TRS TuoTian LLM Integration Platform SPD

1. Product Overview

(1) Product name

TRS TuoTian LLM Integration Platform.

(2) Development companies

拓爾思資訊技術股份有限公司（TRS Information Technology Co.,Ltd）

(3) Product introduction

TRS TuoTian LLM Integration Platform is an intelligent application development platform built based on advanced large model technology, aiming to provide users with efficient and convenient artificial intelligence Q&A and writing application development solutions. By integrating core functions such as knowledge base management, instruction base construction, and scenario customization, the platform can quickly create personalized AI applications, support meeting minutes and AI summarization, document management and processing, writing and content creation, intelligent customer service, data analysis, etc., to meet the intelligent needs of various government fields and scenarios.

Second, the product function module

(1) Knowledge base management

1. Knowledge base type

- Public knowledge base: A knowledge base that can be used by multiple tenants and supports multiple industry classifications.
- Tenant knowledge base: A knowledge base that is only available to the current tenant for more personalized knowledge management.

2. Knowledge base operation

- Create a new knowledge base: You can upload files in multiple formats (such as txt, doc, docx, pdf, etc.) and analyze data asynchronously in the background.
- Edit the knowledge base: You can modify the name, description, classification, and other information of the knowledge base.
- Deleting a knowledge base: After deletion, it cannot be restored.

• Enable/disable the knowledge base: After disabling the knowledge base, it cannot be used by the application, and you need to unbind it from the application first.

(2) Instruction library management

1. Instruction type

• Regular Commands: Manually created commands, which need to be customized.

• Data Command: Docking with the data center, the system automatically generates the instruction settings.

• Capability Command: Docking with the competence center, the system automatically generates command settings.

2. Command operation

• New Commands: Support for customizing command names, descriptions, prompts, etc.

• Edit Order: You can modify the details of the order.

• Delete command: After deletion, it cannot be restored, and you need to unbind it from the app first.

• Enable/Disable Commands: Commands cannot be used by the app when disabled.

• Experience Commands: Quickly verify the settings and effects of commands.

• Copy Instructions: 1:1 copy instructions, which can be modified and adjusted.

(3) Scene customization

1. Scene function

• Complex scene support: Define scenarios and orchestrate multiple instructions to meet the Q&A requirements in complex scenarios.

• Scene library management: easy to query, edit, enable/disable scenes.

2. Scene operation

• Create a scene: You can define the scene name, description, add commands, and arrange it.

• Edit Scene: Modify the scene's details and command arrangement.

• Deletion scenario: Deletion cannot be restored, so you need to be cautious.

• Enable/Disable Scenes: Scenes that are disabled will not be triggered during Q&A.

(4) Model tuning

1. Pre-training

• Corpus management: Support the import of corpus files in multiple formats for classification management.

• Training Task: You can initiate a pre-training task to specify the benchmark model and the generated model version.

2. Instruction fine-tuning

• Corpus management: Support importing Q&A data files for classification management.

• Training Task: You can initiate a command fine-tuning task to specify the benchmark model and the generated model version.

3. Intensive training

• Training data: Intensive training is performed based on the data annotated by the data operation system.

• Training Task: You can initiate an intensive training task to specify the benchmark model and the generated model version.

4. Model evaluation

• Evaluation indicators: Provide professional evaluation indicators such as GSM8K, CMMLU and C-Eval.

• Publish the model: The version of the model that has been evaluated can be published for use.

(5) Application management

1. Application type

• Q&A applications: Conversational interaction, one question and one answer. Q&A is used in scenarios such as customer service bots of enterprises or governments, semantic query of internal documents/knowledge of enterprises, and questions and queries of data analysis results.

• Writing apps: Users create an outline or article after formulating a writing topic and citing materials. Writing scenarios can be used in multiple scenarios, such as writing meeting minutes, creating news releases, generating summary reports, and summarizing and refining document reports.

2. Apply actions

• Create an app: You can set the app icon, name, model version, welcome message, guidance questions, etc.

• Edit App: You can modify the details and configuration of the app.

• Delete an app: You need to be cautious as you cannot restore an app after it is deleted.

• Publish/unpublish an app: Control whether the app is available on the front-end of the Internet.

• Manage apps: View Q&A records, statistics, and more for apps.

(6) Data operation and review

1. Data annotation

• Annotation function: Data operators can annotate Q&A data to improve the training effect of the model.

• Annotation status: Supports multiple annotation statuses such as "Correct Answer", "Incorrect Answer", and "To be discussed".

2. Data review

• Review function: Data reviewers review the labeled data to ensure data quality.

• Audit operations: Support operations such as "Review" and "Revoke Review".

3. Parameter debugging

• Debugging function: Support model parameter debugging and optimize the model output effect.

3. Product features

(1) Efficient development

• Simplified process: Through the all-in-one platform, users can quickly build intelligent applications without complex technical background.

• Rapid go-live: Enables rapid creation, configuration, and release of applications, shortening the development cycle.

(2) Intelligent interaction

• Natural conversations: AI models can have natural and fluent conversations with users to provide accurate answers.

• Contextual memory: Continuous questioning and contextual memory are supported to improve the interactive experience.

(3) Strong customization

• Personalized configuration: Support custom configuration of knowledge base, command library, and scenario to meet the personalized needs of different users.

• Industry adaptation: Support multi-industry classification, provide industry-specific knowledge base and instruction library, and adapt to different industry scenarios.

(4) Data security

• Multi-tenant isolation: The multi-tenant architecture is supported, and the data of each tenant is isolated independently to ensure data security.

• Sensitive word management: Provide sensitive word operation function to effectively block sensitive questions and answers.

Fourth, product advantages

(1) Technical advantages

• Advanced large model technology: Based on advanced large model technology, it provides powerful natural language processing capabilities.

• Efficient training algorithms: Support a variety of training tasks to improve the accuracy and adaptability of the model.

(2) User experience advantages

• Simple and easy to use: Provide a simple and intuitive operation interface, which is convenient for users to get started quickly.

• Smooth interaction: Support natural dialogue and contextual memory to provide a smooth interactive experience.

(3) Market advantage

• Multi-industry application: It supports multi-industry classification, adapts to different industry scenarios, and has a wide range of market application prospects.

• Flexible customization: Provide highly customized functions to meet the needs of different users.

5. Application scenarios

(1) Meeting minutes and artificial intelligence summary

Based on the integrated platform of Tuotian large model, it can develop applications related to meeting minutes and smart summaries, and realize the recording, translation, sorting and summary of meeting content.

1. Meeting minutes generation

• Automatic transcription and summarization: The model can transcribe the speech content of the meeting into text in real time, extract key information through natural language processing technology, and automatically generate structured meeting minutes. For example, in internal meetings, the platform can accurately record the key points discussed in the meeting, decision-making matters, and follow-up action plans, etc., reducing the burden of manual recording and improving meeting efficiency.

• Multi-language support and translation: Support multi-language meeting minutes generation and provide real-time translation capabilities. For multinational enterprises or conferences involving multilingual communication, the platform can translate the content of the meeting into the language required by the participants, breaking the language barrier and ensuring the accurate delivery of information.

• Intelligent recognition and annotation: It can identify different speakers in the meeting and automatically label the content of the speech, which is convenient for subsequent viewing and tracing. At the same time, the meeting minutes can also be classified and sorted according to the keywords and themes of the speech, so that users can quickly find the information they need.

2. Artificial intelligence summary

• Document abstract extraction: For long-form documents, such as research reports, news reports, etc., the platform can quickly extract the core content of the document and generate concise and clear summaries. Users don't need to read the entire document to quickly understand its main content and key messages.

• Video content summarization: In addition to text and audio content, the platform also supports summary extraction of video content. By analyzing the voice, subtitles, and screen content in the video, a video summary is generated to explain that the user

can quickly obtain the key information of the video.

- Customize the length of the summary: Users can set the length and level of detail of the summary according to their needs. The platform will generate a summary of the corresponding length according to the user's settings to meet the needs of different scenarios.

(2) Document management and processing

Based on the integrated platform of Topsky large model, it can develop document management and processing applications, realize the construction of enterprise knowledge base, and facilitate employees to quickly query and learn professional knowledge.

1. Intelligent classification and retrieval of documents

- Automatic classification: Using natural language understanding technology, the platform can automatically classify uploaded documents and classify them into corresponding categories based on their content, topic, and keywords. For example, classify contract documents into the "Legal" category and financial statements into the "Financial" category to make it easy for users to quickly find the documents they need.

- Intelligent search: Support document search by keyword, subject, author and other ways. Users only need to enter relevant search terms, and the platform can quickly return a list of matching documents and sort them according to relevance to help users quickly find the target documents.

- Knowledge Q&A: Based on the natural language understanding ability of large models, the semantic retrieval and summary of knowledge content in documents are realized by question-and-answer applications.

2. Document content extraction and analysis

- Key information extraction: Ability to extract key information from documents, such as date, amount, name, location, etc. For financial documents, financial data can be automatically extracted to generate a summary of financial statements; For contract documents, key information such as contract terms, rights and obligations of both parties can be extracted, so that users can quickly understand the core content of the document.

- Content analysis and statistics: Conduct in-depth analysis of document content and generate statistical reports. For example, analyze the frequency of keywords and topic distribution in the document to indicate that users understand the key content and trends of the document.

3. Document format conversion and editing

• Format Conversion: Support conversion between multiple document formats, such as converting PDF documents to Word documents, or converting Word documents to HTML format. Users do not need to install other software, and the document format conversion can be completed on the platform, which is convenient for further editing and processing of documents.

(3) Writing and content creation

1. Creative Writing Aid

• Writing topic generation: According to the keywords or topics entered by the user, the Tuotian model application can generate relevant writing topics and creative points. For example, if a user enters "Future World", the platform can generate multiple writing topics such as "Technological Development in the Future World" and "Human Life in the Future World" to stimulate the user's creative inspiration.

• Story outline creation: Help users quickly generate a story outline, including the cause, process, and result of the story. Users can further develop their creations according to the generated outline and improve their writing efficiency.

2. Article generation and optimization

• Article generation: Users only need to enter the topic or keywords of the article, and the Tuotian model application can generate a complete article. The generated articles are logical and coherent, and users can make further modifications and refinements as needed.

• Text optimization: Optimize user-written text, including grammar checking, vocabulary replacement, sentence polishing, etc. It shows that users can improve the quality and readability of the text, and make the article more smooth and natural.

3. Content Creation Collaboration

• Multi-person collaborative writing: Support multiple people to write online at the same time, and team members can edit, revise, and comment in the same document. The platform will save the modification record of the document in real time, which is convenient for team members to view and trace.

• Content moderation and feedback: Provide content moderation function, team members can review and feedback on text written by others. With online commenting and annotation, team members can easily exchange ideas and improve the quality of content creation.

(4) Intelligent customer service

Based on the integrated platform of Tuotian large model, it can quickly build intelligent customer service applications for Q&A to help enterprises and governments provide intelligent Q&A services based on large models.

1. Enterprise customer service

- Efficient question answering: Smart customer service can automatically answer customers' common questions, such as account information inquiry, password reset, order status tracking, etc., and solve 70% of the repetitive problems of traditional manual customer service. Through natural language processing technology, intelligent agents can understand the user's natural language input and provide accurate responses, whether spoken or written.

- Multi-round conversations and contextual understanding: Unlike simple auto-replies, intelligent agents can conduct multi-round conversations to understand and respond to more complex user needs. During conversations, intelligent agents are able to remember contextual information to provide more coherent and personalized service.

- Sentiment analysis and personalized recommendations: An advanced intelligent customer service system can analyze the user's sentiment and adjust the tone and content of the reply accordingly. At the same time, based on the user's historical interaction and behavior data, the intelligent customer service can provide personalized product or service recommendations.

- Full-channel integration: Smart customer service can be uniformly connected to WeChat, APP, web page, phone and other channels to build a cross-platform service loop and avoid users from repeatedly describing problems. Enterprises can use intelligent routing and allocation strategies to achieve optimal allocation of service resources.

- Business insight mining: The conversation data accumulated by intelligent customer service can be refined into customer demand heat maps and product improvement suggestions. For example, a retail brand analyzed high-frequency consulting words to identify logistics timeliness pain points and optimize the supply chain, resulting in an 18% reduction in return rates.

2. Government customer service

- Efficiently answer public consultations: Provide intelligent Q&A services for government departments, which can quickly answer public inquiries about policies and regulations, service procedures, service guidelines, etc., and improve service efficiency.

• Multi-language support: Support multi-language Q&A, facilitate the public with different language backgrounds to obtain information, and improve the quality of government services.

• Intelligent guidance and diversion: According to the content of public consultation, intelligent customer service can automatically guide users to the corresponding office or service window to reduce the waiting time of the public.

• Policy interpretation and publicity: Smart customer service can push the latest policy interpretation and publicity content in real time, indicating that the public has a better understanding of the policy.

(5) Data analysis

1. Enterprise data analysis and decision support

• Data integration and analysis: Tuotian model integration platform can integrate a variety of data sources within the enterprise, including ERP, CRM and other system data, to build an enterprise-level knowledge graph. Through intelligent analysis and visualization tools, it explains the data analysis and decision support of enterprises.

• Predictive analytics: Using statistical models and machine learning techniques, the platform can predict future market trends and business performance. For example, by analyzing historical sales data and market trends, the platform can predict future sales demand and help companies plan production and manage inventory in advance.

• Business Intelligence Reporting: Generate visual reports to help management understand the state of the business. The platform supports a variety of visualization components, and users can customize reports according to their own needs, and visually display data and analysis results.

2. Industry-specific data analysis

• Financial industry: In the financial industry, platforms can be used for risk assessment, investment analysis, and customer service. For example, by analyzing customers' trading behavior and credit history, the platform can detect abnormal transactions in real-time and reduce the risk of fraud.

• Manufacturing: In manufacturing, platforms can be used for equipment failure prediction, production process optimization, and supply chain management. For example, by analyzing equipment operating data and historical fault records, the platform can predict equipment failures and automatically dispatch work orders, reducing downtime.

• Healthcare: In the healthcare field, the platform can be used for disease

prediction, treatment recommendations, and patient management. For example, by analyzing patients' medical records and examination results, the platform can provide doctors with treatment recommendations, improving the efficiency of diagnosis and treatment.