Fall Detection and Speech Emotion Detection for elderly people living alone

Bug Killers
Report: Saniya Pandita
Generative AI Workflow Diagram: Jayadithya Nalajala
Application Interface Wireframe: Sai Jahnavi Devabhakthuni
LLM & Transformer Visualization: Hui Jin

Content

LLM & Transformer Visualization

This section shows the attention weights of the last layer of the Transformer, revealing how the model focuses on different words, such as love focuses on i and learning focuses on deep. It also uses t-SNE to visualize text embeddings, showing the semantic similarity of different sentences, such as LLMs like GPT-3 and Transformers are amazing for NLP are close in position, indicating that they are semantically related. These visualizations help explain the attention mechanism and embedding space distribution of LLM, which is suitable for text retrieval, sentiment analysis, and question-answering systems.

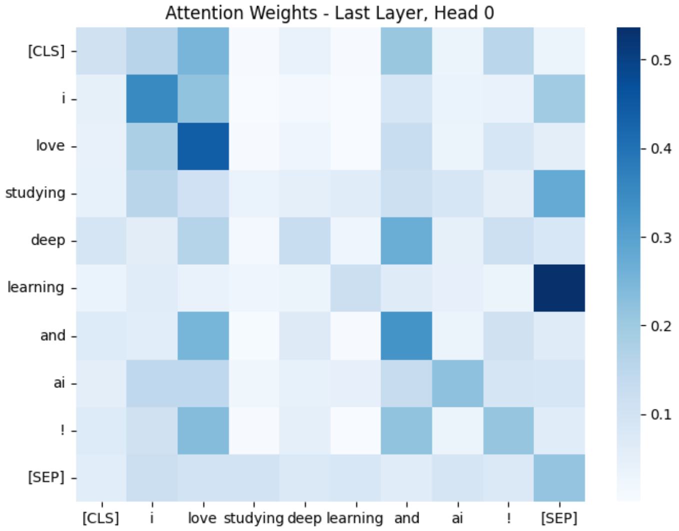


Figure 1: Attention Weights

As shown in Figure 1:

- This heatmap shows the attention weights of the last layer (Head 0) of the Transformer.
- The rows and columns correspond to different tokens (such as [CLS], i, love, deep learning).
- Dark blue represents a higher attention weight, indicating that the token receives more attention at a certain position.
- love has a high attention to i (dark blue block), indicating that the model believes that these two words are highly related.
- learning also has a high attention to deep, which is consistent with semantic logic.
- Special tokens such as [CLS] and [SEP] also have a certain degree of attention and are usually used for sentence-level feature representation.
- It can be used for sentiment classification (check which sentiment words the Transformer focuses on), text generation (analyze how the model chooses words), and reading comprehension (see which parts the model focuses on)

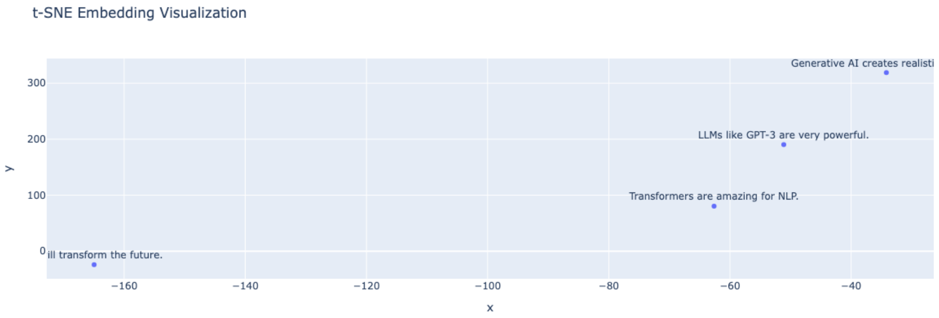


Figure 2: t-SNE Embedding Visualization

As shown in Figure 2:

- This scatter plot shows sentence embeddings after dimensionality reduction using t-SNE.
- Each dot represents a sentence, and similar sentences are closer in space.
- Generative AI creates realistic images is located in the upper right corner, far away from AI will transform the future., indicating that the model thinks they are semantically different.
- Transformers are amazing for NLP. and LLMs like GPT-3 are very powerful. are close to each other, indicating that the model thinks they are semantically related.
- AI will transform the future. is far away from other points, indicating that its semantic features may be unique.
- It can be used for text clustering (analyzing semantic similarity, such as news classification), information retrieval (RAG: Retrieval-Augmented Generation) and question answering systems (checking the similarity between user input and sentences in the database)

Generative AI Workflow Diagram

LoRA efficiently fine-tunes LLM by inserting a low-rank adaptation layer, reducing computational costs and making it suitable for specific tasks. RAG combines retrieval and generation, and uses embedding indexes such as FAISS to improve the factual answering ability of LLM. Stable Diffusion achieves high-quality text-to-image generation through CLIP text encoding, latent space mapping, and U-Net denoising.

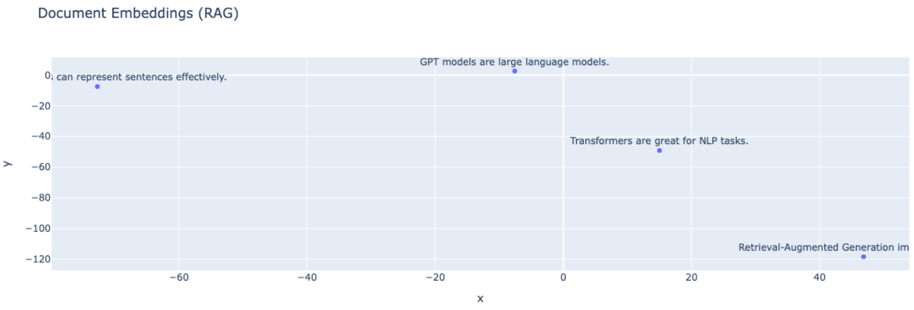


Figure 3: RAG workflow

As shown in Figure 3:

- Sentence embeddings are indexed and stored by FAISS/Pinecone.
- Retriever is responsible for finding the most relevant documents.
- Generator (such as T5, GPT-3) combines the retrieved information to generate accurate answers.
- It can be used for open-domain question answering (QA), legal, academic and financial document retrieval and improving the factual accuracy of LLM answers.

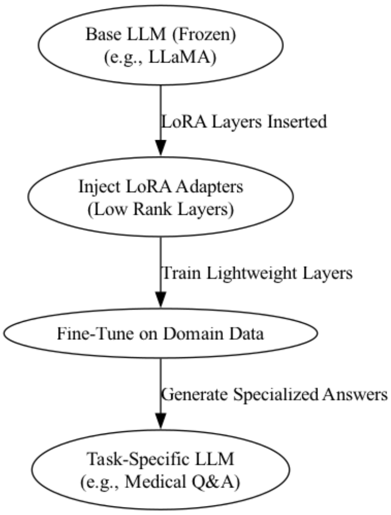
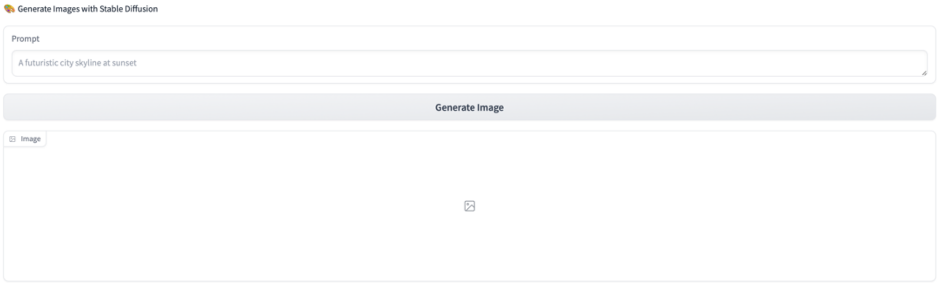


Figure 4: LoRA workflow

As shown in Figure 4:

- The basic LLM (Frozen) (such as LLaMA) remains unchanged, and only the LoRA adaptation layer (low-rank layer) is inserted.
- During training, only the LoRA adaptation layer is optimized to reduce computational overhead.
- The final result is a task-specific LLM (for example, for medical question answering).
- Can be used for domain adaptation (such as finance, medical NLP) and to reduce fine-tuning costs (more efficient than full LLM training)



A futuristic city skyline at sunset

Figure 5: Stable Diffusion Architecture

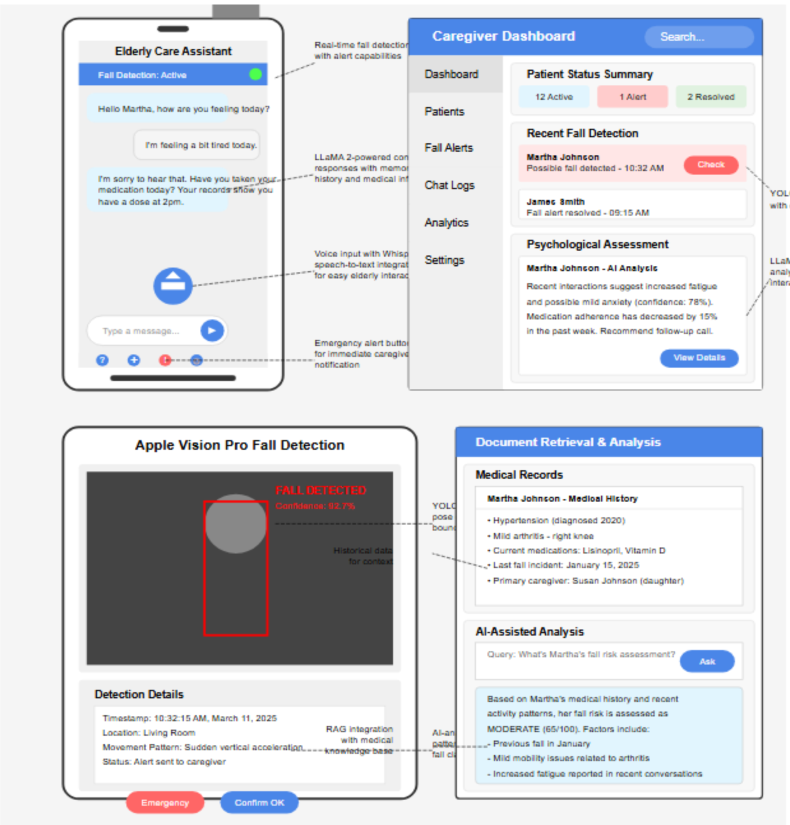
As shown in Figure 5:

- The text encoder (CLIP) converts the input prompt into a latent space vector.
- The latent space stores compressed representations to improve computational efficiency.
- The U-Net denoiser gradually generates clear images.
- The final output is a high-definition image.

- Can be used for text-to-image generation, AI art creation and personalized content generation

Application Interface Wireframe

The application interface wireframe illustrates the user interaction flow for elderly users and caregivers. It includes a messaging interface for text/voice input, a dashboard for psychological analysis and fall detection alerts, and app for real-time detection.



Hello Martha, how are you feeling today?

I'm feeling a bit tired today.

I'm sorry to hear that. Have you taken your medication today? Your records show you have a dose at 2pm.

Martha Johnson
Possible fall detected - 10:32 AM

James Smith
Fall alert resolved - 09:15 AM

Martha Johnson - AI Analysis

Recent interactions suggest increased fatigue and possible mild anxiety (confidence: 78%). Medication adherence has decreased by 15% in the past week. Recommend follow-up call.

Timestamp: 10:32:15 AM, March 11, 2025
Location: Living Room
Movement Pattern: Sudden vertical acceleration
Status: Alert sent to caregiver

Martha Johnson - Medical History

- Hypertension (diagnosed 2020)
- Mild arthritis - right knee
- Current medications: Lisinopril, Vitamin D
- Last fall incident: January 15, 2025
- Primary caregiver: Susan Johnson (daughter)

Query: What's Martha's fall risk assessment?

Based on Martha's medical history and recent activity patterns, her fall risk is assessed as MODERATE (65/100). Factors include:
- Previous fall in January
- Mild mobility issues related to arthritis
- Increased fatigue reported in recent conversations

Real-time fall detection with alert capabilities

LLaMA 2-powered con responses with memor history and medical inf

Voice input with Whisp speech-to-text integrat for easy elderly interac

Emergency alert butto for immediate caregive notification

YOLOv5 detection event with confidence scores

LLaMA 2 psychological analysis from chat interactions

YOLC pose bound

Historical data for context

RAG integration with medical knowledge base

AI-an patte fall cl

Figure 3: Application Interface Wireframe

As shown in Figure 3:

1. Core Screens

- Chat Interface (Elderly Care Assistant)
  - Allows voice and text-based interactions.
  - Provides AI-generated responses for psychological assessment.
  - Displays medication reminders and emergency alerts.
- Caregiver Dashboard
  - Patient Status Summary: Displays active, alert, and resolved statuses.
  - Recent Fall Detection: Logs detected falls with timestamps and confidence scores.
  - Psychological Assessment: Summarizes emotional state and medication adherence based on AI analysis.
- Apple Vision Pro Fall Detection Screen

  - Shows fall detection results with confidence scores.
  - Includes detection details such as location, timestamp, and movement analysis.
- Document Retrieval & Analysis
  - Medical Records: Retrieves patient history, medications, and caregiver details.
  - AI-Assisted Analysis: Generates risk assessment insights based on past records and patient interactions.

2. User Actions

- Submitting Prompts: Users interact via the chat interface to report conditions or ask about medications.
- Uploading Files: The system fetches and analyzes medical history and other stored records.
- Viewing Results:
  - Caregivers view fall alerts, psychological assessments, and AI-assisted risk analysis.
  - Users receive fall detection results and emergency alert options.

3. Outputs

- AI-Generated Text Responses:
  - LLaMA 2-powered responses for psychological assessment.
  - Sentiment-based analysis from chat interactions.
- Images:
  - Fall detection visual output with YOLOv5 confidence scores.
- Retrieved Documents:
  - Patient’s medical history and AI-generated risk assessments.

GitHub Repository Link

https://github.com/nanxuanhui/DSCapstone.git