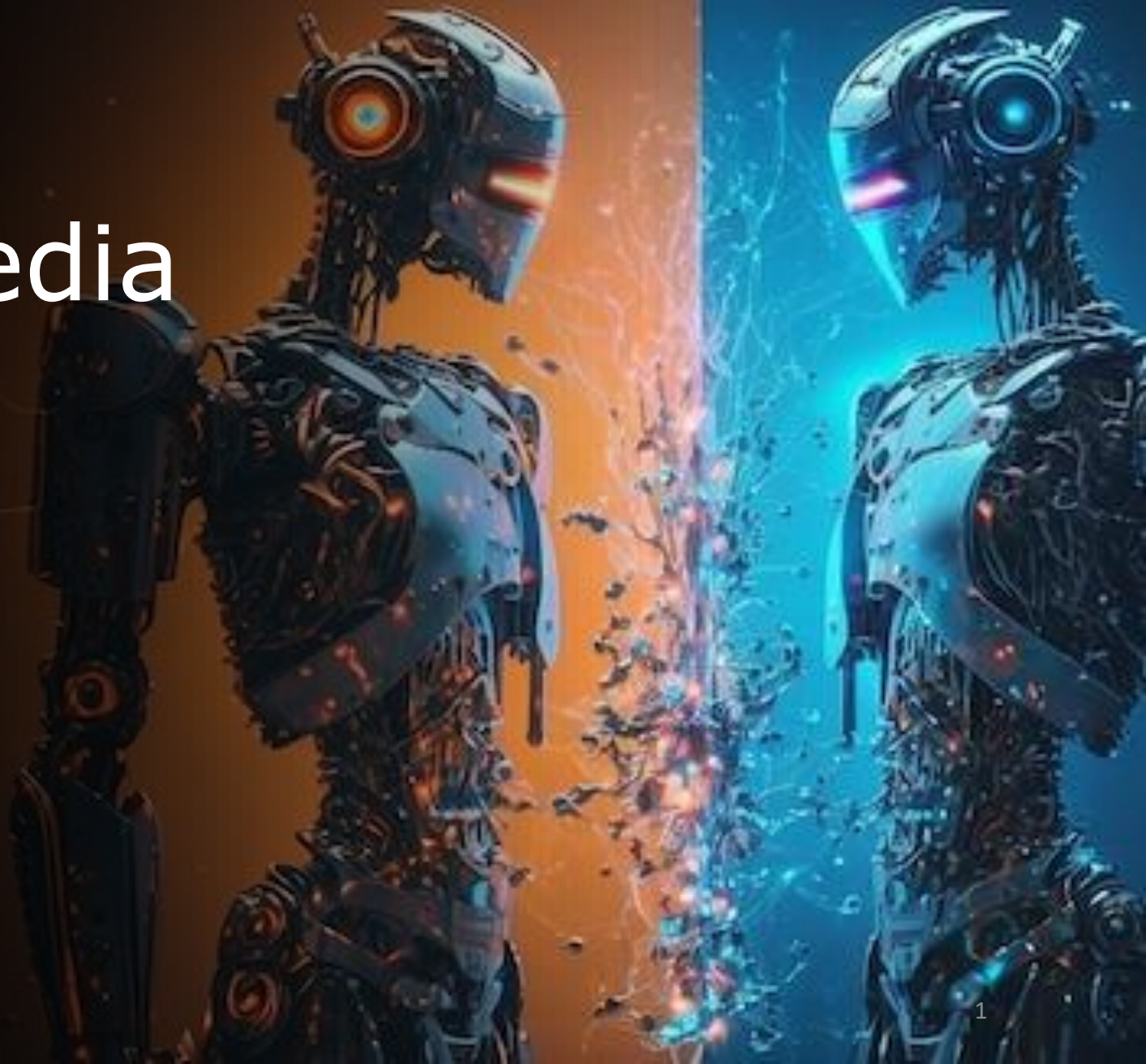
AI-Generated Media Detection



Under Supervision :

Dr. Eman Abdel-Latef .

Eng. Sahar.

Team Members

- Abd-Allah Mohamed Abdel-monem.
- Ahmed Mohamed Ali Abo el-Kassem.
- Mohamed Abd-Allah Abdel-salam Abdel-Dayem.
- Mohannad Ayman Salah Abdel-Fattah.
- Rawan Abdel-Aziz Ahmed Mahmoud.
- Reham Moustafa Ali Abdel-Moati.
- Romani Nasrat Shawqi Gerges.
- Sara Reda Moatamed Hassan.
- Zeyad Elsayed Abdel-Azim Ali Yousif.

Agenda

- Problem Definition.
- Problem Solution .
- Project Design.
- Models .
- UI Design.
- Tools.

Problem Definition

Problem Definition

- **Problem Statement:**
  - Addressing challenges in media authenticity due to deepfake technology.
  - Risks include misinformation, privacy breaches, and malicious intent.

- **Project Overview:**
  - AI Detection System for Multiple Media.
  - Aims to tackle manipulation in text, images, and deepfake photos.

Problem Definition

- **Key Objectives:**
  - Text Analysis: Uncovering deceptive language and misinformation.
  - Image Detection: Identifying manipulated elements in images.
  - Deepfake Photo Recognition: Detecting and mitigating the impact of deepfake photos.

Problem Definition

- **Technical Aspects:**
  - Leveraging cutting-edge technologies for comprehensive solutions.
  - Advanced algorithms for image and deepfake analysis.

- **Methodologies:**
  - Detailed analysis of textual content for deception.
  - Image analysis techniques for identifying manipulation.
  - State-of-the-art algorithms for deepfake detection.

Problem Solution

Project Introduction

- AI Detection System for Multiple Media ,Our graduation project focuses on developing an advanced AI detection system capable of discerning authenticity across various media types text, images, and deepfake photos. Leveraging cutting-edge technologies, our system aims to provide a comprehensive solution to the escalating concerns surrounding manipulated content.

Project Introduction

- **Key Objectives:**
  - **Text Analysis**: Uncovering deceptive language and misinformation in textual content.
  - **Image Detection**: Identifying manipulated elements within images through advanced image analysis.
  - **Deepfake Photo Recognition**: Employing state-of-the-art algorithms to detect and mitigate the impact of deepfake photos.

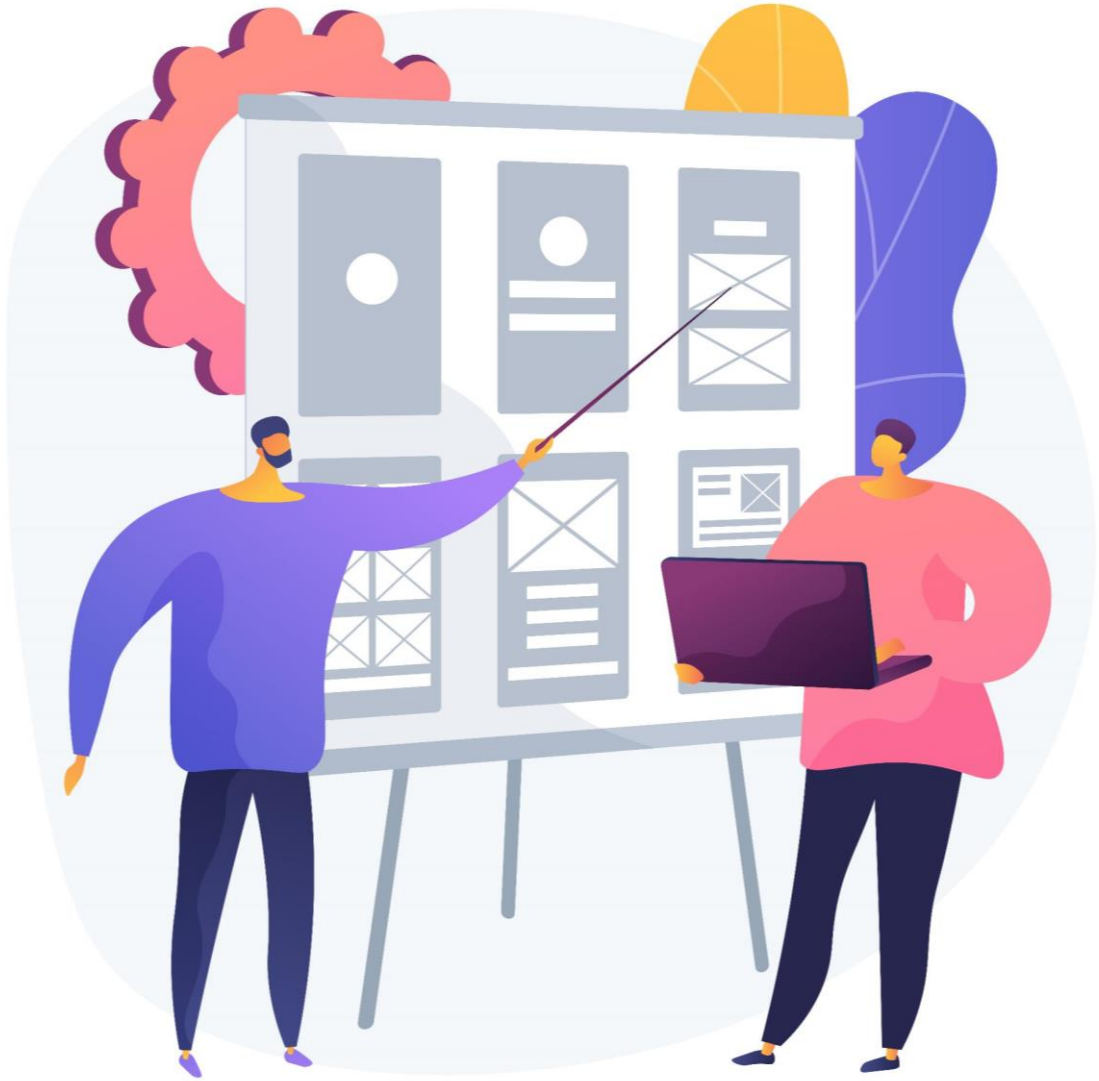
Project Introduction

- **Significance:**
  - **Preserving Trust**: Enhancing trust in digital media by ensuring authenticity.
  - **Mitigating Harm**: Preventing the potential harm caused by misinformation and manipulated content.
  - **Contributing to AI Ethics**: Upholding ethical standards in the development and deployment of AI technologies.

Project Introduction

- **Goals**:
  - Create a powerful detection system.
  - Contribute to ethical AI use in media.
  - User Education: Creating awareness programs to educate users about media manipulation risks.
  - Multimodal Integration: Integrating analysis across multiple modalities for enhanced accuracy.

Let's delve into the technical aspects and methodologies in the subsequent sections.



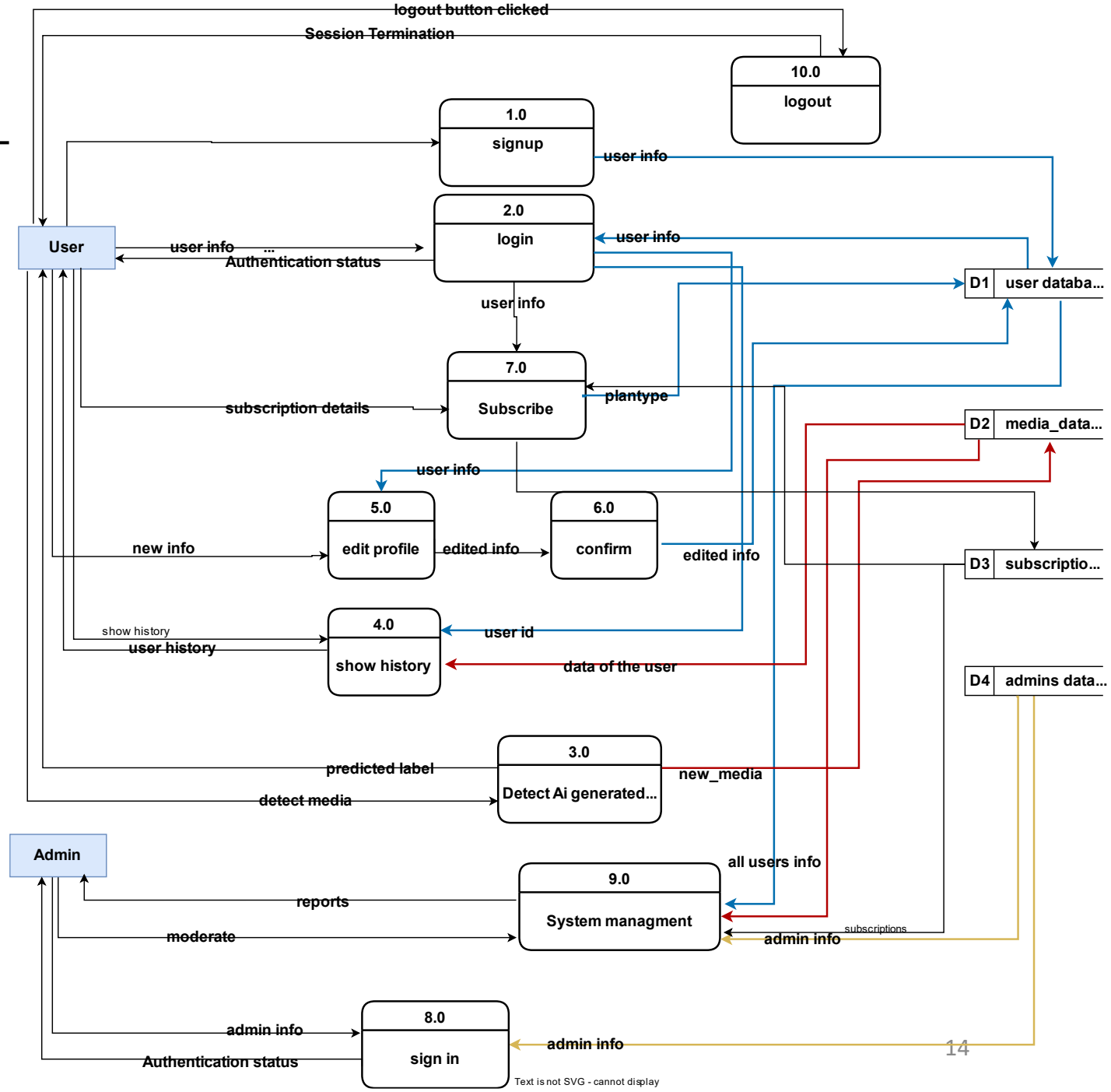
Project Design

- Data Flow Diagram.
- Entity Relationship Diagram.

Data Flow Diagram

- It is a graphical representation of the flow of data through a system, illustrating how data is input, processed, and output.

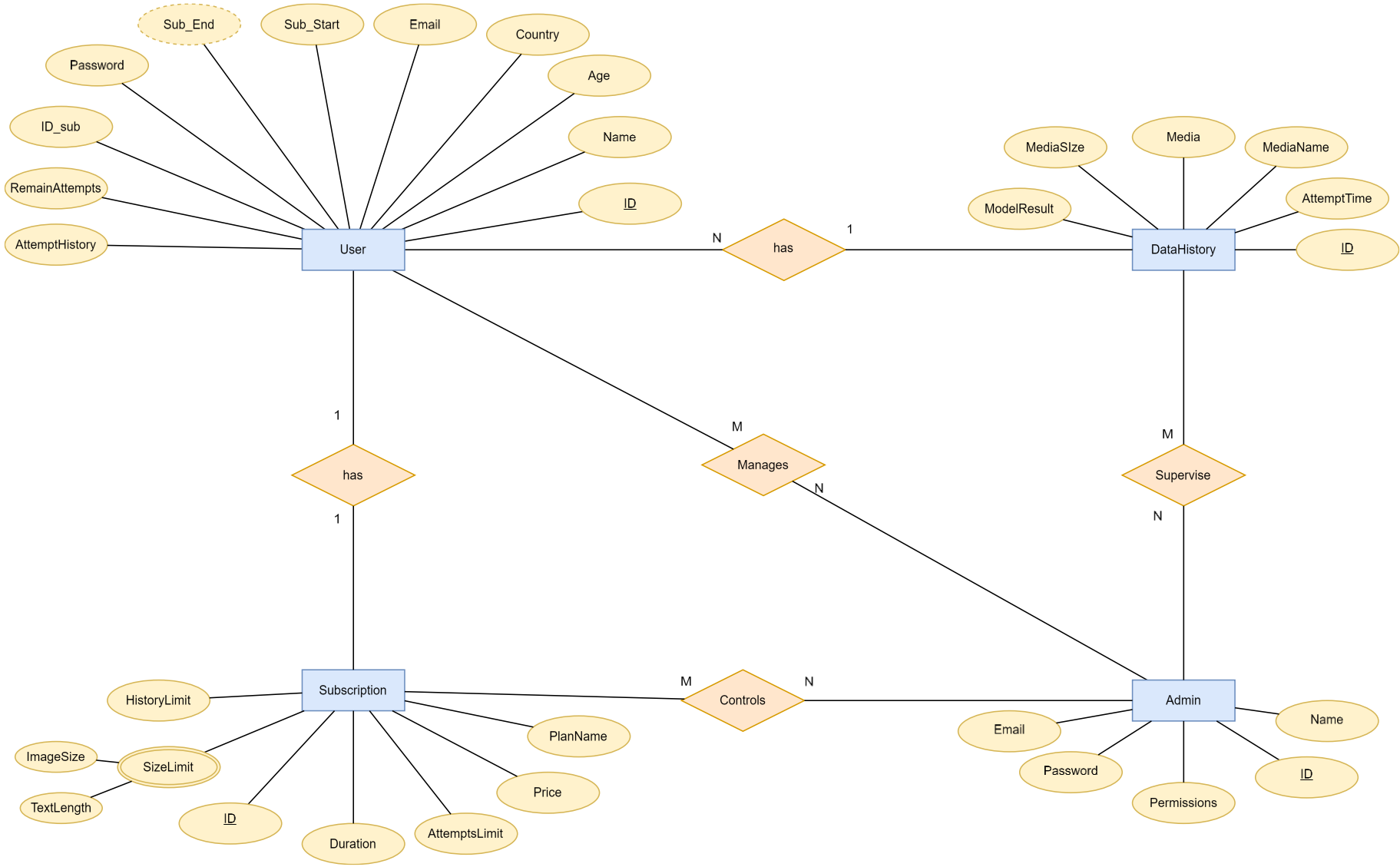
- data flow diagram contains :
  - 10 Process .
  - 4 Data Stores.
  - 2 External Entity.



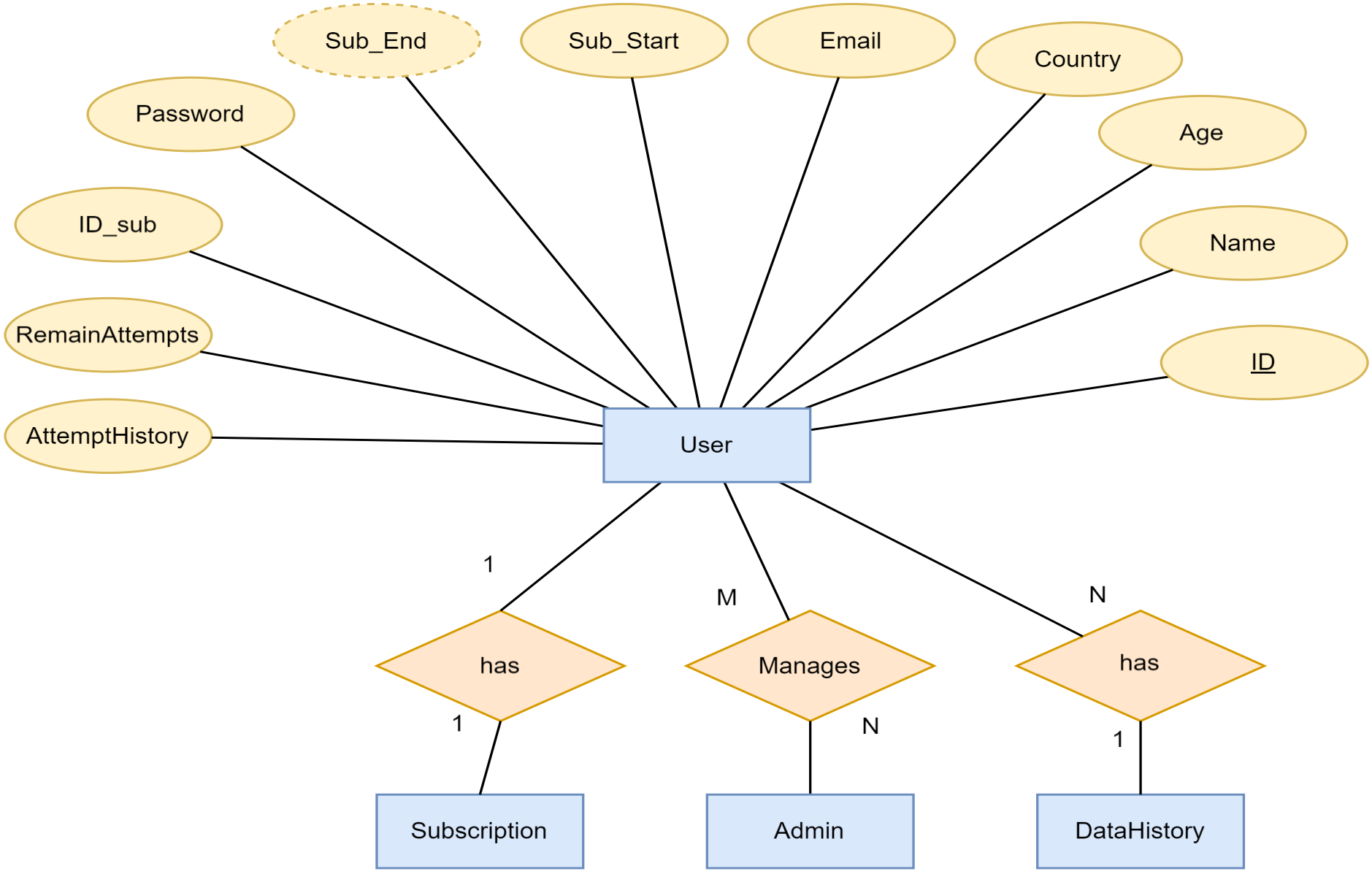
Entity Relationship Diagram

- It is a visual representation of the data model that depicts the entities within a system and the relationships between them.

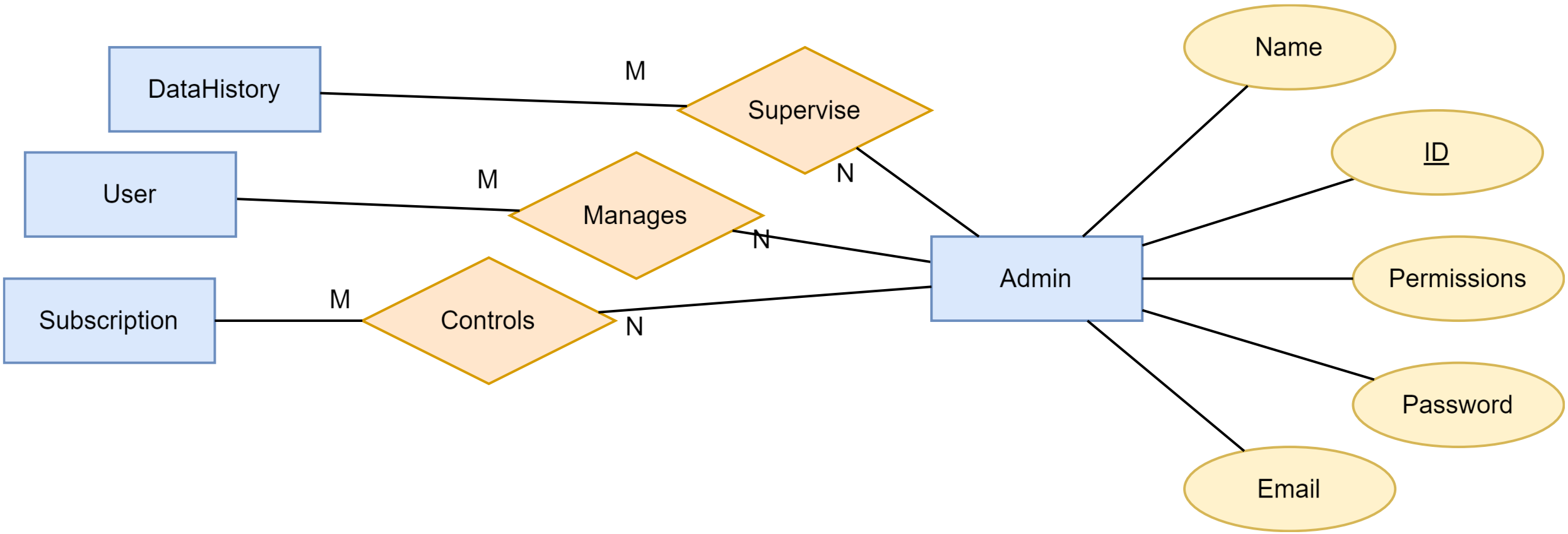
- **ERD** contains 4 entities :
  - User.
  - Admin.
  - Subscription.
  - Data History.



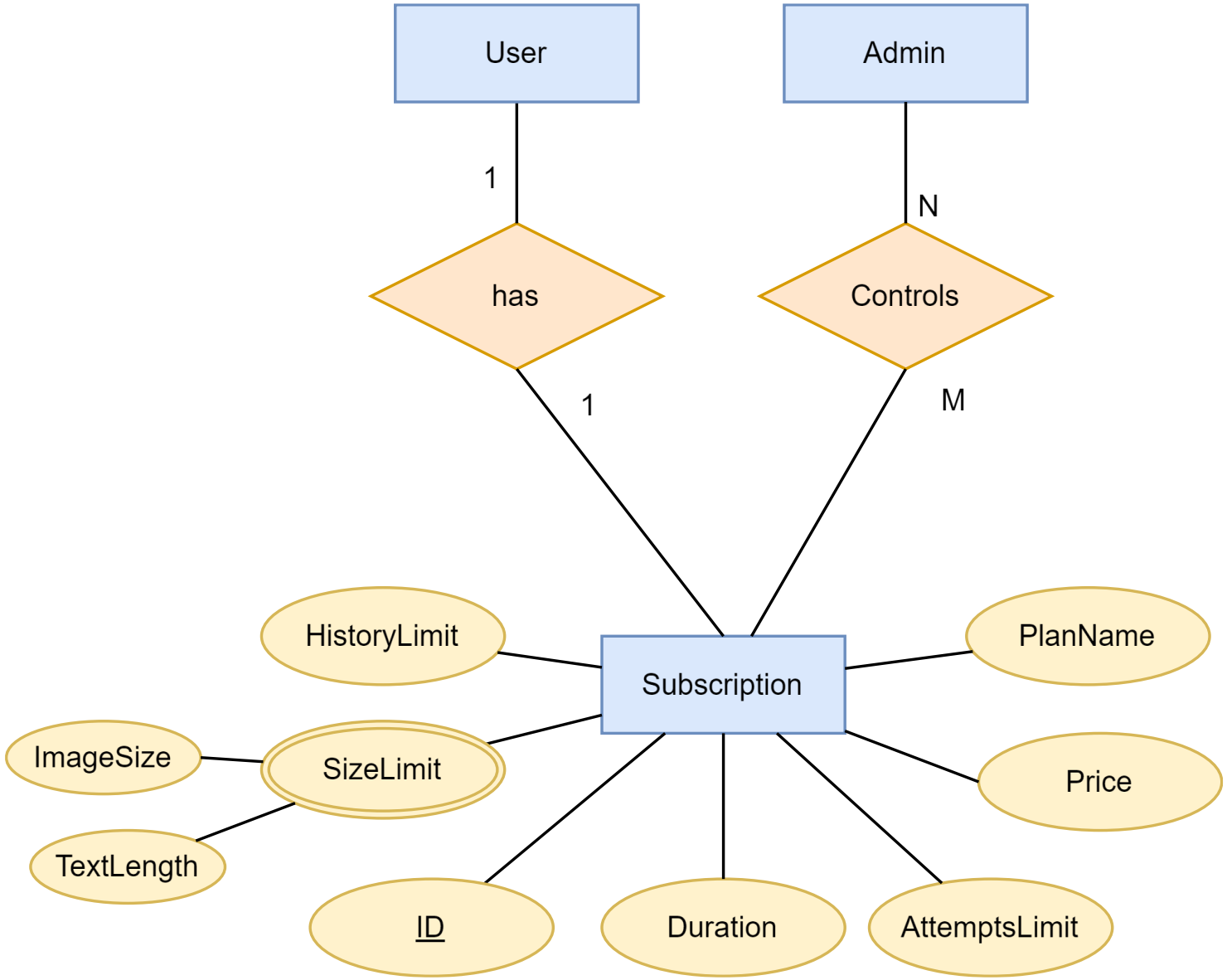
ERD(User)Cont...



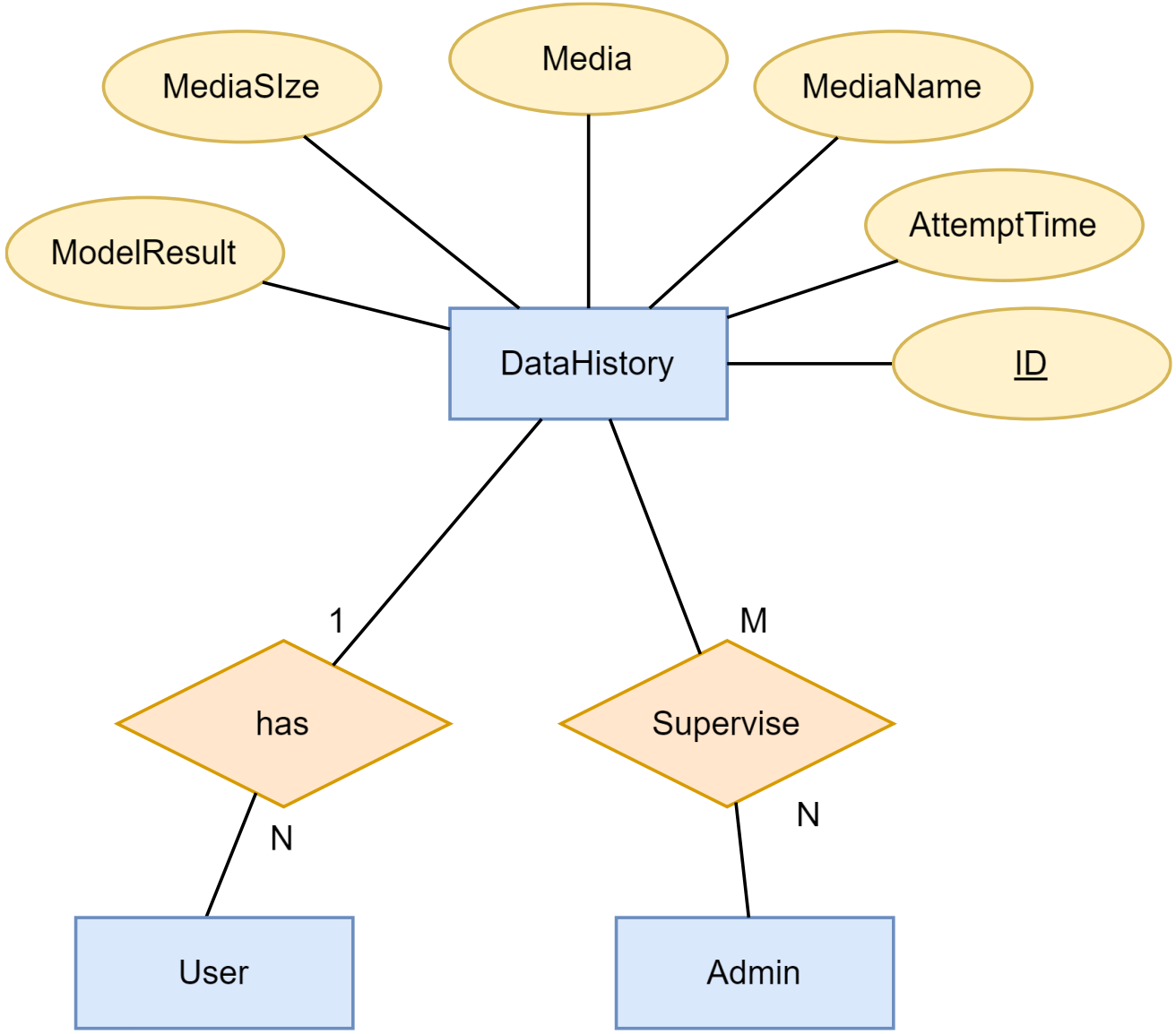
ERD(Admin)Cont...



ERD(Subscription)Cont...



ERD(Data History) Cont...



Models of Detection

- Text detector .
- Image generated detector .
- Deepfake detector.

Text Detection Model

- Large language Models (LLMs) can write things that are exactly like what people write, but it is difficult to know whether words come from a person or a machine.
- This can be a problem for teachers who are checking whether students have done their own work or used these forms to finish assignments.

- A text Detection model is a classification model that separates machine-generated text from human-written ones.

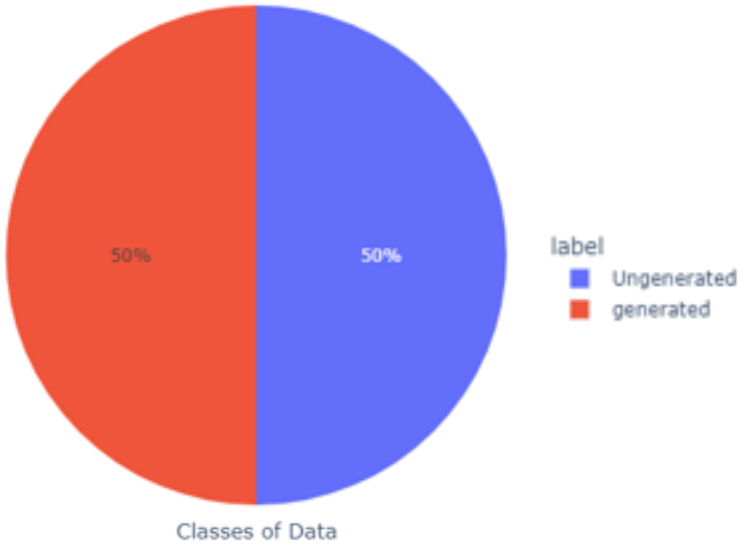
- We are trying, using advanced neural network techniques and language models, to build an architecture capable of distinguishing between what was written by an AI and a human, and to be more robust in the face of adversarial attacks.

Text Detection **Prototype** Model

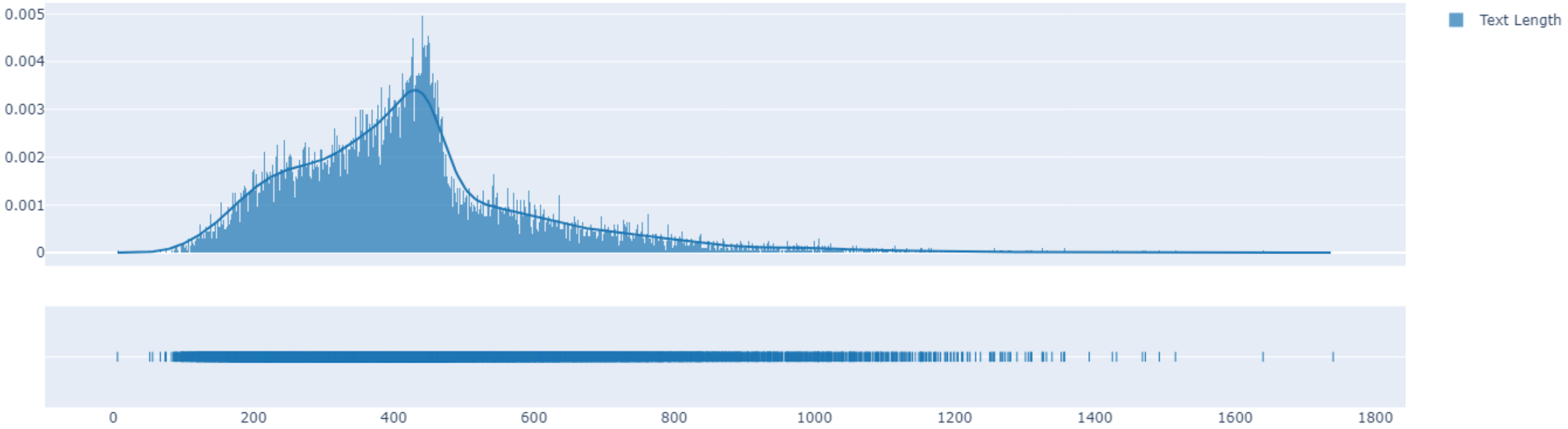
- A prototype model was created using some deep learning and NLP techniques, including LSTM Bidirectional, on this data:

  - LLM - Detect AI Generated Text : It is a competition dataset comprises about 10,000 essays, some written by students and some generated by a variety of large language models (LLMs). But the number of data available to train the model is 1378 , all of which is written by students.
  - DAIGT Proper Train Dataset : We used some of this data files to train the model

- So we have 19,986 essays, divided into:
  - 9,993 essays by students .
  - 9,993 essays of Text generated by ChatGPT, Llama and Falcon.



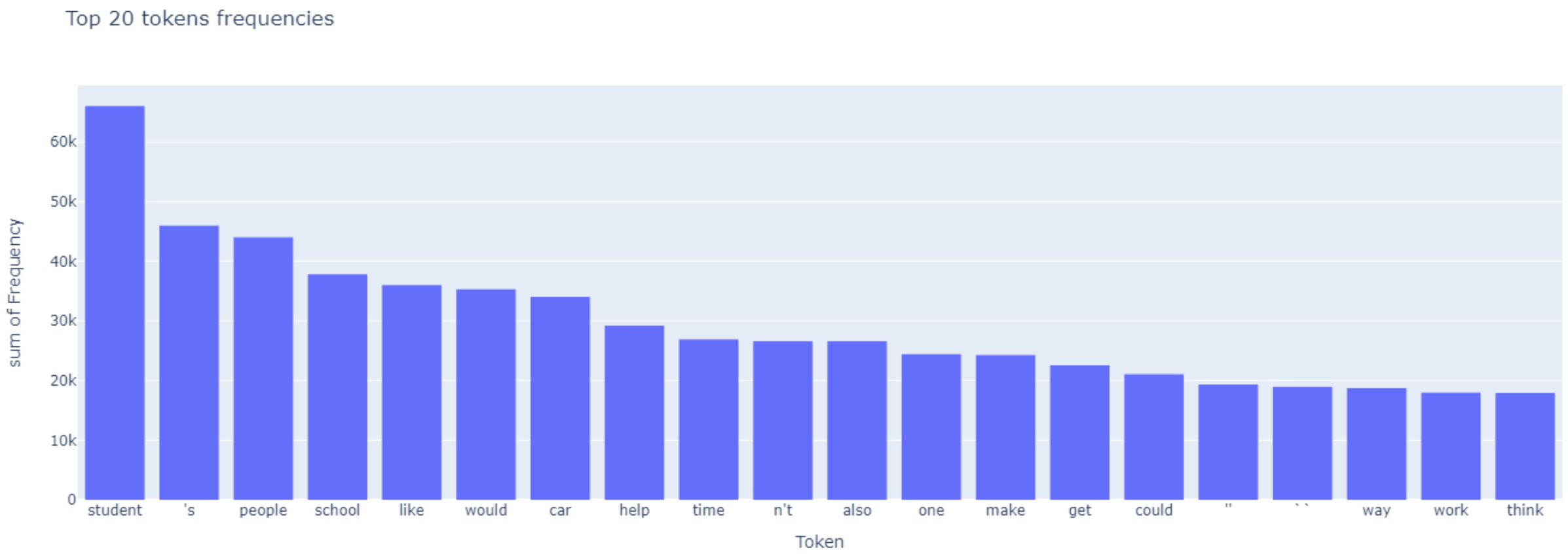
Text Detection **Prototype** Model(cont…)



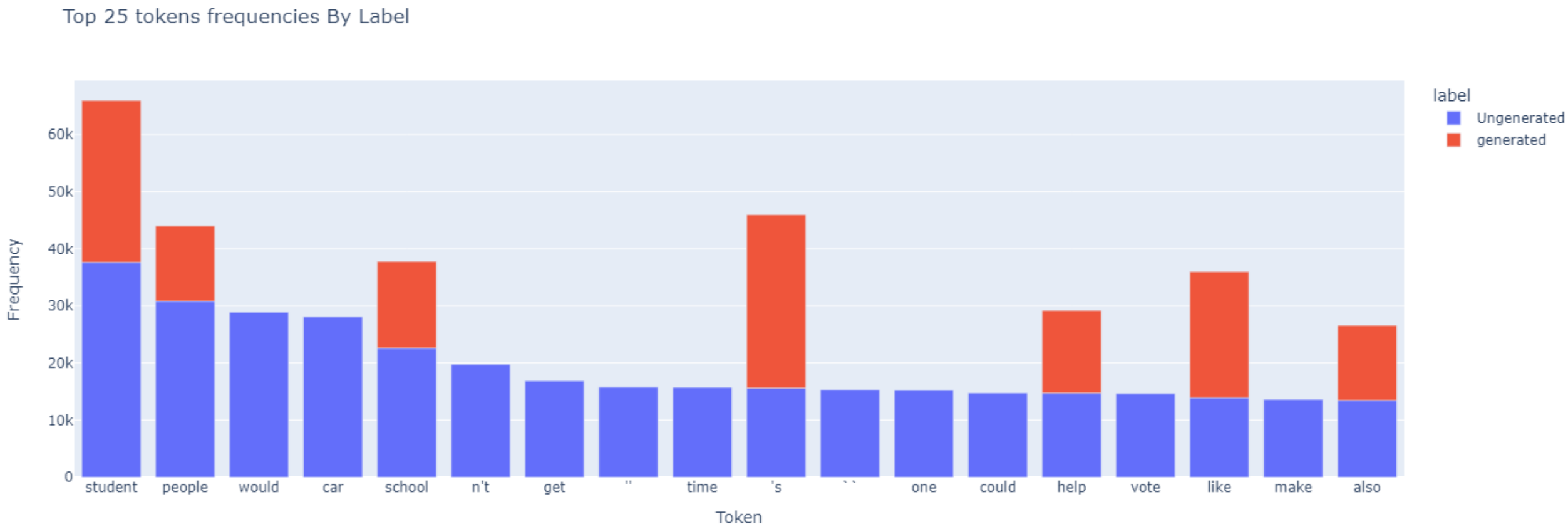
This is a **distribution of data** across text length , from plot:

- The Maximum length is **1740**
- The Minimum length is **6**
- Most of the lengths are between **427** and **450**

Text Detection **Prototype** Model(cont…)



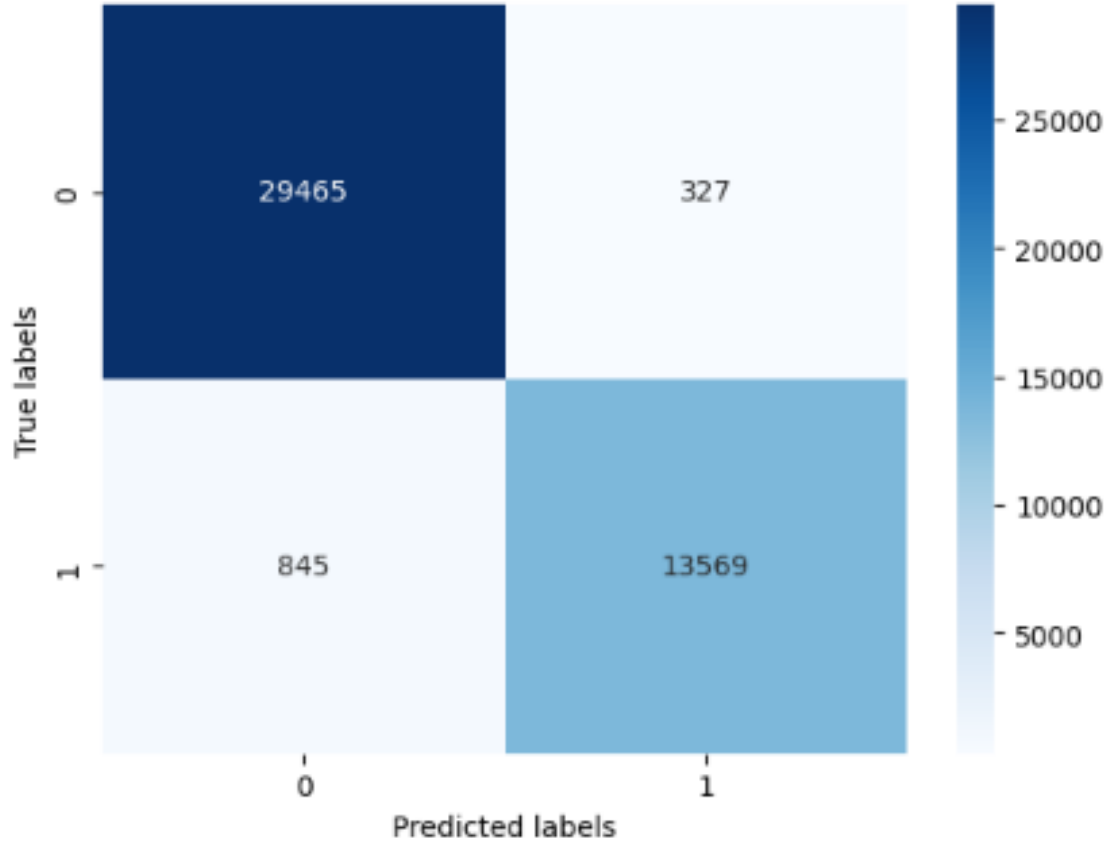
Text Detection **Prototype** Model(cont…)



Text Detection **Prototype** Model(cont…)

Result of model :

	precision	recall	f1-score	support
0	0.97	0.99	0.98	29792
1	0.98	0.94	0.96	14414
accuracy			0.97	44206
macro avg	0.97	0.97	0.97	44206
weighted avg	0.97	0.97	0.97	44206



Models of Detection

- Text detector .
- Image generated detector .
- Deepfake detector.

Image generated detector .

Models of Detection

- Text detector .
- Image generated detector .
- Deepfake detector.

Deepfake detector

Deepfake detector

UI Design

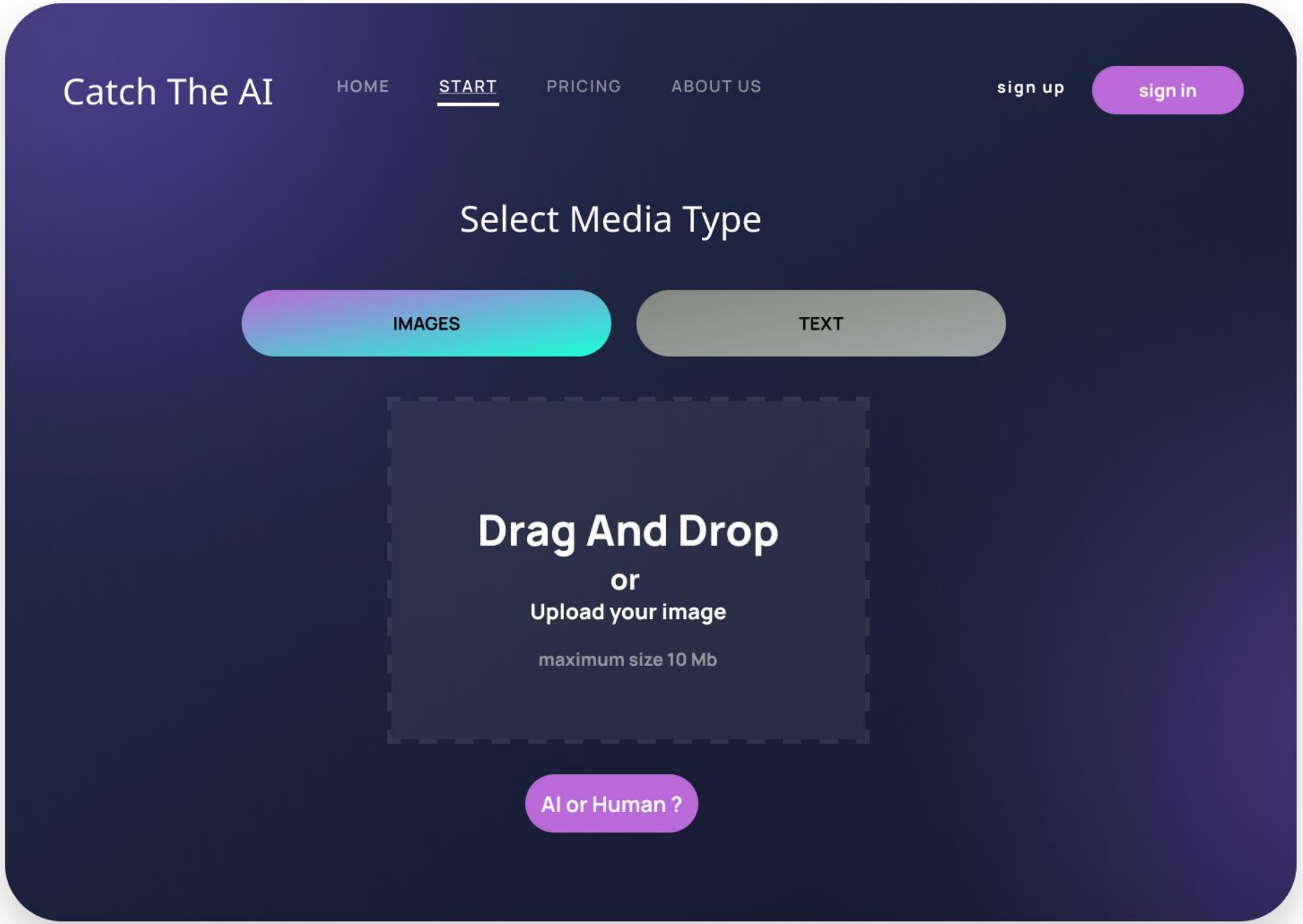
Design for web pages

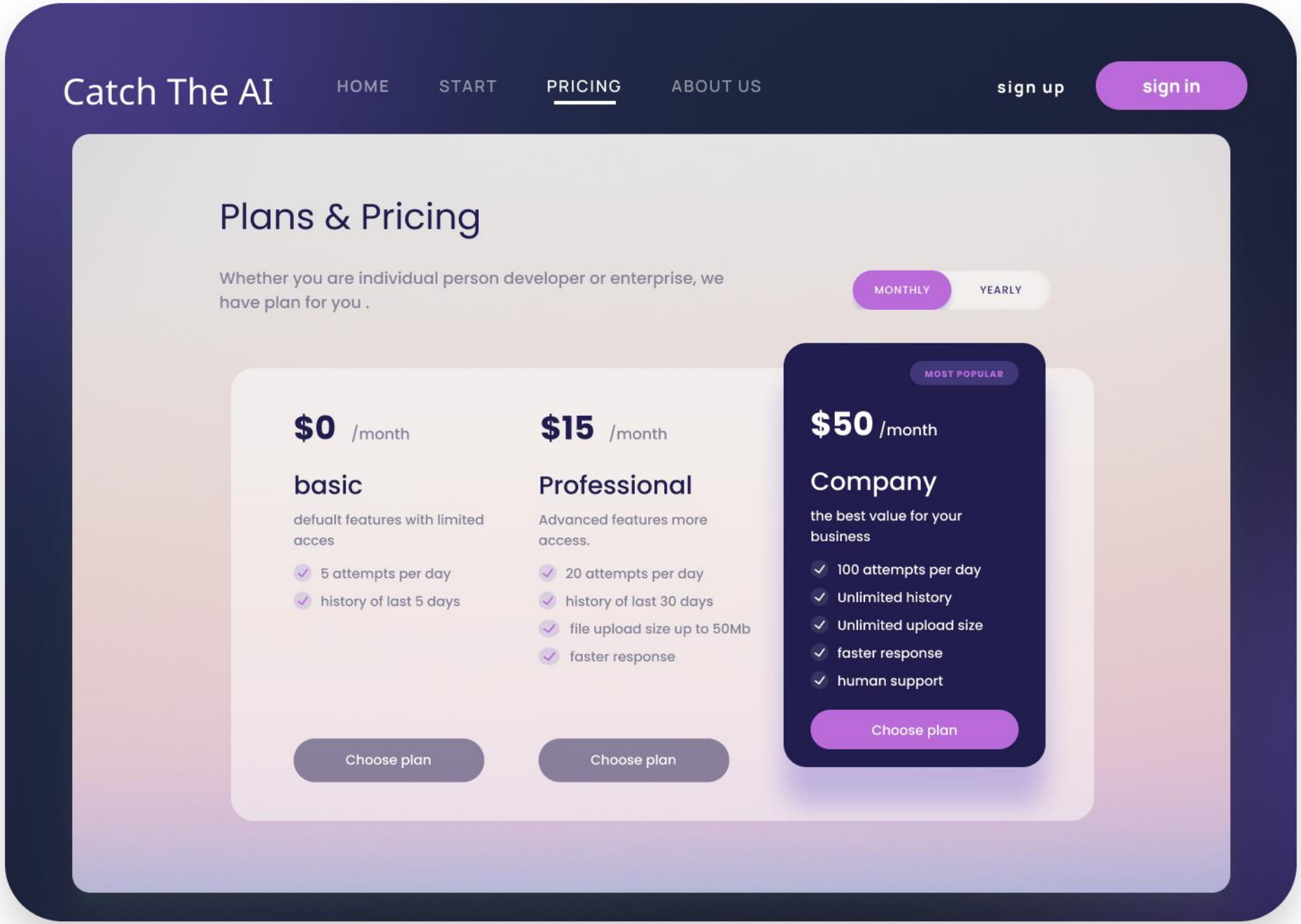
Catch The AI HOME START PRICING ABOUT US sign up sign in

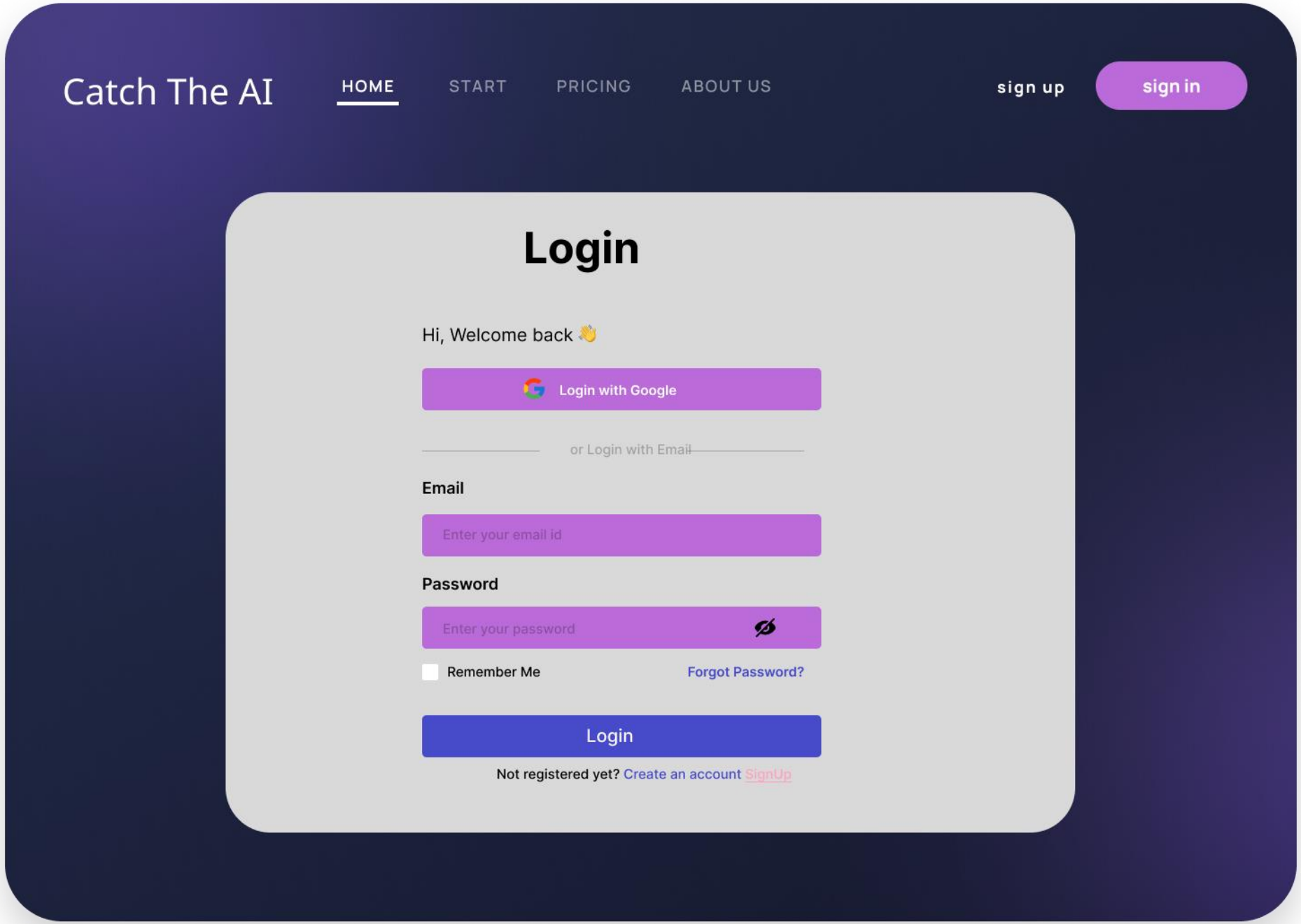
Start Detecting
AI-Generated Media

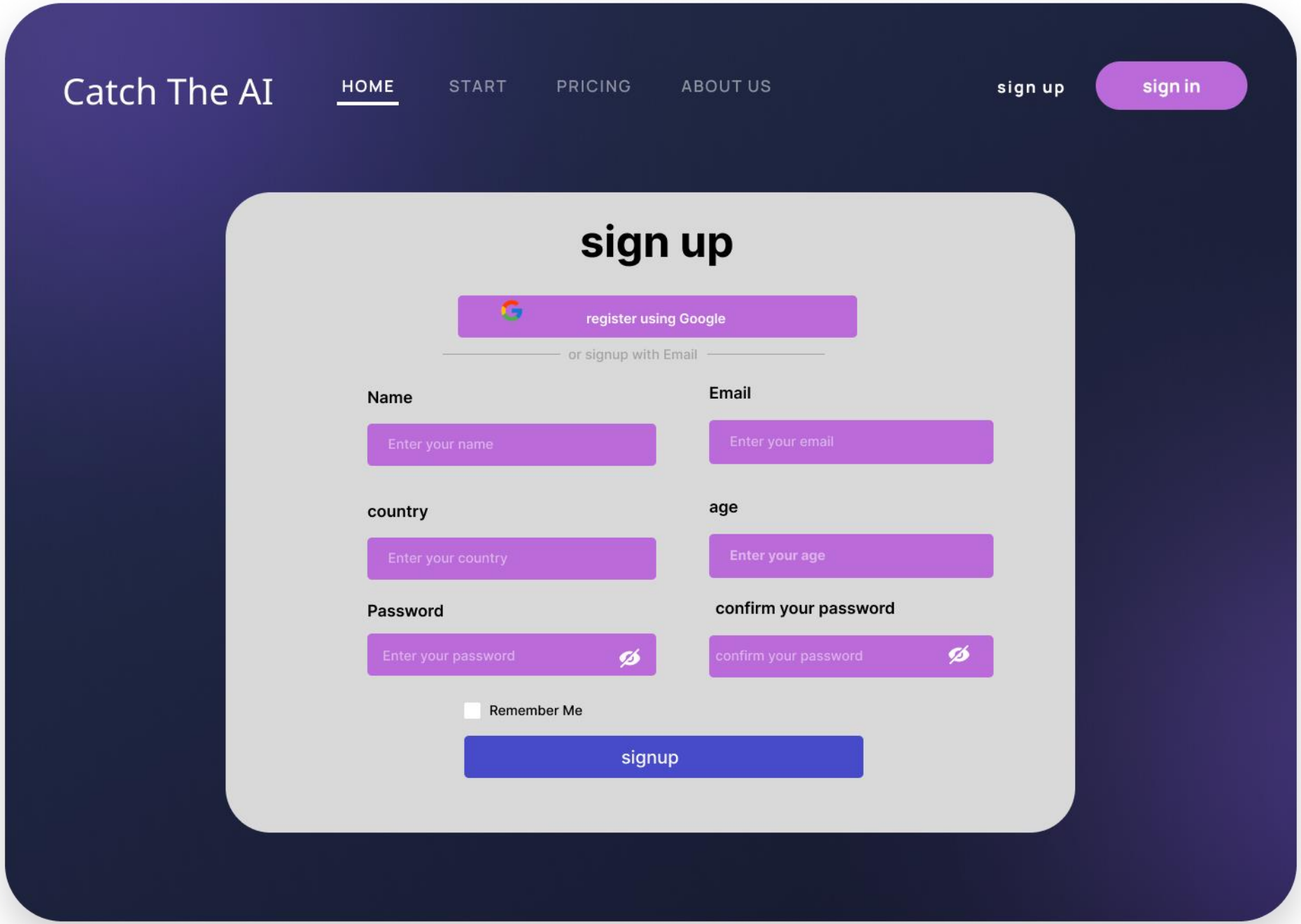
Swiftly and accurately detect AI-generated content with our advanced models. Our deep learning technology distinguishes between AI and human-authored media in images and text.

Get Started

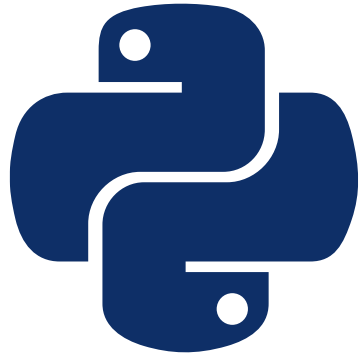

Catch The AI
HOME
START
PRICING
ABOUT US
sign up
sign in
Select Media Type
IMAGES
TEXT
Drag And Drop
or
Upload your image
maximum size 10 Mb
AI or Human ?


Catch The AI
HOME
START
PRICING
ABOUT US
sign up
sign in
Plans & Pricing
Whether you are individual person developer or enterprise, we have plan for you .
MONTHLY
YEARLY
$0 /month
basic
defualt features with limited acces
5 attempts per day
history of last 5 days
Choose plan
$15 /month
Professional
Advanced features more access.
20 attempts per day
history of last 30 days
file upload size up to 50Mb
faster response
Choose plan
MOST POPULAR
$50 /month
Company
the best value for your business
100 attempts per day
Unlimited history
Unlimited upload size
faster response
human support
Choose plan


Catch The AI
HOME
START
PRICING
ABOUT US
sign up
sign in
Login
Hi, Welcome back 👋
Login with Google
or Login with Email
Email
Enter your email id
Password
Enter your password
Remember Me
Forgot Password?
Login
Not registered yet? Create an account SignUp


Catch The AI
HOME
START
PRICING
ABOUT US
sign up
sign in
sign up
register using Google
or signup with Email
Name
Enter your name
Email
Enter your email
country
Enter your country
age
Enter your age
Password
Enter your password
confirm your password
confirm your password
Remember Me
signup





Tools