AIFORGE: A UNIFIED SAAS PLATFORM FOR MULTI-MODEL AI SERVICES

[1] Dr. Shudhodhan Bokefode, [2] Shubham Thange, [3] Swapnil Naik, [4] Kanad Zadpe, [5] Vrushabh Kamble

[1]Associate Professor, [2,3,4,5]Student,
[1,2,3,4,5]Computer Engineering Department,
[1,2,3,4,5]Terna Engineering College, Navi Mumbai, India

Abstract : The rapid advancement of Artificial Intelligence has created a demand for accessible and cost-effective tools that can be used by individuals and businesses. However, most AI solutions are fragmented, complex, and require technical expertise. This paper presents AIForge, an intelligent Software as a Service (SaaS) platform that integrates multiple AI functionalities such as text generation, image generation, code generation, and resume analysis into a single unified system. The platform is designed with a user-friendly interface that enables non-technical users to leverage powerful AI capabilities without requiring specialized knowledge. The system uses modern web technologies and API- based AI integration to deliver real-time results. The resume analysis module achieves an accuracy of 76.46% with a macro F1 score of 66.14%, demonstrating effective performance across multiple job categories. AIForge aims to reduce cost, improve efficiency, and simplify access to AI tools through a centralized platform.

Keywords: Artificial Intelligence, SaaS Platform, Multi-Model AI, Text Generation, Code Generation, Image Generation, Resume Analysis, Machine Learning, AIaaS, Natural Language Processing

INTRODUCTION

Artificial Intelligence (AI) is transforming various industries by enabling automation, intelligent decision-making, and advanced content generation. From education and software development to business automation, AI-driven tools are increasingly being adopted to improve efficiency and productivity. Technologies such as natural language processing, computer vision, and machine learning have made it possible to generate human-like text, create images, write code, and analyze complex data.

Despite its widespread adoption, most AI tools are distributed across multiple platforms and require technical expertise to operate. Users often need to switch between different services for tasks such as generating text, creating images, writing code, or analyzing resumes. This fragmentation leads to inefficiency, increased cost, and a poor user experience. Additionally, managing multiple subscriptions and interfaces creates complexity, especially for non-technical users.

To address these challenges, AIForge is proposed as a unified Software-as-a-Service (SaaS) platform that combines multiple AI capabilities into a single system. The platform is designed to provide a seamless and user-friendly experience by allowing users to access different AI functionalities through a centralized interface. By integrating various AI models and services, AIForge simplifies the process of using AI and reduces the dependency on multiple tools.

NEED OF THE STUDY.

The rapid growth of Artificial Intelligence technologies has led to the emergence of numerous AI tools and platforms, each focusing on specific functionalities such as text generation, image creation, or code assistance. However, these tools are often distributed across different platforms, requiring users to manage multiple applications, subscriptions, and interfaces. This fragmentation increases operational complexity, cost, and time consumption, especially for users who require multiple AI services simultaneously. Moreover, many existing AI platforms are designed primarily for technical users and lack intuitive interfaces for beginners or non-technical individuals. This creates a barrier to entry and limits the accessibility of advanced AI capabilities for a broader audience. In addition, the absence of integration between different AI services leads to inefficient workflows and reduced productivity.
Therefore, there is a strong need for a unified, cost-effective, and user-friendly platform that integrates multiple AI functionalities into a single system. AIForge addresses this need by providing a centralized solution that simplifies access to AI tools, reduces dependency on multiple platforms, and enhances overall user experience and efficiency.

LITERATURE SURVEY

Artificial Intelligence has significantly evolved in recent years, enabling automation and intelligent content generation across multiple domains such as text, image, and code generation. Several research studies have contributed to the development of AI-based systems in these areas. This section reviews important research works along with their methodologies, advantages, limitations, and their relevance to the proposed AIForge system.

1. **Text Generation Using LSTM:** I. Dhall and S. Vashisth proposed the use of Long Short-Term Memory (LSTM) networks to improve text generation. Traditional Recurrent Neural Networks (RNNs) suffer from vanishing gradient problems, which limit their ability to learn long-term dependencies. LSTM overcomes this issue by introducing memory cells that can retain information over long sequences. The model demonstrated improved coherence and context awareness. However, it requires high

computational resources and is not efficient for large-scale real-time applications. *Impact on AIForge:* This work highlights the importance of sequence modeling in text generation, which is further enhanced in AIForge using advanced transformer-based models.

2. **Text-to-Image Generation Using Deep Learning:** S. Ramzan and M. M. Iqbal explored deep learning techniques such as RC-GAN for generating images from textual descriptions. The system combines text encoding and image generation models to produce meaningful visuals. Despite its effectiveness, it struggles with fine details and requires large training datasets. *Impact on AIForge:* This research forms the foundation for integrating image generation in AIForge using more advanced models like Stable Diffusion.
3. **Code Generation Using Machine Learning:** K. Lano and Q. Xue proposed a method for automatic code generation using symbolic machine learning techniques. This approach reduces manual coding effort and improves productivity. However, its performance depends on input quality and struggles with complex logic. *Impact on AIForge:* AIForge improves upon this by using modern large language models that generate more accurate and context-aware code.
4. **AI-Based Image Generation in Design:** E. A. Setiawan studied the use of AI tools such as DALL·E and Midjourney in design applications. These tools enhance creativity and visualization but raise concerns regarding originality and feasibility. *Impact on AIForge:* This motivates the inclusion of image generation features in AIForge for creative and practical use cases.
5. **CLIP-Based Image Generation Models:** R. Ganz introduced CLIPAG, which improves text-to-image generation by aligning visual and textual representations. This enhances image quality but requires high computational power. *Impact on AIForge:* AIForge leverages similar advancements through Stable Diffusion to generate high-quality images efficiently.
6. **Generative AI for Multi-Domain Applications:** P. Zhou and L. Wang explored integrating generative AI across domains such as text, image, and code generation. Their work highlights the need for unified systems, but most implementations remain fragmented. *Impact on AIForge:* This directly motivates AIForge, which integrates multiple AI functionalities into a single platform.
7. **Artificial Intelligence as a Service (AIaaS):** Recent studies in ACM Computing Surveys describe AIaaS as a model where AI services are delivered via cloud-based platforms, enabling accessibility without requiring deep technical expertise. While the study provides a strong theoretical framework, it lacks practical implementation. *Impact on AIForge:* AIForge serves as a real-world implementation of the AIaaS concept by offering multiple AI services through a unified web platform.
8. **AIaaS for Business and Information Systems:** Frick et al. (2021) discussed AIaaS as a combination of cloud computing and AI to democratize access for users regardless of expertise. The study emphasizes enterprise-level platforms such as AWS and Google but highlights the lack of user-centric solutions. *Impact on AIForge:* AIForge addresses this gap by providing a lightweight, user-friendly AI platform tailored for students and developers.

Summary: Most existing platforms such as ChatGPT and Poe focus on individual AI functionalities and lack integration of multiple services in a single system. Additionally, they do not provide domain-specific features such as resume analysis. This highlights the gap addressed by AIForge, which integrates multiple AI capabilities into one unified platform.

From the above studies, it is evident that while significant progress has been made in individual AI domains, most systems focus on a single functionality and lack integration. Additionally, issues such as high computational cost, poor scalability, and lack of user-friendly interfaces are common. The concept of AIaaS provides a theoretical solution, but practical implementations remain limited. AIForge addresses these gaps by integrating multiple AI services into a single, scalable, and user-friendly platform.

PROBLEM STATEMENT

Existing AI tools are fragmented and require users to interact with multiple platforms to perform different tasks such as text generation, image creation, code generation, and resume analysis. This fragmentation forces users to switch between different applications, leading to increased complexity, higher operational costs, and significant time consumption. Managing multiple tools and subscriptions also creates inconvenience, especially for users who need access to several AI functionalities simultaneously.

Additionally, many AI systems are designed with a focus on technical users and require prior knowledge of programming or AI concepts. This makes them less accessible to non-technical users, limiting the widespread adoption of advanced AI technologies. The absence of a unified and intuitive interface further results in inefficient workflows, reduced productivity, and a poor overall user experience.

Moreover, the lack of integration between different AI services leads to inconsistency in outputs and difficulty in comparing results across platforms. This reduces the effectiveness of decision-making and prevents users from fully utilizing the potential of AI technologies.

Therefore, there is a need for a unified platform that simplifies access to multiple AI functionalities while maintaining efficiency, usability, scalability, and cost-effectiveness. Such a platform should provide a centralized and user-friendly environment that integrates diverse AI capabilities into a single system.

PROPOSED SYSTEM

AIForge is designed as a unified Software-as-a-Service (SaaS) platform that integrates multiple AI services into a single application. The system allows users to register and log in to access a centralized interface, where they can select modules such as text generation, image generation, code generation, and resume analysis.

After selecting a module, the user provides input in the form of a prompt or uploads relevant data. This input is sent to the backend through API requests. The backend processes the request using AI models such as Mistral and Stable Diffusion. The generated output is then returned to the frontend and displayed in real time.

Users can save or download the generated results, and all data is securely stored using PostgreSQL to ensure scalability and efficient data management.

This workflow ensures efficient processing, ease of use, and seamless integration of multiple AI functionalities within a single platform.

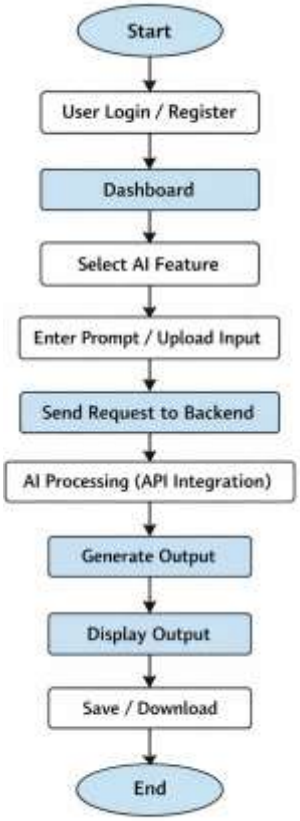

Fig. 1. Flowchart of AIForge SaaS Platform Workflow

SYSTEM ARCHITECTURE

AIForge system architecture is designed as a modular and scalable framework consisting of multiple interconnected components that work together to efficiently process user requests and generate outputs in real time. The architecture follows a client-server model, ensuring a clear separation between the frontend, backend, AI processing layer, and data storage. The frontend is developed using React and Vite, providing a responsive and interactive user interface that allows users to access multiple AI functionalities such as text generation, code generation, image generation, and resume analysis through a centralized interface. The backend is implemented using FastAPI, which handles API requests, manages authentication, processes user inputs, and routes them to appropriate AI services, ensuring secure and efficient communication between components. The AI processing layer integrates multiple AI models and external APIs to perform various tasks, including NVIDIA-hosted models for text and code generation and Stable Diffusion for image generation. In addition, the resume analysis module is implemented using a locally trained machine learning pipeline that includes TF-IDF feature extraction, Logistic Regression for career path prediction, Random Forest for ATS score estimation, and a DistilBERT-based classifier for improved semantic understanding and accuracy. The database layer uses PostgreSQL to store user data, authentication details, token usage, purchase history, and generated outputs, ensuring scalability, consistency, and efficient data management. The system also maintains a history module that allows users to access previously generated results. Furthermore, a billing module integrated with Stripe is used for managing token-based usage and payments, where users receive initial free tokens and can purchase additional tokens as required. Overall, this architecture ensures efficient processing, seamless integration of multiple AI functionalities, scalability, maintainability, and an enhanced user experience within a unified platform.

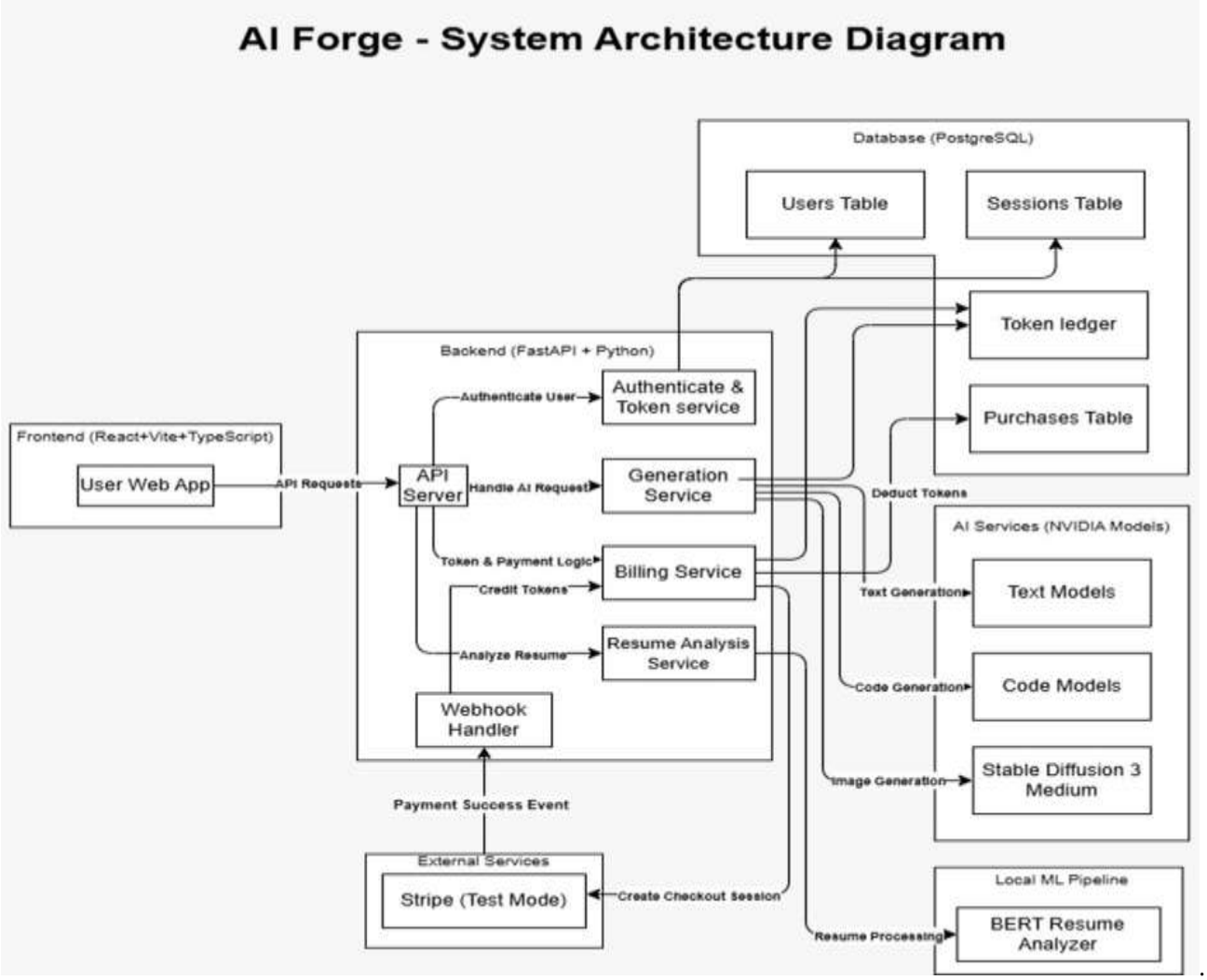

Fig. 2. Architecture of AIForge SaaS Platform

MODEL DETAILS AND IMPLEMENTATION

The AIForge platform integrates multiple state-of-the-art AI models and machine learning techniques to provide diverse functionalities such as text generation, code generation, image generation, and resume analysis.

A. Text Generation Models

The system utilizes multiple NVIDIA-hosted large language models to generate high-quality text outputs. These include Qwen 2.5 7B, Mistral Nemotron, and Seed OSS 36B models. The platform allows users to either compare outputs from multiple models or use a single model for generating context aware and coherent text.

B. Code Generation Models

For code generation tasks, the system employs specialized NVIDIA-hosted coding models such as Qwen2.5 Coder 7B, Mamba Codestral 7B, and Qwen3 Coder 480B. These models are capable of generating syntactically correct and optimized code for various programming tasks.

C. Image Generation Model

The image generation module is powered by Stable Diffusion 3 Medium, which generates high-quality images from textual prompts. This enables users to create visual content dynamically.

D. Resume Analysis Model

The resume analysis module is implemented using a hybrid machine learning pipeline. It combines traditional machine learning techniques and transformer-based models for better accuracy.

The pipeline includes TF-IDF for feature extraction, Logistic Regression for career path prediction, Random Forest for ATS score estimation, and a DistilBERT-based classifier for enhanced semantic understanding. The model is trained on a resume dataset consisting of 2484 samples with 24 labels, achieving an accuracy of 76.46% and a macro F1 score of 66.14%.

TABLE I
RESUME ANALYSIS MODEL PERFORMANCE

Parameter	Value
Dataset Source	Kaggle Resume Dataset (Snehaan Bhawal)
Number of Samples	2484
Number of Classes	24
Train/Test Split	1987 / 497
Model Used	DistilBERT
Accuracy	76.46%
Macro F1 Score	66.14%
Weighted F1 Score	72.75%

Table I summarizes the dataset characteristics and performance metrics of the resume analysis model used in the system.

CONFUSION MATRIX ANALYSIS:

The confusion matrix provides a detailed evaluation of classification performance across all job categories. It shows correct predictions along the diagonal and misclassifications in off-diagonal elements.
The matrix indicates that the model performs well for major categories such as HR, Engineering, and Information Technology, while some misclassifications occur in closely related job roles due to class imbalance in the dataset.

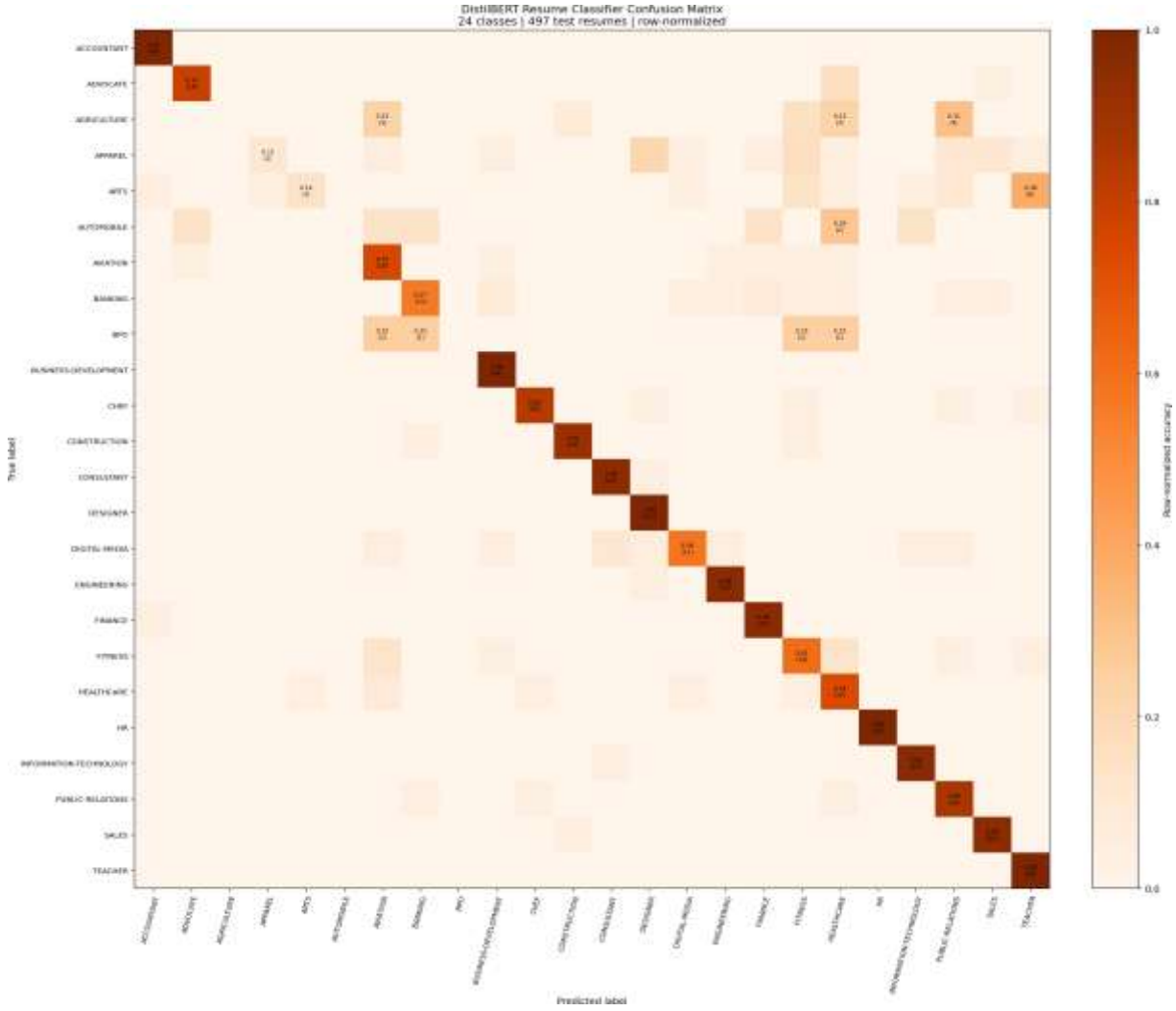

Fig. 3. Confusion Matrix for Resume Classification Model

E. Database and Storage

The system uses PostgreSQL as the primary database to store user data, authentication details, token usage, purchase history, and generated outputs.

F. Billing Integration

The platform integrates Stripe for handling payments and token purchases. Users receive free trial tokens upon registration and can purchase additional tokens using the billing system.

G. Model Selection Justification

The selection of models in AIForge is based on a balance between performance, response quality, speed, and computational cost.

TABLE II
COMPARISON OF SELECTED TEXT GENERATION MODELS

Model	Accuracy	Speed	Cost
Qwen 2.5 7B	High	Fast	Medium
Mistral Nemotron	Very High	Medium	Medium
Seed OSS 36B	High	Slow	High

The comparison shows that Mistral Nemotron provides the highest accuracy and better reasoning capabilities, making it suitable for high-quality text generation. Qwen 2.5 offers balanced performance with faster response time, while Seed OSS 36B is useful for generating detailed long-form content. Therefore, multiple models are selected to balance performance, speed, and output quality.

TABLE III
COMPARISON OF AIFORGE WITH EXISTING PLATFORMS

Feature	AIForge	ChatGPT	Poe
Multi-Model Comparison	Yes	No	Yes
Text Generation	Yes	Yes	Yes
Code Generation	Yes	Yes	Yes
Image Generation	Yes	Limited	Limited
Resume Analysis	Yes	No	No
Token System	Yes	No	Yes
Unified Platform	Yes	No	No

AIForge provides a more comprehensive and unified platform compared to existing systems such as ChatGPT and Poe. While these platforms offer strong capabilities in individual tasks, AIForge integrates multiple functionalities including resume analysis and multi-model comparison within a single system. This unified approach enhances usability, reduces dependency on multiple tools, and improves overall user productivity.

SYSTEM FEATURES

AIForge provides a wide range of features designed to enhance user productivity and simplify access to AI tools.

- **Text Generation:** Supports writing-based prompts with the ability to compare multiple AI models or use a single model.
- **Code Generation:** Generates optimized code outputs and allows comparison between different coding models.
- **Image Generation:** Produces images from textual descriptions using advanced AI models.
- **Resume Analysis:** Provides ATS scores, career path predictions, strengths, and improvement suggestions.
- **Billing System:** Offers free trial tokens and supports token purchases using Stripe.
- **History Management:** Stores all user-generated outputs for future access and reuse.

ADVANTAGES AND DISADVANTAGES

ADVANTAGES:

- Provides multiple AI tools such as text, image, code generation, and resume analysis in a single unified platform
- Reduces overall cost by eliminating the need for multiple subscriptions to different AI services
- User-friendly interface that allows even non-technical users to easily interact with advanced AI features
- Real-time output generation improves efficiency and saves time for users
- Scalable and modular architecture that can support additional AI functionalities in the future
- Centralized dashboard for managing all AI tasks and outputs in one place
- Enhances productivity for students, developers, and professionals by automating repetitive tasks

DISADVANTAGES:

- Requires a stable internet connection as the system relies on cloud-based AI services
- Performance and response time depend on external AI APIs and server availability
- Limited customization options for advanced users who require fine-grained control over AI models
- Potential latency issues during high traffic or heavy processing tasks

APPLICATIONS

1. **Education:**
Students can use AIForge for content generation, assignment assistance, and understanding concepts through AI-generated explanations.
2. **Software Development:**
Developers can utilize the code generation module for writing, debugging, and optimizing code, thereby improving development efficiency.
3. **Business Automation:**
Organizations can automate tasks such as content creation, resume screening, and data analysis, reducing manual effort and increasing productivity.
4. **Freelancing and Content Creation:**
Freelancers and content creators can generate high-quality text and images for blogs, social media, and marketing content.
5. **Resume Analysis and Recruitment:**
The platform can assist in evaluating resumes, providing ATS scores, and suggesting career paths, which is useful for both job seekers and recruiters.
6. **General Productivity Enhancement:**
Users can access multiple AI tools within a single platform, improving workflow efficiency and reducing the need for multiple applications.

RESULTS AND OUTPUT

The system successfully generates outputs across multiple modules. The platform provides functionalities such as text generation, image generation, code generation, and resume analysis. Each module is designed to deliver accurate and context aware results based on user input.

The outputs are generated in real-time, demonstrating the efficiency and effectiveness of the system. The platform ensures a smooth and responsive user experience, allowing users to interact with different AI features seamlessly. Additionally, the system maintains consistency and reliability in output generation across all modules, making it suitable for practical and real-world applications. Furthermore, the system is scalable and can be extended to support additional AI functionalities in future enhancements.

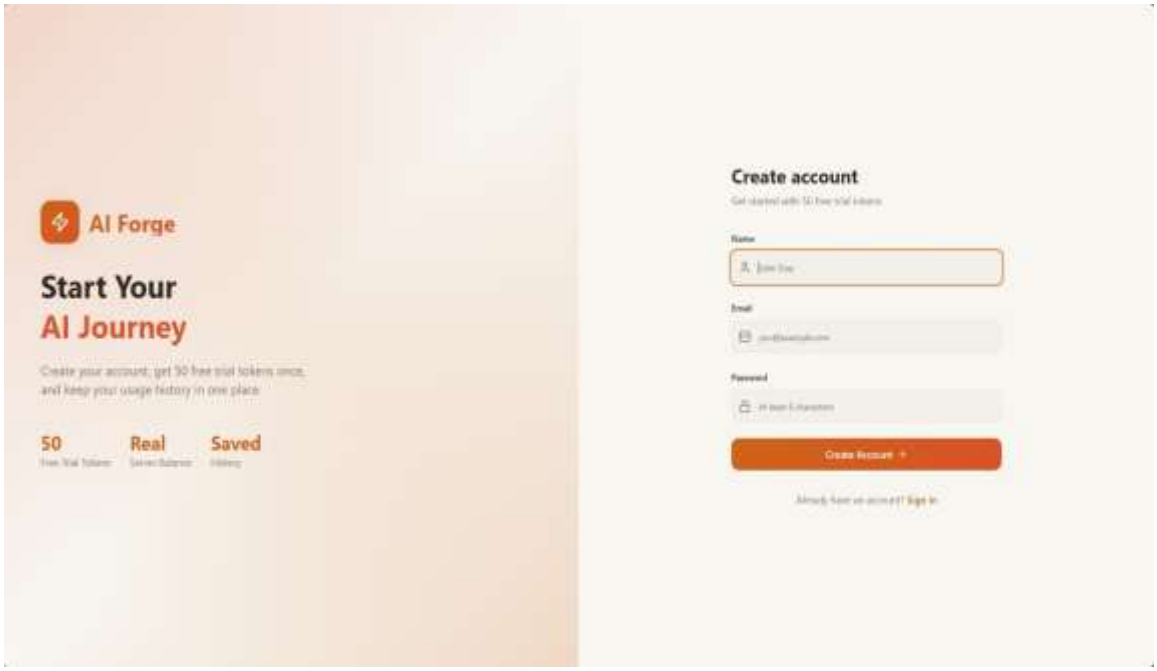

Fig. 4. User Registration Interface

The above figure shows the registration page where users can create an account to access AIForge services.

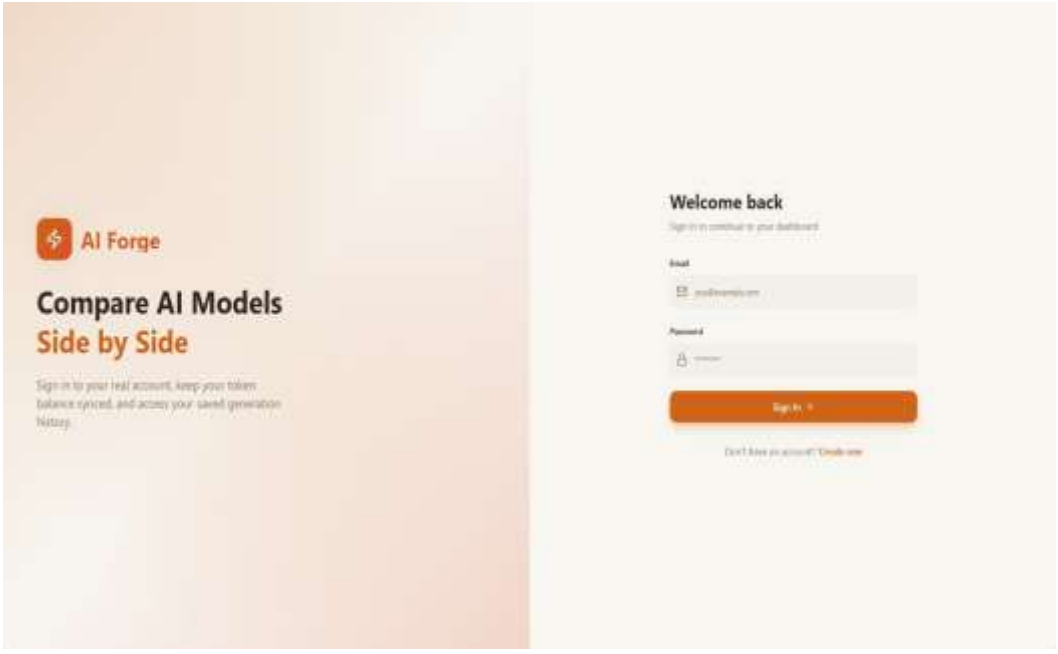

Fig. 5. User Login Interface

This interface allows users to securely log in and access all available AI modules.

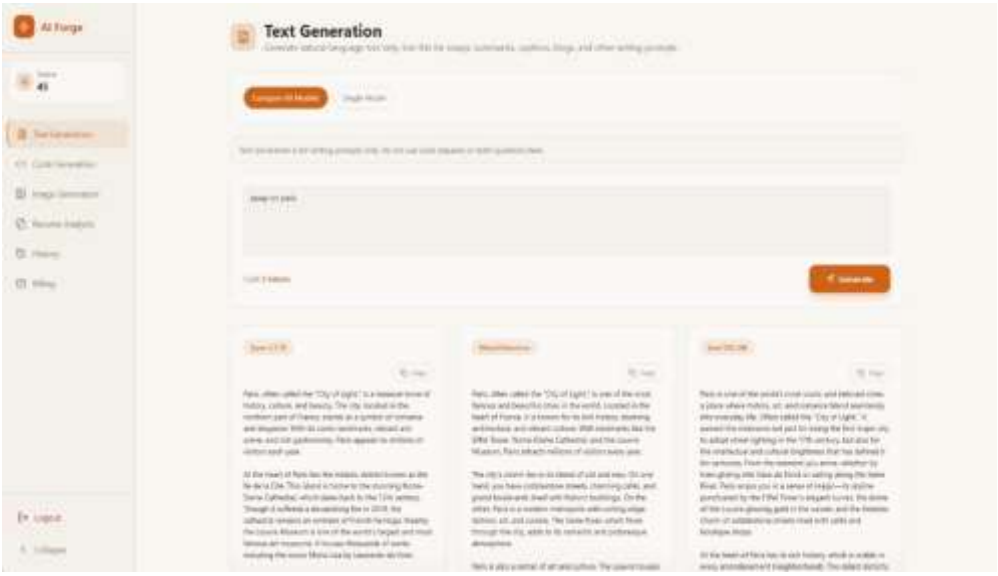

Fig. 6. Text Generation Output

The system generates meaningful and context-aware text responses based on user input prompts. It ensures coherence, relevance, and grammatical accuracy in the generated content.

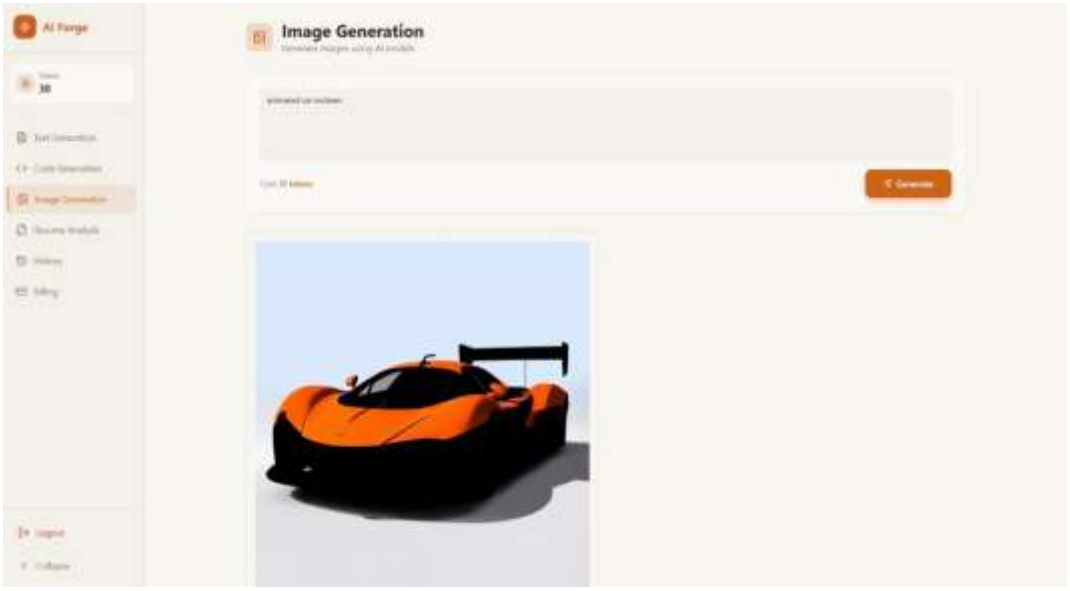

Fig. 7. Image Generation Prompt

The image generation module creates visuals from textual descriptions using AI models.

Fig. 8. Code Generation Output

This module generates functional and optimized code for different programming requirements.

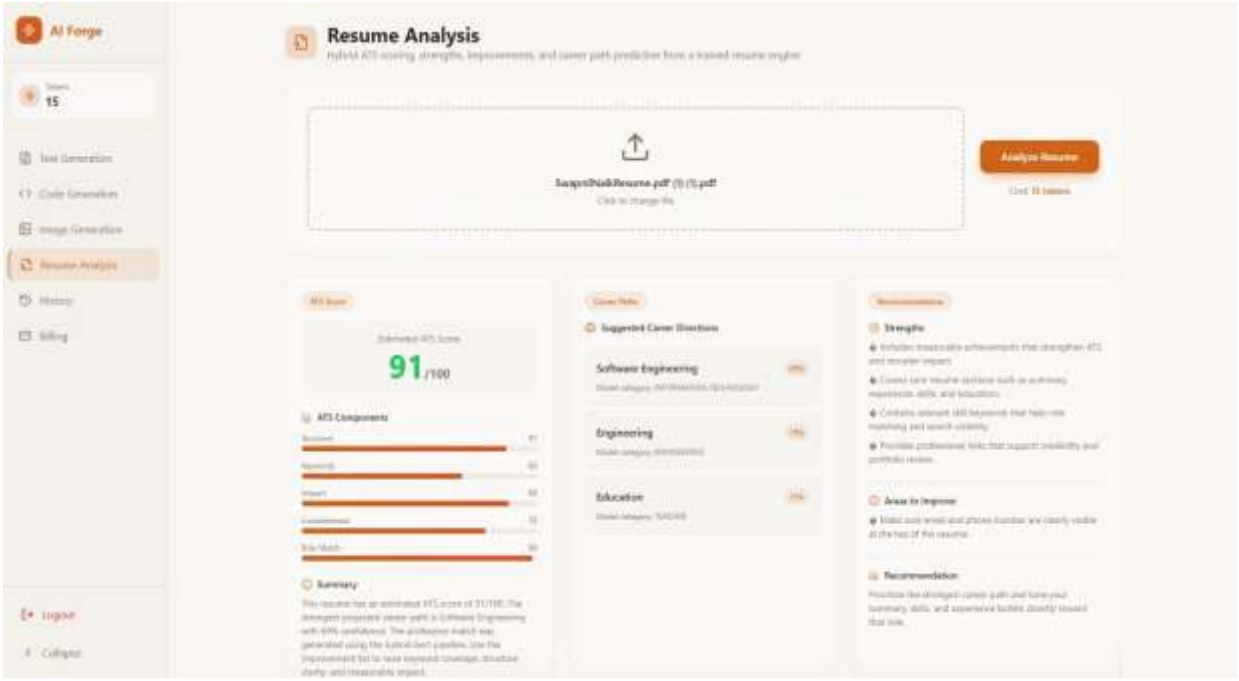

Fig. 9. Resume Analysis Result

The system analyzes resumes and provides ATS scores along with improvement suggestions.

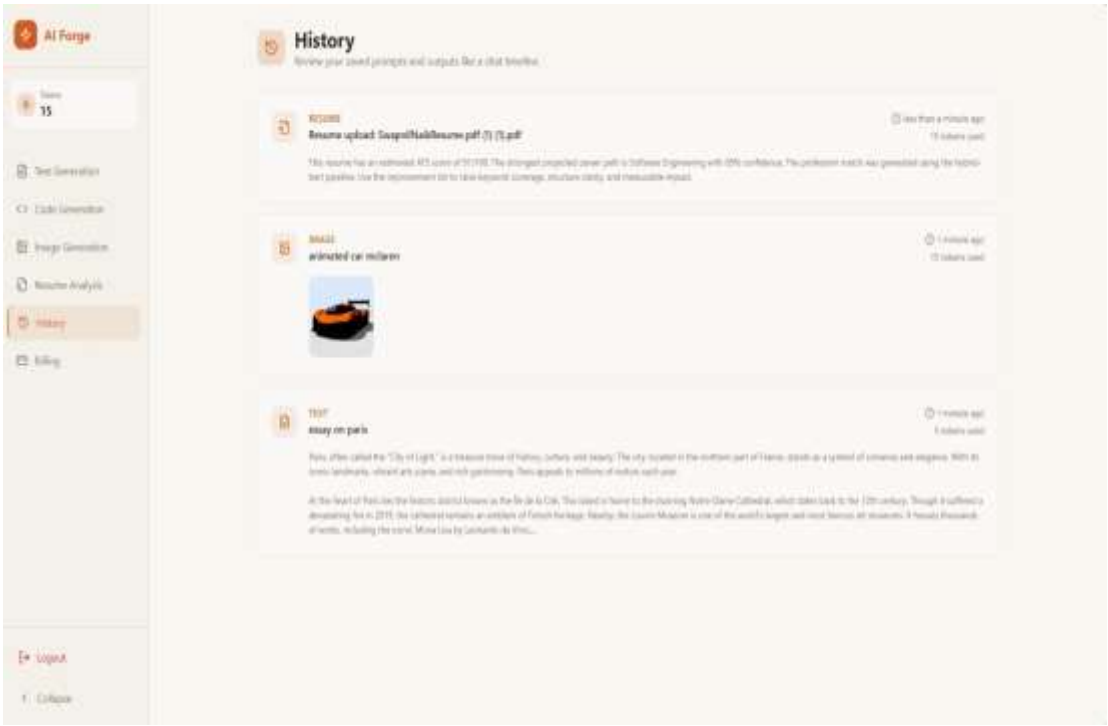

Fig. 10. User History Interface

The history module stores previously generated outputs for easy access and reuse.

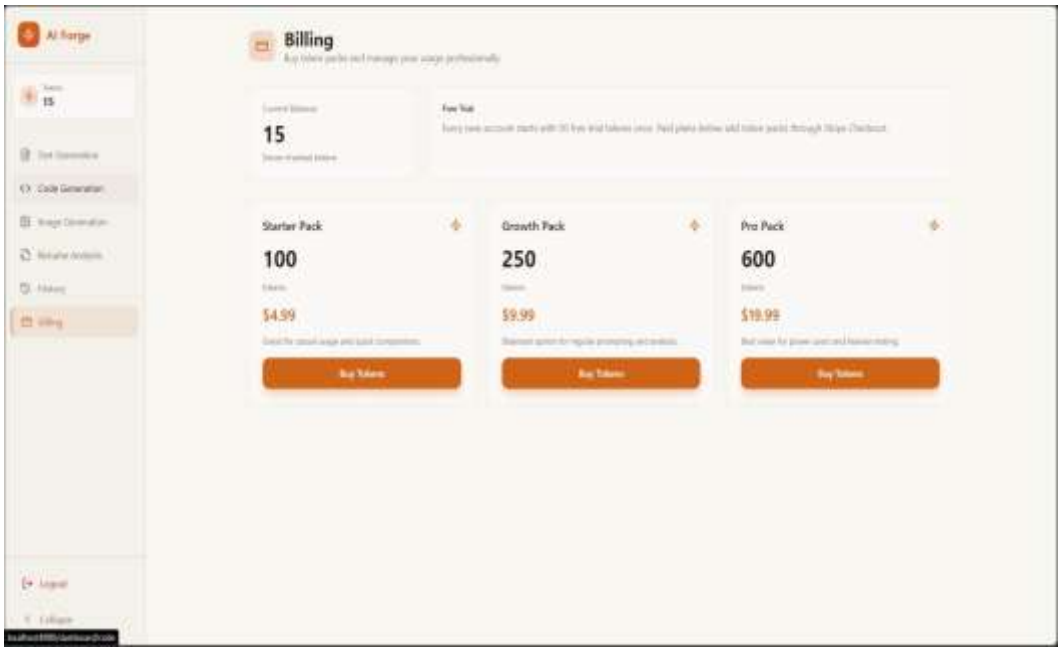

Fig. 11. Subscription and Billing Interface

This module allows users to purchase tokens and manage their subscription plans.

CONCLUSION

AIForge provides a unified and efficient solution for accessing multiple AI tools within a single platform. By integrating functionalities such as text generation, image generation, code generation, and resume analysis, the system eliminates the need for multiple standalone applications. This integration significantly improves usability, reduces operational complexity, minimizes cost, and enhances overall productivity for users across different domains.

The system demonstrates strong performance by generating accurate and real-time outputs across various modules. The inclusion of multiple AI models allows users to compare results and select the most suitable output based on their requirements, thereby improving decision-making and flexibility. Additionally, the resume analysis module, supported by a hybrid machine learning pipeline, achieves reliable performance and provides meaningful insights for career guidance.

AIForge is designed with a user-friendly interface, making it accessible to both technical and non-technical users. The modular and scalable architecture ensures efficient processing, easy maintenance, and the ability to integrate additional AI functionalities in the future. Furthermore, the use of a centralized platform simplifies workflows and enhances the overall user experience.

In conclusion, AIForge highlights the potential of combining multiple AI technologies into a single scalable and accessible platform. It contributes towards making AI more practical, efficient, and widely usable for real-world applications across education, software development, business automation, and other domains.

FUTURE SCOPE

AIForge offers significant potential for future enhancements and expansion. One of the key directions includes the integration of voice-based interaction, which can enable more natural, intuitive, and interactive user experiences. By incorporating speech recognition and voice command capabilities, the platform can become more accessible and user-friendly for a wider audience.

The system can also be improved by integrating more advanced AI models to enhance the accuracy, quality, and efficiency of generated outputs. Continuous model updates and optimization can further improve performance across text, image, and code generation modules. Additionally, the platform can be extended with personalized recommendation systems and intelligent analytics features to provide users with more customized and insightful results.

Further developments may include the implementation of collaborative features that allow multiple users to work on projects simultaneously, improving teamwork and productivity. The system can also be expanded to support additional AI functionalities such as video generation, speech processing, and advanced data analysis, making it more versatile and comprehensive.

Moreover, future enhancements can focus on improving scalability, security, and real-time performance to support a larger user base. With these advancements, AIForge has the potential to evolve into a more powerful, scalable, and intelligent platform, capable of addressing a broader range of real-world applications across various industries.

REFERENCES

[1] I. Dhall and S. Vashisth, "Text Generation Using Long Short-Term Memory Networks," Springer, 2020.

[2] S. Ramzan and M. M. Iqbal, "Text-to-Image Generation Using Deep Learning," Engineering Proceedings, 2022.

[3] K. Lano and Q. Xue, "Code Generation by Example Using Symbolic Machine Learning," SN Computer Science, 2023.

[4] N. Yadav, "Generation of Images from Text Using AI," International Journal of Engineering and Management, 2024.

[5] E. A. Setiawan, "A Review of AI Image Generator in Architecture," Journal of Architecture, 2023.

[6] R. Ganz, "CLIPAG: Towards Generator-Free Text-to-Image Generation," IEEE CVPR Workshops, 2022.

[7] "Artificial Intelligence as a Service (AIaaS) for Cloud, Fog and the Edge: State-of-the-Art Practices," ACM Computing Surveys, 2025.

[8] N. Frick et al., "Artificial Intelligence as a Service," Business & Information Systems Engineering, Springer, 2021.