



AI Under the Microscope

By **HOT Bot Team**



Team Members



Omnia Gamal



Hazem Medhat



Toka Magdy

Table of contents

01 Problem Statement

02 Project Objective

03 Data Set

04 Visuals & Insights

05 Predictions & Results

06 Recommendation & Conclusion

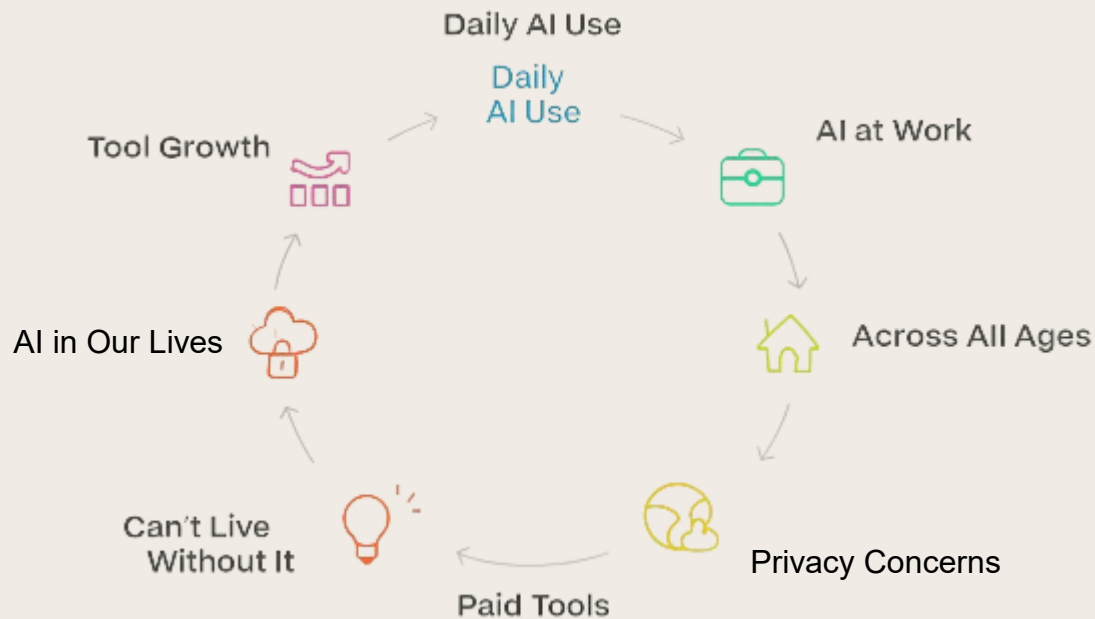




01

Problem Statement

Introduction



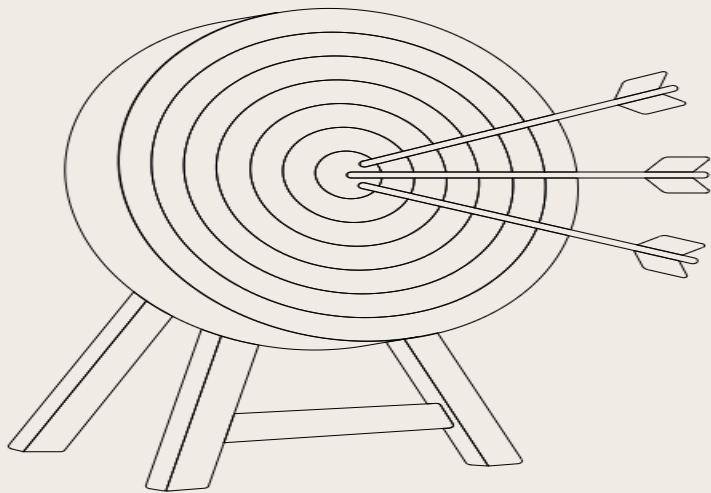


02

Project Objective



Objectives



Predict AI performance

Forecast future AI capabilities



Interpret user behavior patterns

Understand user interactions with AI



Examine user reliance on AI

Assess user dependency on AI



03

Data Set



Meta Data

- **This dataset contains information about individuals' AI usage patterns, including demographic details, AI tool usage, purposes, and related metrics.**
- **Number of records: 10,000**
- **Number of Features : 22**
- **Challenges : Handle Missing Values and Outliers**

Meta Data

Column Name	Type	Description
Age	Integer	Age of the individual
Gender	Categorical	Male, Female
Occupation	Categorical	Job Title
Daily_AI_Usage_Hours	Integer	Daily hours spent using AI tools
AI_Usage_Purpose	Categorical	Purpose to use AI
Specialization	Categorical	Medicine, Law, Education, Engineering, Art, Business, Computer Science
AI_Tools_Used	Categorical	Midjourney, ChatGPT, Notion AI, Copilot,

Meta Data

Column Name	Type	Description
AI_Usage_Frequency	Categorical	Daily, Often, Sometimes
AI_Skill_Level	Categorical	Beginner, Intermediate, Advanced
Device_Used	Categorical	Mobile, Laptop, Desktop, Tablet
Data_Privacy_Concern	Categorical	Low, Medium, High
AI_Trust_Level	Categorical	Low, Moderate, High
Location	Categorical	Urban, Suburban, Rural
Language_Preference	Categorical	English, Arabic, Spanish

Meta Data

Column Name	Type	Description
Internet_Speed	Categorical	Slow, Medium, Fast
AI_Usage_Change_in_6_Months	Categorical	Increased, Decreased, Same
Uses_Paid_AI_Tools	Categorical	Yes, No
Satisfaction_Level	Integer	1 to 10, indicating satisfaction with AI tools
Can_Live_Without_AI	Categorical	Yes, No
Task_Complexity	Integer	1 to 5, indicating complexity of tasks performed with AI
Time_to_Respond_Second	Integer	Time in seconds to respond
Performance	Integer	Performance percentage

Data Cleaning

- Apply Forward-fill approach in missing values in (Gender , Uses_Paid_AI_Tools)
- Fill Missing Values in “AI_Tools_Used” with “other” value
- Filling Missing Values in Multiple Columns with Mode: ['AI_Usage_Purpose', 'Specialization', 'AI_Usage_Frequency', 'AI_Skill_Level', 'Device_Used', 'AI_Trust_Level', 'Internet_Speed']



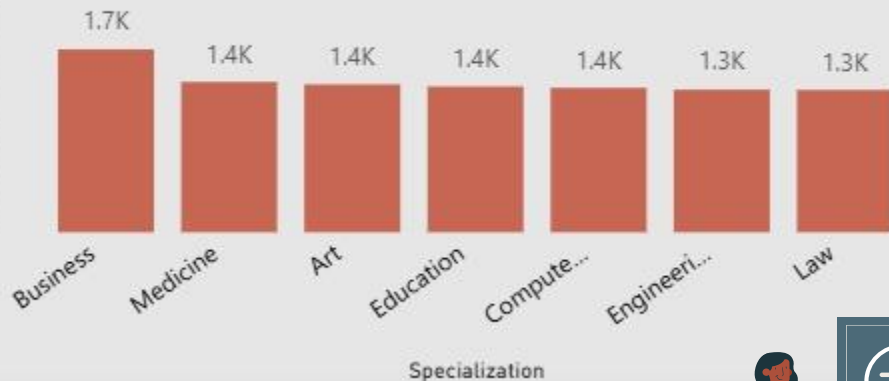
04

Visuals & Insights



Count of AI_Tools_Used by Specialization

Count of AI_Tools_Used



AI Trust, Satisfaction & Performance

Average Satisfaction Level

5.52

% of Users Concerned About Data Privacy

33.84

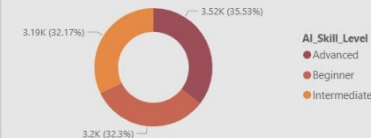
Average Performance %

73.33

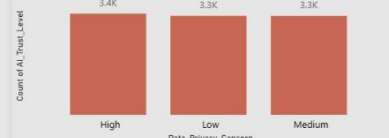
% of Users Who Live Without AI

49.97

AI Skill Level Distribution

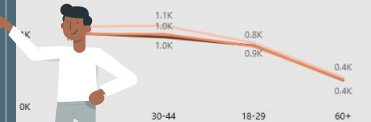


Count of AI_Trust_Level by Data_Privacy_Concern

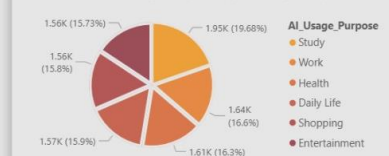


AI Usage Frequency By Age

AI_Trust_Level — High — Low — Moderate



Count of AI_Usage_Purpose by AI_Usage_Purpose



AI Usage Behavior & User Demographics

Average Daily AI Usage (hrs)

5.99

Most Common AI Tool Used

Midjourney

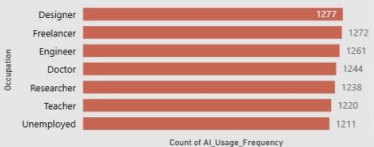
Most Common Purpose

Study

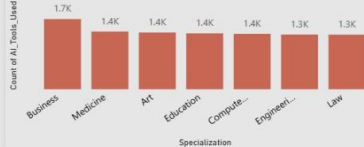
% of Users Who Use AI Daily

27.40

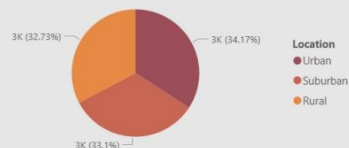
Count of AI_Usage_Frequency by Occupation



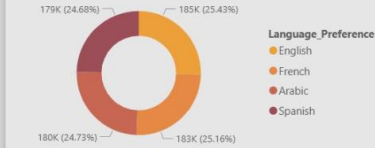
Count of AI_Tools_Used by Specialization



Number of Users by Location

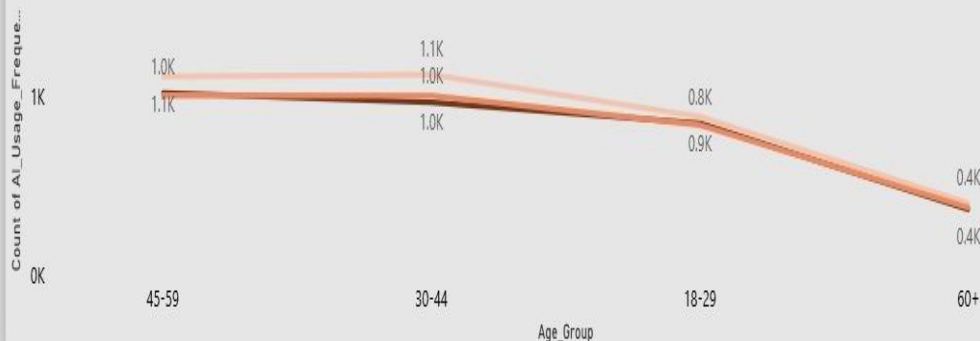


Sum of Performance by Language_Preference



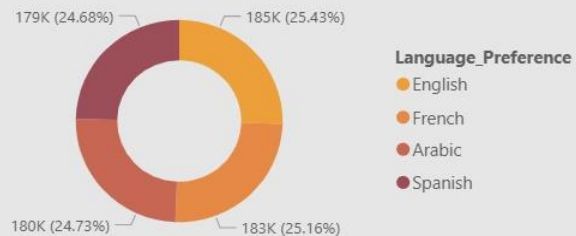
AI_Usage_Frequency By Age

AI_Trust_Level — High — Low — Moderate

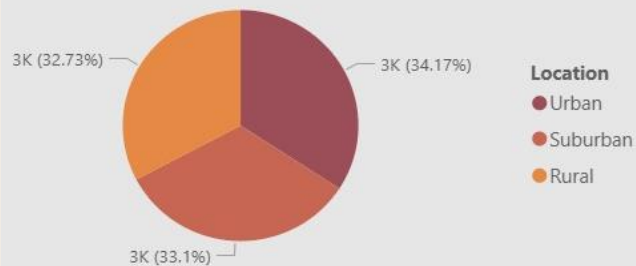




Sum of Performance by Language_Preference



Number of Users by Location





05

Predictions & Results



Objective

AI Dependency and Performance Assessment

AI Dependency Classification

Classify if life is
possible without AI



AI Performance Prediction

Forecast AI model
effectiveness in
scenarios








Pre Processing

- Split data for Classification and Regression Model
- Apply Label Encoding For Target Column in Classification Problem
- Encode Features using 2 approaches: Target and Label Encoding
- Scale Data Using Different Approaches : Standard Scaling, Robust Scaler

Classification Models

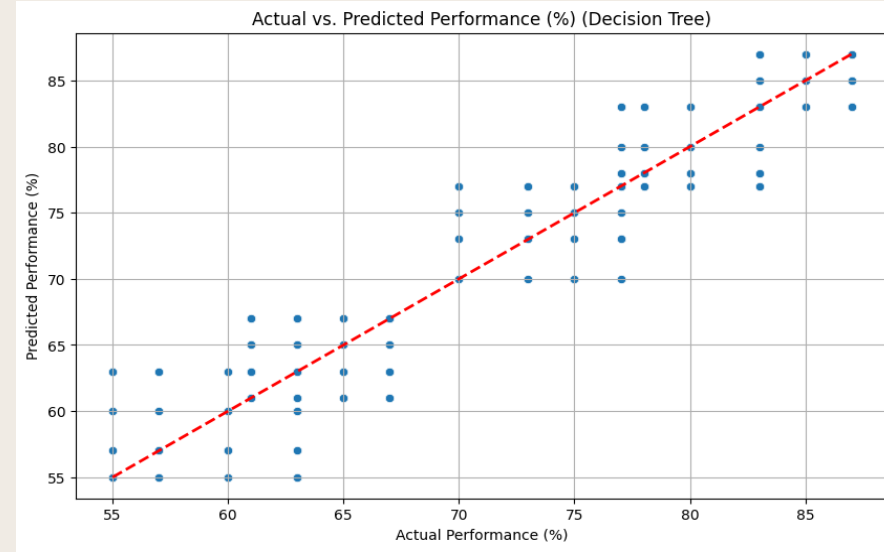
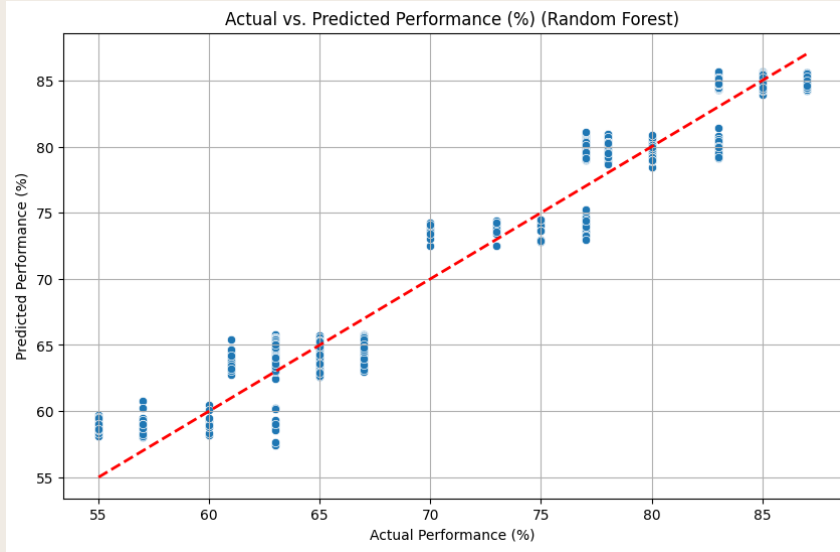
Model	Accuracy
Random Forest	100%
Decision Tree	100%
Logistic Regression	100%
SVM	99%
Naïve Bayes	100%

Regression Models



Model	Accuracy
Random Forest	95%
Decision Tree	92%
Linear Regression	49%
SVR	62%

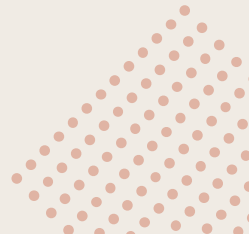
Regression Models





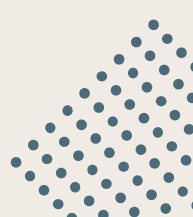
06

Recommendation & Conclusion





Let's Stay in Control

- Let's not let AI take over our lives. Reduce how much you depend on it.
 - AI was created by a human , a brain just like ours is what led to these powerful tools.
 - We must remember that AI tools should go back to their proper place in our lives.
 - They are helpful? yes, but they should not become the center of everything.
 - Let's use AI wisely, without letting it replace the value of our own thinking, creativity, and control.
- 



Thanks!

Data Gathering

- **We searched for AI usage datasets but found most were either too small or already clean, limiting preprocessing work.**
- **After extensive searching, we decided to generate the dataset using AI, guided by reference data and custom column definitions.**
- **We explained each column's purpose and type in detail to ensure accurate generation.**
- **Then we reviewed the data row by row, cleaned illogical entries, and refined some columns through multiple iterations.**
- **The final dataset was rich in features, suitable for both prediction and classification, and ideal for applying what we learned.**