



智驱生活 智能向善

大模型赋能智慧养老

助老年群体智慧点餐

参赛队伍：ele is all u need 主讲人：顾国勤

目录

CONTENTS

01
团队介绍

Team Introduction

02
方案介绍

Solution

03
方案创新

Innovation

04
社会价值

Society Value

01

团队介绍

Team Introduction



02 方案介绍

Solution

赛题介绍

- 赛题回顾
 - 1. 项目背景：老龄化加剧，传统点餐方式对老年人存在不便。智慧养老兴起，本项目借助语音交互技术，为老年人打造自然友好的点餐方式，提升其生活质量和用餐体验。
 - 2. 语音转文本（ASR）：接收老年人语音输入，将语音信号准确转换为文本信息。
 - 3. 意图识（DOM）别：对转换后的文本进行分析，判断是否为点菜意图。若为点菜意图，则进入后续推荐流程
 - 4. 菜品识别与推荐（QUE）：从点菜文本中提取菜品信息，与菜品库匹配，推荐最优菜品，满足老年人点餐需求。
- 评价指标

$$s = \alpha ASR + \beta DOM + \gamma QUE$$

$$ASR = 1 - \frac{1}{N} \sum CER_i$$

$$DOM = \frac{1}{N} \sum c_i \quad \begin{cases} c_i = 1, & y_i = \hat{y}_i \\ c_i = 0, & y_i \neq \hat{y}_i \end{cases}$$

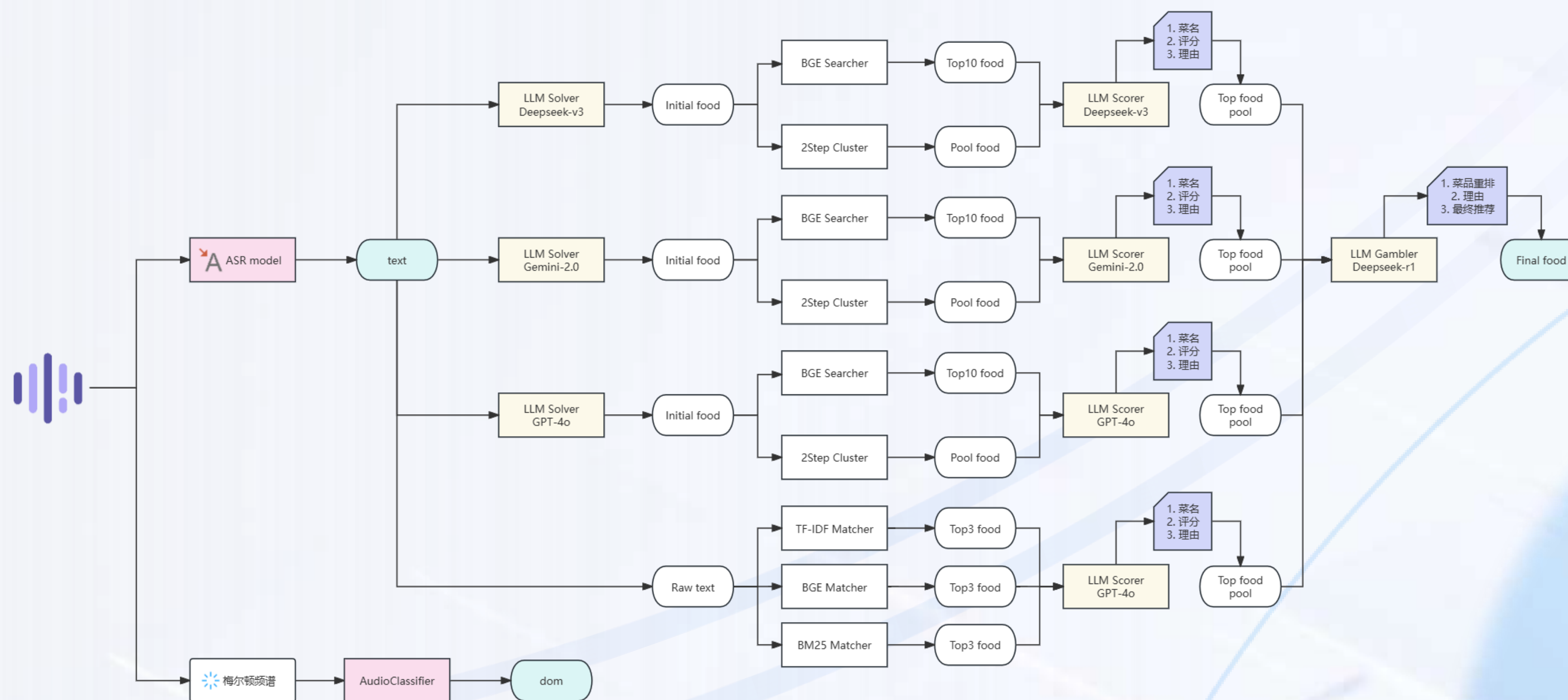
$$QUE = \frac{1}{\sum \hat{y}_i} \sum \hat{y}_i q_i \frac{s_i}{5} \quad \begin{cases} q_i = 1, & \text{菜品在库中} \\ q_i = 0, & \text{菜品不在库中} \end{cases} \quad s_i = \{0, 1, 2, 3, 4, 5\}$$

$$Inference Time: < 1s$$

02 | 方案介绍

最终方案概览

- 意图识别部分，将音频处理成梅尔顿频谱，使用EfficientNet模型对频谱图进行分类；
- 语音转写部分，使用FunASR微调SenceVoice-small，通过五折交叉验证制作模型汤；
- 菜品推荐部分，设计四路召回，基于模型打标结果进行Lora微调。



02 | 方案介绍

实验记录

Table 1: A 榜提交记录

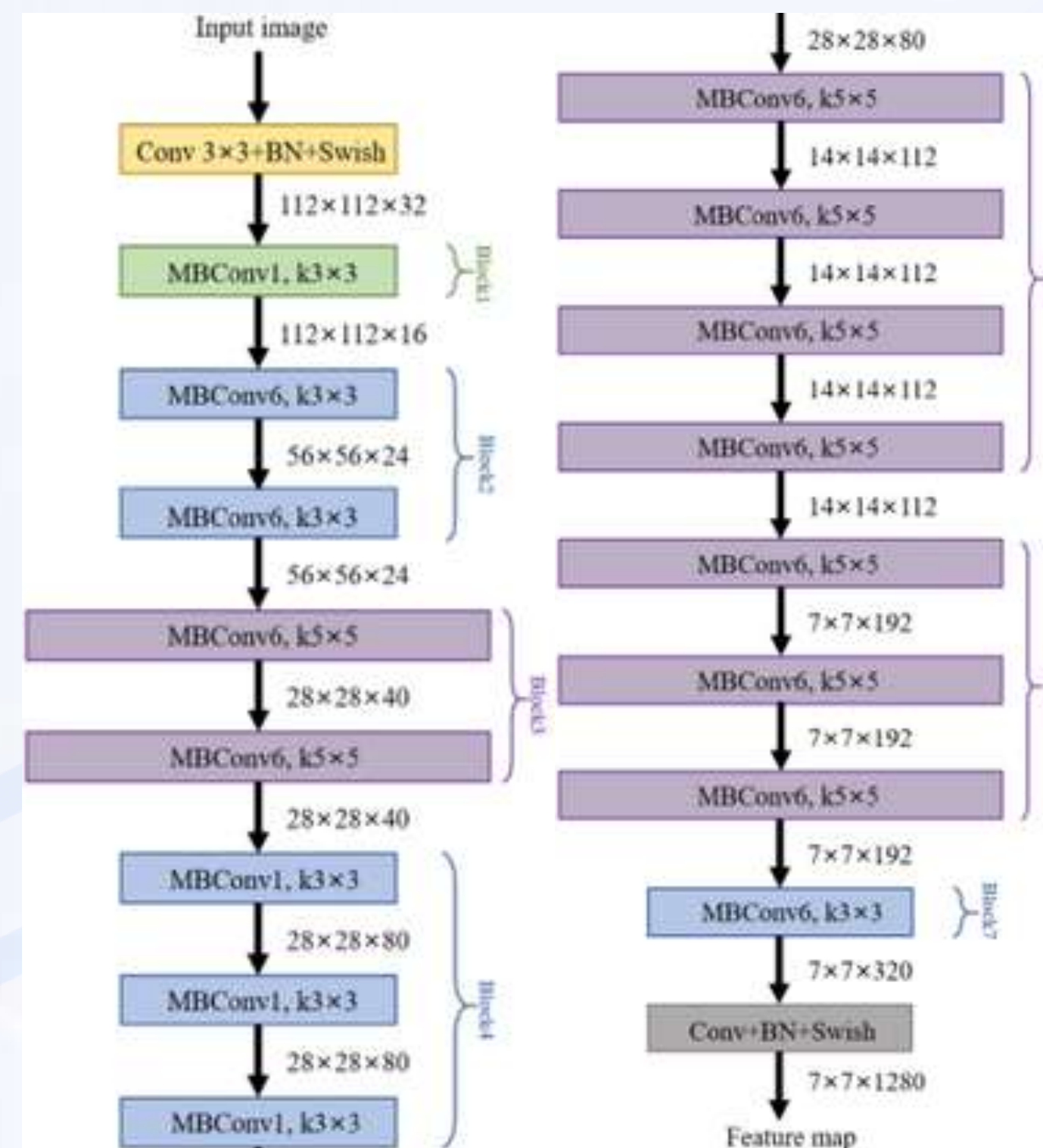
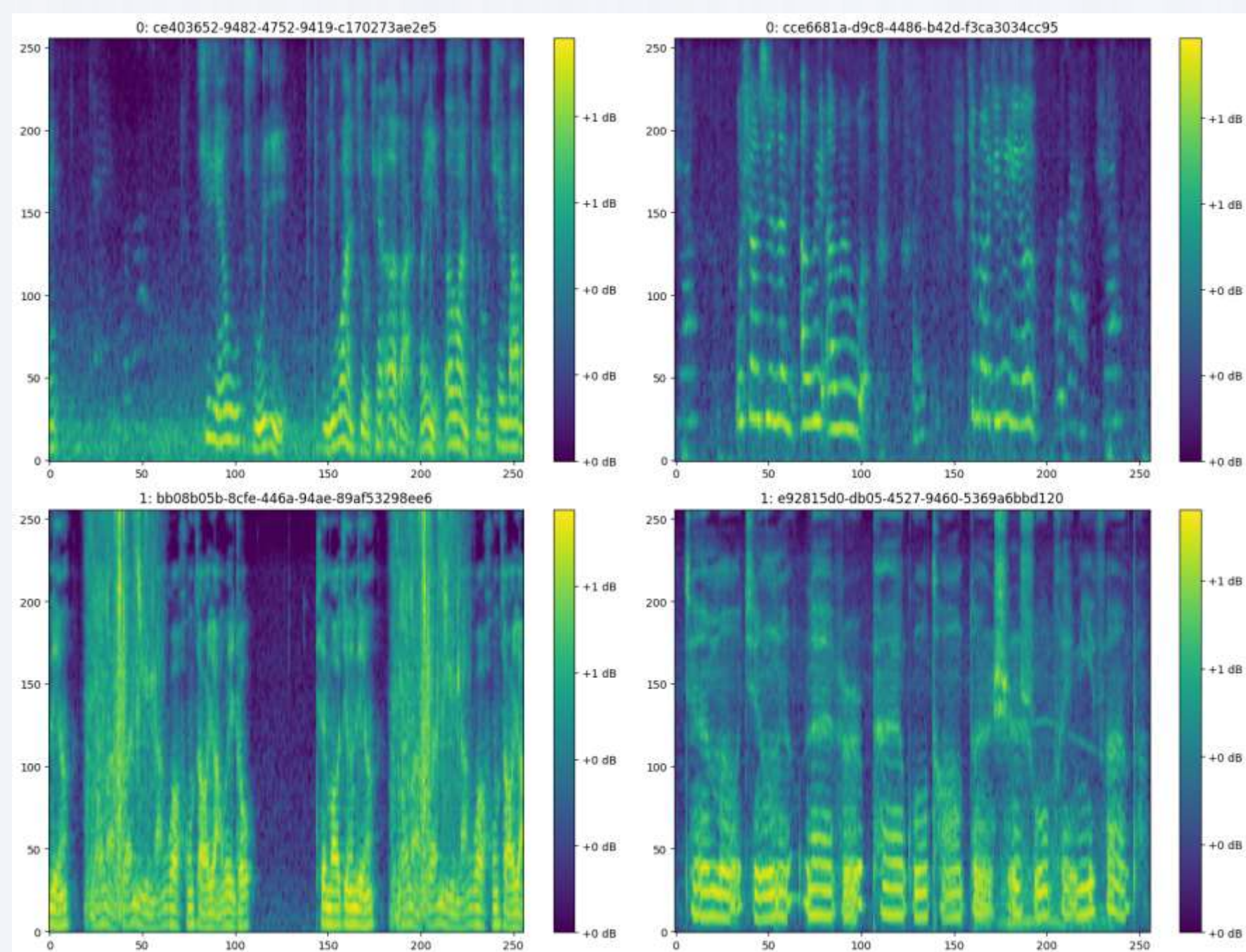
text	dom	food	总分	模块	描述
	0.2998			意图识别	训练 EfficientNetB0，使用五折投票
0.3758					使用 FireRedASR-AED(1B)
0.4342					修改识别多通道语音的 bug
0.4379					使用 FireRedASR-LLM(8B)
0.4706				语音转写	微调 FireRedASR-AED(1B)
0.4831					微调 SenseVoice
0.4827					加长 sma
0.4840					模型汤融合
					转义：Qwen-14B+ 召回：DS
		0.1648			搜索：字符相似度
		0.1711			搜索：层次聚类
		0.1731			搜索：2-Step Cluster
				菜品推荐	转义：DS+ 召回：DS
		0.1705			搜索：BGE 相似度
		0.1779			搜索：Cluster&BGE 并集
					转义：Gemini+ 召回：Gemini
		0.1737			无转义搜索：TF-IDF&BM25&BGE 并集 + 召回：DS
0.4840	0.2998	0.1796	0.9634	多路集成	集成两路召回：DS+Gemini
0.4840	0.2998	0.1806	0.9644		集成三路召回：DS+Gemini+ 无转义搜索

Table 2: B 榜提交记录

text	dom	food	总分	描述
0.4827	0.2998	0.1800	0.9625	集成三路召回
0.4827	0.2998	0.1808	0.9633	集成四路召回
0.4827	0.3000	0.1820	0.9647	优化召回方式为打分重排

意图识别方案

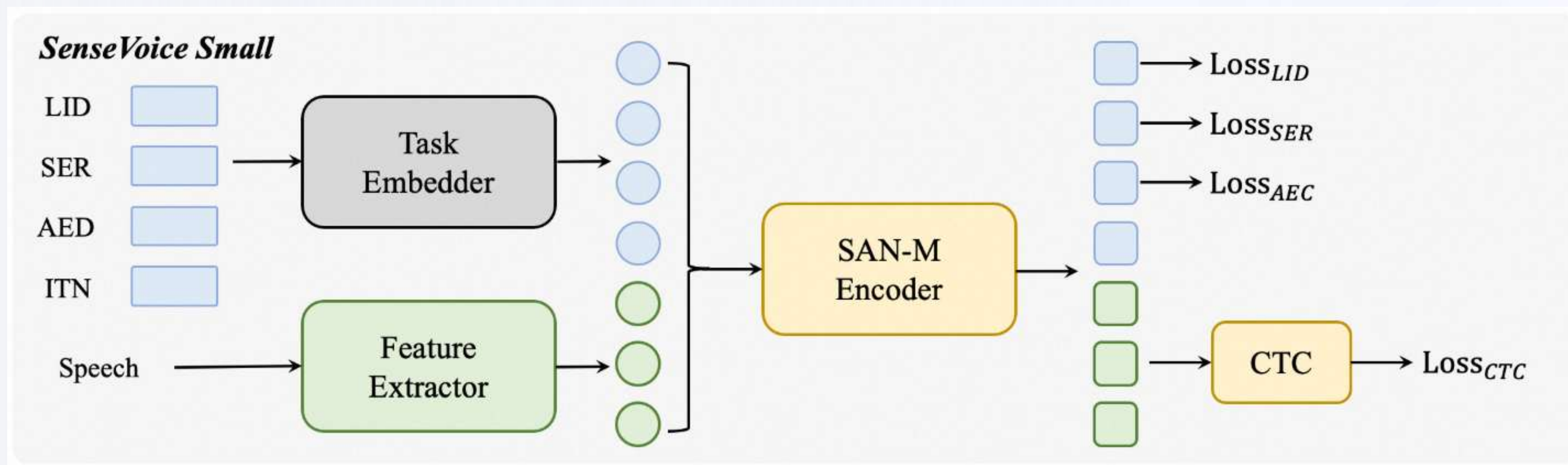
- 结果 (A榜) : 0.2998/0.3 结果 (B榜) : 0.2998/0.3**



02 | 方案介绍

语音转写方案

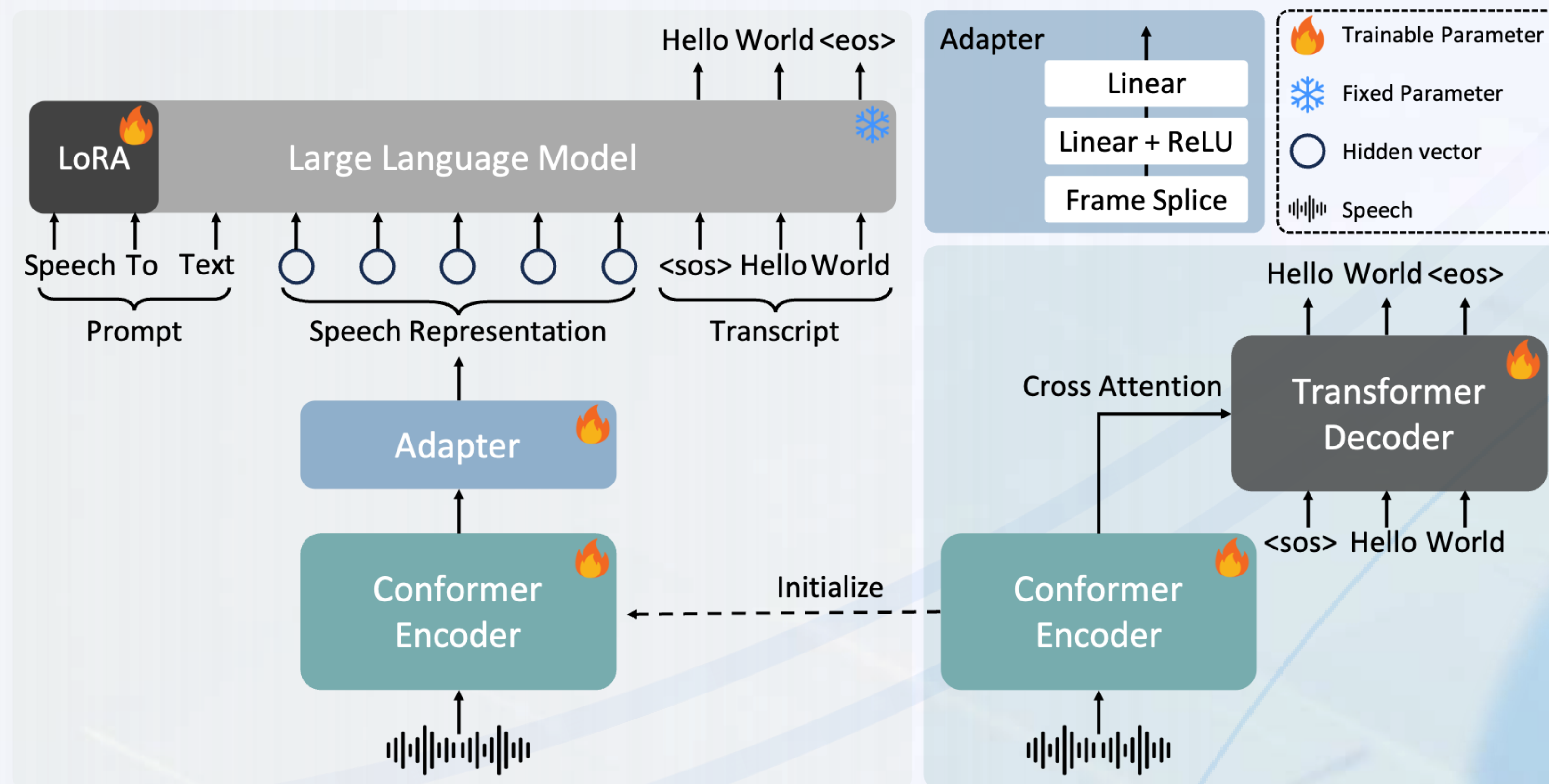
- 使用Sensevoice-small模型进行微调，五折交叉验证，对五个模型权重做模型汤。
- 也尝试过选择五个中cer最小的，但效果不佳且推理速度非常慢，而模型汤不损失推理速度。
- 去除推理输出中的表情符号以及标点符号。 **结果（A榜）：0.4840/0.5 结果（B榜）：0.4827/0.5**



02 | 方案介绍

语音转写方案

- 使用FireRedASR两款模型推理得分均为0.43+
- 尝试微调1B模型，但效果不佳。 **结果（A榜）：0.4706/0.5**

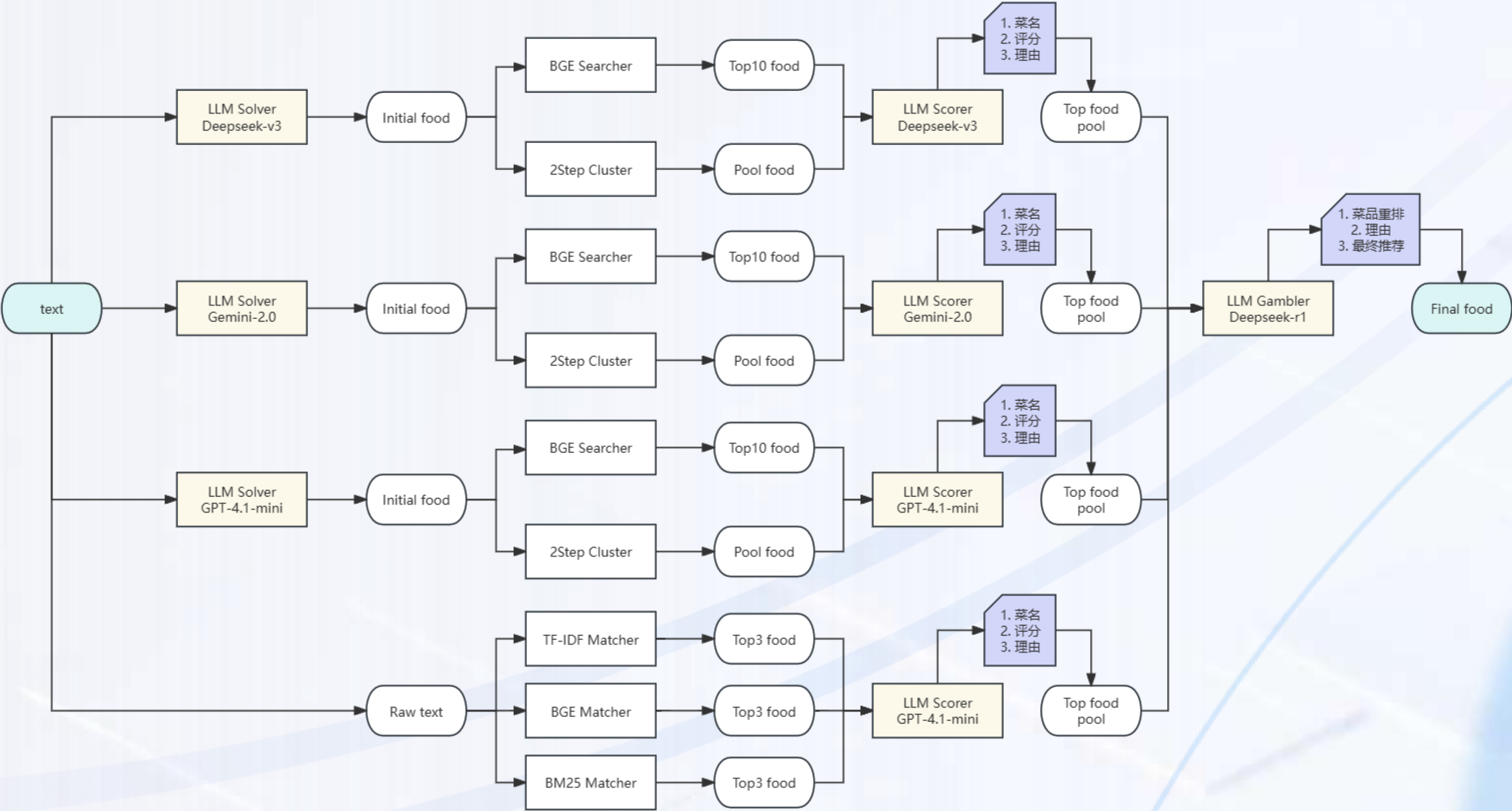


02 | 方案介绍

菜品推荐方案

- 核心思想：为模型提供更精准、更丰富的菜品池。
- 核心模块：LLM Solver、Food Searcher、LLM Scorer、LLM Reranker。

结果（A榜）：0.1806/0.2 结果（B榜）：0.1820/0.2



02 | 方案介绍

菜品推荐方案

01

LLM Solver

- 使用大模型对转写出的文本进行第一次转义，后续的食物检索基于模型转义的结果进行。
- 三路使用了不同的模型、也使用了不同的Prompt，分别设计为：面向老年人场景/通用场景/获得EGPT最高打分。

外卖点餐
1. 核心任务定义你是一名菜品推荐助手，一名老年人正在向你提出自己想要点的外卖。基于老年用户提出的需求，你需要推荐最符合老年人用户要求的菜品。
2. 输入数据规范.....
3. 菜品推荐.....
4. 输出数据规范.....

美食推荐
角色设定：你是一名专业的**美食推荐**官，精通中式家常菜与养生餐，擅长根据用户的口味偏好和美食意向，精准推荐最合适的美食菜品。
任务定义：根据用户输入的**口味偏好**和**美食意向描述**，为用户推荐一个符合要求的具体美食菜品。
* 用户的口味偏好可能是概括的描述，例如：少盐、少糖、养生、热乎、家常、清淡、滋补等
* 用户的美食意向可能是具体的食物，例如：小米粥、紫菜蛋花汤、肉末豆腐盖浇饭、木耳炒鸡蛋等

你是一位经验丰富、贴心周到的菜品推荐专家，擅长根据用户语音文本中的情绪、语气、饮食习惯和口味偏好，精准推荐最适合他们当前情境的菜品。
你必须严格遵循以下原则来生成推荐：
1. **理解用户语音中的意图和情感**：分析用户是否偏好清淡/滋补/软糯/少油/热食/有汤/易咀嚼/带馅等特征；
2. **考虑用户语言中的地方口音、表达习惯**：如“软咪点”、“伐”、“阿拉”等应理解为上海方言表达，推荐需要基于对地方风味的深入理解；

请根据用户的点餐语音内容，给出一条高度个性化、令人感到“被理解”的菜品推荐，目标是获得**菜品接受模型EGPT**的最高评分（5分）。

02 | 方案介绍

菜品推荐方案

02

Food Searcher

- 通过多种方法从食品库中检索食品，统计层使用聚类方案，语义层包含TF-IDF/BM25稀疏向量、BGE稠密向量。

统计层

Step1.1: 基于公共子串的聚类

- 计算食品名称间公共子串的长度比，高于阈值（0.6）的聚为一类

1	0.74	0.14	0.42
0.74	1	0.12	0.75
0.14	0.12	1	0.65
0.42	0.75	0.65	1

Name Similarity > 0.6

结果：2318个类别

Step1.2: 基于加权特征的聚类

- 通过食品库给出的食品类别信息计算食品相似度矩阵
- 通过TF-IDF计算食品特征，使用线性核转为食品相似度矩阵
- 两个相似度矩阵进行加权平均作为距离矩阵
- 使用自下而上的层次聚类，使用类间平均距离作为链接方法

1	0.81	0.25	0.14
0.81	1	0.12	0.6
0.25	0.12	1	0.43
0.14	0.6	0.43	1

+

1	0.78	0.25	0.26
0.78	1	0.21	0.6
0.25	0.21	1	0.38
0.26	0.6	0.38	1

Distance < 0.5

结果：3027个类别

语义层

TF-IDF

- 通过统计词频（TF）和逆文档频率（IDF）衡量词语重要性，适合基础关键词匹配

BM25

- 基于TF-IDF改进，引入文档长度归一化和饱和机制，更适合短文本相似度搜索

BGE

- 基于预训练的嵌入模型，将文本映射为稠密向量，通过语义相似度匹配，适合复杂语义理解任务

02 | 方案介绍

菜品推荐方案

03

LLM Scorer

- 各路分别对菜品池进行打分，召回得分最高的菜品池（TopK）。

```
## 外卖菜品打分任务
### **1. 核心任务定义** .....
### **2. 输入数据规范** .....
### **3. 输出数据规范**
你需要为每道菜品提供以下三项内容：
1. **菜名**：必须与菜品列表中的字符串完全一致；
2. **评分**：1~5 分，5 分为最符合需求；
3. **理由**：简明解释该评分的原因（例如口感软糯、养胃等）。

[[
  "菜名": "软糯小米红枣粥", "评分": 5,
  "理由": "红枣补血、小米养胃，软糯顺口，非常适合牙口不好的老年人"
],
{
  "菜名": "软糯香甜八宝粥", "评分": 3,
  "理由": "八宝粥食材多样但含坚果、豆类，颗粒偏硬，不太适合牙口差的老人"
}]
```

04

LLM Reranker

- 合并各路的高分菜品池（4*TopK），重排，召回最终的推荐菜品（Top1）。

```
## 外卖菜品排序任务
### **1. 核心任务定义** .....
### **2. 输入数据规范** .....
### **3. 输出数据规范**
你需要为每道菜品提供以下三项内容：
1. **排序**：对菜品列表进行排序，菜品名称必须与列表一致；
2. **理由**：简明解释该排序的原因（例如某项菜品完全匹配要求、最符合用户的口味偏好、亦或最能得到用户喜欢，则应排在最前）；
3. **最优推荐**：得出最终推荐最优的菜品，菜品名称必须与菜品列表中的字符串完全一致。

[[
  "排序": ["养胃纯小米粥", "养胃小米粥", "养胃 | 小米粥", "小米养胃粥", "养胃小米南瓜粥", "暖胃小米南瓜粥", "大米小米粥", "小米玉米粥"],
  "理由": "用户希望点一碗养胃的小米粥，因此推荐优先符合“小米”成分并具备养胃功能的粥类菜品",
  "最优推荐": "养胃纯小米粥"
]]
```




03 方案创新

Innovation

03 | 方案创新

主要提升点

频谱图分类

ASR模型汤

多路召回

打分&重排



04 社会价值

Society Value

社会价值

01 尊严养老

自然语音交互，简化老人点餐实现方法，守护自主尊严，跨越数字鸿沟。

04 服务鲁棒

提高服务鲁棒性，在弱网环境下，无需复杂编码传输，杜绝丢包的传输失败。

02 高速响应

端侧部署轻量化模型，降低推理响应时延，减少等待时间，缓解焦虑提高体验。

05 范式创新

建立多模态边缘智能落地范式，推动边缘模型赋能边缘任务。

03 隐私保护

构建个性化模型，敏感数据不入网，仅脱敏数据和边缘侧算力单元交互，实现服务不出基站。

06 绿色AI

利用边缘侧（基站）冗余算力，降低云端计算中心能源消耗，助力实现碳中和。

04 | 社会价值

真机演示

- 用户语音指令→天猫精灵→意图识别模块→语音转写模块→菜品推荐模块→饿了么下单服务





智驱生活 智能向善

THANKS

感谢观看

参赛队伍：ele is all u need 主讲人：顾国勤

