

5.1 方法论

我们的方法旨在将语义相似的图像和图注归为一组，并观察其中是否有些组的平均互动度较高。

嵌入。为了定义图像和文本语义相似度，我们采用嵌入这一概念。嵌入是对象（例如图像、文本）的向量表示，其中语义相似的对象具有相似的向量轮廓 [22]。嵌入通常通过提取执行分类任务的深度神经网络倒数第二层的输出获得。在我们的实验中，我们使用基于 ImageNet 数据集预训练的 ResNet50 [13] 提取图像嵌入，并使用 Sentence-Bert [42] 提取文本嵌入（具体而言，使用其 *all-mpnet-base-v2* [9]版本）。在提取文本嵌入之前，我们借助 Google Cloud Platform [12], 翻译非英语文本，以自动进行语言检测和翻译。

语义相似邻域。作为第一种方法，我们可以将相似的图像或图注聚类，并观察某些聚类的互动度是否高于其他聚类。然而，正如文献所示 [35], 目前的聚类算法需要预先决定聚类数量。此外，寻找热门话题也颇具挑战 [55], 因为它们可能规模很小，因而被淹没在大型聚类中。因此，我们更倾向于采用最近邻方法，寻找互动度相似的点所构成的邻域。具体来说，我们首先根据百分位数将帖子划分为五个互动度类别 [0-20、20-40、40-60、60-80、80-100], , 并保存每个百分位数对应的阈值。然后，对于每个点，我们搜索其 N 个最近邻，计算这些邻居的平均互动度，并判断该平均值是否与待考察点属于同一互动度类别。如果是，则该邻域被视为"纯"邻域，之后落入其中的新帖子很可能也会产生这一特定类别的互动度。为寻找最近邻，我们首先使用主成分分析将嵌入的维度降至 100 维，然后采用 Scikit-Learn 实现的最近邻算法 [38]，并使用欧几里得距离作为距离指标。图 5 展示了中型层级下不同 N 值对应的纯邻域百分比 = [1, 3, 5, 10, 20, 30, 50] 。平均而言，约有 20%（N=50）的点处于纯邻域中，这表明某些话题比其他话题更能（或不那么能）引发互动。宠物类别的平均值最高，这可能是因为相关话题可以是动物的物种和品种（视觉上相似），而人们对其中某些话题的点赞可能多于（或少于）其他话题。

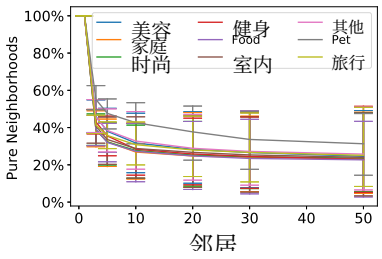
5.2 一般热门话题与用户特定热门话题

§中纯邻域所占的比例5.1 涵盖了仅由单个网红构成的邻域，即并非通用热门话题的邻域。在这种情况下，我们确定了一类可以称为用户特定热门话题的话题，这对于了解哪些话题会（或不会）吸引特定网红非常有用。因此，一般热门话题与用户特定热门话题的区别在于：有多少网红参与了一个纯邻域，以及他们发布了多少帖子。我们将这一参数称为用户多样性。为了计算该参数，我们参考了辛普森多样性指数 [47], , 该指数用于生态学中量化栖息地的生物多样性。它会同时考虑存在的物种数量以及每个物种的丰富度。多样性指数 D 表示为：

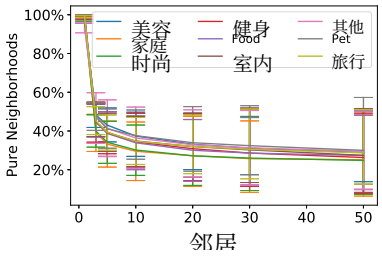
$$D = 1 - \frac{\sum_{i=1}^{K} n_i(n_i - 1)}{N(N-1)}, \tag{2}$$

其中 N 是样本总量， K 是物种数量，n_i 是第 i^{th} 个物种的生物体数量。 D 的取值范围为 0 到 1，其中 0 表示多样性最低，1 表示多样性最高。在我们的场景中，物种指网红，生物体指帖子。同样，我们可以定义互动多样性，

[12]所有结果均可在我们的代码仓库中查看。



（a）点赞



（b）评论

图 5：中型层级中互动指标的纯邻域占比。

该指标用于衡量帖子在互动方面的多样性。我们需要这一指标，因为我们仅根据帖子互动的平均值创建了纯邻域，因此一个纯邻域中可能包含属于不同互动度类别的帖子。总而言之，通过衡量我们的纯邻域的互动和用户多样性，我们可以像图6所示定义主题。显然，我们更关注绿色区域，因为互动多样性高的邻域可靠性较低。我们将低与高之间的阈值设为 0.5。

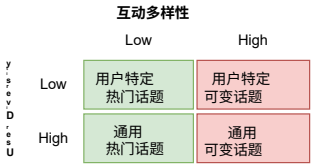


图 6：互动和用户多样性的话题类型。

研究发现。 我们的热门话题研究仅聚焦于互动最高的类别（即按百分位排名位于前 20% 的帖子），并针对不同层级和类别分别进行分析。在自动搜索过程中，我们移除了点位重叠超过 80% 的邻域。我们探索了由此得到的邻域，发现了若干在特征提取阶段未曾想到、且最先进模型无法检测到的热门话题，这证实了这项无监督研究的优势。所有邻域均可在我们的代码仓库中浏览，而图 7展示了每个类别的一个示例。例如，我们在家庭类别中发现了“母亲与孩子”，在时尚类别中发现了“两个（或更多）穿比基尼的女孩”，在宠物类别中发现了“女孩/孩子在马匹附近/骑马”。对于配文方面，我们发现类别特定的热门话题较少，更多的是普遍适用的策略。例如，赠品活动能吸引大量评论，因为参与者通常需要发表评论并标记其他朋友。此外，我们还发现一种有效策略：询问用户对某个话题的看法、提出通用问题，或征求他们对后续内容的建议。更多通用热门话题见第 §6。

> **要点：** *Instagram* 提供了视觉和文本热门话题，这些话题很可能带来高水平的互动。尽管视觉热门话题因类别而异，但图注往往会在不同类别中采用相似的策略。

6 指南与洞见

本节根据决策树互动分类器（§4）和热门话题检测（§5），为各类别提供了一些获取更多点赞和评论的指南。我们还提供了一些撰写吸引人的配文的建议。

6.1 点赞、评论和话题指南

各类别在提升互动方面各有特点。下面介绍一些指南，帮助大家围绕我们发现的热门话题获得更多点赞和评论。

美容。 裸露臀部、脚、较高的图像愉悦度和积极表情符号情感，主要能带来点赞。裸露臀部也会引发大量评论，此外，较低的平均年龄、设置位置以及大量提及也有助于获得评论。波浪发备受青睐，情侣以及搭配完美眉形的眼妆也属于热门话题。用户通常喜欢网红收到新的化妆品、仿妆重现知名妆容（例如电影中的妆容），以及谈论个人问题。

家庭。 点赞和评论受到相似因素的驱动。人们的特征，如年龄和性别，起着主导作用。女性主体的平均年龄应较低，且标准差较高。这表明，正如 §5.2 中所检测到的，带孩子的母亲是热门话题。室内或户外自然环境更受青睐，而位置、色彩和提及次数的影响很大。我们发现，怀孕、分娩和身体变化是热门话题；在这些时期，粉丝会觉得自己与网红更加亲近。

时尚。 就点赞而言，小层级账号适合增加提及次数，而色彩相关特征（例如主导度、唤醒度）则对高层级网红影响很大。裸露臀部起着主导作用，既能带来点赞，也能引发评论。裸露脚部会带来大量评论，户外图片、带有许多话题标签的简短图注以及积极的表情符号情感也同样如此。穿比基尼的女孩和拥有六块腹肌的男性是热门话题，也很受欢迎。讨论特殊活动的穿搭也十分吸引人，例如旅行、参加音乐会、庆祝生日、出席正式晚宴，或仅仅是开启新的一周。

健身。 暖色调以及具有高支配性的女性主体更容易获得点赞，且这些女性主体的年龄标准差较低，最好穿着运动服。简短且带有积极情感的配文有助于获得点赞和评论。根据“身体转变”这一热门话题，低唤醒度会带来更多评论。配文应激励人们在健身时更加努力。

美食。 在这一类别中，男性更常见，也能获得更多点赞。夸张的汉堡和烈酒也备受青睐，



图7：我们在各类别中发现的热门话题示例。

精美且色彩丰富的食物构图也很受欢迎。在厨房或餐厅外拍摄的照片有助于获得点赞。位置同样有助于获得评论，此外，高唤醒度以及配有许多表情符号的正面配文也很重要。配文应简要描述餐盘中的食物，并询问大家最喜欢的美食。披萨日、巧克力和纯素食品常常引发激烈讨论。

室内。 要获得点赞，室内环境（如客厅）和冷色调更受青睐，画面中有儿童和女性也有帮助。位置对点赞和评论都有影响，但应避免商业建筑和美食图片。明亮悦目的图片能带来更多评论，画面中有动物也有帮助。好的配文应结合积极表情符号情感、少量话题标签和一般性问题，例如周末做什么。

宠物。 要获得大量点赞，图片应拍摄于室内或自然户外环境，采用暖色调，并传达积极情感。添加位置和提及有助于获得评论，动物的可爱程度非常高或非常低也有助于引发评论。在最受喜爱的动物中，我们发现了马、珍奇动物、暹罗猫，以及穿着衣服、系着蝴蝶结的狗狗。家里添了新成员，自然会收获许多评论！

旅行。 点赞主要来自女性主体，男性出现较少，且年龄标准差通常较低，最低年龄也较低。这表明，年轻朋友或同龄群体的旅行照片极具吸引力。若想获得大量评论，除了户外相关特征很重要之外，文本传达的情感也应倾向于正面。此外，话题标签和提及至关重要；图片应令人愉悦、富有感染力，且支配性较低。海边的照片广受欢迎，用户也更愿意与夏季和假期相关的帖子互动。

其他。 在室内场所拍摄、色彩偏冷且支配性较低的女性照片能获得更多点赞。使用许多表情符号和较短的图注也有帮助。单张照片中出现年龄相仿的男性，配文采用中性情感，能获得许多评-论。热门话题很多，例如足球、网络迷因或超级英雄。配文通常风趣且相当简短。

6.2 提升图注互动效果的指南

即使每个类别和层级都需要特定的配文来提升互动，我们还是发现了一些能有效提升互动的通用策略。通过探索，我们发现大多数热门配文都有一种典型模式，即向受众提问。这些问题可以非常宽泛（例如“你今天感觉如何？”），也可以与主题相关（例如“你更喜欢哪套穿搭？”），有助于提升用户互动。此外，创作者经常会要求人们采取特定行动，例如标记朋友或分享内容给朋友。这种做法称为行动号召，通常能带来大量互动。最后，话题标签通常放在配文末尾，常常与其他文本分开，每行用一个句点或破折号隔开。这种做法会促使用户点击“查看更多”按钮来查看完整配文，从而提升互动。

6.3 局限性

我们的指南源自对迄今发布的最大 Instagram 数据集进行的分析。该数据集包含约 1,000 万条由 3.4 万名网红创作的帖子。不过，其中并未涵盖许多值得关注的类别（例如体育、电影、音乐），而且人们的喜好以及热门话题可能会随时间发生变化。然而，我们介绍了两种方法（有监督方法 §4 和无监督方法 §5），这些方法可以应用于任何类别（必要时还可扩展特征集），并可在任何时间，通过获取目标类别和层级的网红所创作的 Instagram 内容“快照”来实施。此外，读者可能会担心，Instagram算法已于2021年开始将评论纳入推荐帖子时所依据的因素，而我们的数据集来自 2020年。我们需要说明的是，本文旨在解释哪些帖子的特征会促使用户发表更多评论（评论如今已成为一个关键因素），而不是解释 Instagram 目前如何向用户推荐帖子。事实上，与点赞一样，用户发表评论的主要原因取决于帖子本身 [1]，，而不是用户是否看到了这些帖子。

7 结论

在本研究中，我们旨在弥补以往研究的不足，阐释 Instagram 互动背后的机制，并聚焦于可解释模型。如此一来，便可以依据预先制定的指南，有计划地创作更具吸引力的 Instagram 内容，从而节省时间和资金。通过审慎而全面的特征提取流程，我们训练出的预测器最高达到了 94% 的 F1分数。具体而言，我们的结果表明，点赞主要受图像驱动，而评论主要由说明文字引发。此外，我们还证明，随着层级提升，网红的行为会变得更具类别特定性。最后，我们提出了一种新颖的无监督方法，用于检测和分析热门话题，从而更好地理解各类别的内在动态。

未来，我们计划通过整合热门话题提取的模型来改进预测。此外，还应研究更多类别，并引入结合点赞和评论的指标，以便更好地了解它们之间的关系。关于这些指标，可以通过剔除虚假互动等方式加以完善。不过，目前我们没有证据表明 Instagram算法会考虑这些差异。最后，应将地理因素纳入考量，以了解其是否会影响互动机制。

参考文献

[1] Kholoud Khalil Aldous、Jisun An 和 Bernard J Jansen。2019年。《浏览、点赞、评论、发帖：分析 5 个社交媒体平台上 53 家新闻机构在 4 个层级按主题划分的用户参与度》。载于《国际 AAAI 网络与社交媒体会议论文集》第 13 卷。47–57。[2] Victor Campos、Brendan Jou 和 Xavier Giro-i Nieto。2017年。《从像素到情感：微调卷积神经网络以进行视觉情感预测》。《图像与视觉计算》65（2017），15–22。[3] Salvatore Carta、Alessandro Sebastian Podda、Diego Reforgiato Recupero、Roberto Saia 和 Giovanni Usai。2020年。《Instagram 帖子的流行度预测》。《信息》11卷9期（2020），453。[4] Sharath Chandra Guntuku、Daniel Preotiuc-Pietro、Johannes C. Eichstaedt 和 Lyle H. Ungar。2019年。《Twitter 个人资料和发布的图像揭示了哪些有关抑郁症和焦虑症的信息》。《国际 AAAI 网络与社交媒体会议论文集》13卷01期（2019年7月），236–246。[5] Mauro Conti、Jenil Gathani 和 Pier Paolo Tricomi。2022年。《在线社交媒体中的虚拟网红》。《IEEE 通信杂志》（2022）。[6] Shaunak De、Abhishek Maity、Vritti Goel、Sanjay Shitole 和 Avik Bhattacharya。2017年。《利用深度学习预测生活方式杂志 Instagram 帖子的流行度》。载于《2017 年第二届国际通信系统、计算与 IT 应用会议（CSCITA）》。IEEE，174–177。[7] Jiankang Deng、Jia Guo、Evangelos Ververas、Irene Kotsia 和 Stefanos Zafeiriou。2020年。《Retinaface：野外场景中的单次拍摄多层级人脸定位》。载于《IEEE/CVF 计算机视觉与模式识别会议论文集》。5203–5212。[8] Keyan Ding、Kede Ma 和 Shiqi Wang。2019年。《图像固有受欢迎度评估》。载于《第 27 届 ACM 国际多媒体会议论文集》。1979–1987。[9] Hugging Face。2022年。https://huggingface.co/sentence-transformers/all-mpnet-base-v2 访问日期：2022年8月。[10] Mehmetcan Gayberi 和 Sule Gunduz Oguducu。2019年。《基于用户、帖子和图像特征的社交网络帖子流行度预测》。载于《第 11 届国际数字生态系统管理会议论文集》。9–15。[11] Werner Geyser。2021年。《2021 年网红营销现状：基准报告》。https://influencermarketinghub.com/influencer-marketing-benchmark-report-2021/ 访问日期：2022年8月。[12] 谷歌。2022年。Google Cloud Platform。https://cloud.google.com/ [13] Kaiming He、Xiangyu Zhang、Shaoqing Ren 和 Jian Sun。2016年。《用于图像识别的深度残差学习》。载于《IEEE 计算机视觉与模式识别会议论文集》。770–778。[14] Ming Shan Hee、Roy Ka-Wei Lee 和 Wen-Haw Chong。2022年。《关于多模态仇恨表情包检测模型的解释》。载于《ACM Web 会议 2022 论文集》。3651–3655。[15] HypeAuditor。2022年。《2022 年网红营销现状》。
h
t
t
p
s://
hypeauditor.com/blog/wp-content/uploads/2022/01/US-State-of-Influencer-Marketing-2022.pdf 访问日期：2022年8月。[16] Insiderintelligence。2021年。《Instagram 的美国广告收入将超过 Facebook 主平台》。
h
t
t
p
s://
w
w
w.
i
n
s
i
d
e
rintelligence.co
m/content/instagram-will-net-more-us-ad-revenues-than-core-facebook-platform 访问日期：2022年8月。[17] Insiderintelligence。2022年。《2022 年 Instagram：值得关注的全球用户统计、人口统计和营销趋势》。https://www.insiderintelligence.com/insights/instagram-user-statistics-trends/ 访问日期：2022年8月。[18] Jaakko Kangasharju。2022年。《可爱度检测器》。https://github.com/asharov/cute-animal-detector。[19] Seungbae Kim、Jyun-Yu Jiang、Masaki Nakada、Jinyoung Han 和 Wei Wang。2020年。《面向网红营销的多模态帖子注意力画像》。载于《Web 会议 2020 论文集》。2878–2884。[20] Shu Kong、Xiaohui Shen、Zhe Lin、Radomir Mech 和 Charless Fowlkes。2016年。《结合属性和内容适配的照片美学评分网络》。载于《欧洲计算机视觉会议》。施普林格，662–679。[21] Petra Kralj Novak、Jasmina Smailović、Borut Sluban 和 Igor Mozetič。2015年。《表情符号情感》。《PLOS ONE》10卷12期（2015年12月），1–22。[22] Quoc Le 和 Tomas Mikolov。2014年。《句子和文档的分布式表示》。载于《国际机器学习会议》。PMLR，1188–1196。[23] Guillaume Lemaître、Fernando Nogueira 和 Christos K. Aridas。2017年。《Imbalanced-learn：用于应对机器学习中不平衡数据集难题的 Python 工具箱》。《机器学习研究杂志》18卷17期（2017），1–5。http://jmlr.org/papers/v18/16-365.html [24] Tsung-Yi Lin、Michael Maire、Serge J. Belongie 等。2014年。《Microsoft COCO：上下文中的常见物体》。CoRR abs/1405 0312 2014 1405 0312。（）。arXiv：。[25] Daniel Maposa、Edinah Mudimu 和 Olina Ngwenya。。《津巴布韦不同农业生态区高粱品系表现的多变量方差分析（多元方差分析）》。《非洲农业研究杂志》5期（2010年2月），196–203。[26] Roberto Martinez-Pecino 和 Marta Garcia-Gavilán。2019年。《点赞与问题性 Instagram 使用：自尊的调节作用》。《网络心理学、行为，

4 Tricomi 和 Chilese 等人 2010 社交网络 12,8(2019),411–412。[27] Yelena Mejova、Su Jung Kim、Edward C Malthouse 和 Vijay Viswanathan。2019。《在线评论：客户与品牌之间及围绕品牌展开的对话》。载于《客户参与研究手册》。Edward Elgar 出版社。[28] Masoud Mazloom、Iliana Pappi 和 Marcel Worring。2018。特定类别帖子受欢迎度预测。载于国际多媒体建模会议。施普林格，594–607。[29] Masoud Mazloom、Robert Rietveld、Stevan Rudinac、Marcel Worring 和 Willemijn Van Dolen。2016。品牌相关社交媒体帖子的多模态受欢迎度预测。载于第 24 届 ACM 国际多媒体会议论文集。197–201。[30] Albert Mehrabian 和 James A Russell。1974。《环境心理学方法》。麻省理工学院出版社。[31] Christoph Molnar。2020。《可解释机器学习》。Lulu.com。[32] mowshon。2022。年龄和性别。https://github.com/mowshon/age-and-gender [33] J.L. Myers、A. Well 和 R.F. Lorch。2010。《研究设计与统计分析》。劳特里奇。[34] Christina Newberry。2022。2022 年 Instagram 增长策略：12 个万无一失的方法。https://blog.hootsuite.com/instagram-growth/ 访问日期：2022年8月。[35] Chuang Niu、Hongming Shan 和 Ge Wang。2021。Spice：用于图像聚类的语义伪标签。arXiv 预印本 arXiv:2103.09382（2021）。[36] notAI.tech。2022。NudeNet：用于裸露内容分类、检测和选择性审查的神经网络。https://github.com/notAI-tech/NudeNet 访问日期：2022年8月。[37] Adam Paszke、Sam Gross、Massa 等。2019。PyTorch：一种命令式风格、高性能的深度学习库。载于《神经信息处理系统进展》第 32 卷，H. Wallach、H. Larochelle、A. Beygelzimer、F. d'Alché-Buc、E. Fox 和 R. Garnett（编）。Curran Associates, Inc.，8024–8035。[38] F. Pedregosa、G. Varoquaux 等。2011。Scikit-learn：Python 机器学习。《机器学习研究杂志》12（2011），2825–2830。[39] Kristo Radion Purba、David Asirvatham 和 Raja Kumar Murugesan。2021。使用话题标签、图像评估和用户历史特征分析并预测 Instagram 帖子热度趋势。《阿拉伯信息技术国际期刊》18，1（2021），85–94。[40] J. Ross Quinlan。1986。决策树的归纳。《机器学习》1，1（1986），81–106。[41] Rosalyn Ransaw。2021。我们分享社交媒体内容背后的心理学。
h
t
t
p
s://
w
w
w.shutterstock.com/blog/the-psychology-behind-why-we-share-on-social-media 访问日期：2022年8月。[42] Nils Reimers 和 Iryna Gurevych。2019。Sentence-BERT：使用孪生 BERT 网络生成句子嵌入。（2019），3982–3992。[43] Christoffer Riis、Damian Konrad Kowalczyk 和 Lars Kai Hansen。2020。Instagram 多模态受欢迎度预测的局限：一种新的稳健、高效且可解释的基线。arXiv 预印本 arXiv:2004.12482（2020）。[44] Cynthia Rudin。2019。在高风险决策中，停止解释黑箱机器学习模型，改用可解释模型。《自然·机器智能》1，5（2019），206–215。[45] Shivam Sharma、Firoj Alam、Md Akhtar、Dimitar Dimitrov、Giovanni Da San Martino、Hamed Firooz、Alon Halevy、Fabrizio Silvestri、Preslav Nakov、Tanmoy Chakraborty 等。2022。检测和理解有害模因：综述。arXiv 预印本 arXiv:2205.04274（2022）。[46] Claire Sherman 和 Pascale Quester。2006。产品与裸露内容的一致性对广告效果的影响。《促销管理杂志》11，2 3 2006 61 89 -（），–。[47] Edward H Simpson。1949。多样性的测量。《自然》163，4148（1949），688–688。[48] Statusbrew。2021。Instagram 算法 2022：如何掌握它。https://statusbrew.com/insights/instagram-algorithm/ 访问日期：2022年8月。[49] Barbara G Tabachnick、Linda S Fidell 和 Jodie B Ullman。2007。《多变量统计》。第 5 卷。培生，马萨诸塞州波士顿。[50] Pier Paolo Tricomi、Sousan Tarahomi、Christian Cattai、Francesco Martini 和 Mauro Conti。2022。我们是否都身处《楚门的世界》？ 通过自训练识别 Instagram 刷量行为。arXiv 预印本 arXiv:2206.12904（2022）。[51] Patricia Valdez 和 Albert Mehrabian。1994。颜色对情绪的影响。《实验心理学杂志：综合》123，4（1994），394。[52] Koosha Zarei、Reza Farahbakhsh 和 Noël Crespi。2020。冒充者如何利用 Instagram 制造虚假互动？ 载于 ICC 2020-2020 IEEE 国际通信会议（ICC）。IEEE，1–6。[53] Fattane Zarrinkalam、Stefano Faralli、Guangyuan Piao、Ebrahim Bagheri 等。2020。从社交媒体中提取、挖掘和预测用户兴趣。《信息检索基础与趋势》14，5（2020），445–617。[54] Min Zeng、Beiji Zou、Faran Wei、Xiyao Liu 和 Lei Wang。2016。结合 SMOTE 与 Tomek 链接技术，有效预测三种常见疾病，以处理不平衡医疗数据。载于 2016 年 IEEE 在线分析与计算科学国际会议（ICOACS）。IEEE，225–228。[55] Lei Zhang、Paul Jones、Kent Aaron Otis、Jonathan Gale 和 Evelyn Chan。2018。社交媒体热门话题提取。美国专利 10,095,686。[56] Zhongping Zhang、Tianlang Chen、Zheng Zhou、Jiaxin Li 和 Jiebo Luo。2018。如何成为 Instagram 知名用户：采用双重注意力的帖子受欢迎度预测。

载于*2018* 年 *IEEE* 国际大数据会议（大数据）。IEEE，2383–2392。[57] Bolei Zhou、Agata Lapedriza、Aditya Khosla、Aude Oliva 和 Antonio Torralba。2017。Places：用于场景识别的千万级图像数据库。*IEEE* 模式分析与机器智能汇刊（2017）。

[58] Alireza Zohourian、Hedieh Sajedi 和 Arefeh Yavary。2018 年。Instagram 上图像和视频的流行度预测。载于*2018* 年第四届国际网络研究会议（*ICWR*）。IEEE，111–117。[59] Daniel Zwillinger 和 Stephen Kokoska。1999 年。*CRC*标准概率与统计表及公式。CRC出版社。