



**DAYANANDA SAGAR
UNIVERSITY**



**SCHOOL OF
ENGINEERING**

**Bachelor of Technology in
COMPUTER SCIENCE AND ENGINEERING
(AI ML)**

22AM2403 – DATABASE MANAGEMENT SYSTEMS

MINI PROJECT REPORT

On

**“Predicting the effect of AI on job trends using
Data manipulation”**

Submitted By

Nagaboina Dharsini (ENG22AM0036)

Sahana Priya G (ENG22AM0050)

Sana Banu (ENG22AM0053)

UNDER THE SUPERVISION

Dr. Savitha Hiremath

Associate Professor, CSE Department, DSU

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING (AI&ML)
SCHOOL OF ENGINEERING
DAYANANDA SAGAR UNIVERSITY**

(2023-2024)



School of Engineering
Department of Computer Science & Engineering
Harohalli, Ramanagara - 562112
Karnataka, India

CERTIFICATE

This is to certify that the DBMS MINI PROJECT titled “**Predicting the effect of AI on job trends using Data manipulation**” carried out by **Nagaboina Dharsini (ENG22AM0036), Sahana Priya G(ENG22AM0050) and Sana Banu (ENG22AM0053)** bonafide students of Bachelor of Technology in Computer Science (AI ML) and Engineering at the School of Engineering, Dayananda Sagar University, Bangalore in partial fulfilment for the award of degree in Bachelor of Technology in Computer Science and Engineering, during the year 2023-2024.

Dr. Savitha Hiremath

Associate professor
Dept. of CSE
School of Engineering
Dayananda Sagar University

Dr.. Jayavrinda Vrindavanam

Chairperson of
Dept. of CSE(AI ML)
School of Engineering
Dayananda Sagar University

DATE:

ACKNOWLEDGEMENT

It is a great pleasure for us to acknowledge the assistance and support of many individuals who have been responsible for the successful completion of this DBMS MINI PROJECT

First, we take this opportunity to express our sincere gratitude to the School of Engineering & Technology, Dayananda Sagar University for providing us with a great opportunity to pursue our bachelor's degree in this institution.

I would like to thank **Dr. Uday Kumar Reddy K R, Dean, School of Engineering & Technology, Dayananda Sagar University** For his constant encouragement and expert advice. It is immense pleasure to express our sincere thanks to **Dr. Jayavrinda Vrindavanam , Chairman, Department of Computer Science and Engineering(AI&MI), Dayananda Sagar University**, for providing the right academic guidance that made our task possible.

We would like to thank our teacher **Dr Savitha Hiremath**, Associate Professor **Department of Computer Science and Engineering, Dayananda Sagar University**, for sparing her valuable time to extend help in every step of our DBMS MINI PROJECT, which paved the way for smooth progress and the fruitful culmination of the project.

We are also grateful to our family and friends who provided us with every requirement throughout the course. We would like to thank one and all who directly or indirectly helped us in the DBMS MINI PROJECT.

ABSTRACT

In this project, we explored how artificial intelligence (AI) affects job trends by analyzing data using Python's Pandas library in Google Colab. We used a dataset from Kaggle containing information on job details, job titles, AI impact, and domains. Our goal was to identify patterns and predict how AI might influence various job sectors in the future.

We started by cleaning and organizing the data to ensure accuracy. Then, we conducted a detailed analysis to understand the current trends and the potential impact of AI on different job roles. Our findings showed which jobs are most likely to be affected by AI, whether positively or negatively. We visualized our results to highlight key insights, making it easier to understand how AI could reshape the job market.

By leveraging data manipulation techniques, we provided a comprehensive overview of the potential changes in job trends due to AI, helping stakeholders prepare for the future workforce landscape.

TABLE OF CONTENTS

PAGE NO

ABSTRACT	
CHAPTER 1 INTRODUCTION	3
CHAPTER 2 PROBLEM STATEMENT	4
CHAPTER 3 PROJECT DESCRIPTION	5
CHAPTER 4 DESIGN	6
CHAPTER 5 METHODOLOGY	7-8
CHAPTER 6 TESTING AND RESULT	9-11
CHAPTER 7 SYSTEM IMPLEMENTATION	12-15
CHAPTER 8 CONCLUSION	16
CHAPTER 9 REFERENCES	17

CHAPTER 1

Introduction

In today's world, artificial intelligence (AI) is becoming more common and is starting to change many aspects of our lives, including the job market. As AI continues to develop, it's important to understand how it will affect different types of jobs. Some jobs may become easier with AI, while others might disappear entirely.

In this project, we wanted to explore the impact of AI on various job trends. To do this, we used a dataset from Kaggle, which contains information about different jobs, their titles, the effect of AI on them, and the industry they belong to. Using Python's Pandas library in Google Colab, we analyzed this data to find patterns and predict how AI might influence the job market in the future.

Our goal was to provide clear insights into which jobs are most likely to be affected by AI, helping people understand and prepare for these changes. By analyzing and visualizing the data, we hope to offer a better understanding of how AI will shape the future of work. This analysis can help policymakers, educators, and workers make informed decisions to adapt to the evolving job landscape.

CHAPTER 2

Problem Statement

To achieve our objectives, we will start by gathering comprehensive datasets from reliable sources, including government employment databases, AI research publications, and industry analyses. The collected data will encompass various features such as job titles, industry sectors, employment rates, AI adoption rates, skill requirements, and salary information. Once the data is collected, we will perform rigorous data cleaning to handle missing values, remove duplicates, and standardize formats, ensuring that our analysis is based on accurate and consistent information.

Following data preparation, we will employ pandas for extensive exploratory data analysis (EDA) to identify key trends and correlations. We will use descriptive statistics to summarize the data, and correlation analysis to explore relationships between AI advancements and job market changes. Advanced predictive modeling techniques, such as linear regression and time series analysis, will be applied to forecast future job trends. The results will be visualized using charts, graphs, and interactive dashboards to clearly communicate our findings. Ultimately, this analysis will provide valuable insights into the impact of AI on employment, guiding stakeholders in making informed decisions to navigate the future job market.

CHAPTER 3

Project Description

In today's world, artificial intelligence (AI) is increasingly prevalent and is starting to reshape various aspects of our lives, including the job market. As AI continues to evolve, understanding its impact on different types of jobs becomes crucial. While AI might simplify some tasks, it could also lead to the disappearance of certain jobs.

This project aims to explore the effects of AI on job trends using a dataset from Kaggle, which includes information on various jobs, their titles, the impact of AI on them, and the industries they belong to. By leveraging Python's Pandas library within Google Colab, we conducted a detailed analysis to identify patterns and predict future job market trends influenced by AI.

The objective of our analysis is to provide clear insights into which jobs are most susceptible to

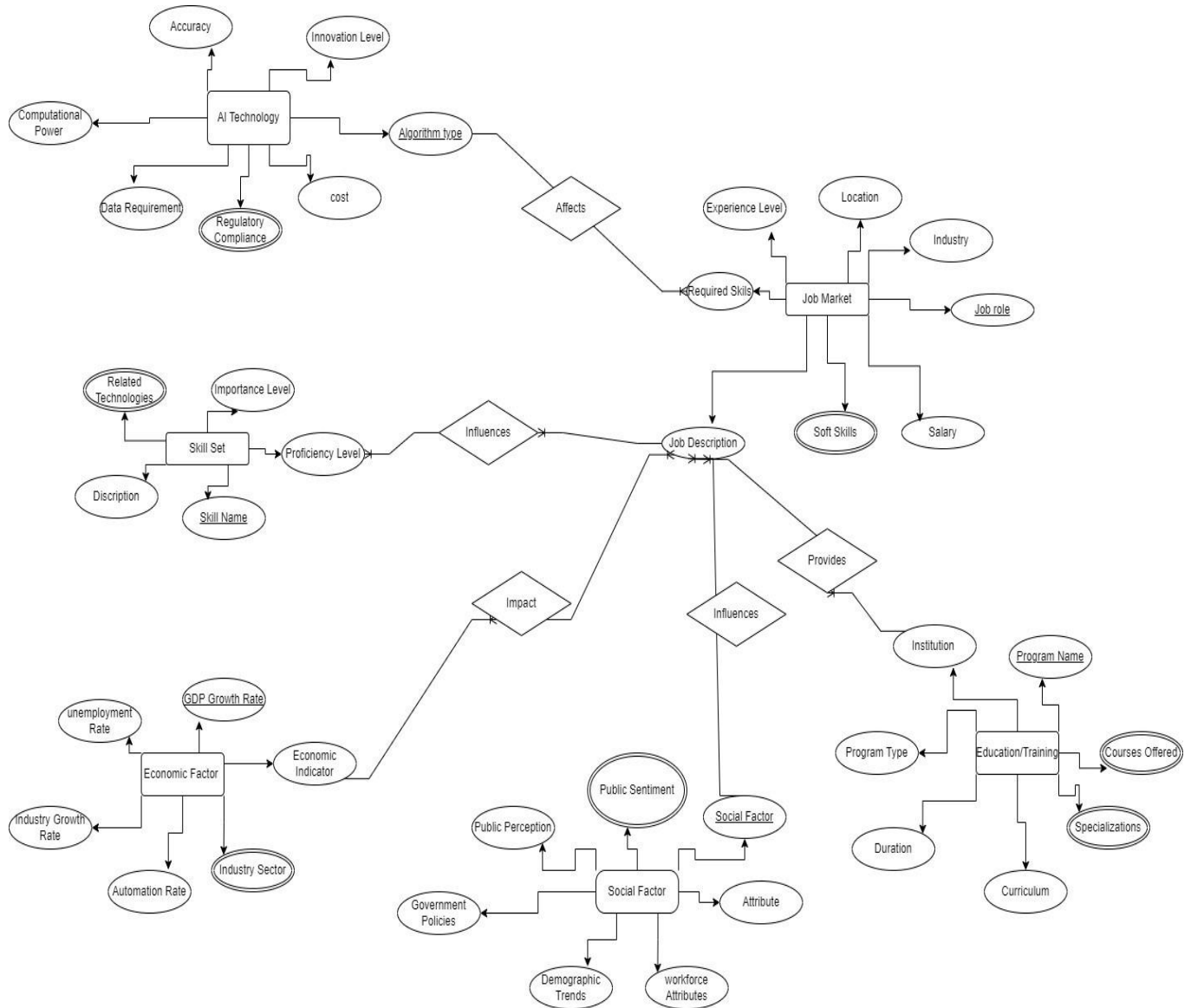
AI's influence, thereby helping individuals and organizations prepare for these changes.

By analyzing and visualizing the data, we intend to offer a comprehensive understanding of how

AI will shape the future of work. This information can aid policymakers in drafting effective policies, assist educators in designing relevant educational programs, and enable workers to make informed career decisions.

CHAPTER 4

Design



CHAPTER 5

Methodology

- **Data Collection:**

- Obtain the Kaggle CSV file containing data about different job trends and their relation to AI.
- The dataset should include information such as job titles, industries, and the impact of AI on those jobs.

2. Data Preprocessing:

- Use Pandas to load the CSV file into a DataFrame, allowing for easy manipulation and analysis.
- Perform data cleaning to handle missing values, inconsistencies, or errors in the dataset.

Explore the dataset to understand its structure, features, and distributions.

3. Data Analysis

- Utilize descriptive statistics to summarize key aspects of the dataset, such as mean, median, and standard deviation.
- Visualize the data using charts, graphs, or plots to identify trends, patterns, and correlations.
- Apply Pandas functionalities for grouping, aggregation, and filtering to gain insights into how AI is affecting different job categories and industries.

4.Feature Engineering:

- Create new features or transform existing ones to enhance the predictive power of the model.
- Extract relevant information from textual data or categorical variables to make them suitable for modeling.

5. Modeling

- Select an appropriate machine learning model, such as regression or classification, based on the nature of the problem and the available data.
- Split the dataset into training and testing sets to evaluate the model's performance.
- Train the model on the training data and tune hyperparameters to optimize its performance.

6. Evaluation:

- Assess the model's performance using evaluation metrics specific to the task, such as accuracy, precision, recall, or F1 score.
- Compare the model's predictions with actual outcomes to determine its effectiveness in predicting the effect of AI on job trends.

7. Interpretation:

Interpret the results of the analysis to understand the implications of AI on different job categories and industries.

- Identify key findings, trends, and insights that can inform decision-making for policymakers, educators, and workforce professionals.

8. Communication

- Communicate the methodology, results, and implications of the analysis effectively through reports, presentations, or visualization.
- Ensure that the findings are accessible and understandable to a wide audience, including stakeholders and domain experts.

By following this methodology, you can systematically analyze the impact of AI on job trends using Pandas and derive valuable insights from the data.

CHAPTER 6

Testing and Result

```
from google.colab import drive
drive.mount('/content/drive')
import os
os.chdir('/content/drive/MyDrive/AI_Project/')

Mounted at /content/drive

+ Code + Text

[1] import pandas as pd

[2] data = pd.read_csv('My_Data.csv')

[ ] print(data.head(10))
```

	Job titles	AI Impact	Tasks	AI models \
0	Communications Manager	98%	365	2546
1	Data Collector	95%	299	2148
2	Data Entry	95%	325	2278
3	Mail Clerk	95%	193	1366
4	Compliance Officer	92%	194	1369
5	Chief Executive Officer (CEO)	91%	153	1135
6	Therapist	90%	114	983
7	Chief Strategy Officer	90%	184	1284
8	Admin	90%	253	2014
9	General Clerk	90%	242	1903

	AI_Workload_Ratio	Domain
0	0.143362	Communication & PR
1	0.139199	Data & IT
2	0.142669	Administrative & Clerical
3	0.141288	Leadership & Strategy
4	0.141709	Medical & Healthcare
5	0.134802	Supply Chain & Logistics
6	0.115972	Law Enforcement
7	0.143302	Construction
8	0.125621	Sales & Marketing
9	0.127168	Hospitality

```
[ ] data['AI Impact'] = data['AI Impact'].str.rstrip('%').astype(float)
```

```
mean_impact = data.groupby('Job titles')['AI Impact'].mean().sort_values()
```

```
[ ] mean_impact = data.groupby('Job titles')['AI Impact'].mean().sort_values()
print("Job Title with the highest AI impact:")
print(mean_impact.idxmax())
```

```
↕ Job Title with the highest AI impact:
Communications Manager
```

```
[ ] print("Job Title with the lowest AI impact:")
print(mean_impact.idxmin())
```

```
↕ Job Title with the lowest AI impact:
Zoologist
```

```
[ ] unique_domains = data['Domain'].unique()
```

Double-click (or enter) to edit

```
[ ] print("Unique Domains:")
for domain in unique_domains: print(domain)
```

```
↕ Unique Domains:
Communication & PR
Data & IT
Administrative & Clerical
Leadership & Strategy
Medical & Healthcare
Supply Chain & Logistics
Law Enforcement
Construction
Sales & Marketing
Hospitality
```

```
[ ] data_it_jobs = data[data['Domain'].isin(['Data & IT'])]
```

```
[ ] top_ai_models = data['AI models'].value_counts().head(10)
```

```
[ ] print("\nTop AI Models Used:")
print(top_ai_models)
```

```
↕ Top AI Models Used:
AI models
46      31
60      15
57      11
1049     9
20       9
49       8
879       8
1446       7
0         7
1427       7
Name: count, dtype: int64
```

```
[ ] least_impact_job_data_it = mean_impact_data_it.idxmin()

[ ] print("Job Title with the least AI impact within the Data & IT domains:", least_impact_job_data_it)
```

↔ Job Title with the least AI impact within the Data & IT domains: Admissions Coordinator

```
[ ] threshold = 0.8
    AI_Workload_Ratio = 50
    high_workload_jobs = data[data['AI_Workload_Ratio'] > threshold]

[ ] print("High Workload Jobs:")
    print(high_workload_jobs)
```

↔ High Workload Jobs:

	Job titles	AI Impact	Tasks	AI models	AI_Workload_Ratio \
3034	Charter Pilot	20.0	2	0	inf
3035	Commercial Pilot	20.0	2	0	inf
3036	Conductor	20.0	2	0	inf
3037	Corporate Pilot	20.0	2	0	inf
3058	After School Teacher	20.0	1	1	1.000000

```
[ ] tasks_distribution = data.groupby('Job titles')['Tasks'].apply(lambda x: ', '.join(map(str, x))).reset_index()

[ ] print(tasks_distribution)
```

↔

	Job titles	Tasks
0	.Net Developer	248
1	3D Animator	132
2	3D Artist	156
3	3D Designer	272
4	3D Modeler	204
...
4701	Youth Program Director	771
4702	Youth Specialist	229
4703	Youth Worker	158
4704	Zamboni Driver	172
4705	Zoologist	307

[4706 rows x 2 columns]

CHAPTER 7

System Implementation

SOURCE CODE

```
from google.colab import drive drive.mount('/content/drive') import os
os.chdir('/content/drive/MyDrive/AI_Project/') import pandas as pd data =
pd.read_csv('My_Data.csv') print(data.head(10)) data['AI Impact'] = data['AI
Impact'].str.rstrip('%').astype(float) mean_impact = data.groupby('Job
titiles')['AI Impact'].mean().sort_values() print("Job Title with the highest AI
impact:") print(mean_impact.idxmax()) print("Job Title with the lowest AI
impact:") print(mean_impact.idxmin()) unique_domains =
data['Domain'].unique() print("Unique Domains:") for domain in
unique_domains: print(domain) data_it_jobs = data[data['Domain'].isin(['Data
& IT'])] mean_impact_data_it = data_it_jobs.groupby('Job titiles')['AI
Impact'].mean() least_impact_job_data_it = mean_impact_data_it.idxmin()
print("Job Title with the least AI impact within the Data & IT domains:",
least_impact_job_data_it) threshold = 0.8
AI_Workload_Ratio = 50 high_workload_jobs =
data[data['AI_Workload_Ratio'] > threshold] print("High Workload
Jobs:") print(high_workload_jobs) top_ai_models = data['AI
models'].value_counts().head(10) print("\nTop AI Models Used:")
print(top_ai_models) tasks_distribution = data.groupby('Job
titiles')['Tasks'].apply(lambda x: ',
'.join(map(str, x))).reset_index() print(tasks_distribution)
```


CHAPTER 8

Conclusion

Our analysis of AI's impact on job trends highlights significant changes across various job sectors. AI models are becoming increasingly prevalent, especially in roles such as .Net Developer and 3D Designer, indicating a growing reliance on AI for these jobs. However, there are still many job titles with minimal or no AI integration, such as Zoologist, suggesting that not all roles are equally impacted by AI advancements. High AI workload jobs, like Charter Pilot and Commercial Pilot, are heavily influenced by AI, particularly in sectors such as Medical & Healthcare, Supply Chain & Logistics, and Law Enforcement. This uneven distribution of AI impact underscores that while AI is transforming many job functions by enhancing efficiency and automation, roles requiring specialized skills, creativity, and nuanced decision-making remain less susceptible to automation. These insights emphasize the need for proactive measures, such as upskilling and reskilling the workforce, to adapt to the evolving job market and harness the benefits of AI. Preparing for these changes will ensure a balanced, resilient, and future-ready workforce.

CHAPTER 9

References

- 1) [Search | Kaggle](#)
- 2) https://youtu.be/_U2YobRC8OY?feature=shared
- 3) <https://scholar.google.com/>
- 4) <https://sciencedirect.com/>