

## F题要求：

问题要求选择三类职业（STEM、贸易和艺术职业），并设计数据驱动模型，分析 **生成性AI（Gen-AI）** 对这些职业的未来影响。然后，基于这个分析，为不同类型的教育机构（大学、职业学校、艺术学校）提出针对性建议。

## 使用数据集的步骤与方法：

### 1. 职业选择与分类

根据题目要求，需要选择三类职业：

- **STEM职业**：例如，软件工程师、数据科学家等。
- **贸易职业**：例如，电工、厨师等。
- **艺术职业**：例如，舞蹈家、音乐家等。

可以根据数据集中的 **O\*NET任务映射** 和 **职业类别** 来选择这些职业类别，并根据每个职业的任务自动化与增强的可能性来做进一步的分析。

### 2. 数据集介绍与分析目标

在 **Anthropic Economic Index** 数据集中，最相关的数据文件是：

- **automation\_vs\_augmentation.csv**：该文件包含了关于 **AI任务自动化与增强** 的数据，您可以用来分析哪些任务会被AI完全取代，哪些任务会得到AI的增强。
- **bls\_employment\_may\_2023.csv**：该文件提供了美国劳工统计局的就业数据，可以用来分析不同职业的就业趋势。
- **onet\_task\_mappings.csv**：该文件映射了任务与 **O\*NET** 类别，帮助您将职业任务与自动化程度进行对比。

### 3. 数据清洗与预处理

首先，您需要加载并清洗数据，确保数据没有缺失或异常值。

```
import pandas as pd

# 加载数据集
automation_data = pd.read_csv('automation_vs_augmentation.csv')
wage_data = pd.read_csv('wage_data.csv')
task_mapping = pd.read_csv('onet_task_mappings.csv')

# 查看数据结构
print(automation_data.head())
print(wage_data.head())
```

### 4. 数据分析与建模

**目标：** 使用该数据集来预测 **Gen-AI** 对不同职业的影响，特别是如何改变这些职业的就业趋势、技能需求和收入。

### 任务自动化与增强分析：

- 使用 `automation_vs_augmentation.csv` 文件中的数据，查看哪些任务会被AI完全自动化，哪些任务会得到增强。
- 分析 **STEM、贸易和艺术职业** 的任务自动化程度，预测AI对这些职业的影响。

例如，您可以计算每种职业的任务自动化比例，并将其与 **BLS就业数据** 结合，看看在这些任务的影响下，哪些职业的就业需求可能下降。

```
# 计算任务自动化比例
automation_data_grouped = automation_data.groupby('task_name')['pct'].mean()

# 查看某些职业任务的自动化比例
print(automation_data_grouped.head())
```

### 构建预测模型：

- 您可以使用 **回归模型** 或 **机器学习模型**，例如 **逻辑回归** 或 **随机森林** 来预测 **Gen-AI** 对这些职业的影响。例如，预测哪些职业任务会被AI取代，哪些职业的技能要求会发生变化。

```
from sklearn.ensemble import RandomForestRegressor

# 假设我们有一个关于职业自动化的数据
x = automation_data[['interaction_type', 'pct']] # 输入特征
y = wage_data['avg_wage'] # 标签（工资）

# 训练模型
model = RandomForestRegressor()
model.fit(x, y)

# 预测
predictions = model.predict(x)
```

### 分析影响：

- 使用 `wage_data.csv` 数据分析 **Gen-AI** 对不同行业和职业收入的影响。可以将 **职业的自动化程度** 与 **工资水平** 对比，看看高自动化职业的薪资变化趋势。

## 5. 生成教育建议

根据分析结果，您可以为每种职业的教育和培训提供具体建议。例如：

- 对于 **STEM职业**：强化 **AI技术相关课程**，培养学生能够在AI领域进行深入工作的能力。
- 对于 **贸易职业**：提供**AI工具的使用**课程，帮助学生适应AI辅助的工作环境。
- 对于 **艺术职业**：强调 **AI与创意的结合**，让学生了解如何使用AI进行创作，但保持独立的艺术表现力。

## 6. 结果可视化与报告

使用 `plots.ipynb` 文件中的代码，生成与任务自动化、职业影响、工资变化等相关的图表。这样可以帮助您清晰地呈现 **Gen-AI** 对各职业的影响，并使报告更具说服力。

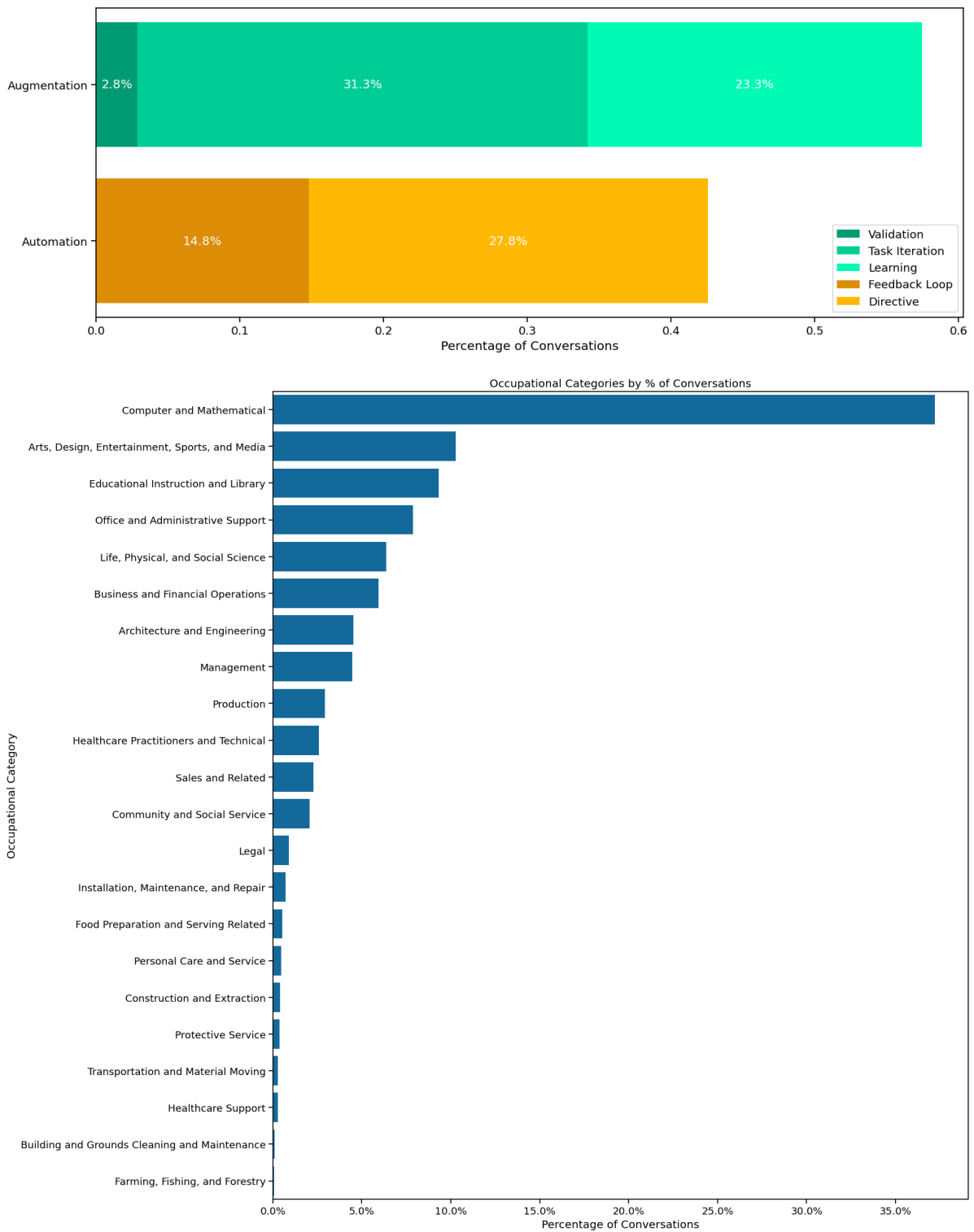
7. 提出改进的投票系统

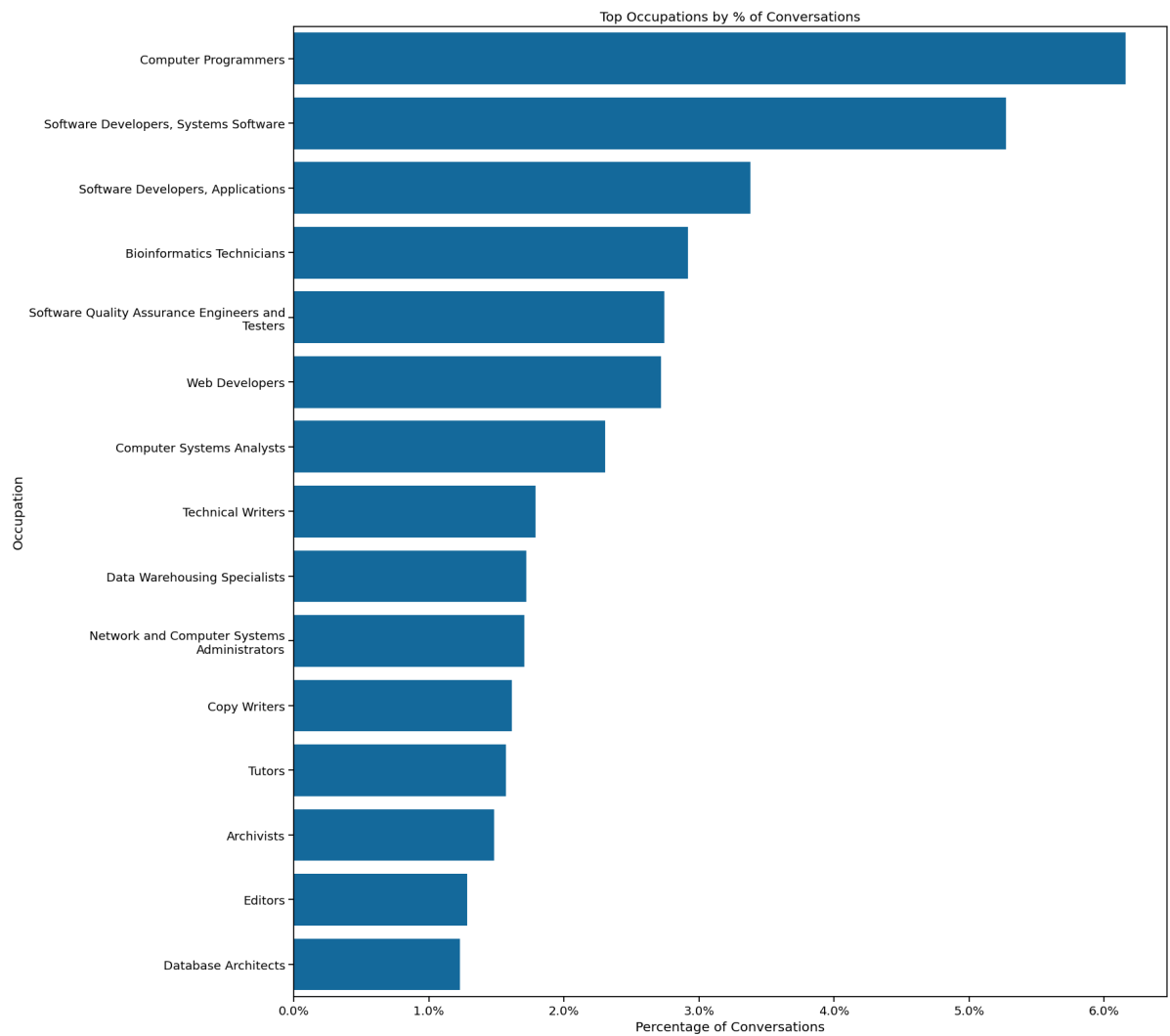
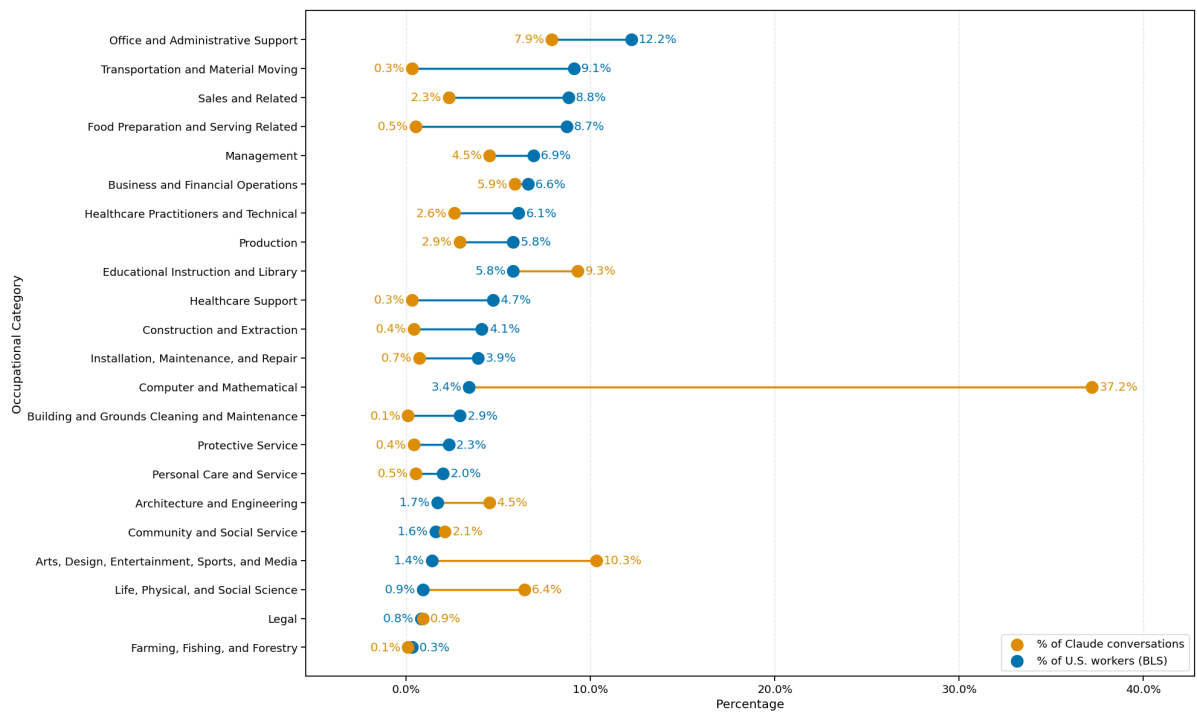
除了分析教育建议，还可以探索其他关于 **职业任务自动化与增强** 的系统，如建议高校如何调整课程结构，提供更加针对 **Gen-AI** 的培训。

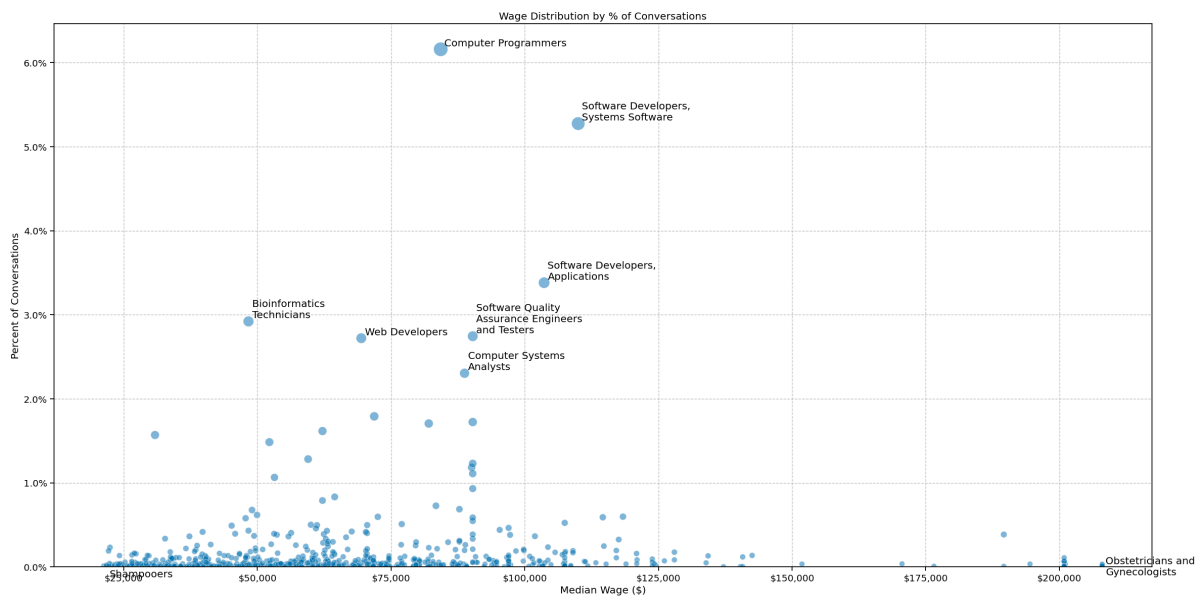
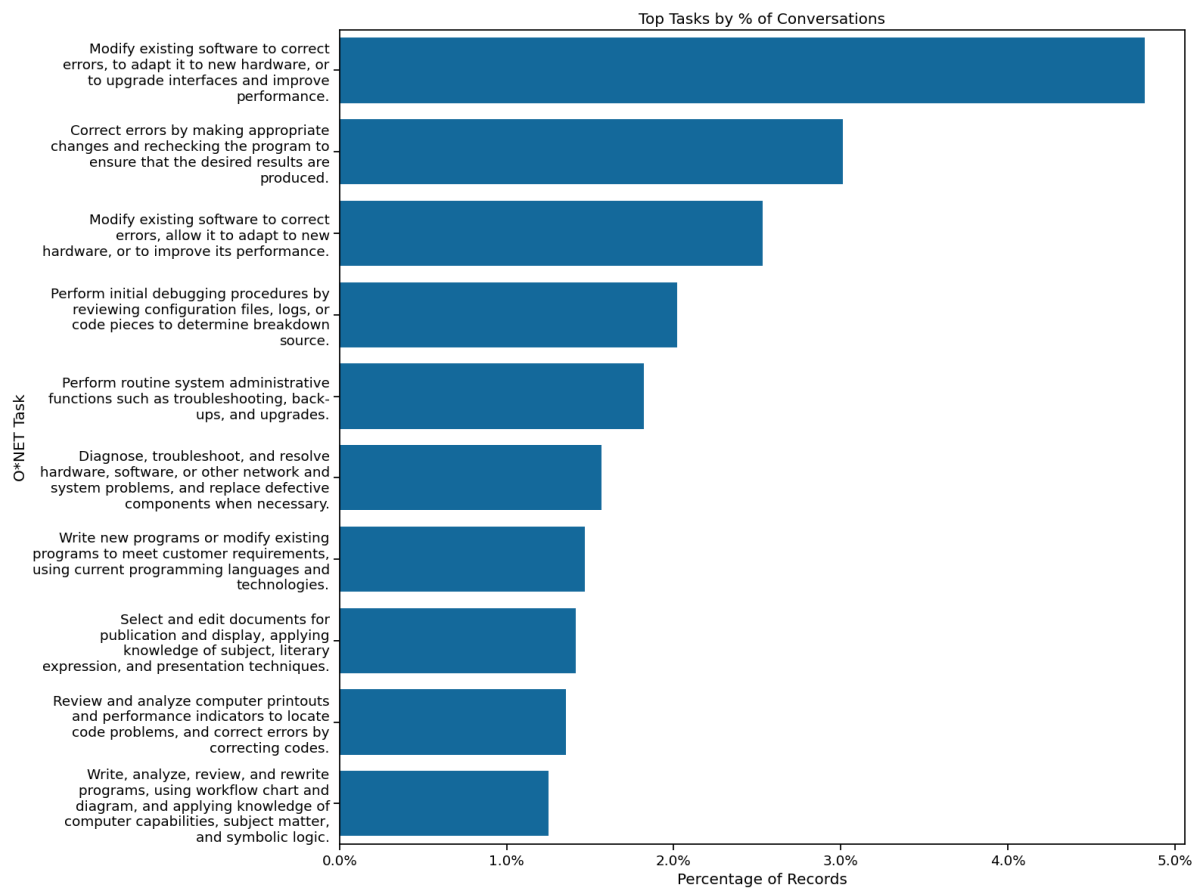
总结：

使用 **Anthropic Economic Index** 数据集，可以：

- 1. **分析AI对职业的影响：** 特别是哪些职业任务会被AI自动化，哪些任务会得到增强。
- 2. **构建数据驱动的预测模型：** 预测哪些职业将受到Gen-AI影响，以及AI对就业趋势、技能要求和收入的影响。
- 3. **为教育机构提供建议：** 根据您的分析结果，向教育机构提供如何调整课程和培训策略的建议，以适应未来AI的变化。







[Economic Tasks AI Paper.pdf](#)

Despite widespread speculation about artificial intelligence's impact on the future of work, we lack systematic empirical evidence about how these systems are actually being used for different tasks. Here, we present a novel framework for measuring AI usage patterns across the economy. We leverage a recent privacy-preserving system [Tamkin et al., 2024] to analyze over four million Claude.ai conversations through the lens of tasks and occupations in the U.S. Department of Labor's O\*NET Database. Our analysis reveals that AI usage primarily concentrates in software development and writing tasks, which together account for nearly half of all total usage. However, usage of AI extends more broadly across the economy, with  $\sim 36\%$  of occupations using AI for at least a quarter of their associated tasks. We also analyze *how* AI is being used for tasks, finding 57% of usage suggests augmentation of human capabilities (e.g., learning or iterating on an output) while 43% suggests automation (e.g., fulfilling a request with minimal human involvement). While our data and methods face important limitations and only paint a picture of AI usage on a single platform, they provide an automated, granular approach for tracking AI's evolving role in the economy and identifying leading indicators of future impact as these technologies continue to advance.

【要点】：论文通过分析数百万Claude.ai对话，提出了一个测量AI在经济各领域中任务使用模式的新框架，发现AI主要集中于软件开发和写作任务，并识别了AI使用的扩展趋势和未来影响的领先指标。

【方法】：研究者利用一个隐私保护的系统，结合美国劳工部O\*NET数据库中的任务和职业分类，分析了Claude.ai平台上的AI使用情况。

【实验】：实验分析了超过四百万次Claude.ai对话，通过任务和职业的视角来研究AI的使用模式，结果显示AI在软件开发和写作任务中使用最为集中，共涉及约36种相关任务，其中有57%是用于学习或迭代输出，而43%则用于最小化人类参与的请求满足。