

華南師範大學

《人工智能导论》课程项目  
开 题 报 告

项 目 题 目：基于概率扩散模型的图像生成算法

所 在 学 院：计算机学院

项 目 组 长：温志森

小 组 成 员：邓实诚、张亮、施楷鑫、许培炫

开 题 时 间：2023 年 4 月 2 日

## 一、选题背景

Diffusion model 是一种广泛应用于市场研究和预测的数学模型。它可以帮助企业预测和掌握新产品或服务在市场中的表现和潜力，并制定最佳的营销策略和竞争策略。这种模型基于社会学和心理学的理论，探究了消费者在采纳新产品或服务时的决策过程。它将市场分成几个阶段，例如市场启蒙期，快速增长期，成熟期和衰退期，以描述产品或服务的采纳速度和总体采纳率。

Diffusion model 通常使用 S 型曲线（S 曲线）来描述产品或服务在市场中的采纳速度和总体采纳率。S 曲线的初始阶段呈平坦状态，说明产品或服务在市场中受众数量很少。然后随着时间的推移，采纳率开始加速增长，进入市场的采纳速度呈现指数级增长。最后，采纳率达到饱和状态，产品或服务的市场份额趋于稳定。这种模型可以帮助企业预测市场需求和趋势，确定产品定价和市场占有率等策略，进而提高企业的市场竞争力和盈利能力。

Diffusion model 将市场分成五个不同的采纳阶段：创新者、早期采纳者、早期多数采纳者、晚期多数采纳者和滞后者，每个阶段的消费者都有不同的特征和行为。创新者是最早采纳新产品或服务的人，他们通常对创新性和新颖性更感兴趣，而滞后者则是最后采纳新产品或服务的人，他们往往更为保守和谨慎。

通过使用 Diffusion model，企业可以了解产品在市场中的表现和潜力，并预测市场趋势和规模，以及确定最佳的营销策略。例如，企业可以通过了解不同采纳阶段的消费者特征和行为，制定不同的营销计划，以更好地吸引和满足不同阶段的消费者需求。此外，Diffusion model 还可以帮助企业优化产品设计和定价策略，以确保产品在市场中的成功。

## 二、相关研究综述

Diffusion Model 是一种生成模型，它可以用来生成高质量的图像、视频、音频等多媒体数据。其主要思想是将噪声数据逐渐“扩散”或“漂移”到真实数据，从而生成高质量的样本。Diffusion Model 的核心理论是基于 Langevin Dynamics，这是一种模拟随机过程的方法，可以用于模拟自然界中的扩散和漂移过程。在 Diffusion Model 中，Langevin Dynamics 被用来模拟数据的“扩散”，并生成

高质量的数据样本。

随着深度学习的发展，生成模型的研究也日益受到关注。传统的生成模型如 VAE 和 GAN 已经在许多领域得到了广泛应用。然而，这些模型在某些任务上的表现仍然有限，例如生成高质量的图像和视频。Diffusion Model 则提供了一种全新的生成模型，具有生成高质量数据的能力。自从 2020 年开始，Diffusion Model 在机器学习领域受到越来越多的关注，并被证明在图像生成和图像修复任务中具有出色的性能。

不同于传统的生成模型，Diffusion Model 将生成过程分解成了多个步骤，每个步骤都是一个微分方程。该模型通过从低噪声水平开始，逐渐将数据漂移到高噪声水平，以产生高质量的数据。这种逐步扩散的方式，使得 Diffusion Model 在生成高质量数据方面表现出色。此外，Diffusion Model 还采用了一种新的训练方法，称为 score matching，这种方法可以更好地处理概率密度估计和梯度计算问题。

### 三、拟解决的问题和研究内容

#### 1. 在 PC 上配置深度学习模型运行环境

配置深度学习模型运行环境是使用 diffusion 模型进行研究所必须要做的第一步。需要安装适合的操作系统、GPU 驱动以及对应版本的 CUDA 和 cuDNN 等依赖库。同时，还需要安装 Python 及相关深度学习框架，如 PyTorch、TensorFlow 等。此外，为了保证训练过程的稳定性，还需要确保系统硬件设备（如 CPU、GPU、内存等）的兼容性和稳定性。在配置环境的过程中，需要考虑不同软件之间的版本兼容性，以及不同硬件之间的性能差异等问题。如果出现环境配置问题，需要及时查找资料解决。

#### 2. 理解 diffusion 模型的原理

理解 diffusion 模型的原理是掌握其在图像生成中的应用的关键。diffusion 模型是通过在噪声图像上迭代执行随机漫步扩散来生成新样本的模型。在迭代过程中，噪声逐渐被扩散并变得更加平滑，最终生成高质量的图像。在理解模型原理的过程中，需要掌握扩散步数、噪声水平、时间步长等重要概念，以及模型训练和生成的整个流程。

### 3. 编写 diffusion 模型的代码

在掌握 diffusion 模型的原理之后，编写代码是将其应用于实际图像生成任务的关键。需要编写包括模型定义、数据预处理、训练循环等在内的完整代码，并进行调试和参数优化。此外，还需要实现评估和测试代码，以评估模型的性能和效果。编写代码时需要注意代码的规范性和可读性，避免出现代码冗余、错误或低效问题。

### 4. 寻找合适的数据集

数据集是训练深度学习模型的重要资源，也是 diffusion 模型研究不可或缺的一部分。寻找合适的数据集需要考虑多方面因素，如数据集规模、样本多样性、标注准确性等。可以使用公开数据集，如 ImageNet、CIFAR 等，也可以采用自己收集的数据集。数据集的质量对模型的性能影响较大，需要注意数据集的质量控制，例如去除异常数据和噪声数据，保证数据集的多样性和代表性。

### 5. 评价 diffusion 模型生成的照片质量以及不同参数对照片生成效果的影响

评价 diffusion 模型生成的照片质量是研究的重要内容之一。除了使用 FID 以外，还可以使用其他指标如 Inception Score、SSIM 等。此外，还需要探究不同超参数对模型表现的影响，如扩散步数、噪声水平、时间步长等。通过对模型进行参数优化，可以提高生成图像的质量和生成速度。

### 6. 探究如何使用 diffusion 模型在数据增强中的应用

在深度学习中，数据增强可以通过改变样本的旋转角度、剪裁比例等方式增加数据集的多样性，提高模型泛化能力。使用 diffusion 模型进行数据增强可以进一步丰富数据集，提高模型性能。具体而言，可以通过生成更多的样本数据，或者对训练数据进行扰动来实现数据增强。

### 7. 研究如何将 diffusion 模型与其他图像处理算法结合使用

将 diffusion 模型与其他图像处理算法，如 GAN 或者 VQ-VAE 等模型结合使用，可以进一步优化和修复生成的图像。通过结合不同模型的优势，可以实现更好的效果和更高的效率。例如，可以使用 GAN 来优化 diffusion 模型生成的图像，并减少噪声和伪影。同时，还可以使用 VQ-VAE 对生成结果进行修复，以提高图像质量和可视性。在研究过程中需要探究不同模型的结合方式，并分析不同方法之间的优劣。

## 四、可行性分析

## 01 小组成员具备的能力

在专业知识方面，大部分小组成员已开始涉足人工智能和深度学习领域，熟悉常见的 Python 语言和 PyTorch 框架，并接触过常见的网络模型和知识蒸馏项目。他们对人工智能项目的完成过程有一定的了解。然而，还需要进一步加强对相关领域的学习和深入研究，以提高项目完成的质量和效率。

在团队协作方面，小组成员具备良好的团队协作能力，分工明确。小组组长认真负责，团队成员的执行能力也很出色。这对于项目的顺利开展非常重要。然而，还需要注重团队成员之间的沟通和协调，以确保项目进度和质量的控制。

## 02 初步分工

选题背景：施楷鑫

相关研究综述：许培炫

拟解决的问题和研究内容：邓实诚

可行性分析：张亮

计划进度安排、参考文献：温志森

在项目初期，我们的目标是确保开源代码代码的正确运行，并且每个项目成员都清楚地了解每行代码的功能和实现目的。

## 03 已有学习资源

1. 《模式识别与机器学习》(西瓜书)。这是一本关于机器学习的经典教材，其中涵盖了很多基于概率模型的算法，包括 Diffusion Model 在内。
2. 李沐《动手深度学习》及其配套视频。《动手深度学习》是由李沐、Aston Zhang、Zack Lipton 合作编写而成的一本开源的深度学习教材，旨在帮助人们深入了解深度学习的原理和应用，并指导如何通过代码实现深度学习的经典算法。
3. diffusion model 相关文献（见第六部分）。

## 04 项目开源代码

dome272/Diffusion-Models-pytorch:Pytorch implementation of Diffusion Models

(<https://arxiv.org/pdf/2006.11239.pdf>) ([github.com](https://github.com))

## 05 可借助的工具

1. python 语言及其丰富的科学计算库和人工智能工具包。
2. pytorch 开发框架。
3. GitHub 开源代码仓库。

## 五、计划进度安排

在第 7-9 周，我们的目标是确保代码正确运行，并且每个项目成员都清楚地了解每行代码的功能和实现目的。这是项目成功的关键之一。接下来的第 10 周，我们将挖掘项目中的可创新点，并确定实验的创新方向。在接下来的数周中，我们将开始创新并记录实验结果、实验过程和进度。这个过程可能会有挑战，但我们将竭尽全力克服。最后，我们在第 15-16 周总结实验，并制作答辩 PPT。我们将充分准备并参加答辩，以展示我们的成果。

## 六、参考文献

- Centola, D. (2010). The spread of behavior in an online social network experiment. *science*, 329(5996), 1194-1197.
- Aral, S., & Walker, D. (2012). Identifying influential and susceptible members of social networks. *science*, 337(6092), 337-341.
- Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9), 5766-5771.
- Shao, J., Ciampaglia, G. L., Varol, O., Yang, K. C., Flammini, A., & Menczer, F. (2018). Anatomy of an online misinformation network. *PLoS One*, 13(4), e0196087.
- Bruggeman, J., & Dekker, R. (2019). Spread of behavior and opinion dynamics in social networks: a survey of models and new results. *Journal of Artificial Societies and Social Simulation*, 22(3), 1-19.
- Zhang, J., Wu, X., Ma, J., & Zhang, Y. (2019). A survey of epidemic models on complex networks. *Information*, 10(1), 33.
- Granovetter, M. (1978). Threshold models of collective behavior. *American journal of sociology*,

83(6), 1420-1443.

- Centola, D., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American journal of sociology*, 113(3), 702-734.
- Liu, Y. Y., Slotine, J. J., & Barabási, A. L. (2011). Controllability of complex networks. *Nature*, 473(7346), 167-173.
- Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical review letters*, 86(14), 3200-3203.
- Karsai, M., Kaski, K., Barabási, A. L., & Kertész, J. (2012). Universal features of correlated bursty behaviour. *Scientific reports*, 2(1), 1-7.