# AI Jobs Analytics - Comprehensive Documentation

1. Abstract

Artificial Intelligence (AI) is significantly reshaping the job market by introducing new opportunities while also causing job displacement in various sectors. The **AI Jobs Analytics** project is designed to analyze AI's impact on employment trends across diverse industries. This study employs **Machine Learning (ML) techniques** such as **K-Means Clustering** and **Association Rule Mining** to gain insights into AI-created jobs, AI-displaced jobs, and unemployment trends. Additionally, **network graph visualizations** provide a holistic view of AI's influence on different job sectors.

This documentation outlines the **system architecture, database schema, dataset details, and implementation methodologies**, serving as a guide for researchers, policymakers, and industry professionals looking to understand AI's role in shaping the workforce of the future.

2. System Architecture & ER Diagram

The AI Jobs Analytics system is structured around key entities, including **Job Sector, AI Impact, and Economic Factors**. The **Entity-Relationship Diagram (ERD)** below illustrates the relationships between these components:
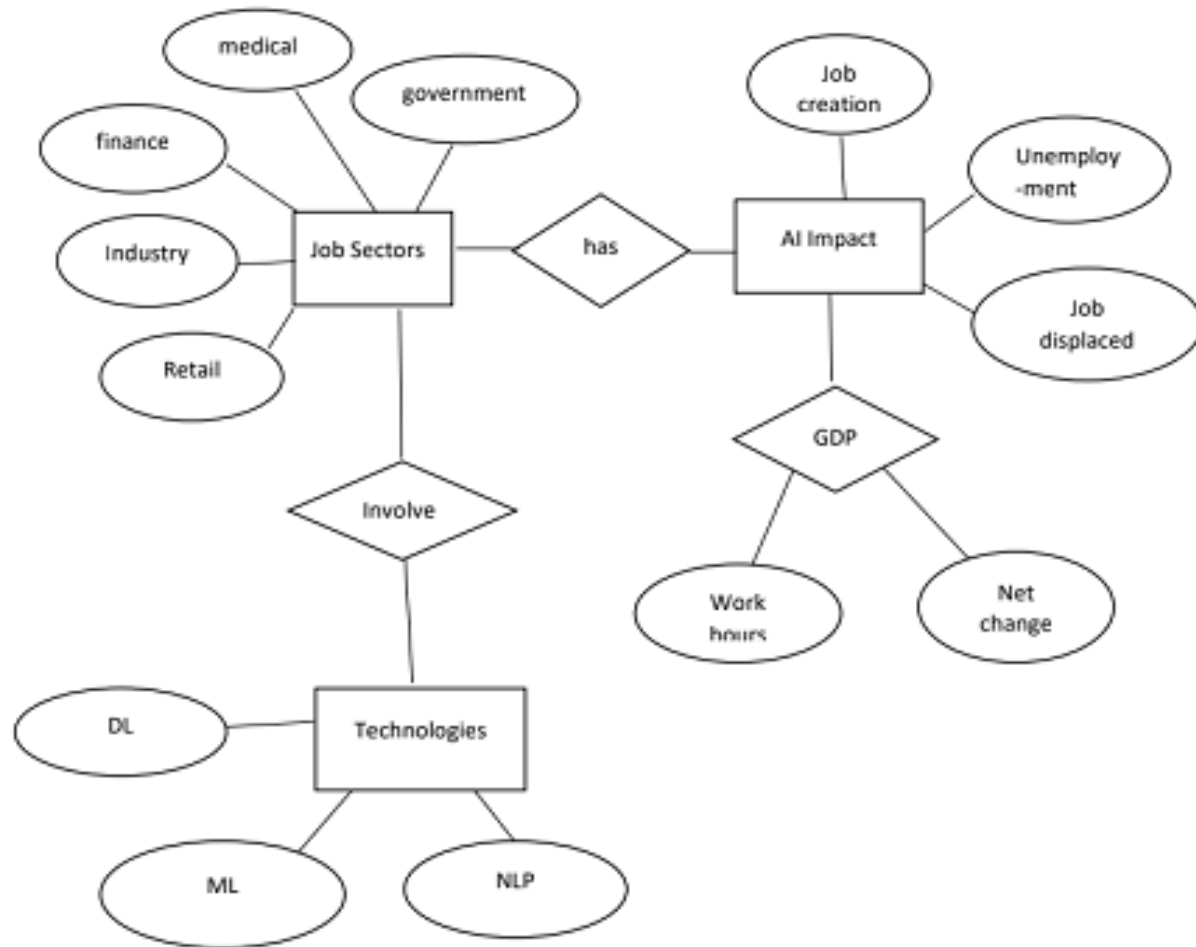
3. Database Schema & Data Model

To facilitate structured data storage and retrieval, the system follows a **relational database schema** with the following key tables:

1. Job_Sectors

Column Name	Data Type	Description
sector_id	INT (PK)	Unique ID for each job sector
sector_name	VARCHAR(50)	Name of the job sector

2. AI_Impact

Column Name	Data Type	Description
impact_id	INT (PK)	Unique impact record ID
sector_id	INT (FK)	Foreign Key referencing Job_Sectors
ai_created_jobs	INT	Number of jobs created due to AI
ai_displaced_jobs	INT	Number of jobs lost due to AI
unemployment_rate	FLOAT	Unemployment rate (%)
net_impact	INT	Net effect of AI in this sector

3. Economic_Factors

Column Name	Data Type	Description
factor_id	INT (PK)	Unique ID for economic factor
sector_id	INT (FK)	Foreign Key referencing Job_Sectors
avg_work_hours	FLOAT	Average weekly working hours
gdp_growth	FLOAT	GDP growth percentage for the sector

4. Dataset Overview

The dataset used in this project provides statistical insights into AI-driven employment shifts. Below is a snapshot of the dataset:

Job Sector	AI Created Jobs	AI Displaced Jobs	Unemployment Rate (%)	Avg Work Hours	GDP Growth (%)	Net Impact
Medical	509	334	5	40	4	78
Finance	453	440	7	35	7	77
Retail	624	475	8	23	5	76
Manufacturing	789	598	6	48	4	45
Government	800	449	4	34	7	65

5. Implementation Methodology

Machine Learning Approaches

- **K-Means Clustering:** Applied to AI-created jobs, AI-displaced jobs, and unemployment rates to group similar job sectors based on AI impact.
- **Association Rule Mining:** Utilizes the **Apriori Algorithm** to discover job sectors with common AI impact patterns and trends.
- **Network Graph Analysis:** Constructs a **graph-based model** to visualize the interconnections between different job sectors based on AI's effect.

Data Preprocessing & Cleaning

- Missing values in AI displacement data are handled using **backward fill (bfill)**.
- Outliers in work hours and net impact are detected and normalized for better clustering accuracy.

Visualization & Insights

- **Scatter Plot Analysis:** Provides an **AI impact clustering view**, mapping net job gains and losses.
- **Graph Network Visualization:** Helps in understanding AI's interconnected effects across multiple industries.
- **Correlation Matrices:** Used to explore relationships between **GDP growth, AI job displacement, and unemployment rates**.

6. Key Findings & Conclusion

AI is bringing about significant **structural changes** in employment, with some industries benefiting from AI-driven job creation while others experience displacement. This analysis provides:

- A **data-driven assessment** of AI's role in workforce transformation.
- Predictions of **high-risk sectors** prone to job displacement.
- Potential strategies for governments and businesses to **mitigate unemployment risks** through AI adaptation.

Future Scope

- Expansion of datasets to **include real-time AI job statistics**.
- Integration of **predictive analytics models** to forecast AI job trends.
- Development of an **AI-powered recommendation system** for workforce upskilling.