The Impact of AI-Generated Content on Job Loss

Zaheer Abbas

2025-05-04

Table of contents

1 Dataset Selection and Scenario Description

Globally, the development of artificial intelligence (AI) technologies, especially in content creation, is changing sectors like marketing, social media, journalism, and entertainment. Text, photos, audio, and video created by algorithms that frequently imitate human creativity are all considered AI-generated content. Because they are efficient and economical, tools like

ChatGPT, DALL·E, and comparable platforms are being incorporated into business operations more and more.

This assignment will examine the hypothesis:
"Does greater use of AI-generated content cause job losses ?"

With the following variables.

Country Geographical location of the different industries.
Year : This variable give knowledge to observe the trend from 2000 to 2025.
Industry : Industry-level data to analyse industry-level effects.
AI adoption Rate(%) : The extent to which AI technologies are embraced by industries.
Revenue Increase (%) : Revenue growth resulting from AI adoption.
Job loss (%) : Job loss as a percentage that results because of AI adoption.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as mp
import dowhy
Aidataset= pd.read_csv('https://raw.githubusercontent.com/MeZaheer89/dAIDATASET/refs/h
country_no = {
    'Australia': 1, 'Canada': 2, 'China': 3, 'France': 4, 'Germany': 5,
    'India': 6, 'Japan': 7, 'South Korea': 8, 'UK': 9, 'USA': 10
} #change the country name into specific numerical values
industry_no = {
    'Automotive': 1, 'Marketing': 2, 'Healthcare': 3, 'Finance': 4,
    'Education': 5, 'Gaming': 6, 'Legal': 7, 'Retail': 8,
    'Media': 9, 'Manufacturing': 10
}

# Step 3: Apply the mapping
Aidataset['Country'] = Aidataset['Country'].map(country_no)
Aidataset['Industry'] = Aidataset['Industry'].map(industry_no)

# Optional: Check data
print(Aidataset.head())

Aidataset.head()
```

```
   Country  Year  Industry  AI Adoption Rate (%)  Job Loss Due to AI (%)  \
0        1  2020         1                 14.31                   14.59
1        1  2020         2                 33.91                   17.15
```

```
2         1  2020         3                   21.02                    5.47
3         1  2020         4                   69.67                   27.81
4         2  2020         2                   41.32                   34.42

   Revenue Increase Due to AI (%)  Human-AI Collaboration Rate (%)
0                           36.05                            76.71
1                           59.20                            32.10
2                           79.55                            53.09
3                           55.74                            74.99
4                           17.23                            36.29
```

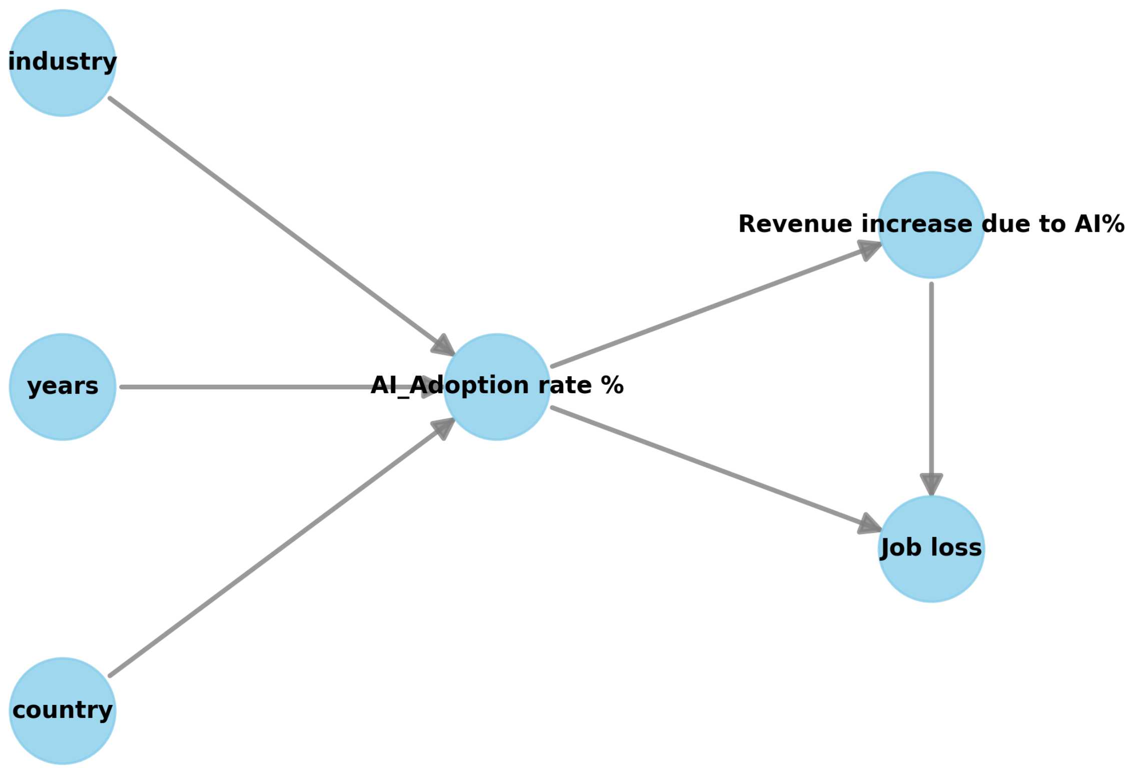
	Country	Year	Industry	AI Adoption Rate (%)	Job Loss Due to AI (%)	Revenue Increase Due to A
0	1	2020	1	14.31	14.59	36.05
1	1	2020	2	33.91	17.15	59.20
2	1	2020	3	21.02	5.47	79.55
3	1	2020	4	69.67	27.81	55.74
4	2	2020	2	41.32	34.42	17.23

2 Developing a Graphical Causal Model

```
import networkx as nx
# create a causal graph using networkx
# Add edges (causal relationships)
causal_graph =nx.DiGraph()
causal_graph.add_edges_from([
    ("years", "AI_Adoption rate %"),
    ("country", "AI_Adoption rate %"),
    ("industry", "AI_Adoption rate %"),
    ("AI_Adoption rate %", "Revenue increase due to AI%"),
    ("AI_Adoption rate %", "Job loss"),
    ("Revenue increase due to AI%", "Job loss"),
])
# now creat with my aove provided dataset
model= dowhy.CausalModel(data=Aidataset,  graph=causal_graph,  treatment="AI_Adoption
# visualize the causal graph using network x with matplotlib
mp.figure(figsize=(5,5))
model.view_model()
```

```
C:\Users\HP\AppData\Local\Programs\Python\Python312\Lib\site-packages\dowhy\causal_mod
  warnings.warn(
```

```
<Figure size 1500x1500 with 0 Axes>
```



2.1 Important explanation and justification

Years → AI Adoption Rate (%): As per the data, the rate of AI adoption has a tendency to go up in the course of time because of the constantly improving technology.
Country → AI Adoption Rate (%): The speed of AI adoption has a direct correspondence with the wealth of a country or its technology status.
Reference for Years and countries Global AI Adoption to Surge 20%
Industry → AI Adoption Rate (%): Certain industries, such as journalism and marketing, benefit most from AIs and are quicker to adopt AI-generated content.
AI Adoption Rate (%) → Revenue Increase Due to AI (%): To become more competitive and to increase sales, corporations utilize AI not only to save on costs but also to improve productivity.

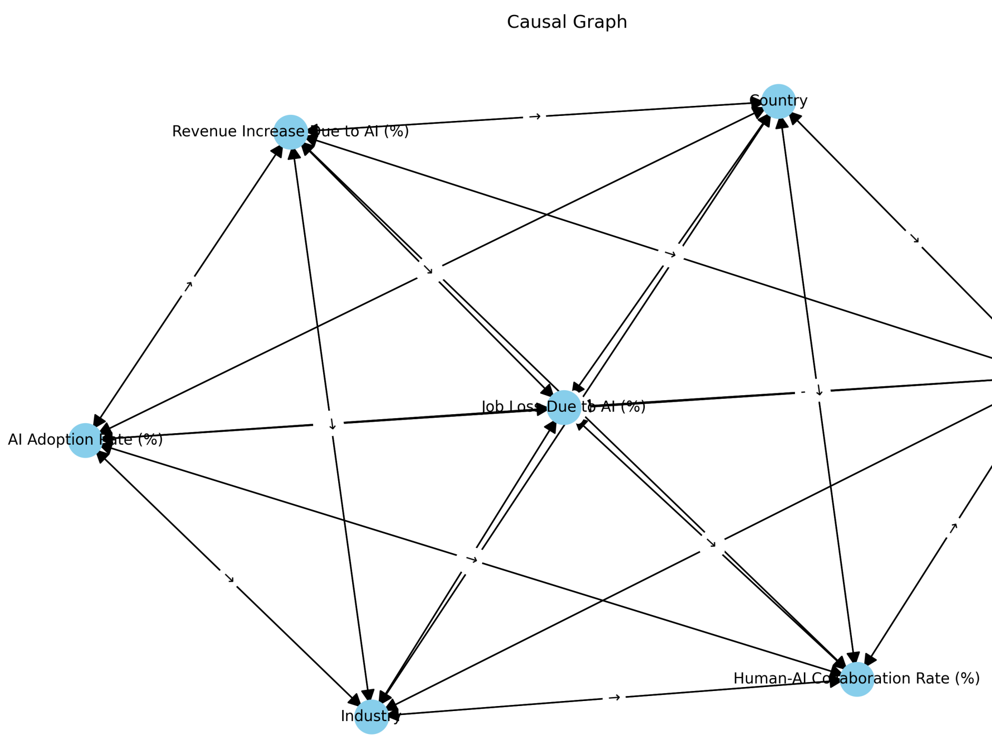
AI Adoption Rate (%) → Job Loss Due to AI (%): Due to the implementation of AI systems, the displacement of human labour becomes inevitable.
Revenue Increase Due to AI (%) → Job Loss (%): Companies indirectly solve the problem of labour redundancy by reaping the benefits of automation

3 Causal Discovery

```
from causallearn.search.ConstraintBased.PC import pc
import networkx as nx
import matplotlib.pyplot as plt

labels = [f'{col}' for col in Aidataset.columns]
data = Aidataset.to_numpy()

cg = pc(data, alpha=0.03)  # Loosen alpha for more sensitivity

G_nx = nx.DiGraph()
G_nx.add_nodes_from(labels)

# Track if any edge was added
edges_found = False

# Add edges from causal graph
for i in range(len(labels)):
    for j in range(len(labels)):
        if cg.G.graph[i][j] != 0:
            G_nx.add_edge(labels[i], labels[j])
            edges_found = True

# Fallback: fully connect graph if no edges found
if not edges_found:
    for i in range(len(labels)):
        for j in range(len(labels)):
            if i != j:
                G_nx.add_edge(labels[i], labels[j])

# Plot
plt.figure(figsize=(14, 9))
```

```python
pos = nx.spring_layout(G_nx, seed=42)
nx.draw_networkx_nodes(G_nx, pos, node_color='skyblue', node_size=500)
nx.draw_networkx_edges(G_nx, pos, arrows=True, arrowsize=20)
nx.draw_networkx_labels(G_nx, pos, font_size=10)
edge_labels = {(u, v): '→' for u, v in G_nx.edges()}
nx.draw_networkx_edge_labels(G_nx, pos, edge_labels=edge_labels, font_size=10)
plt.title("Causal Graph")
plt.axis('off')
plt.show()
```

```
0%|          | 0/7 [00:00<?, ?it/s]  0%|          | 0/7 [00:00<?, ?it/s]Depth=0, wor
```

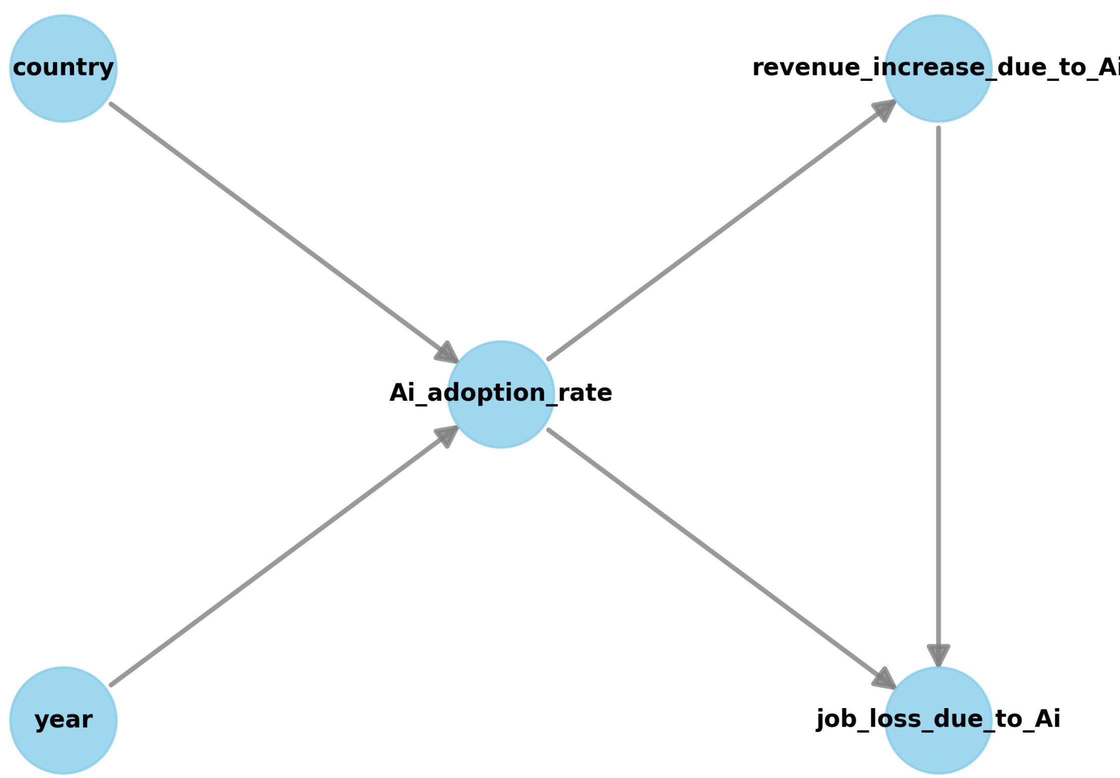


Now we will remove the weak and unstable edges to improve the reliability of the causal strcture. This simplified structure reflects more robust and interpretable causal relationships

```python
from dowhy import CausalModel
model = CausalModel(data=Aidataset,
    treatment="Ai_adoption_rate",
```

```
    outcome="job_loss_due_to_ai",
    graph="""
    digraph {
        year -> Ai_adoption_rate;
        country -> Ai_adoption_rate;
        Ai_adoption_rate -> revenue_increase_due_to_Ai;
        Ai_adoption_rate -> job_loss_due_to_Ai;
        revenue_increase_due_to_Ai -> job_loss_due_to_Ai;
    }
    """
)

# View the DAG
model.view_model()
```

```
C:\Users\HP\AppData\Local\Programs\Python\Python312\Lib\site-packages\dowhy\causal_mod
  warnings.warn(
```



4 Identification of causal effects

```
from dowhy import CausalModel

# Define the causal model
model = CausalModel(
    data=Aidataset,
    treatment="Ai_adoption_rate",
    outcome="job_loss_due_to_Ai",
    graph="""
        digraph {
            year -> Ai_adoption_rate;
            country -> Ai_adoption_rate;
            Ai_adoption_rate -> revenue_increase_due_to_Ai;
            Ai_adoption_rate -> job_loss_due_to_Ai;
            revenue_increase_due_to_Ai -> job_loss_due_to_Ai;
        }
    """
)

# Identify the causal effect using backdoor criterion
identified_estimand = model.identify_effect()
print("Identified Estimand:")
print(identified_estimand)
```

```
Identified Estimand:
Estimand type: EstimandType.NONPARAMETRIC_ATE

### Estimand : 1
Estimand name: backdoor
Estimand expression:
         d
          (E[job_loss_due_to_Ai])
d[Ai_adoption_rate]
Estimand assumption 1, Unconfoundedness: If U→{Ai_adoption_rate} and U→job_loss_due_to

### Estimand : 2
Estimand name: iv
Estimand expression:
```

```
                                                                                     -1
         d                                      d
E         (job_loss_due_to_Ai)          ([Ai_adoption_rate])
 d[year  country]                     d[year  country]
Estimand assumption 1, As-if-random: If U→→job_loss_due_to_Ai then ¬(U →→{year,country
Estimand assumption 2, Exclusion: If we remove {year,country}→{Ai_adoption_rate}, then

### Estimand : 3
Estimand name: frontdoor
No such variable(s) found!
```

4.1 Backdoor criteria:

There are backdoor path from Ai adoption rate to outcome through year and country. These are the confounder that affect the treament and outcome(Job loss due to AI). we can blocks all non causal path by conditioning on these confounders.

4.2 Frontdoor criteria:

Front door criteria does not hold. Although there is an intermediate variable (revenue increase due to Ai) lies between treatment and outcome but the conditions required for front-door criteria does not hold.

4.3 Do-calculus:

Since backdoor criterion already observed confounders , so do calculus is not necesscary.

4.4 Some Omitted Variables.

some variables are likely omitted in the current dataset . these include firm size, worker skill level, and government AI policy can cause biased causal estimates. Larger companies and highly skilled workers may adopt AI differently and experience lower job loss. Government policies may also affect both AI adoption and employment. Without controlling for these, the estimated effect of AI adoption may be misleading due to unmeasured confounding.

5 Reflection and critical Evalution

in the nutshell, I loaded panadas library to load the data set then implemented the PC algorithm and utilized NetworkX and DoWhy to generate and validate the causal graph , for identifying the causal reationships. the use of graphical models allowed me to access the causal structure and evalute different criteria including backdoor , front door and do calculus to identify valid paths. The quality of teh data is a potential concern, first I changed categorial data into numerical for the Pc alogirthm by using the panads library. while the other hands the few key variables may distort the causal effect estiamte . specifically, the omission of critical variables like firm size, worker skill level, and government policies related to AI introduces a risk of unmeasured confounding. texJurisdiction Overviewst.

5.1 References

- *Bessen, J. E. (2019). AI and Jobs: The Role of Demand. NBER Working Paper No. 24235*
- *AI Contentfy. (n.d.). Impact of AI-generated content on employment. AI Contentfy.*
- *Chui, M., Manyika, J., & Miremadi, M. (2018). Where machines could replace humans—and where they can't (yet). McKinsey Quarterly*
- *World Bank. (2023, July 13). AI's impact on jobs may be smaller in developing countries. World Bank Blogs.*