



2025南师大暑期课程

城市关键要素获取与表征

姚尧 博士，教授
中国地质大学（武汉）

Email: yaoy@cug.edu.cn



城市是人类社会高度集中的空间，融合了复杂的人口结构、经济活动、生态环境与技术系统，构成了一个多层次、动态演化的综合体系。

如何更好地理解、规划、管理和优化城市？





目录



- 1 城市要素的简介
- 2 城市功能结构
- 3 基于POI的城市要素挖掘
- 4 基于多模态数据的表征方法
- 5 总结



目录



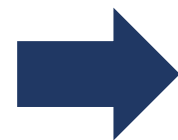
- 1 城市要素的简介
- 2 城市功能结构
- 3 基于POI的城市要素挖掘
- 4 基于多模态数据的表征方法
- 5 总结

- 1.1 城市要素
- 1.2 城市要素表征
- 1.3 重要城市要素获取

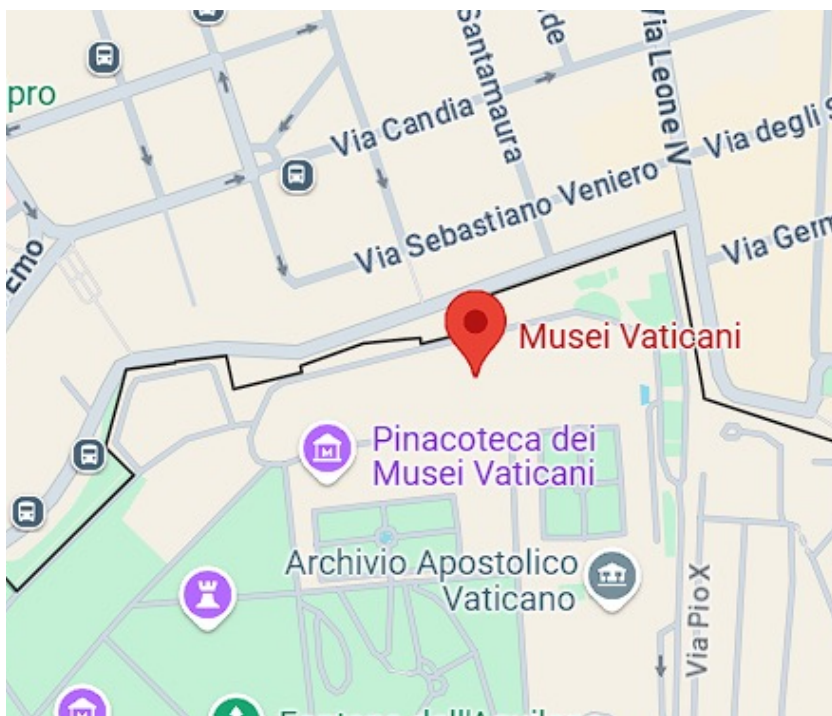
1.1 城市要素及其表征



什么是城市要素？应该如何理解城市要素？
我的周边有那些城市要素？又该如何描述？



分析这些城市要素有什么作用？



位置/地点



轨迹



图像

1.2 城市要素表征



➤ 表示学习



声音和音乐
音素、语调和频率



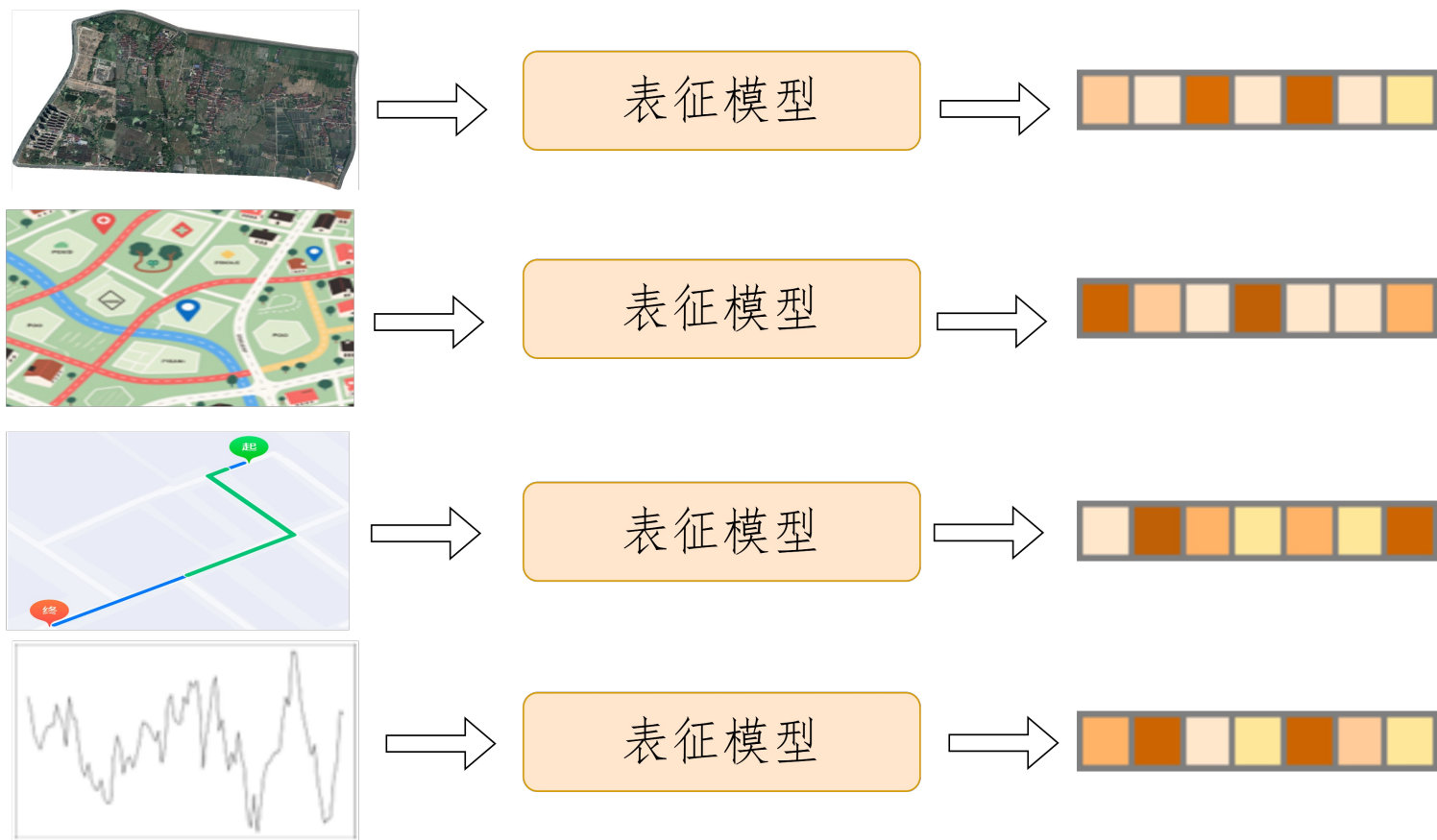
物体识别
边缘、纹理、形状



自然语言处理
音素、语调、频率

- 表示学习：通过自动学习数据中的有效特征表示，将原始数据转换为更易于建模和分析的形式过程。
- 为什么我们要对这些信息进行表示学习？

1.2 城市要素表征



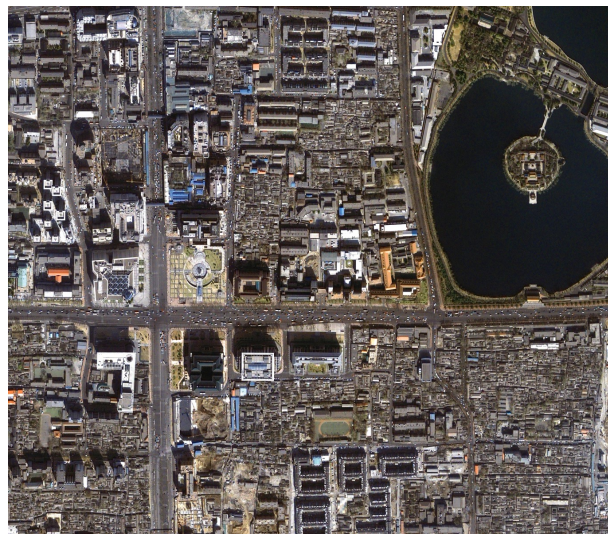
- 地理表征：指通过抽象与编码的方式，将现实世界中的地理现象、空间关系与过程转化为可识别、可分析、可建模的表达形式。
- 地理信息具有多源性、异构性、高冗余性和空间异质性等特点，地理表征已成为必然。

1.3 重要城市要素获取



➤ 遥感影像

- 遥感数据具有覆盖广、更新快、信息丰富等特点。
- 不同影遥感影像适用于不同的应用场景。



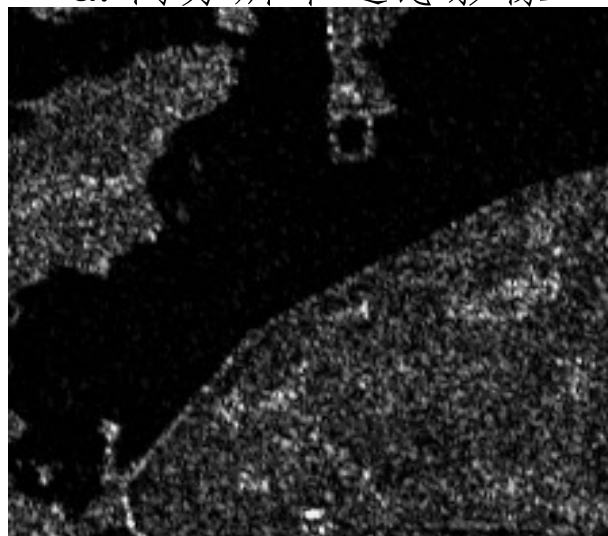
a. 高分辨率遥感影像



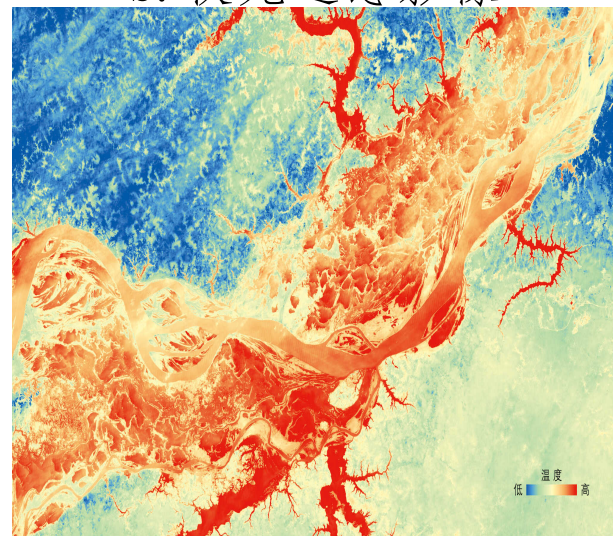
b. 夜光遥感影像



c. 无人机遥感影像



d. 雷达遥感影像



e. 热红外遥感影像

1.3 重要城市要素获取



➤ 遥感影像

a. 高分数据中心

- 高分辨率遥感影像
- 雷达遥感影像

b. USGS Earth Explorer

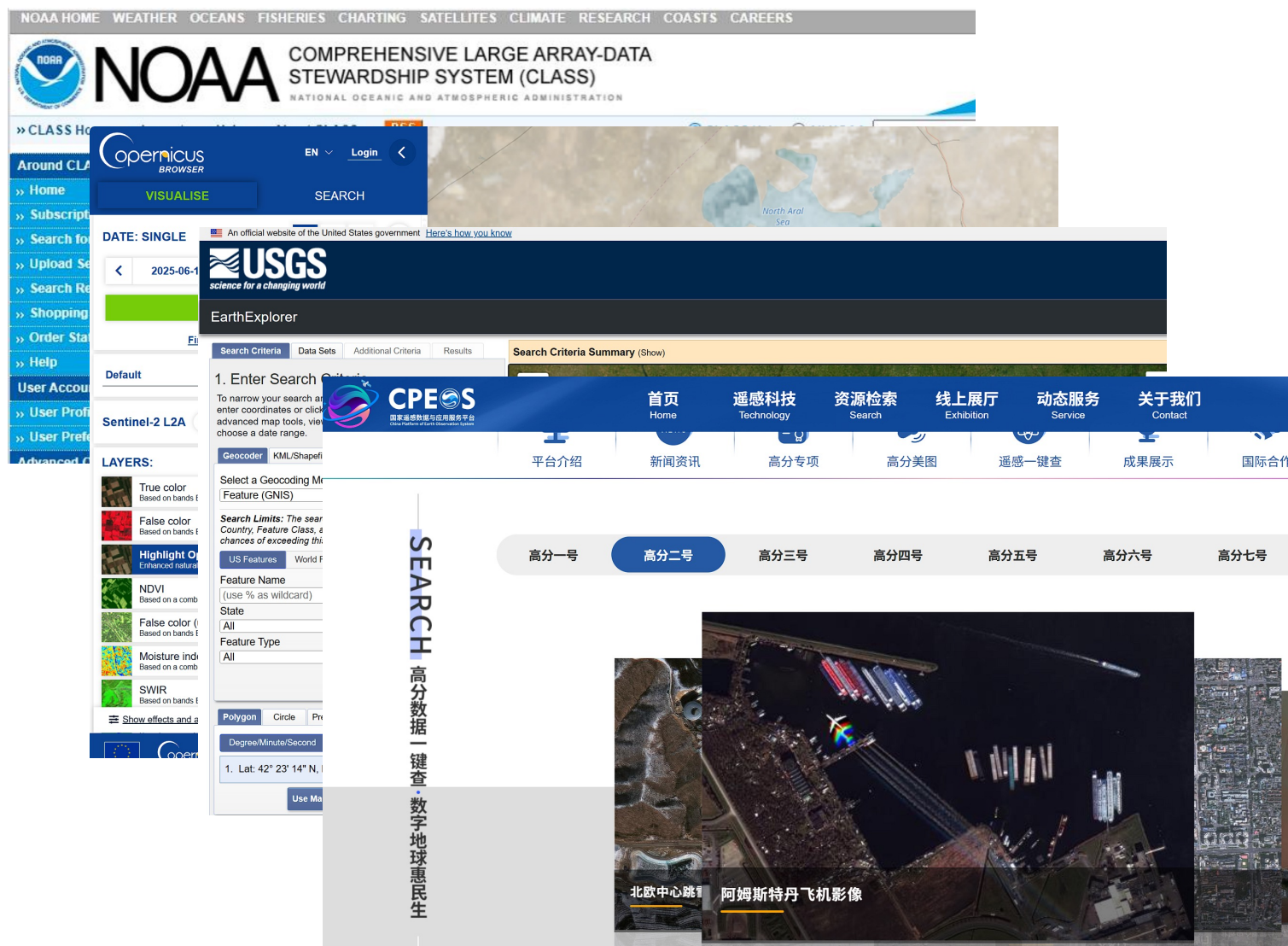
- Landsat系列多光谱影像
- 热红外遥感数据

c. NOAA

- NPP-VIIRS夜光遥感影像
- 历史DMSP夜灯数据

d. Copernicus Open Access Hub

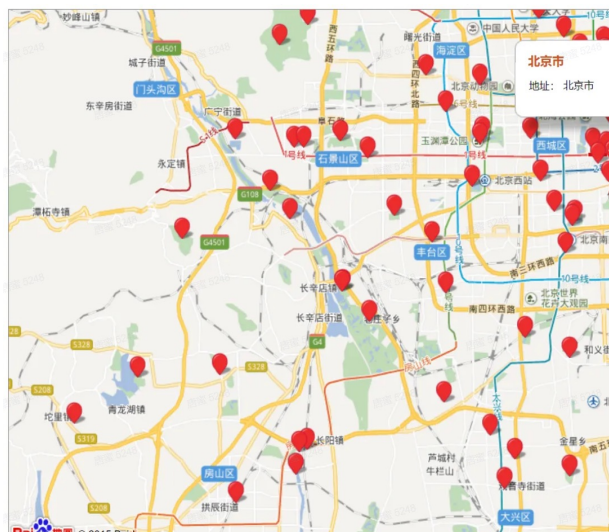
- Sentinel-1雷达影像
- Sentinel-2多光谱影像



1.3 重要城市要素获取



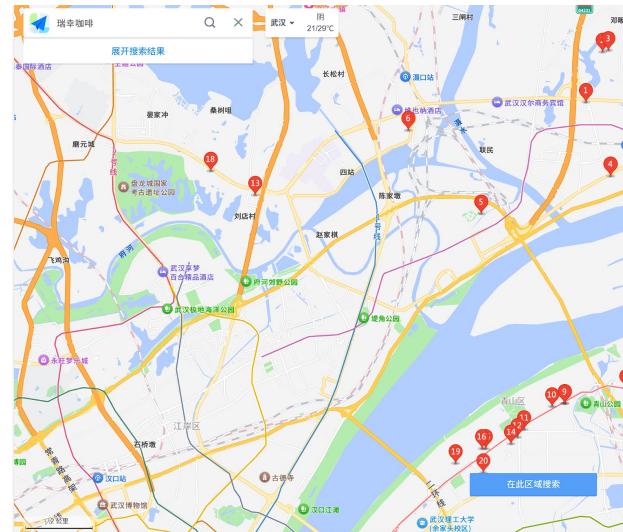
➤ POI数据获取



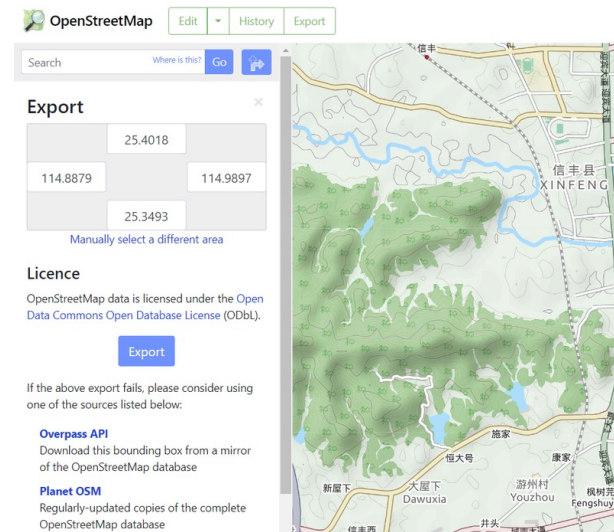
a. 百度地图



b. 腾讯地图



c. 高德地图



d. OpenStreetMap

➤ POI (Point of Interest), 指在地理空间中具有特定意义或功能的地点, 用于描述现实世界中各类设施和服务的位置与功能。

➤ POI富含丰富的语义信息, 可用于用地类型识别、刻画空间结构、支持位置服务等。

1.3 重要城市要素获取

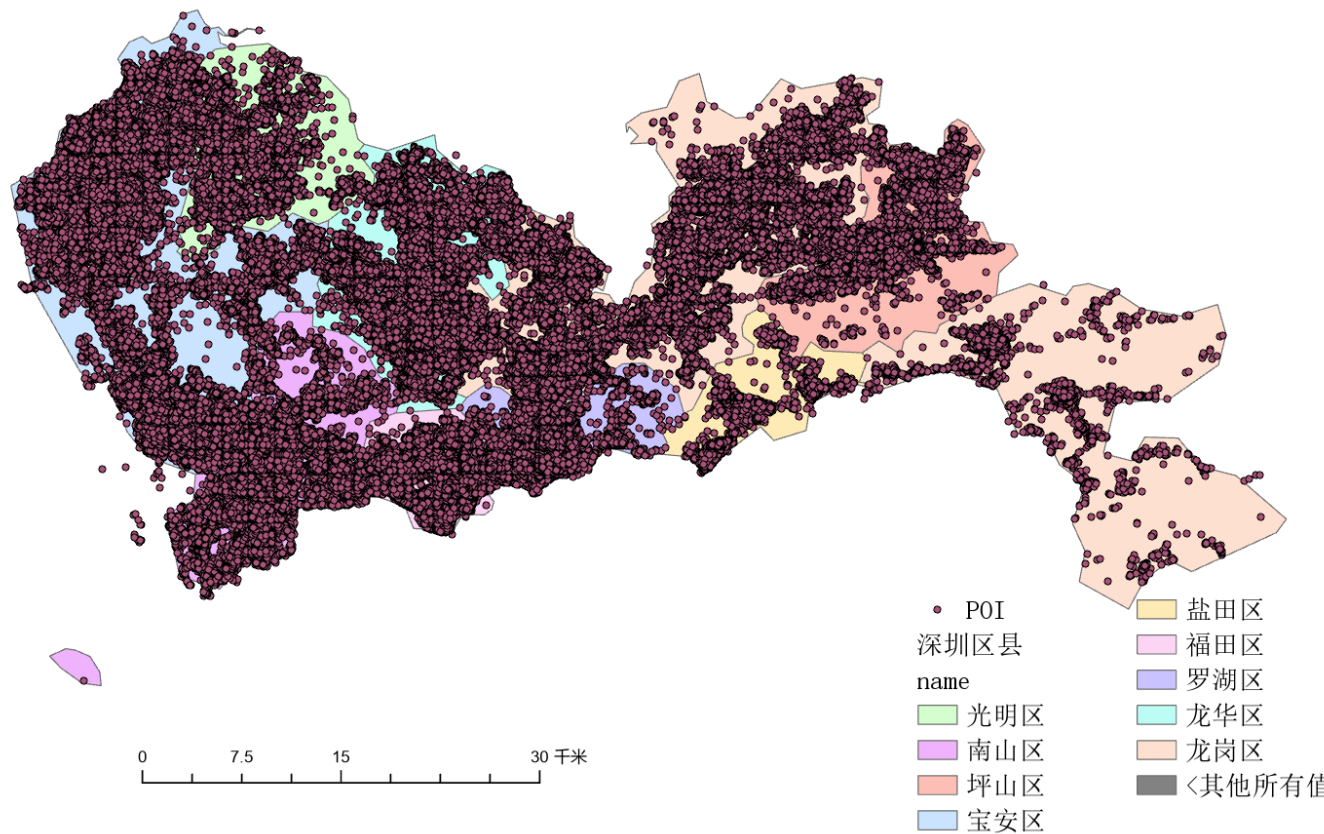


➤ POI数据

- 获取到的POI数据通常只包含POI的经纬度坐标、一级标签、二级标签信息

一级类别	交通设施, 休闲娱乐, 公司企业, 医疗, 名称标注类, 商务大厦, 地产小区, 宾馆, 政府机构, 教育...
二级类别	停车场, 交叉路口, 公交车站, 地铁, 收费处, 出入口, 长途汽车站, 加油站, 飞机场, 出租车站, 港口, 火车站, 桥, 服务区, ktv, 文化馆, 网吧...

深圳市POI分布图





目录



- 1 城市要素的简介
- 2 城市功能结构
- 3 基于POI的城市要素挖掘
- 4 基于多模态数据的表征方法
- 5 总结

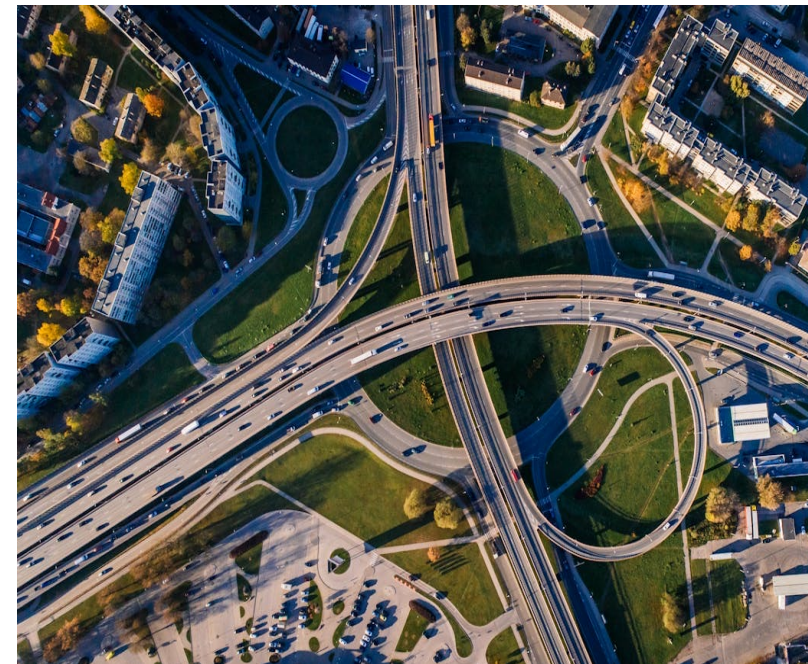
- 2.1 城市功能结构简介
- 2.2 城市功能区的类型与构成
- 2.3 城市功能结构典型特征

2.1 城市功能结构简介



➤ 城市功能结构

城市功能结构是什么？有哪些城市功能结构？为什么要研究城市功能结构？



城市功能结构是指城市内部因不同活动（如居住、商业、工业等）集中分布而形成的、具有典型特征和功能差异的空间区域。

2.2 城市功能区的类型与构成



商业区



住宅区



工业区



行政区



文教区



公共服务区



修养疗养区



交通与物流区

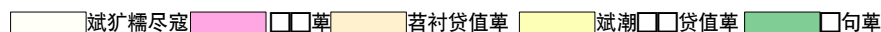


绿地与生态区



混合功能区

- 通过分析一个区域的POI的数量、类型、密度、比例、聚集强度等，可以有效区分不同的功能区。
- 城市功能区并非完全独立，往往存在过渡地带。



- ## 城市空间布局模式

- ## 功能混杂与单一

- ## 核心数据源



目录



- 1 城市要素的简介
- 2 城市功能结构
- 3 基于POI的城市要素挖掘
- 4 基于多模态数据的表征方法
- 5 总结

- **3.1 Skip-gram算法与CBOW算法**
- **3.2 Word2Vec方法**
- **3.3 Place2Vec方法**
- **3.4 Semantic2Vec方法**
- **3.5 Hierarchical Graph Infomax (HGI)方法**

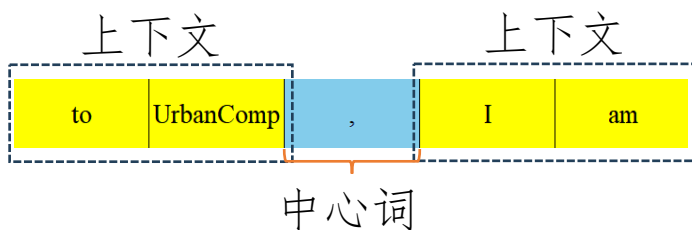
3.1 Skip-gram算法与CBOW算法



➤ One-Hot编码

Welcome to **UrbanComp** , I am also a member of **UrbanComp** .

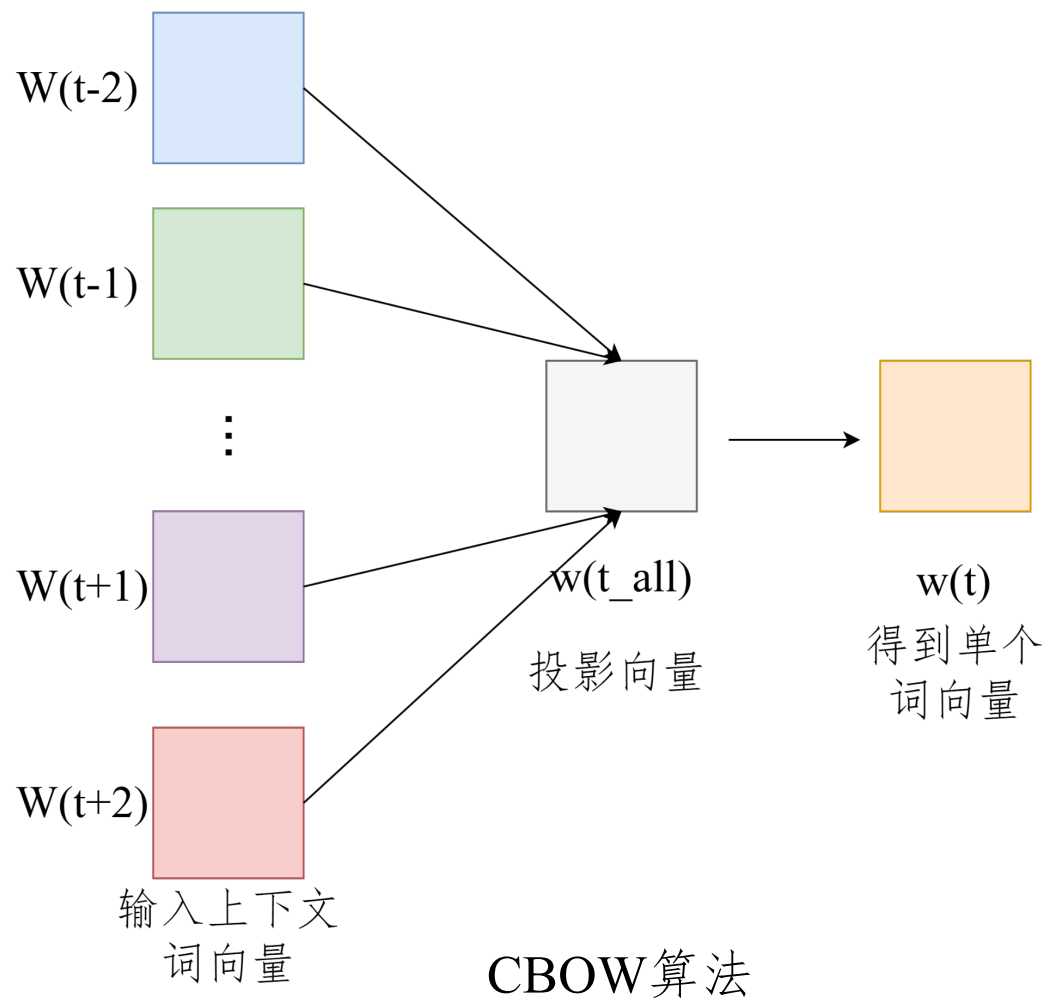
1 2 3 4 5 6 7 8 9 10 3 11



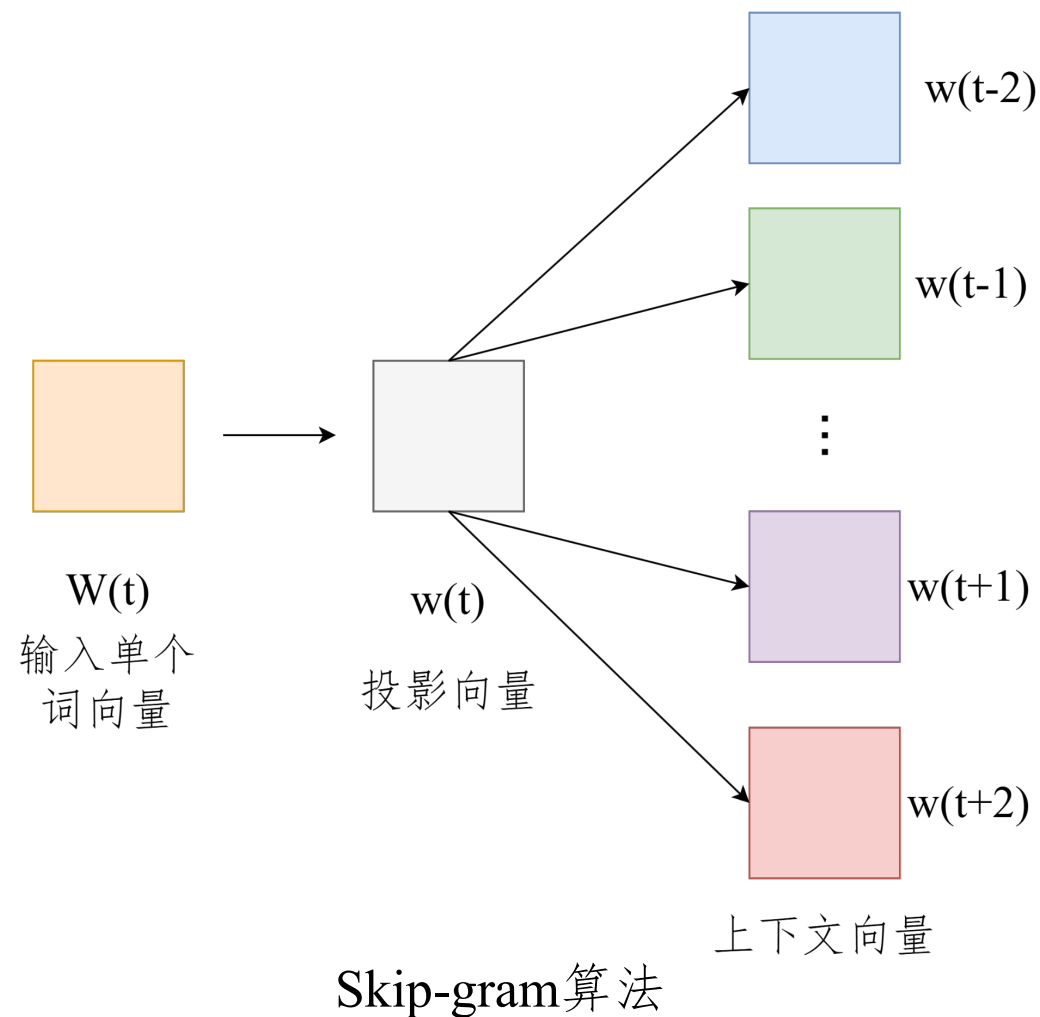
	1	2	3	4	5	6	7	8	9	10	3	11
Step 1	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 2	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 3	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 4	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 5	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 6	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 7	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 8	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 9	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 10	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 11	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.
Step 12	Welcome	to	UrbanComp	,	I	am	also	a	member	of	UrbanComp	.

➤ 通过滑动窗口的方式检索“中心词”和它周围的“上下文”。

3.1 Skip-gram算法与CBOW算法



- 通过输入上下文的词向量检索中心词的词向量



- 通过输入中心词的词向量检索上下文的词向量

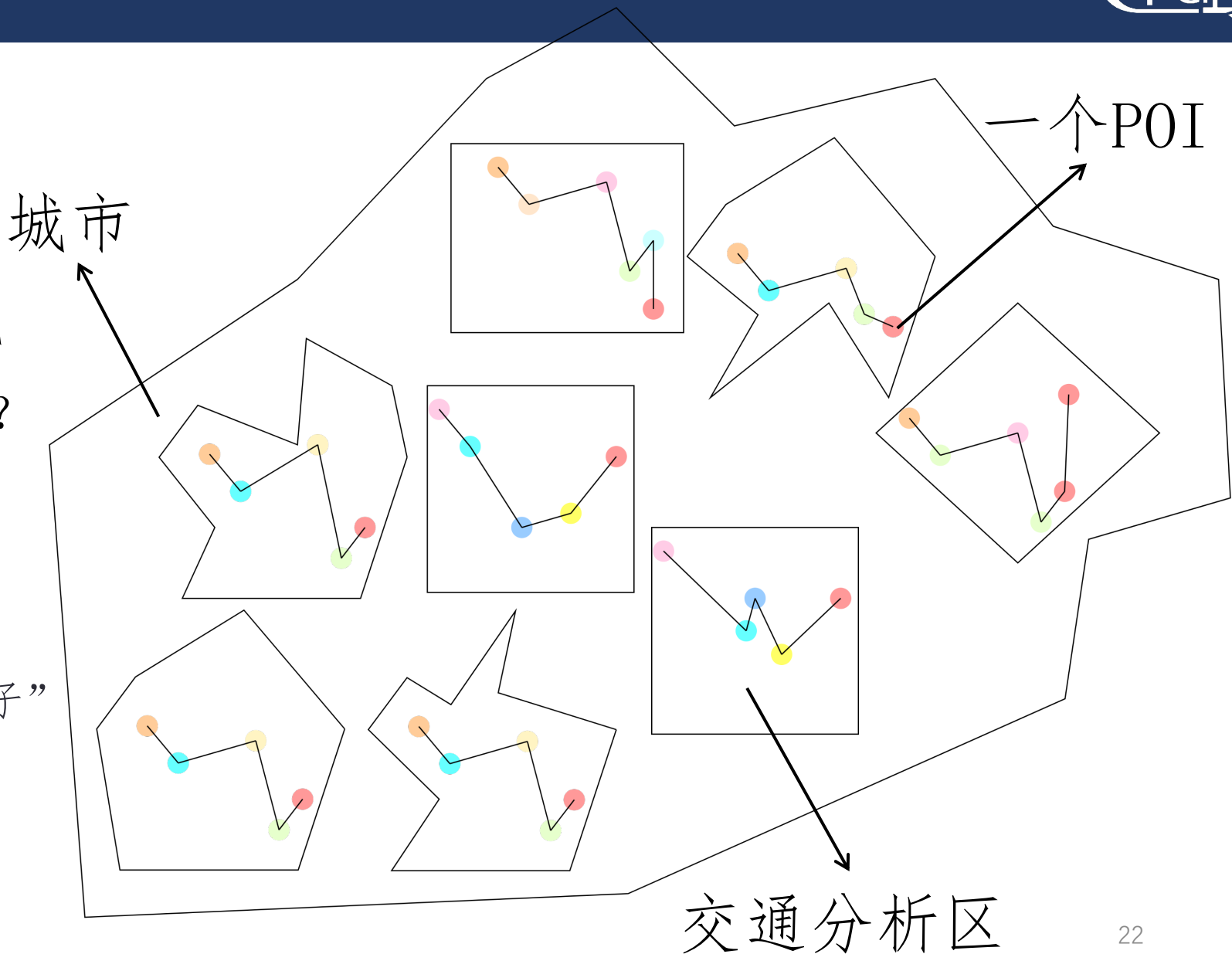
3.2 Word2Vec方法



➤ 构建语料库

如何以自然语言的形式表达城市中POI及其之间的空间关系？

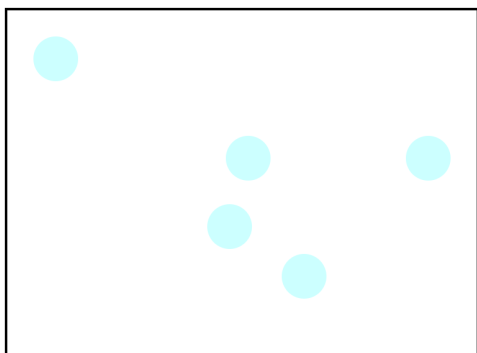
- POI（兴趣点） → 当作“单词”
- TAZ（交通分析区） → 当作“句子”
- 整个城市区域 → 当作“语料库”



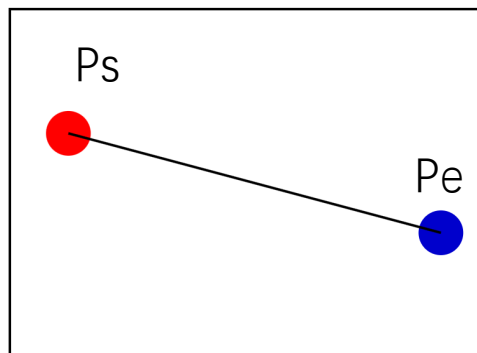
3.2 Word2Vec方法



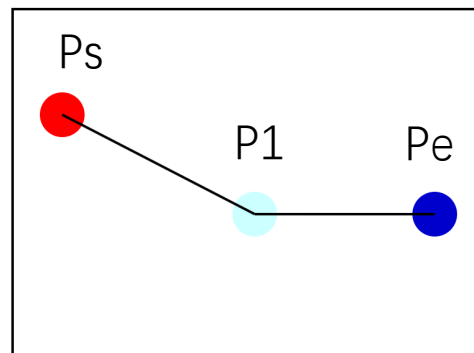
➤ 构建语料库



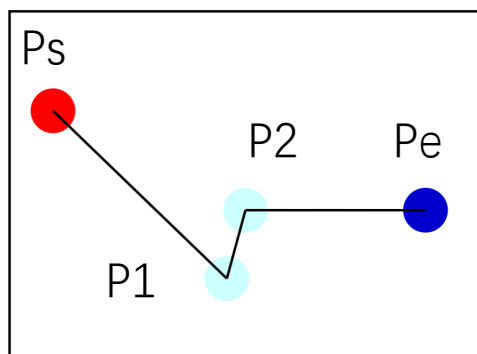
1



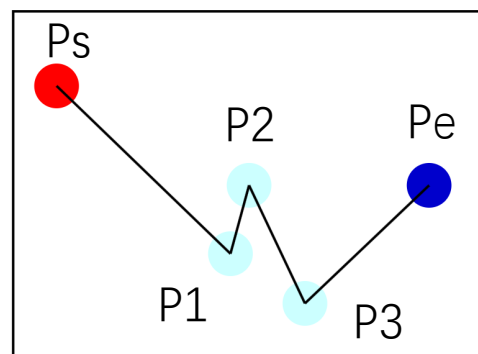
2



3



4



5

贪心算法

➤ 每个兴趣点（POI），比如“超市”、“学校”、“地铁站”，都相当于一个“单词”。

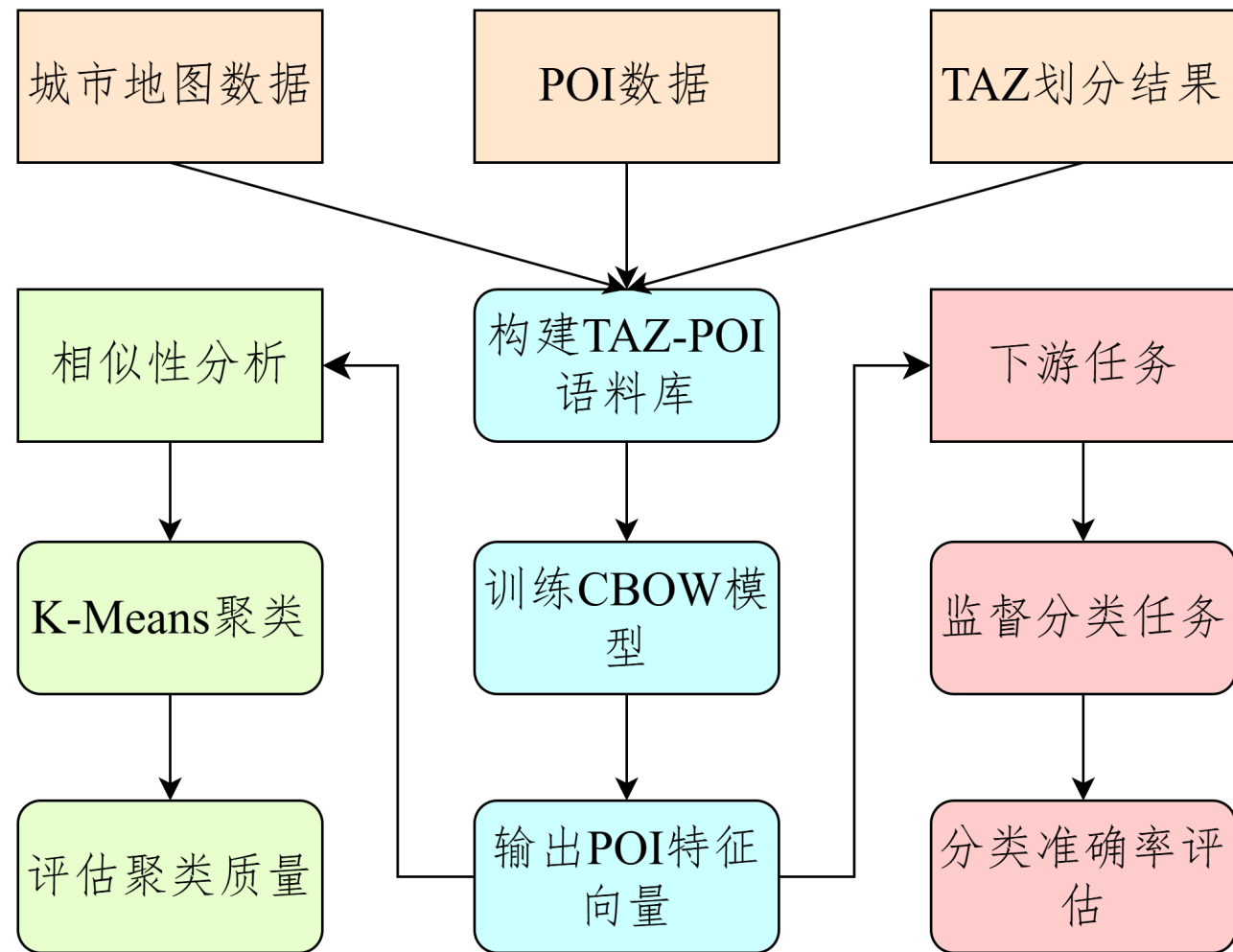
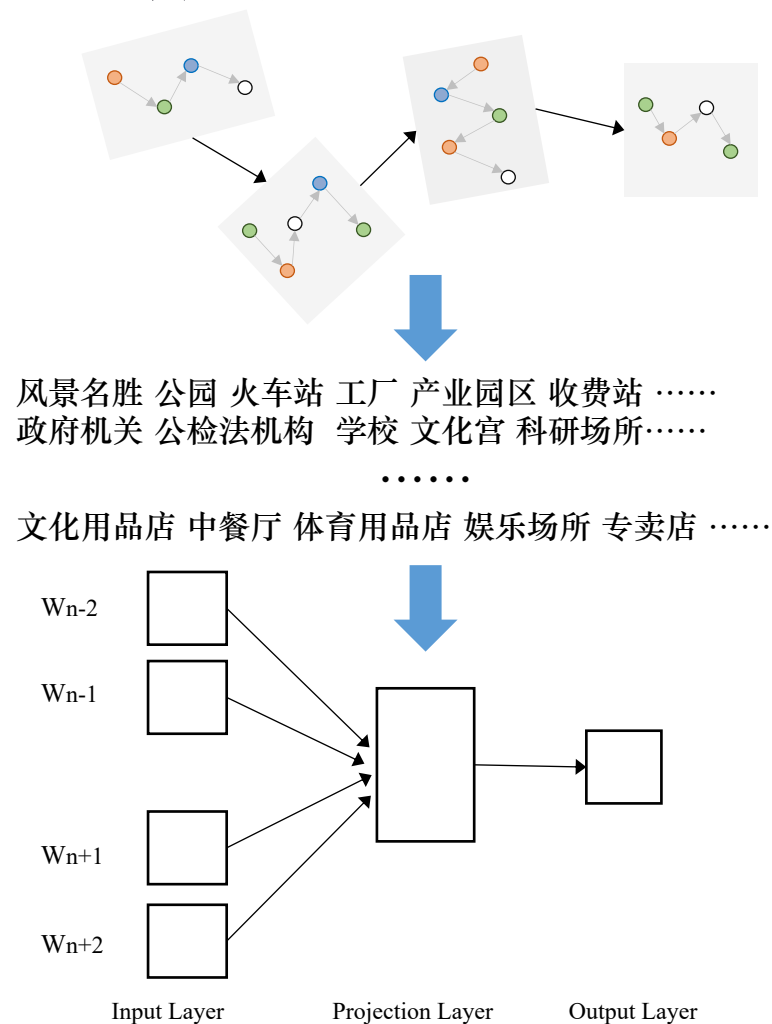
➤ 通过最短路径法确定单词（POI）的顺序来构成句子。

交通分析区（TAZ）中的最短路径构建可采用贪心算法实现。

3.2 Word2Vec方法



➤ Word2Vec方法



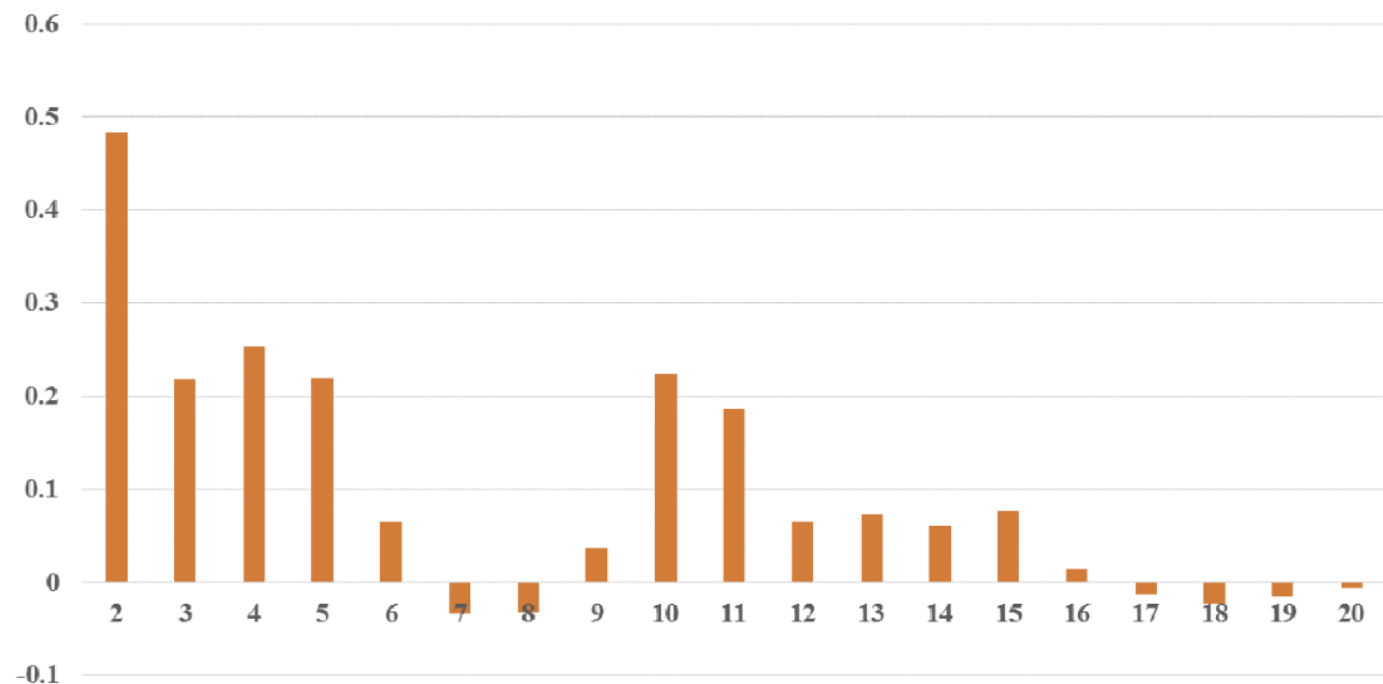
参考网址: <https://www.urbancomp.net/archives/word2vec-for-poi>

3.2 Word2Vec方法



➤ Word2Vec方法

- POI 数据来源：百度地图
- POI 总数：944,698 条
- TAZ 单元数量：37,584 个
- 有效 TAZ 数量：排除无 POI 的单元
- 最终使用 POI 类别数：419 个末级类别
- CBOW 模型输出向量维度：200
- 滑动窗口大小：5
- 训练迭代次数：20



随着 K 值 (x 轴) 的增加, 聚类结果 (POI 向量) 的平均轮廓系数 (y 轴) 的变化。

同质功能的城市微观实体具有相似的空间分布特征, 异质POI之间也存在空间相关性。

3.2 Word2Vec方法



➤ Word2Vec方法

➤ Word2Vec优于传统语义模型（TF-IDF、pLSA、LDA）。

➤ 所提出的方法能够有效量化城市土地利用类型与 POI 空间分布模式之间的相关性。

➤ 适用于大规模城市区域的土地利用分类任务。

W2V	PFL	GSL	IUL	CBF	RUL	APS	RST	LWL	SUL	URC	RTF	RPF	OCL	MUL
PFL	0.81	0.09	0.01	0.01	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GSL	0.00	0.89	0.01	0.01	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IUL	0.00	0.06	0.86	0.01	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CBF	0.00	0.06	0.01	0.82	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RUL	0.00	0.05	0.01	0.01	0.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
APS	0.00	0.06	0.01	0.01	0.09	0.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RST	0.00	0.10	0.01	0.02	0.07	0.01	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LWL	0.00	0.07	0.05	0.01	0.07	0.00	0.00	0.80	0.00	0.00	0.00	0.00	0.00	0.00
SUL	0.00	0.04	0.01	0.01	0.10	0.01	0.00	0.00	0.84	0.00	0.00	0.00	0.00	0.00
URC	0.00	0.08	0.02	0.01	0.09	0.00	0.00	0.00	0.00	0.80	0.00	0.00	0.00	0.00
RTF	0.00	0.06	0.02	0.01	0.09	0.01	0.00	0.00	0.00	0.00	0.82	0.00	0.00	0.00
RPF	0.00	0.29	0.00	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.57	0.00	0.00
OCL	0.00	0.07	0.01	0.01	0.11	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.79	0.00
MUL	0.00	0.08	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.83

(a) Word2Vec

pLSA	PFL	GSL	IUL	CBF	RUL	APS	RST	LWL	SUL	URC	RTF	RPF	OCL	MUL
PFL	0.59	0.16	0.03	0.03	0.19	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GSL	0.00	0.79	0.03	0.02	0.14	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IUL	0.00	0.13	0.70	0.01	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CBF	0.00	0.13	0.02	0.64	0.20	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RUL	0.00	0.10	0.02	0.02	0.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
APS	0.00	0.12	0.02	0.03	0.20	0.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RST	0.00	0.17	0.01	0.04	0.12	0.01	0.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LWL	0.01	0.14	0.07	0.01	0.18	0.01	0.00	0.59	0.00	0.00	0.01	0.00	0.00	0.00
SUL	0.00	0.13	0.03	0.03	0.24	0.01	0.00	0.00	0.56	0.00	0.00	0.00	0.00	0.00
URC	0.00	0.16	0.05	0.01	0.17	0.01	0.00	0.00	0.00	0.60	0.00	0.00	0.00	0.00
RTF	0.00	0.13	0.06	0.02	0.20	0.01	0.00	0.00	0.00	0.00	0.58	0.00	0.00	0.00
RPF	0.00	0.57	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.00	0.00
OCL	0.00	0.16	0.01	0.03	0.23	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.00
MUL	0.00	0.17	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.58

(c) pLSA

TFIDF	PFL	GSL	IUL	CBF	RUL	APS	RST	LWL	SUL	URC	RTF	RPF	OCL	MUL
PFL	0.00	0.51	0.03	0.06	0.37	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
GSL	0.00	0.69	0.02	0.04	0.23	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IUL	0.00	0.32	0.41	0.02	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CBF	0.00	0.28	0.03	0.39	0.28	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RUL	0.00	0.20	0.02	0.03	0.74	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
APS	0.00	0.27	0.02	0.04	0.29	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RST	0.00	0.39	0.03	0.04	0.24	0.02	0.27	0.00	0.00	0.00	0.01	0.00	0.00	0.00
LWL	0.01	0.30	0.08	0.03	0.30	0.00	0.00	0.27	0.00	0.01	0.01	0.00	0.00	0.00
SUL	0.01	0.21	0.04	0.04	0.37	0.01	0.00	0.00	0.32	0.00	0.00	0.00	0.00	0.00
URC	0.00	0.39	0.04	0.03	0.26	0.00	0.00	0.00	0.00	0.27	0.00	0.00	0.00	0.00
RTF	0.00	0.33	0.04	0.06	0.31	0.01	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00
RPF	0.00	0.43	0.14	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.00	0.00
OCL	0.00	0.32	0.04	0.03	0.29	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.30	0.00
MUL	0.00	0.42	0.00	0.00	0.42	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00

(b) TF-IDF

LDA	PFL	GSL	IUL	CBF	RUL	APS	RST	LWL	SUL	URC	RTF	RPF	OCL	MUL
PFL	0.43	0.32	0.03	0.03	0.17	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GSL	0.00	0.78	0.03	0.02	0.16	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IUL	0.00	0.23	0.58	0.02	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CBF	0.00	0.21	0.03	0.58	0.17	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RUL	0.00	0.15	0.02	0.02	0.80	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
APS	0.00	0.21	0.03	0.02	0.18	0.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RST	0.00	0.30	0.03	0.02	0.18	0.03	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LWL	0.00	0.23	0.08	0.02	0.20	0.00	0.00	0.46	0.00	0.00	0.00	0.00	0.00	0.00
SUL	0.00	0.17	0.04	0.03	0.24	0.01	0.00	0.00	0.52	0.00	0.00	0.00	0.00	0.00
URC	0.00	0.33	0.03	0.02	0.18	0.01	0.00	0.00	0.00	0.44	0.00	0.00	0.00	0.00
RTF	0.00	0.26	0.04	0.04	0.24	0.01	0.00	0.02	0.00	0.01	0.39	0.00	0.00	0.00
RPF	0.00	0.71	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.00
OCL	0.00	0.16	0.04	0.02	0.28	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00
MUL	0.00	0.67	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17

(d) LDA

3.3 Place2Vec方法

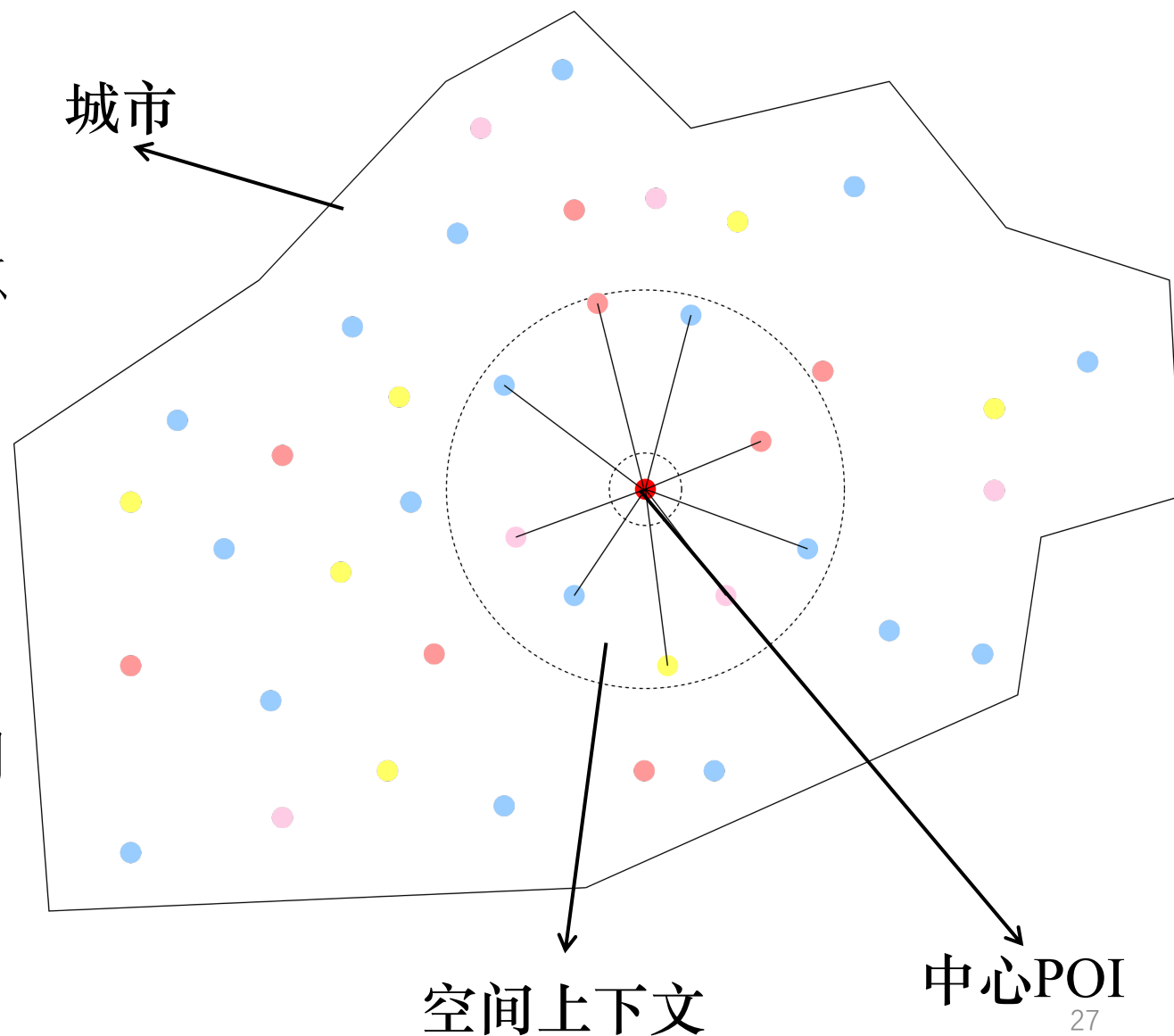


➤ 简单空间上下文

通过选取中心POI点周围的一定距离的点来作为一个POI点的上下文是否可行？

如图所示，以某个POI点作为中心点，以半径 r 的圆形区域内的其他POI点即为其空间上下文。

如果仅仅这样构建POI点的上下文会遇到什么问题？还可以怎样构建？



3.3 Place2Vec方法

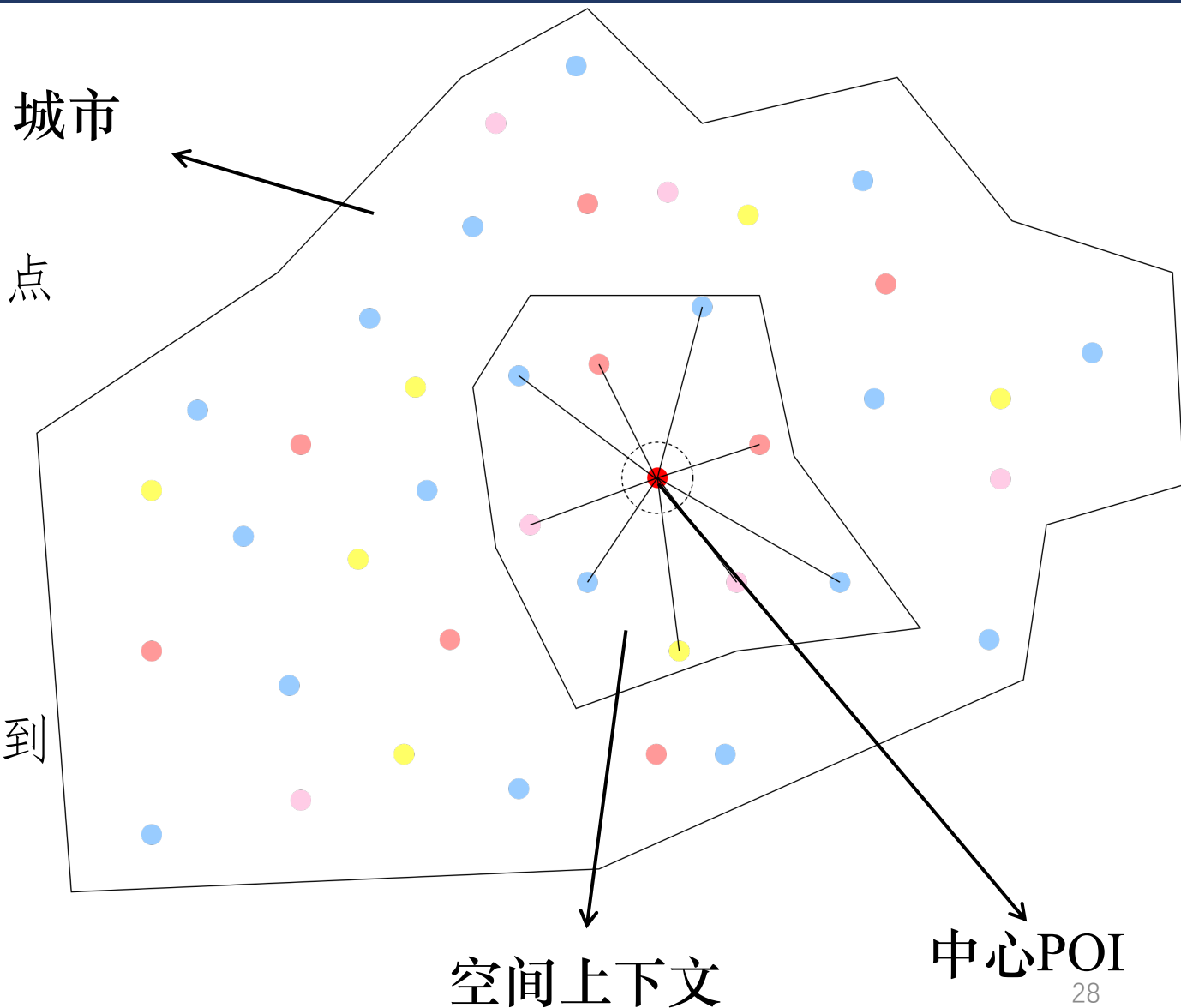


➤ 简单空间上下文

通过选取中心POI点周围的一定数量的点来作为一个POI点的上下文是否可行？

如右图，以一个POI点为中心点，以周围9个POI点为上下文信息。

如果仅仅这样构建POI点的上下文会遇到什么问题？还可以怎样构建？



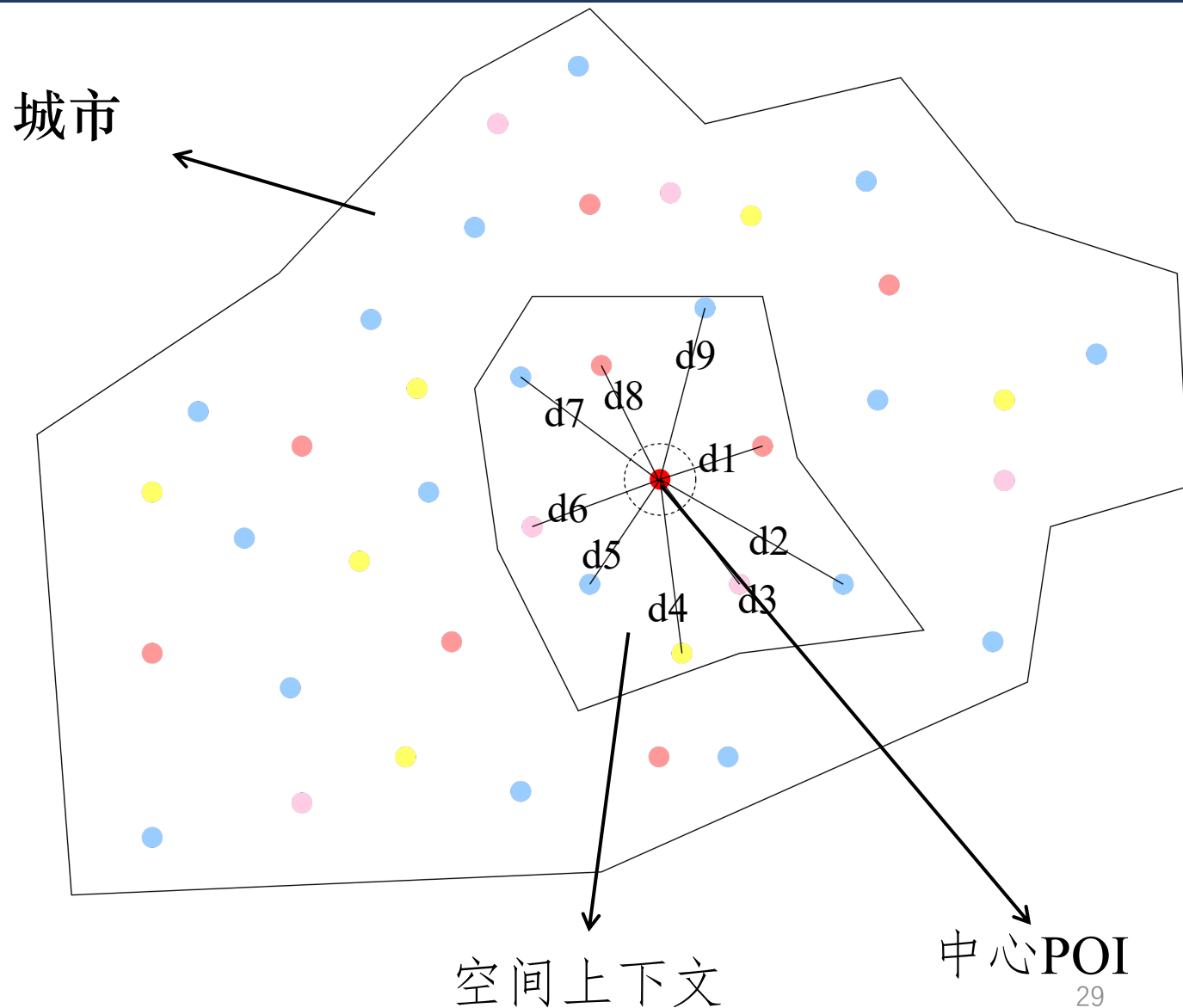
3.3 Place2Vec方法



➤ 增强型空间上下文

- 距离衰减：距离越远的 POI 对中心的影响越小。

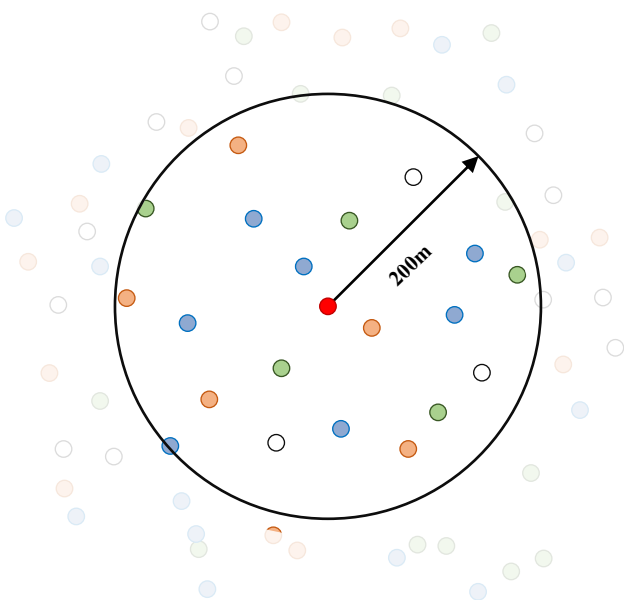
- 人类活动强度（Check-in 数量）：代表 POI 的热度或影响力。



3.3 Place2Vec方法



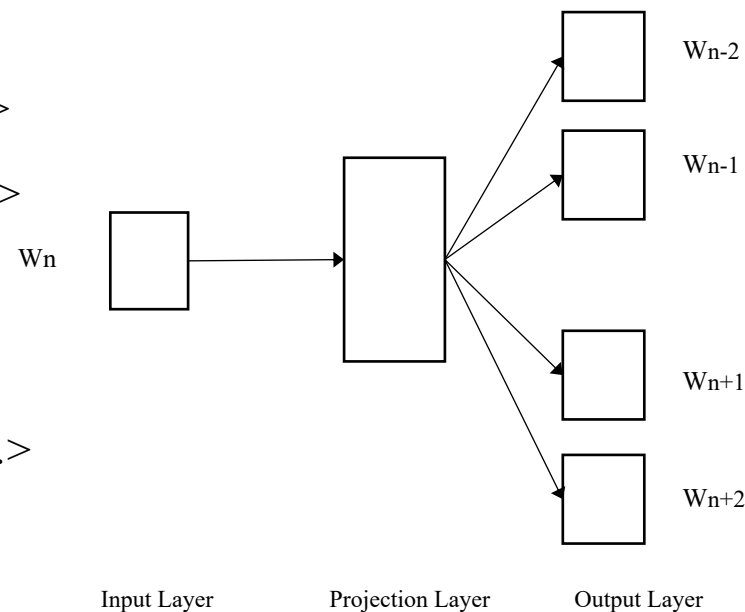
➤ Place2Vec



搜索相邻一定距离POI

<风景名胜 公园 火车站 工厂 产业园区 收费站>
<政府机关 公检法机构 学校 文化宫 科研场所>
.....
<文化用品店 中餐厅 体育用品店 娱乐场所 专卖店>

生成训练词对语料库



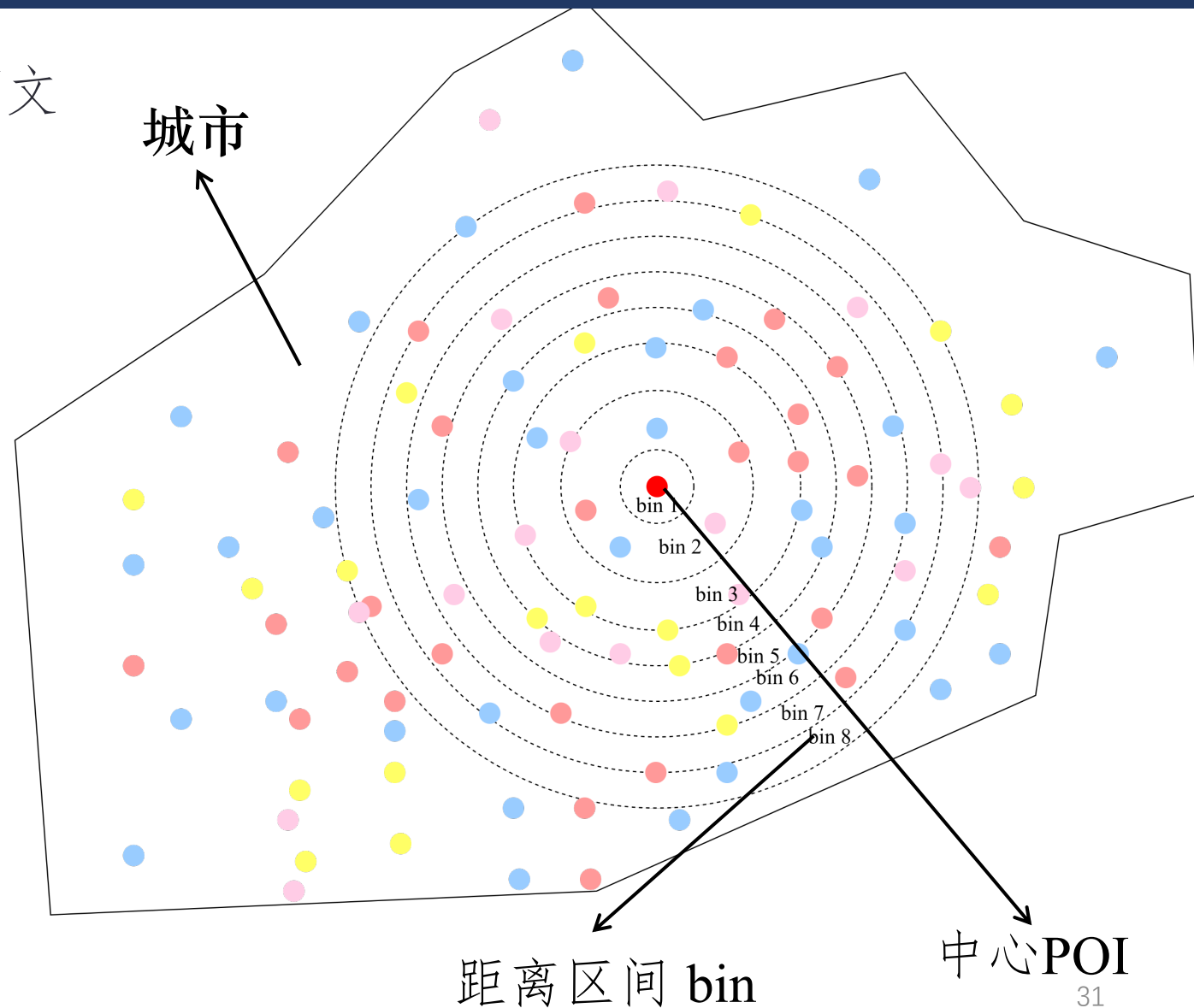
Skip-gram算法

3.3 Place2Vec方法



➤ 信息理论 + 距离滞后增强空间上下文

- 划分距离区间：通过等距离或指定数值划分距离区间 (**bin**)
- 统计同一距离区间 (**bin**) 中的共线关系：统计**POI**类型、人气、出现频率等。
- 计算独特性：出现概率越低的 **POI** 类型，其独特性越高；反之则越低。
- 综合考虑独特性与共现特点，训练模型。

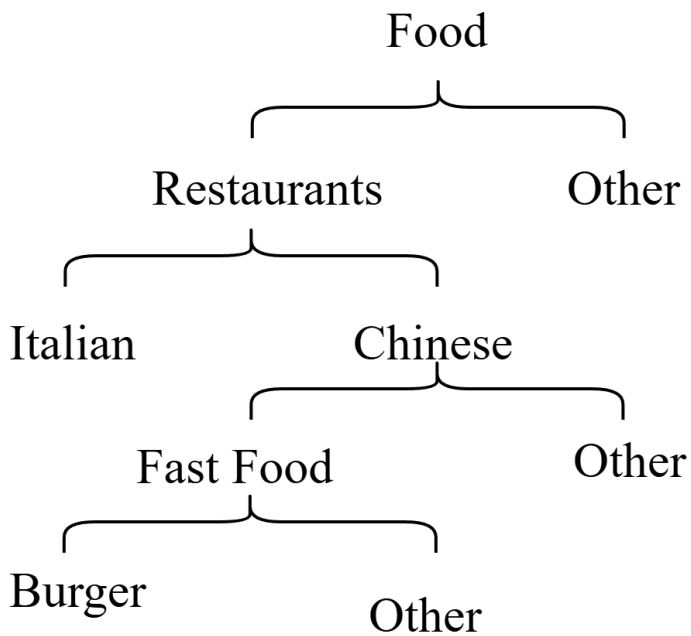


3.3| Place2Vec方法



➤ 基于 Yelp 层级结构的语义相似度评估

Yelp 的 POI 类型是以树状层级结构 (hierarchy) 组织的。



树状层级结构

基于层级结构评估的平均倒数排名

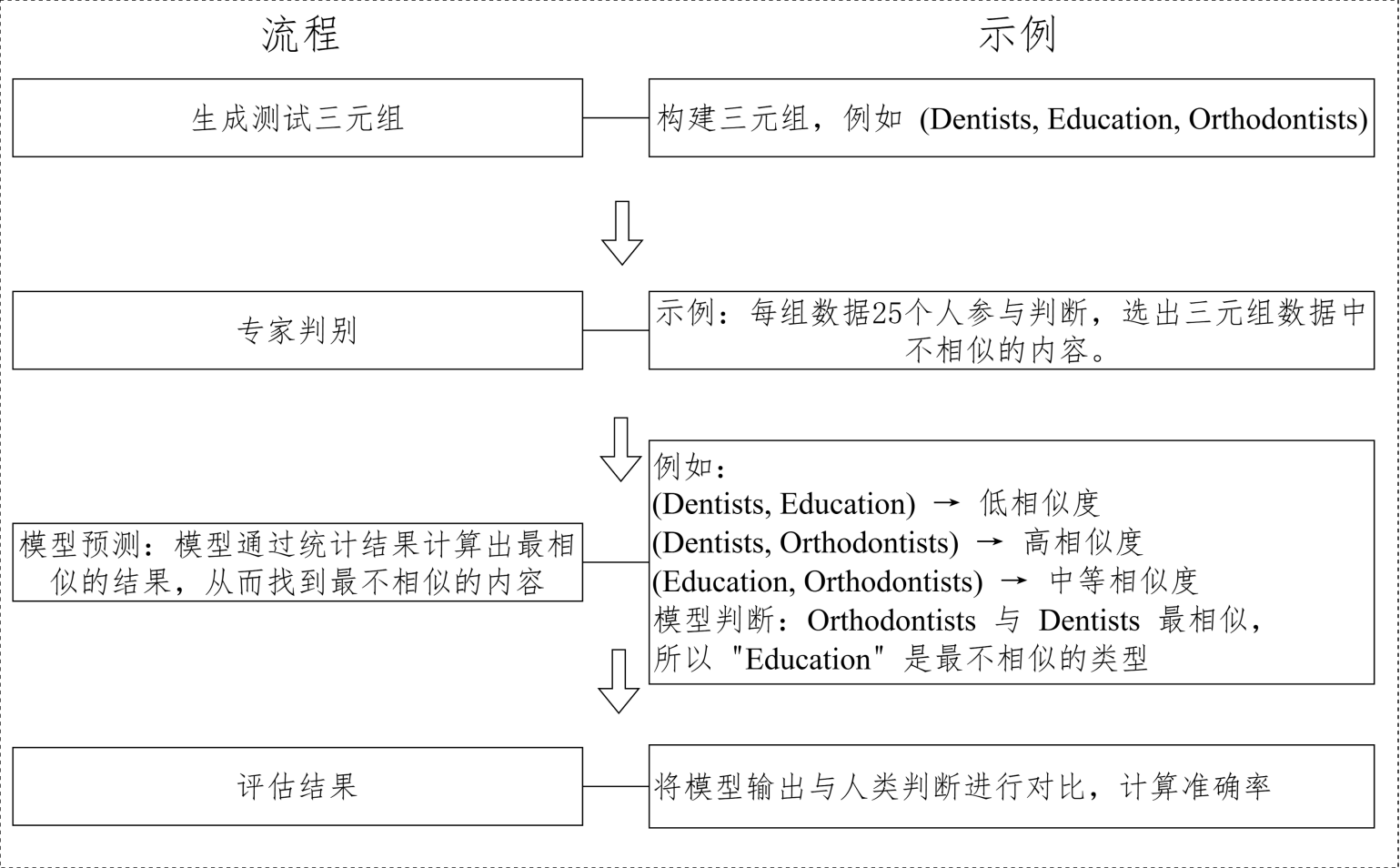
Model	SimWP	SimLC	SimLin (ICSánchez)	SimLin (ICSeco)	SimJC (ICSánchez)	SimJC (ICSeco)
Skip-gram	0.288	0.321	0.354	0.334	0.349	0.333
SC _{naive}	0.412	0.398	0.474	0.442	0.455	0.478
SC _{checkin}	0.385	0.387	0.448	0.428	0.452	0.474
SC _{distance}	0.381	0.396	0.458	0.426	0.443	0.458
SC _{combined}	0.42	0.418	0.478	0.435	0.462	0.482
SCITDL	0.447	0.431	0.498	0.479	0.487	0.483

基于 Yelp 分类体系，我们采用六种语义相似度指标，评估模型在不同信息内容定义下的语义匹配能力。

3.3| Place2Vec方法



➤ 二元 HIT 评估



二元 HIT 评估方案流程图

Model	Accuracy
Word2Vec	0.75
SC _{naive}	0.85
SC _{checkin}	0.7
SC _{distance}	0.875
SC _{combined}	0.875
SC _{ITDL}	0.95

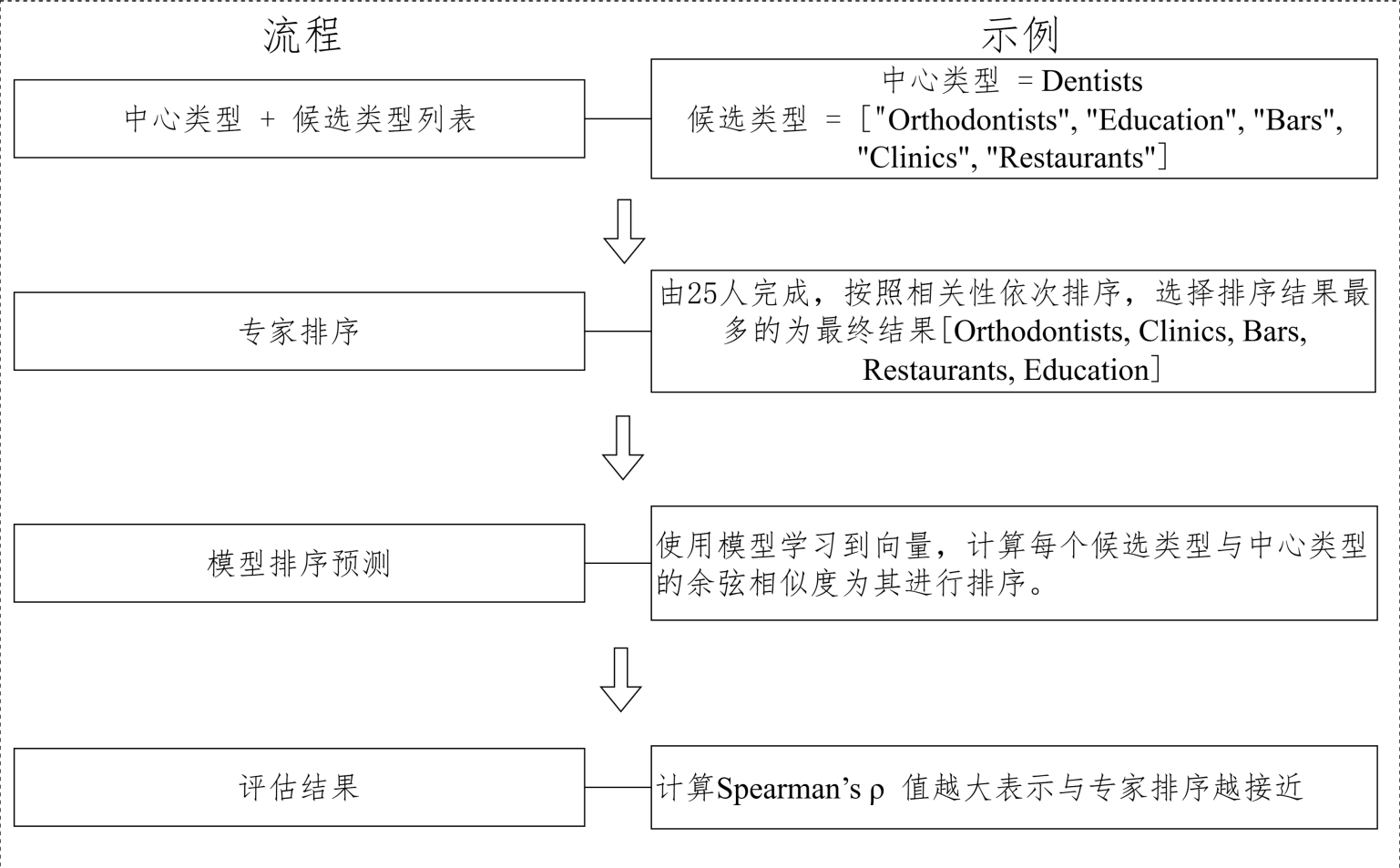
评估结果

准确率 = 判断正确的三元组 / 总数

3.3| Place2Vec方法



➤ 基于排序的 HIT 评估



基于排序的 HIT 评估方案流程图

Model	ρ
SC _{naive}	0.56
SC _{checkin}	0.57
SC _{distance}	0.51
SC _{combined}	0.7
SC _{ITDL}	0.7

评估结果 ρ

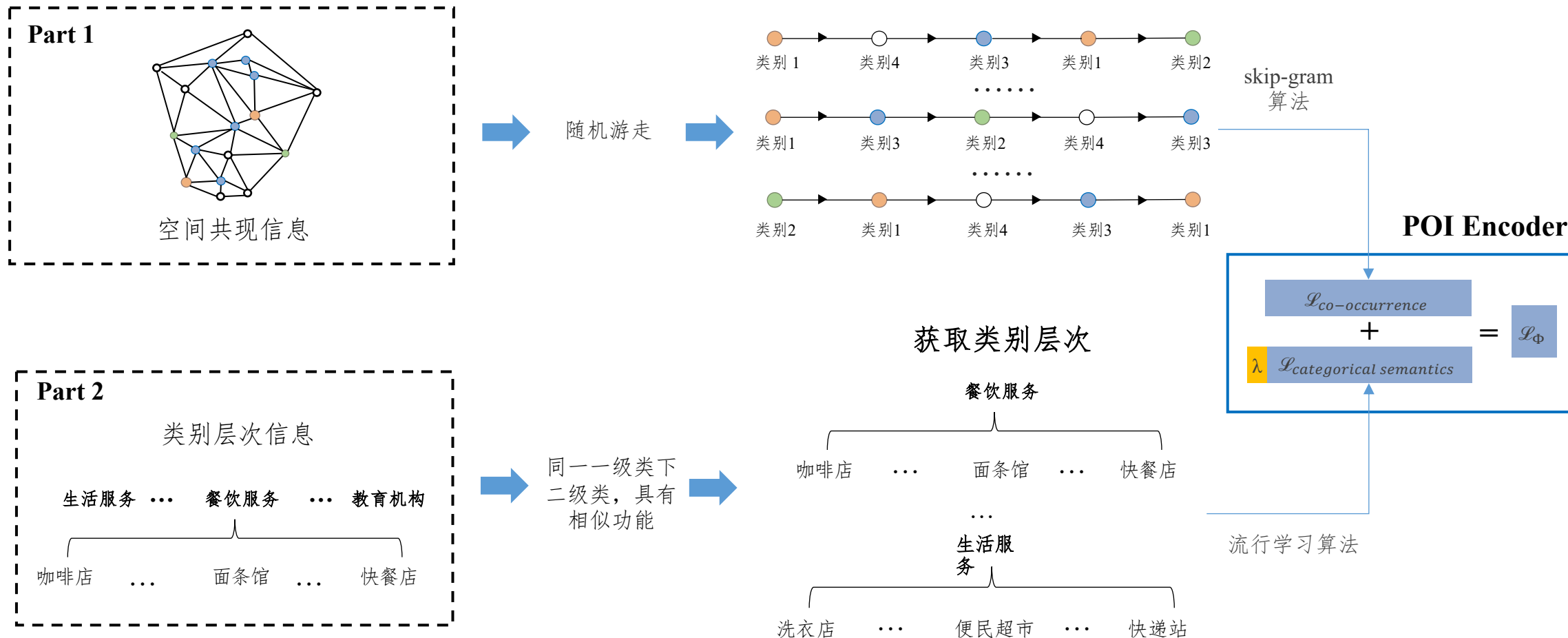
$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

d_i 专家打分与模型预测之间排名差异
 n 进行排序的数据样本

3.4 Semantic2Vec方法



➤ Semantic2Vec



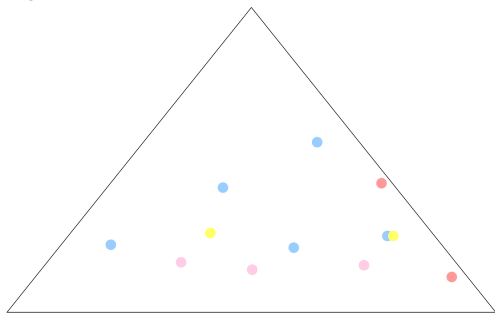
Semantic2Vec方法流程图

3.4 Semantic2Vec方法

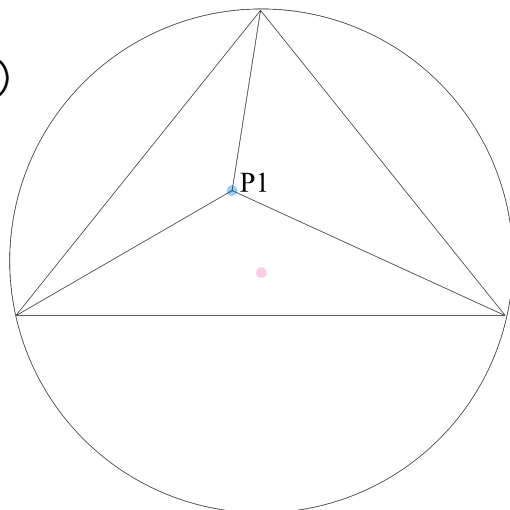


➤ Delaunay 三角剖分

①

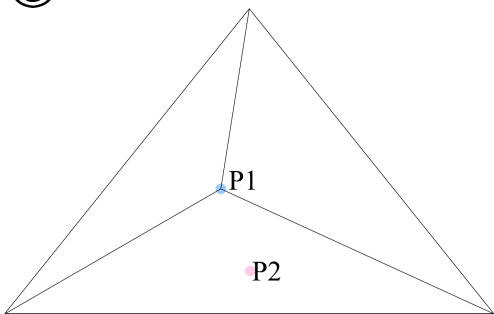


②



③

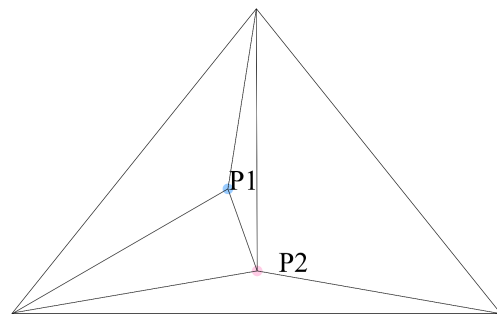
根据点云构建
超级大三角形



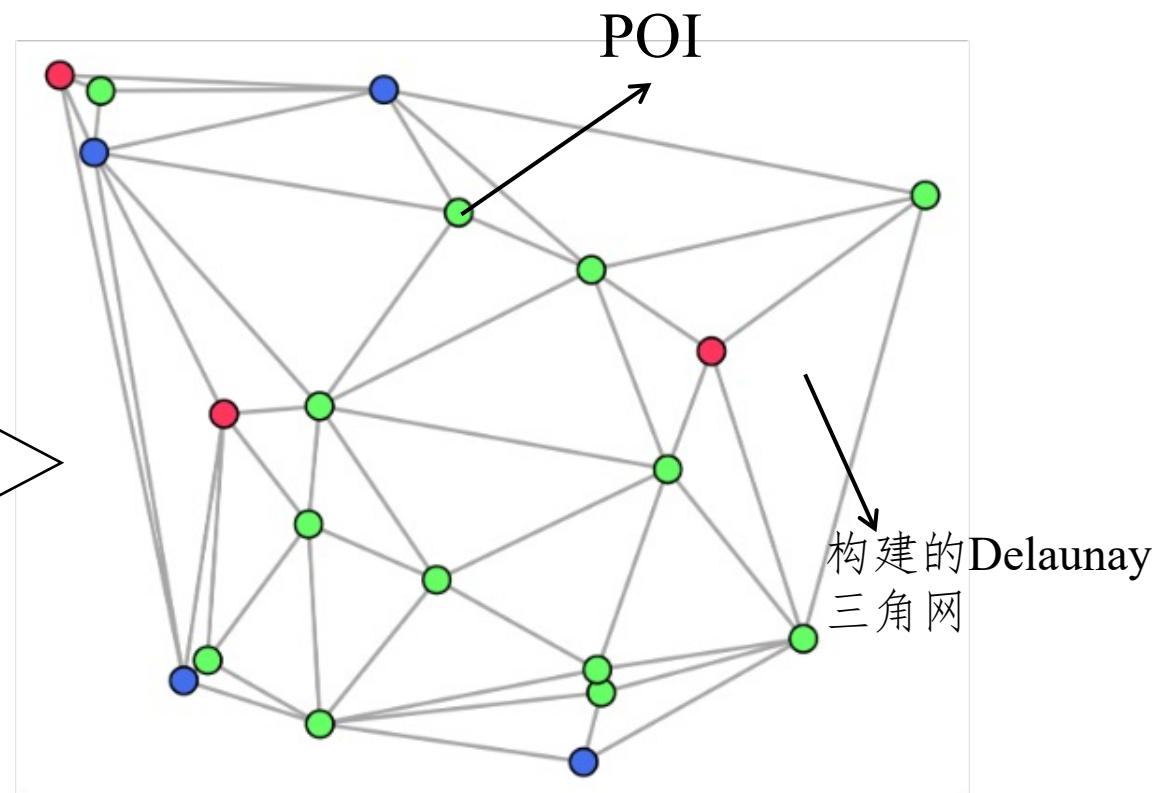
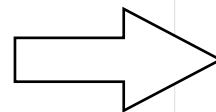
重复步骤插入下一个点

④

计算半径并绘制外接圆



擦除空边，直到
所有点完成插入



3.4 Semantic2Vec方法



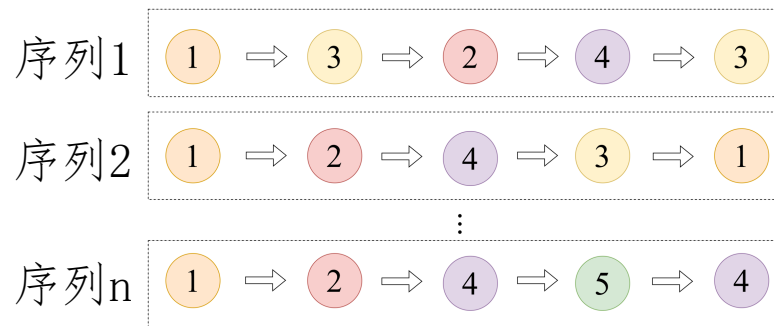
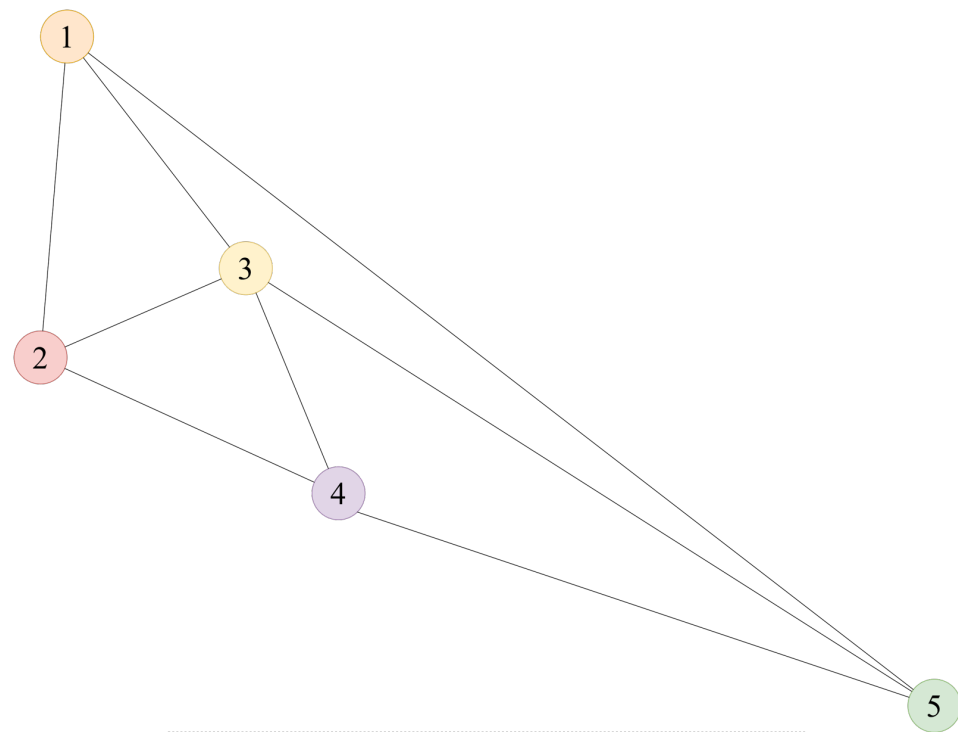
➤ Node2Vec

Node2Vec: 随机游走采样 + Skip-gram 语言模型

通过随机游走生成多个节点序列，将这些序列看作“句子”，每个节点看作“词”，通过Skip-gram进行模型训练。

过渡概率：当前点到下一个候选点的概率。在实际使用过程中，可以通过设置偏置来调整随机游走的概率分布，以满足特定需求。

- 距离衰减偏置
- 局部与长程共现平衡偏置
- 区域内与跨区域共现平衡偏置



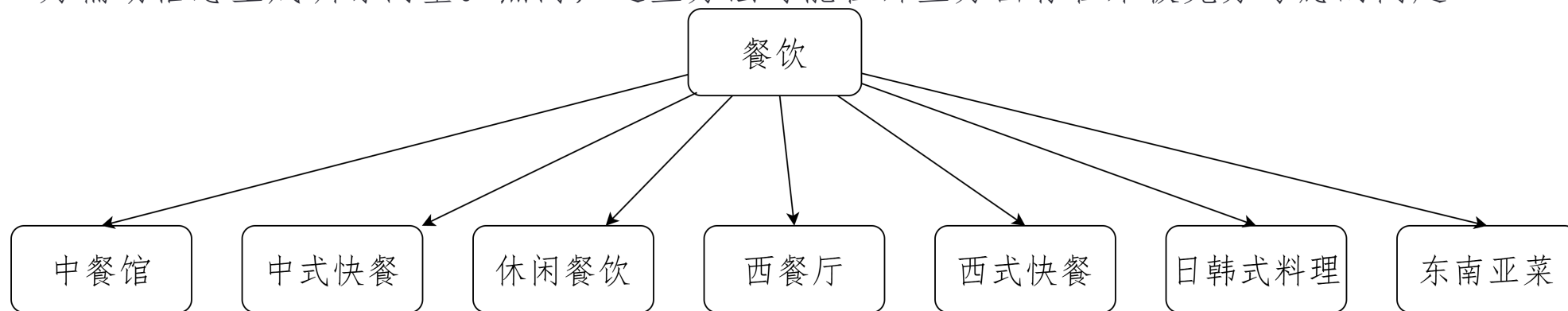
模型随机游走示例

3.4 Semantic2Vec方法



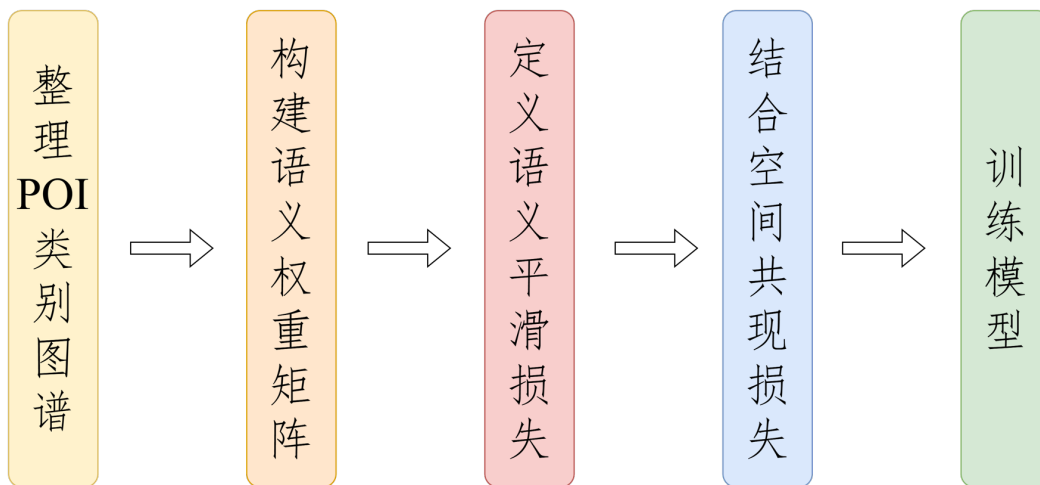
➤ 融入类别语义信息

在前述研究中，部分方法以交通分析区为单位进行表征学习，或借助距离区间的划分作为辅助信息生成训练向量。然而，这些方法可能在哪些方面存在未被充分考虑的问题？



拉普拉斯特征映射 (Laplacian Eigenmaps, LE) :

$$\mathcal{L}_{\text{categorical semantics}} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \|\mathbf{c}_i - \mathbf{c}_j\|_2^2 \cdot w_{ij}$$



①两个POI经常在地理位置上一起出现，它们的嵌入就会更接近。

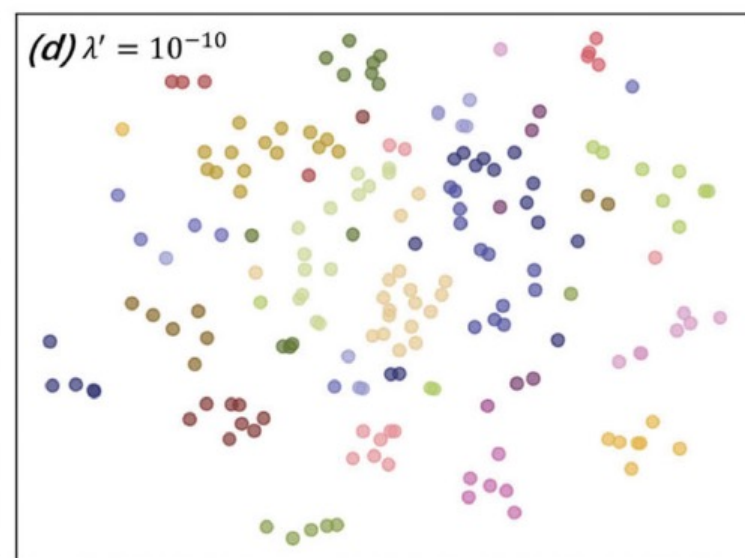
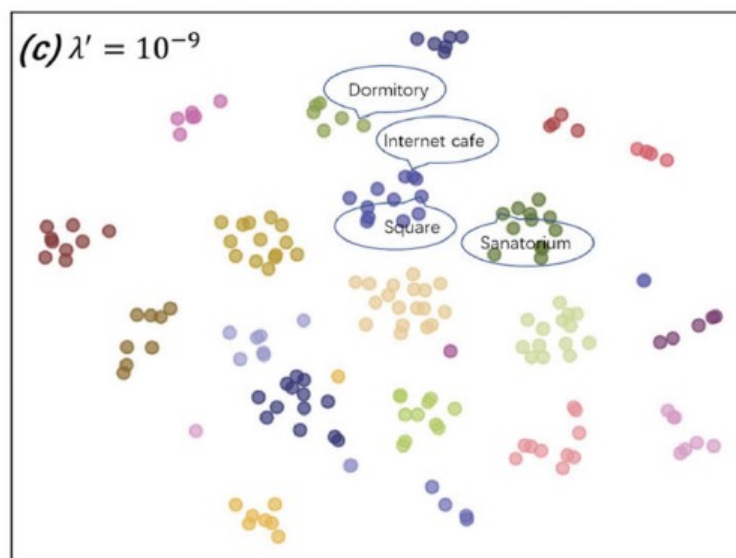
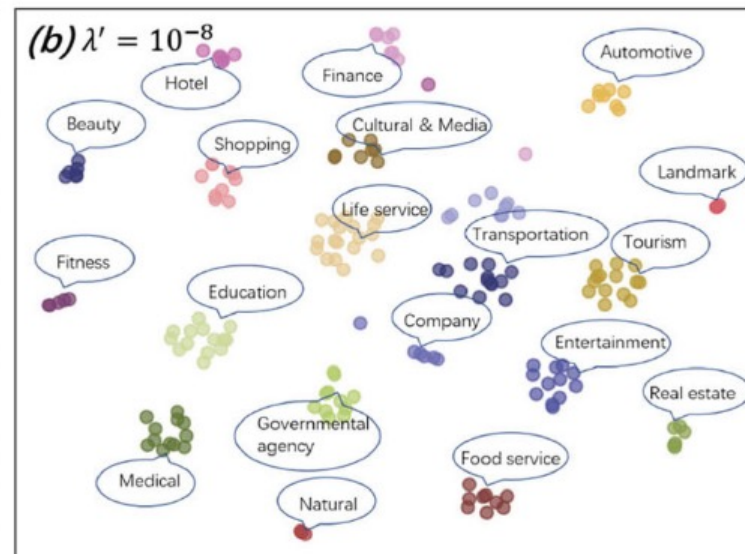
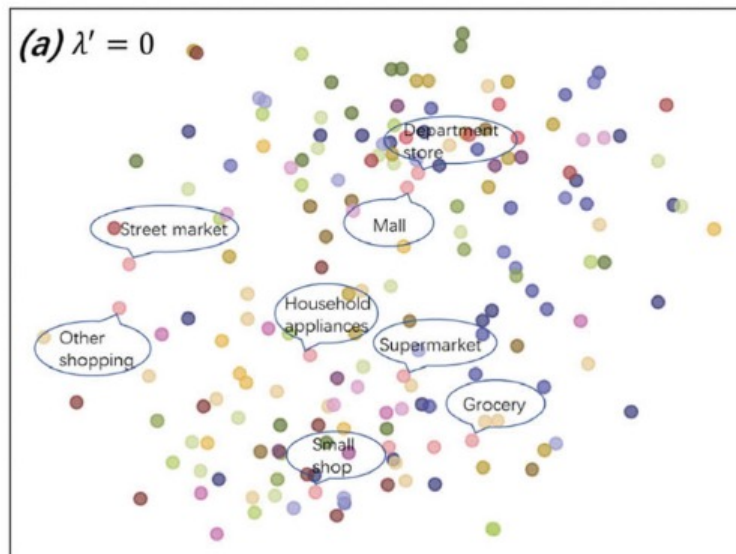
②“超市”、“便利店”虽然可能出现在不同地方，但功能相似

3.4 Semantic2Vec方法



➤ 结果验证

将通过Semantic2Vec方法的结果通过降维到2维平面，同一一级类别的二级类别使用同一颜色。



(a) 完全基于空间共现信息

(b) 同时考虑空间共现信息和类别语义信息

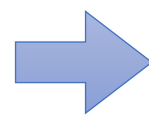
(c) 和 (d) 则是在 (b) 的基础上逐渐削弱对语义信息的约束。

3.5 | HGI方法



➤ 独立POI

尽管我们已知不同类型的POI通常具有不同的功能，但一个值得深入探讨的问题是：同一类别的POI是否在功能或特征上完全一致？如果不一致，其差异又受到哪些因素的影响？



我们设想这样一个场景：小学校门外的小卖部更多售卖玩具和零食，而乡村小卖部则相对较少。



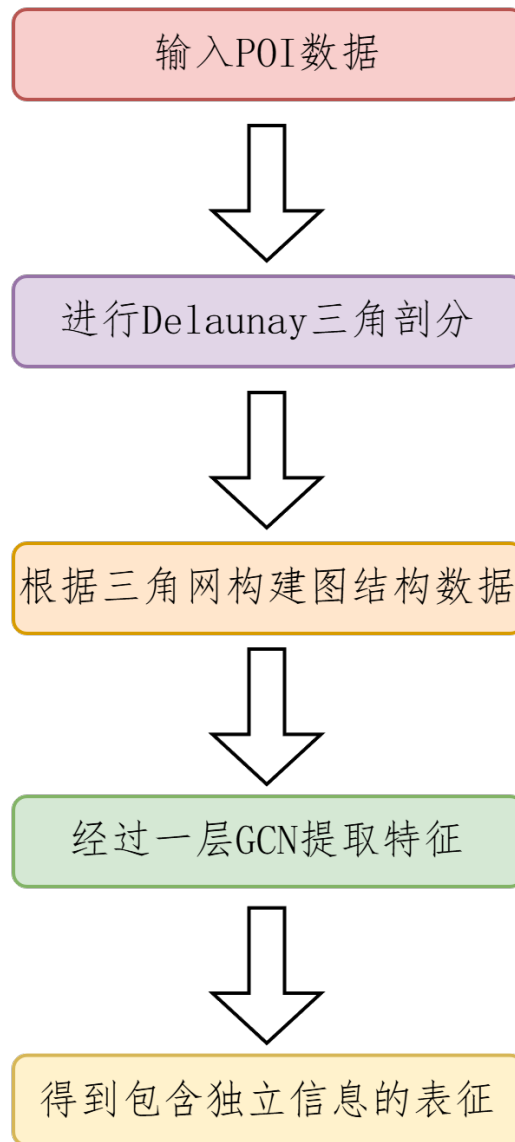
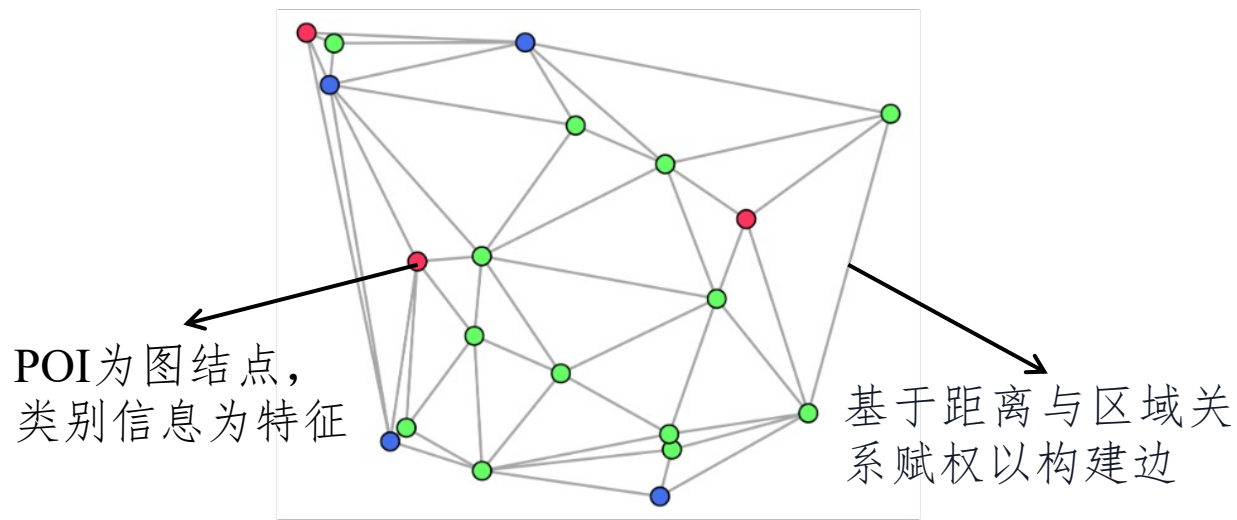
3.5 HGI方法



➤ 独立POI

为什么使用图结构来表示独立POI?

- (1) 图对于 POI 而言在空间上更加紧凑;
- (2) 图在仿射和旋转变换下具有鲁棒性,
- (3) 消息传递机制可以自然地模拟 POI 之间的交互, 从而捕捉单个 POI 的空间上下文。



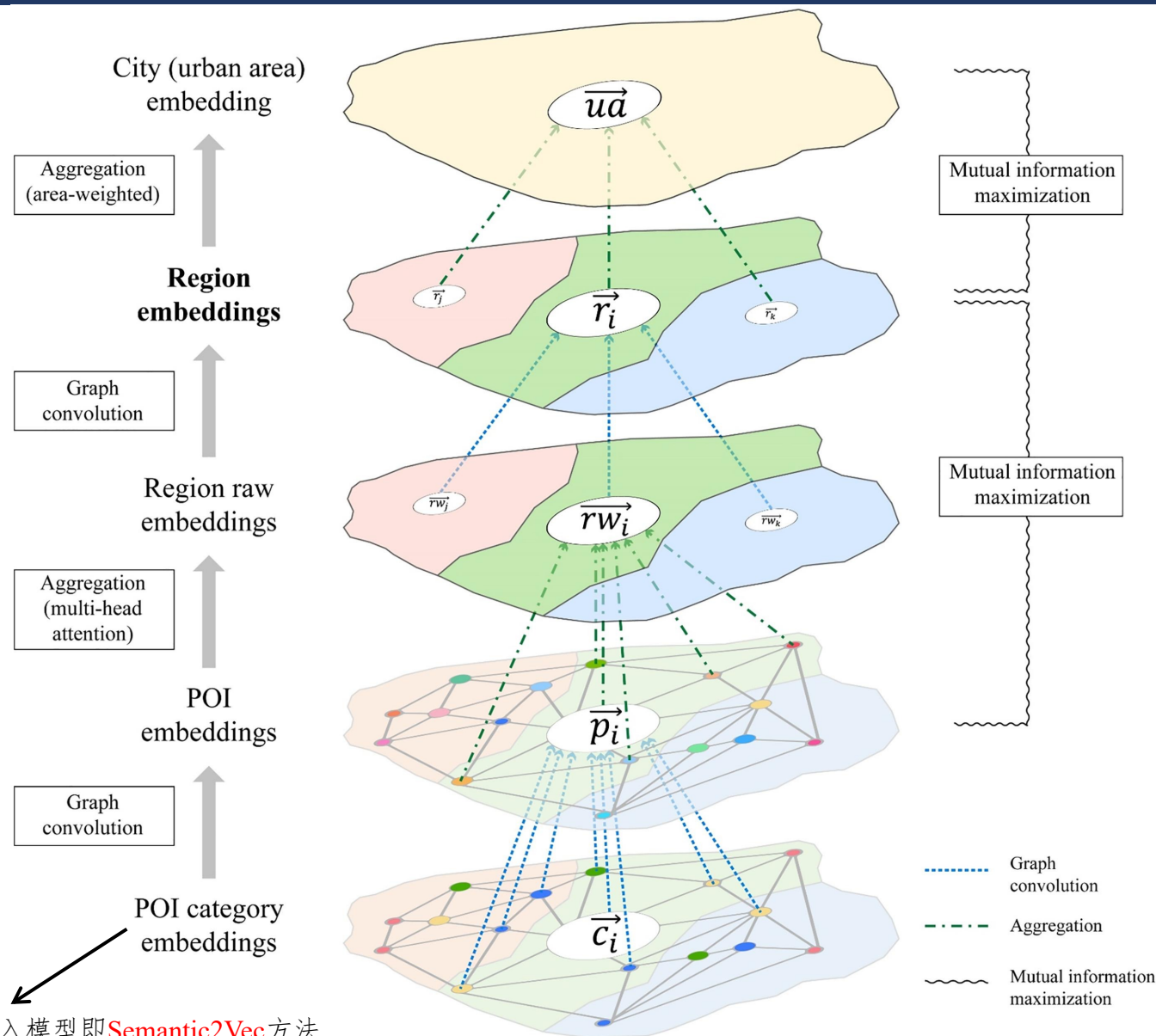
3.5 HGI方法



➤ HGI

Hierarchical Graph Infomax (HGI)方法由Huang等人在2023年提出。

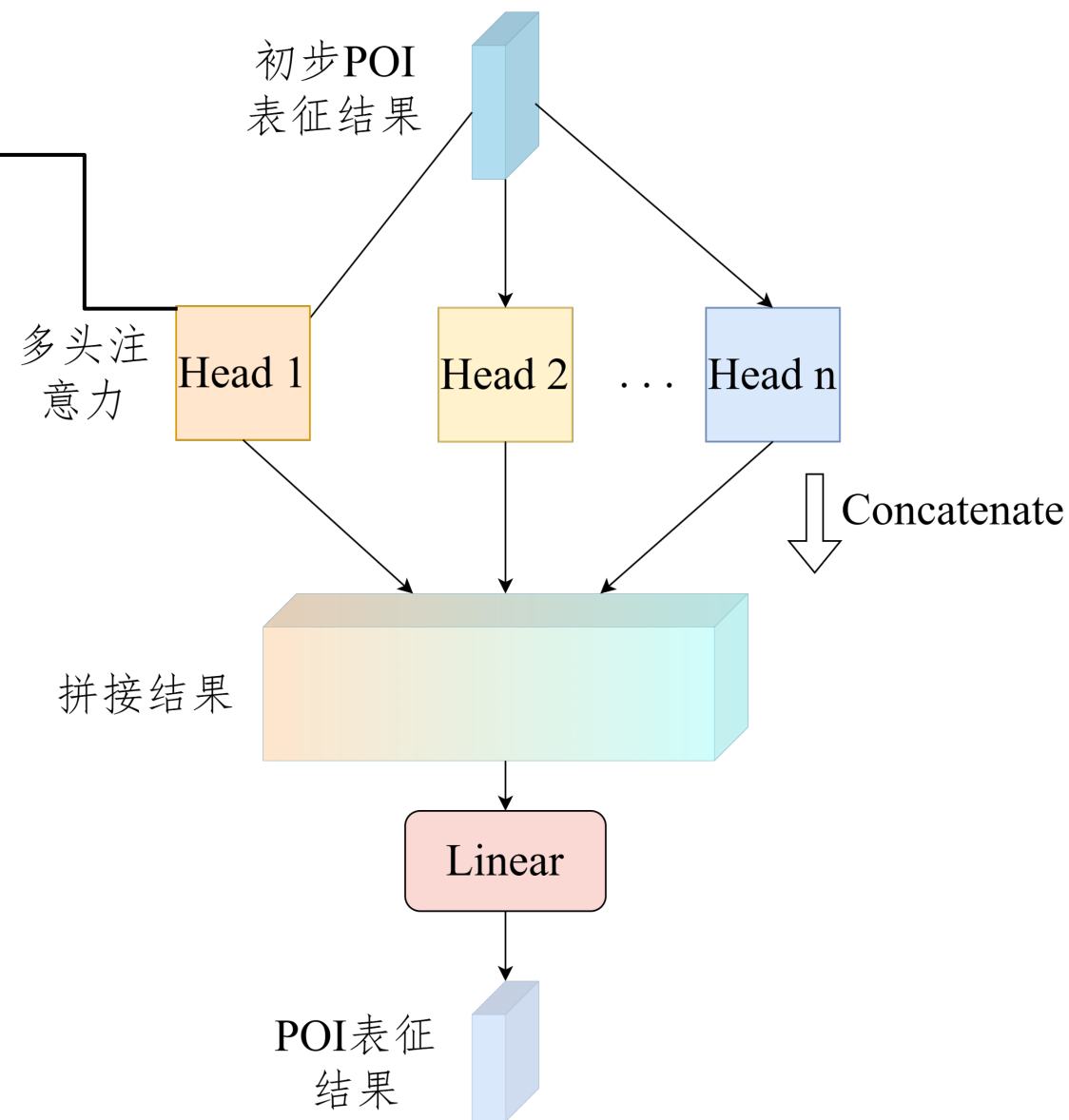
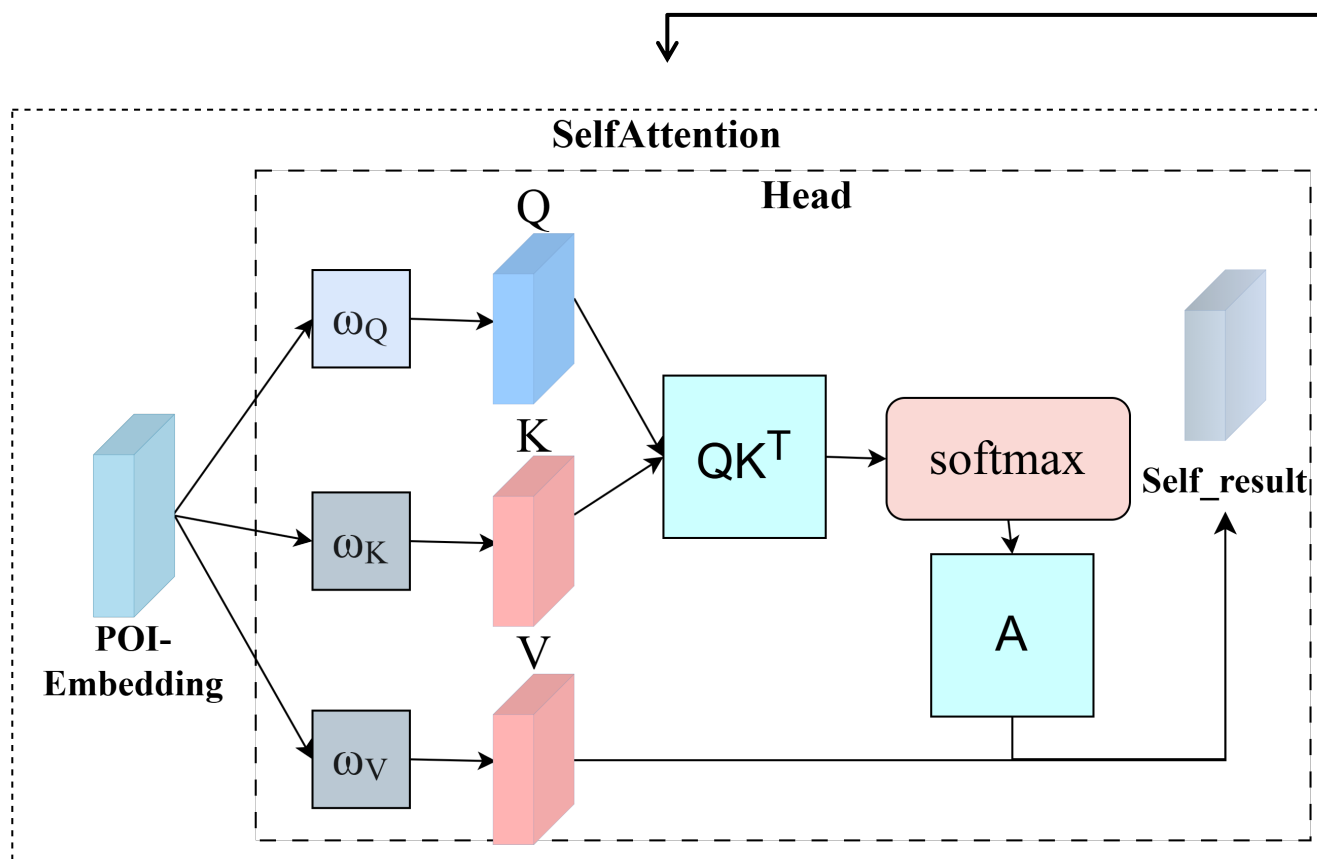
- (1) 预先训练一个 POI 类别编码器来生成初始 POI 特征
- (2) 使用Delaunay 三角剖分 (DT) 构建 POI 的图结构，并将图卷积编码器应用于 POI 图
- (3) 利用基于多头注意机制的聚合函数来聚合 POI 嵌入
- (4) 利用基于多头注意机制的聚合函数来聚合 POI 嵌入



3.5 HGI方法



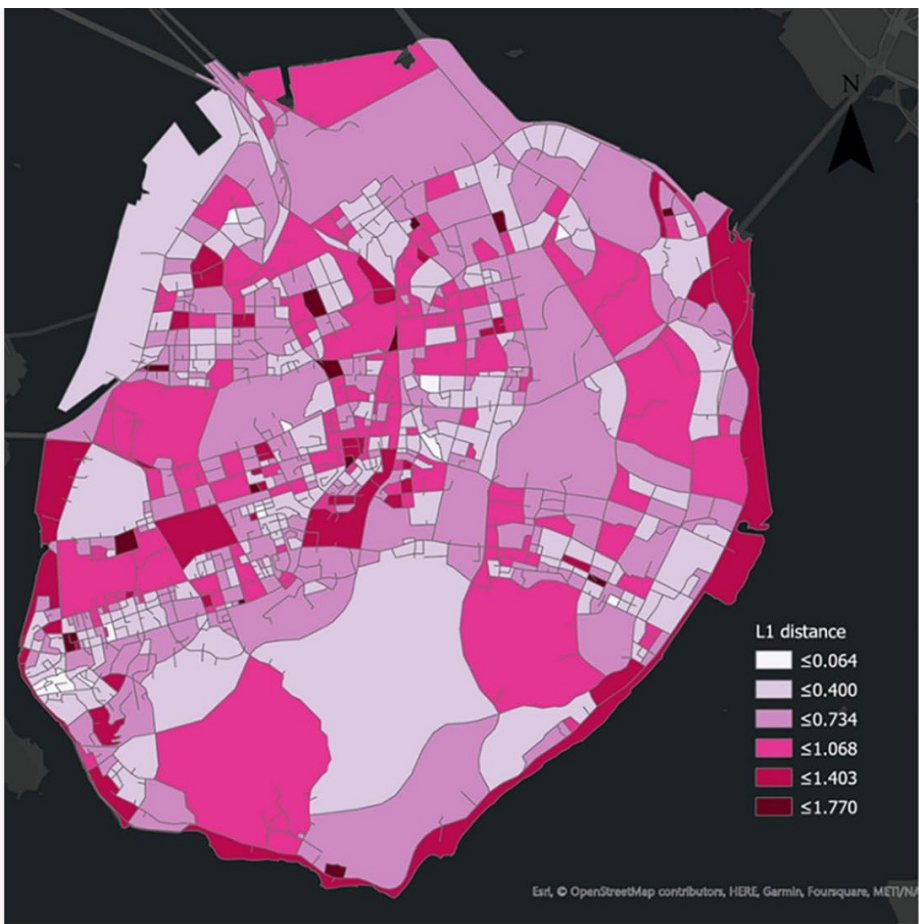
➤ 多头注意力机制



3.5 HGI方法



➤ 城市功能区识别



预测与实际城市功能分布的
L1距离误差的空间分布图

Model	Xiamen Island			Shenzhen		
	L1↓	KL↓	Cosine↑	L1↓	KL↓	Cosine↑
HGI	0.727 ± 0.019	0.587 ± 0.033	0.795 ± 0.017	0.936 ± 0.016	0.833 ± 0.025	0.707 ± 0.009
Semantic-s	0.754 ± 0.029	0.619 ± 0.038	0.789 ± 0.014	0.972 ± 0.011	0.877 ± 0.024	0.695 ± 0.009
Semantic-u	0.784 ± 0.024	0.658 ± 0.032	0.770 ± 0.015	0.980 ± 0.015	0.893 ± 0.024	0.693 ± 0.008
Place2Vec	0.835 ± 0.023	0.724 ± 0.037	0.748 ± 0.015	0.982 ± 0.016	0.895 ± 0.023	0.694 ± 0.007
Doc2Vec	0.833 ± 0.027	0.728 ± 0.047	0.749 ± 0.023	0.985 ± 0.012	0.894 ± 0.015	0.692 ± 0.007
LDA	0.896 ± 0.028	0.848 ± 0.051	0.711 ± 0.017	1.000 ± 0.015	0.924 ± 0.022	0.678 ± 0.008

城市功能分布评估结果

HGI模型在功能分布估计中表现最优，尤其在深圳等复杂城市中优势更明显，显示出其在城市功能识别中的强大潜力。



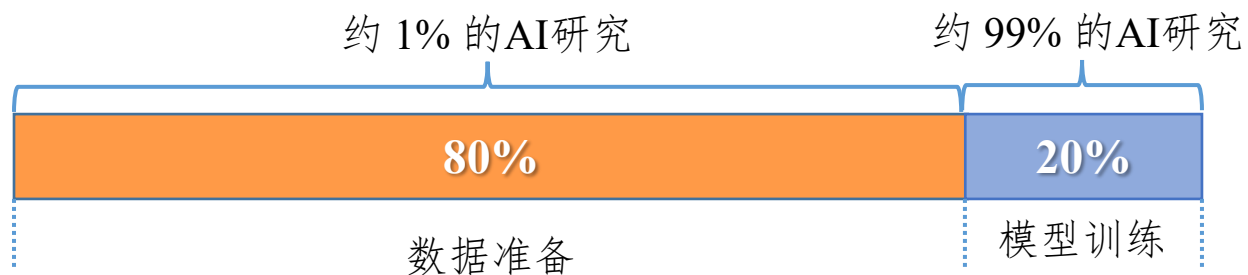
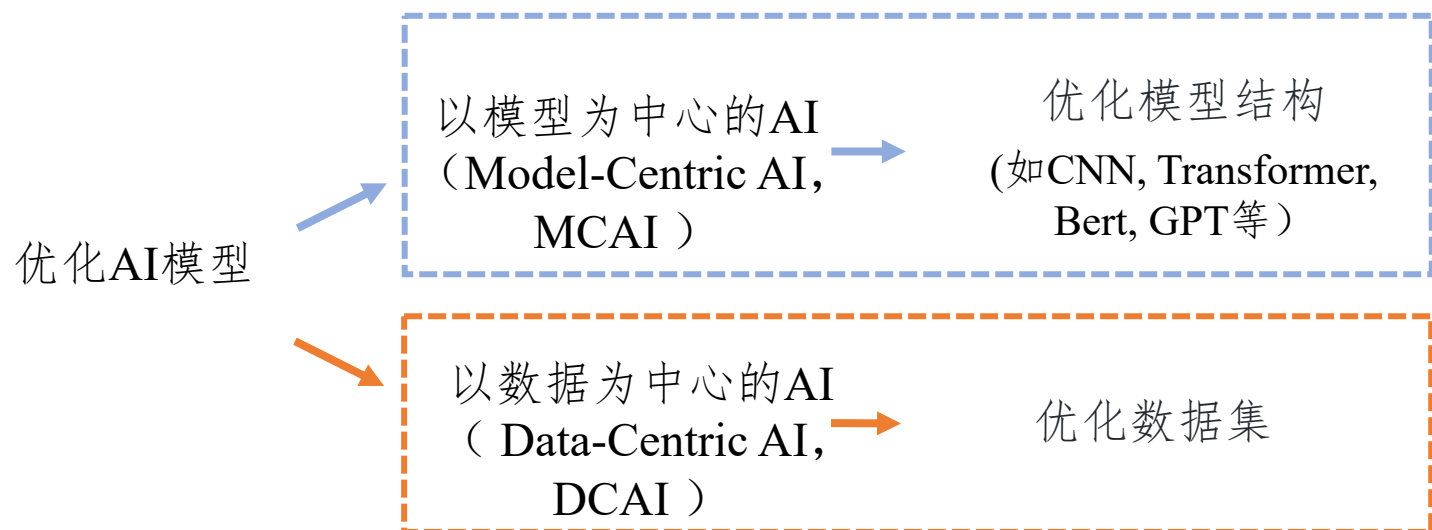
目录



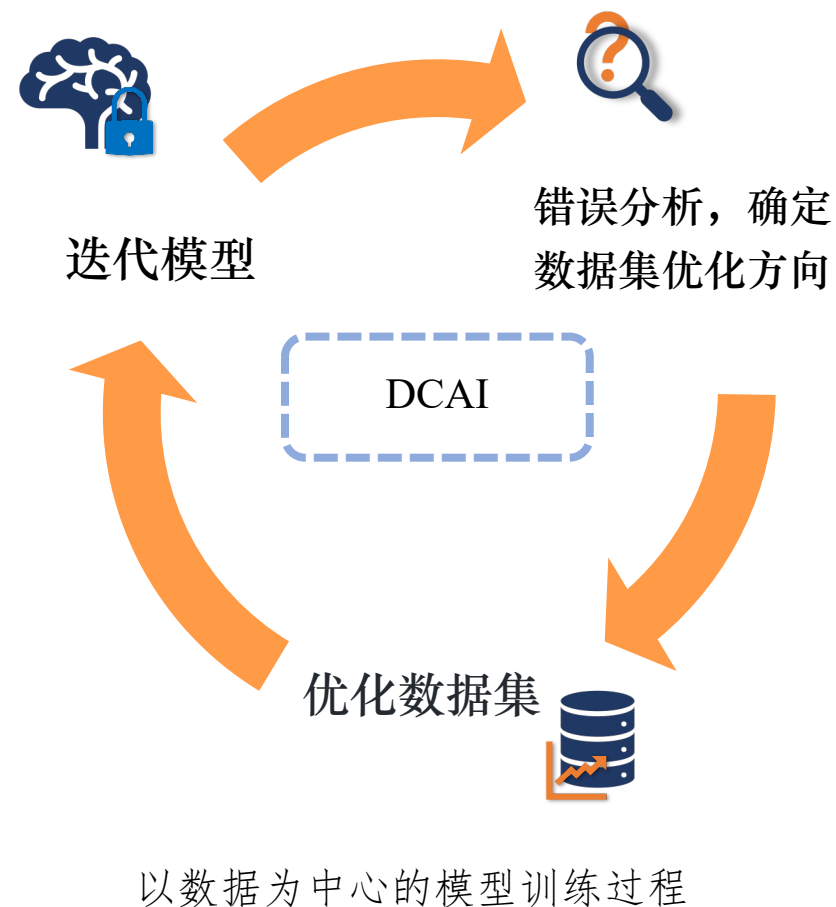
- 1 城市要素的简介
- 2 城市功能结构
- 3 基于POI的城市要素挖掘
- 4 基于多模态数据的表征方法
- 5 总结

- 4.1 “以数据为中心”的土地利用数据集构建框架
(DCAI-CLUD)
- 4.2 全国土地利用数据集质量评价方法
- 4.3 LandGPT: 一个用于地块级土地利用分类的多模态大语言模型
- 4.4 带有思维链的LandAI

4.1 DCAI-CLUD 研究区域与数据



AI模型训练的时间分配

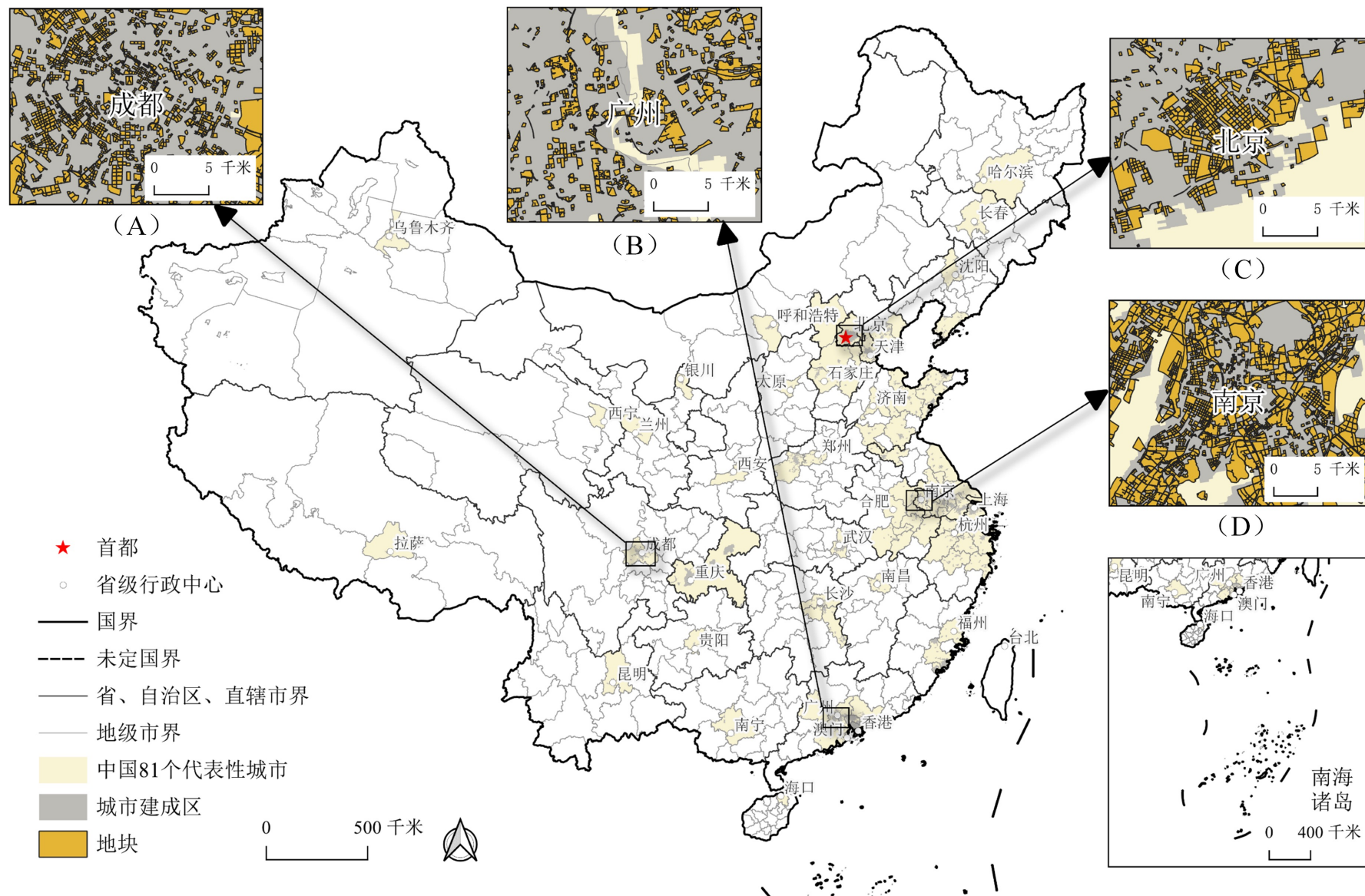


4.1 DCAI-CLUD 研究区域与数据



研究区域

- 中国81个代表性城市
- 具有不同的空间形态、格局和景观
- 总面积 983,215 km²



➤ 研究数据

地块数据：

- 各城市行政区划数据通过阿里云数据可视化平台下载
- 路网数据从OpenStreetMap（OSM）下载
- 参照《公路工程技术标准（JTG B01-2003）》中各级公路路基宽度，对各级路网建立缓冲区后与行政区划数据矢量裁剪得到

土地利用数据标注

- 基于《城市用地分类与规划建设标准（GB 50137-2011）》建立
- 5个一级类别（居住用地、商业服务业设施用地、工业产业用地、公共管理与公共服务设施用地和农业自然用地）和22个二级类别

影像数据

- 下载自谷歌地球地球引擎（Google Earth Engine，GEE）
- 统一重采样为2.5m分辨率

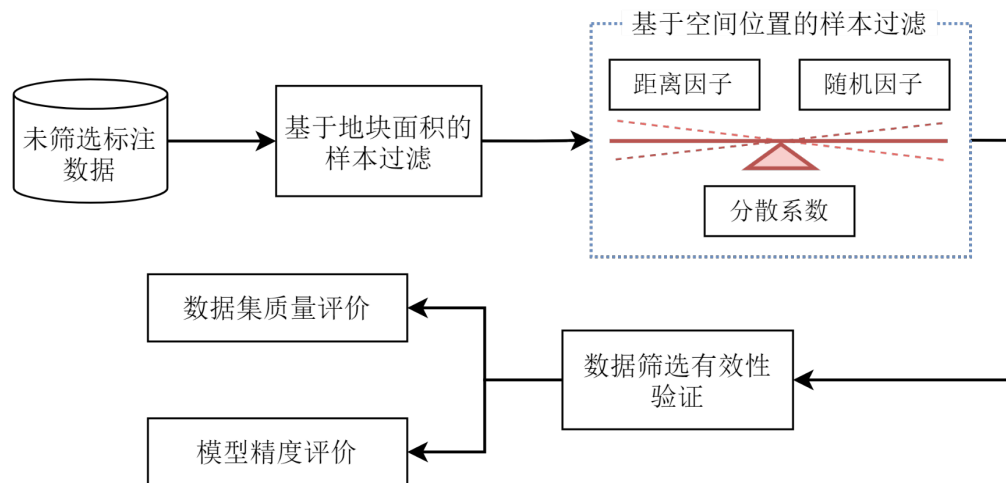
POI数据

- 基于高德地图开放平台提供的API接口，利用网络爬虫技术共收集得到6090.13万条数据
- 包含23个一级类别和261个二级类别。

5.1 DCAI-CLUD 地块属性数据筛选与预标注可行性研究

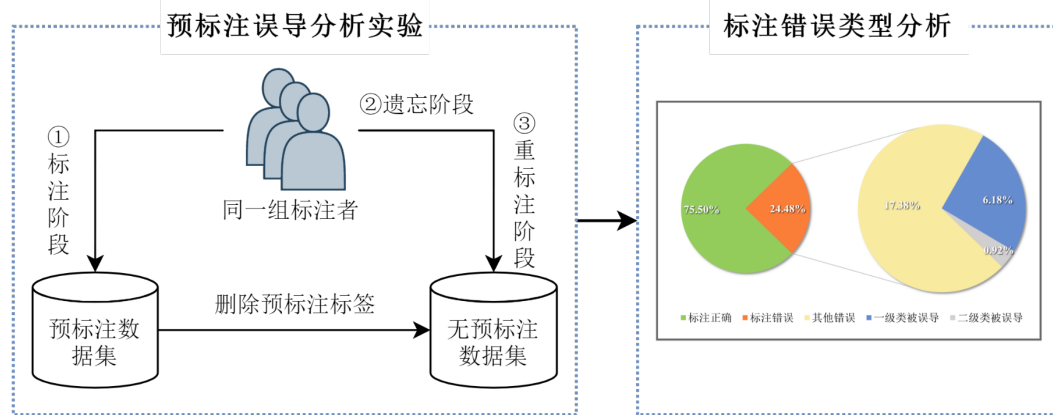


1. 基于地块面积和位置的数据筛选



- 一种基于地块自身属性的数据筛选方法。分析地块的面积和空间位置与数据质量的关系。通过对不同筛选方法进行质量评价，选定最佳的数据筛选参数，最后分别训练模型来验证筛选的有效性。

2. 模型预标注的误导分析



- 设计了一种量化模型预标注对标注者产生的误导作用的方法，用来评估预标注可能的副作用。

4.1 DCAI-CLUD 地块属性数据筛选与预标注可行性研究



➤ 基于地块面积的数据筛选

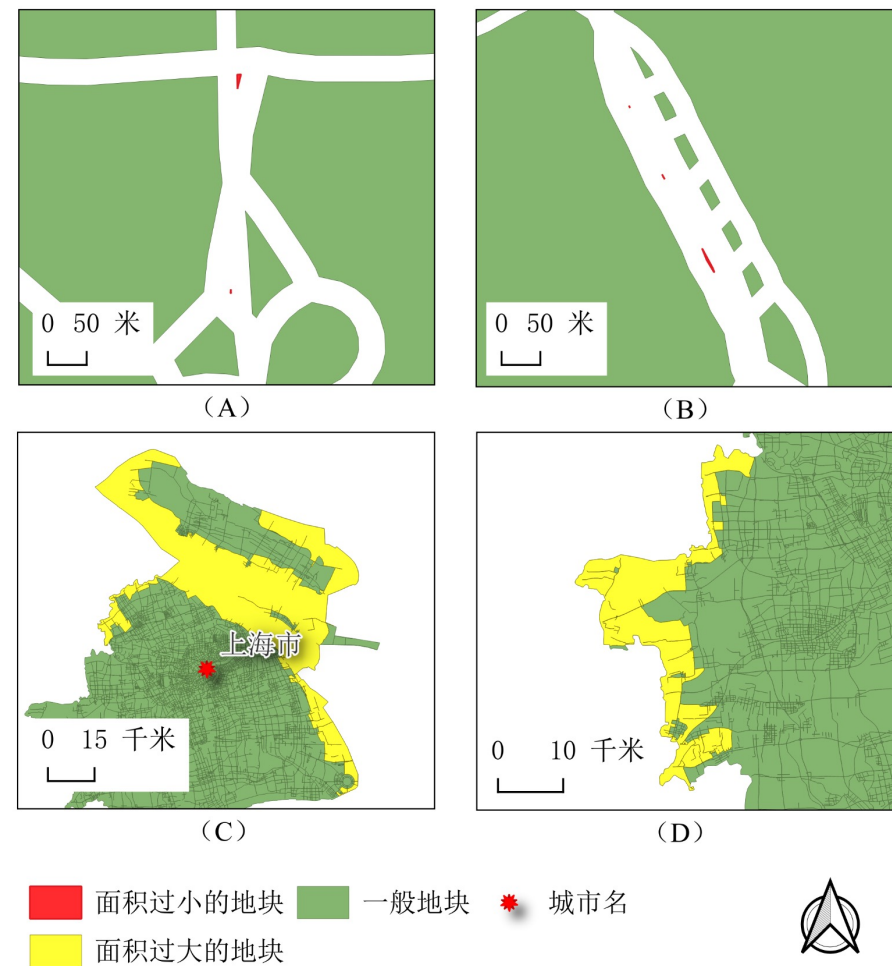
原因与依据

- 地块的大小会影响预期利润，从而影响其土地用途类别。 (*Sinclair et al., 1967*)
- 地块尺寸差异极大，过大或过小的地块在模型训练中被视为噪声，因为它们无法有效地输入模型，或无法给出具体的类别。



➤ 处理方法

- 将地块按面积划分区间；
- 统计数据集质量评价指标随地块面积的变化。



面积过小 (A、B) 和过大 (C、D) 的地块示例

4.1 DCAI-CLUD 地块属性数据筛选与预标注可行性研究



➤ 基于地块位置的数据筛选

按位置抽样方法

$$B_i^p = \frac{(B_i^r)^d}{(N(B_i^d))^{(1-d)}}$$

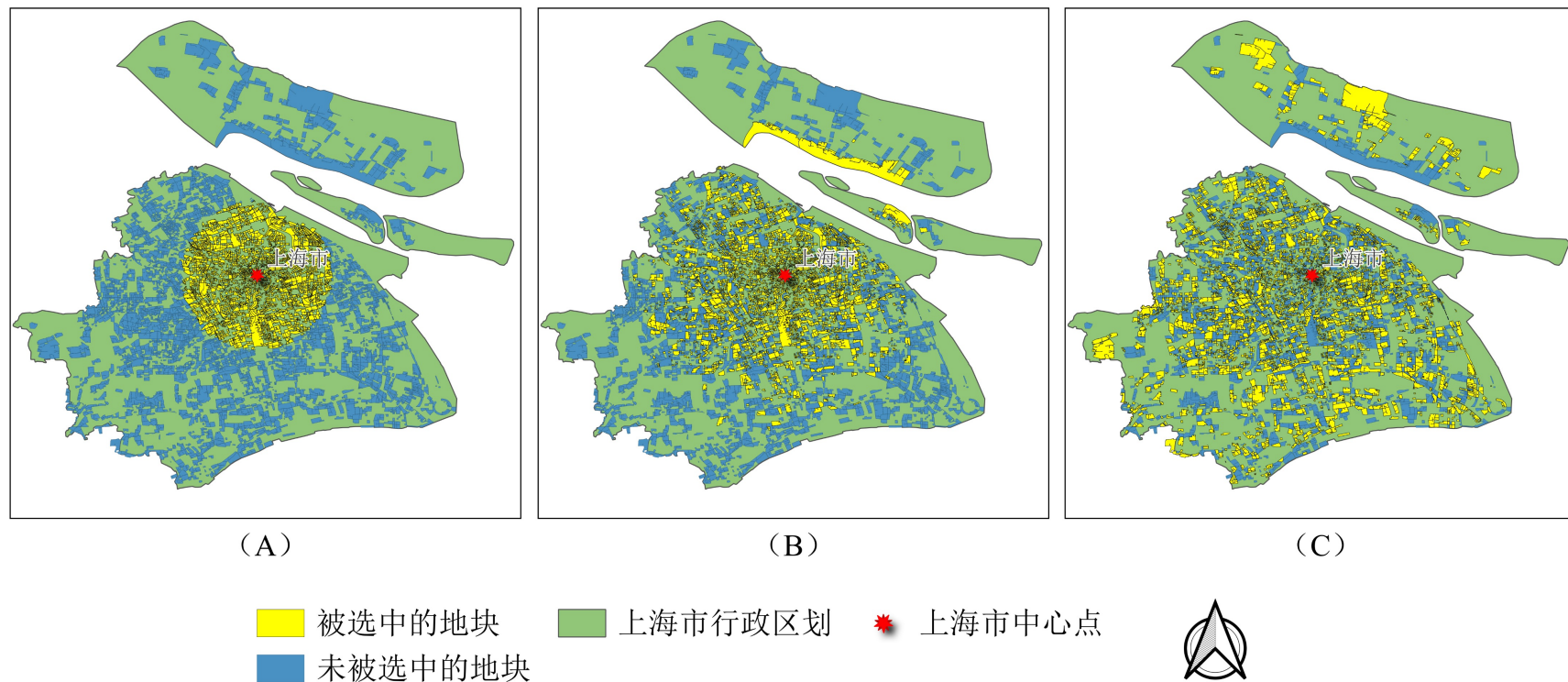
B_i^p : 地块 B_i 被选中的概率

B_i^r : 0~1之间随机分布的随机数。命名为随机因子。

B_i^d : B_i 到城市中心的距离。命名为距离因子。

d : 分散系数, $d \in [0, 1]$

$N(\dots)$: 归一化函数



设置不同分散系数 d 时的按位置筛选效果: (A) $d=0$, (B) $d=0.5$, (C) $d=1$

统计数据集质量评价指标随分散系数 d 的变化, 找到最佳取值。



➤ 数据集与模型评价指标

数据集构建环节	评价指标	描述
数据集构建	效率	<i>Acc</i> 数据标注的准确率
		<i>Rate</i> 数据标注的速度（个/小时）
	数据质量	基尼不纯度（基尼系数）。表示数据集均衡程度。值越大代表类别越均衡。计算公式如下：
		<i>Gini</i> $Gini = 1 - \sum_{i=1}^K \left(\frac{n_i}{N}\right)^2$
		<i>P_{um}</i> 非混合土地利用类型地块的占比
		<i>P_{iu}</i> 城市区域地块的占比
模型训练	<i>S_{avg}</i>	数据质量综合评分： $S_{avg} = [N(Gini) + N(P_{nm}) + N(P_{iu})]/3$
	Kappa	$Kappa = \frac{\sum_{i=1}^K x_{ii}/N - \sum_{i=1}^K x_{ii} (\sum_{j=1}^K x_{ij} \sum_{j=1}^K x_{ji})/N^2}{1 - \sum_{i=1}^K x_{ii} (\sum_{j=1}^K x_{ij} \sum_{j=1}^K x_{ji})/N^2}$
	TA	测试精度： $Test\ Accuracy = \frac{\sum_{i=1}^K x_{ii}}{N}$

4.1 DCAI-CLUD 地块属性数据筛选与预标注可行性研究



➤ 模型预标注的误导分析

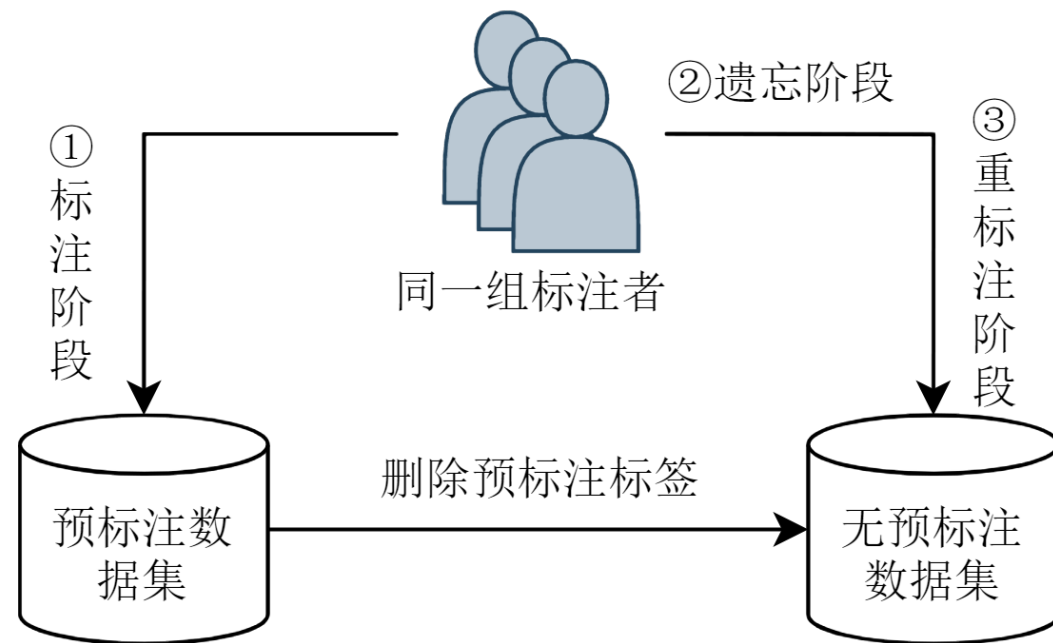
理想实验

- 标注者在有预标注情况下进行数据标注
- **完全忘记**刚才的标注经历
- 在没有预标注的情况下进行数据标注。



➤ 设计实验

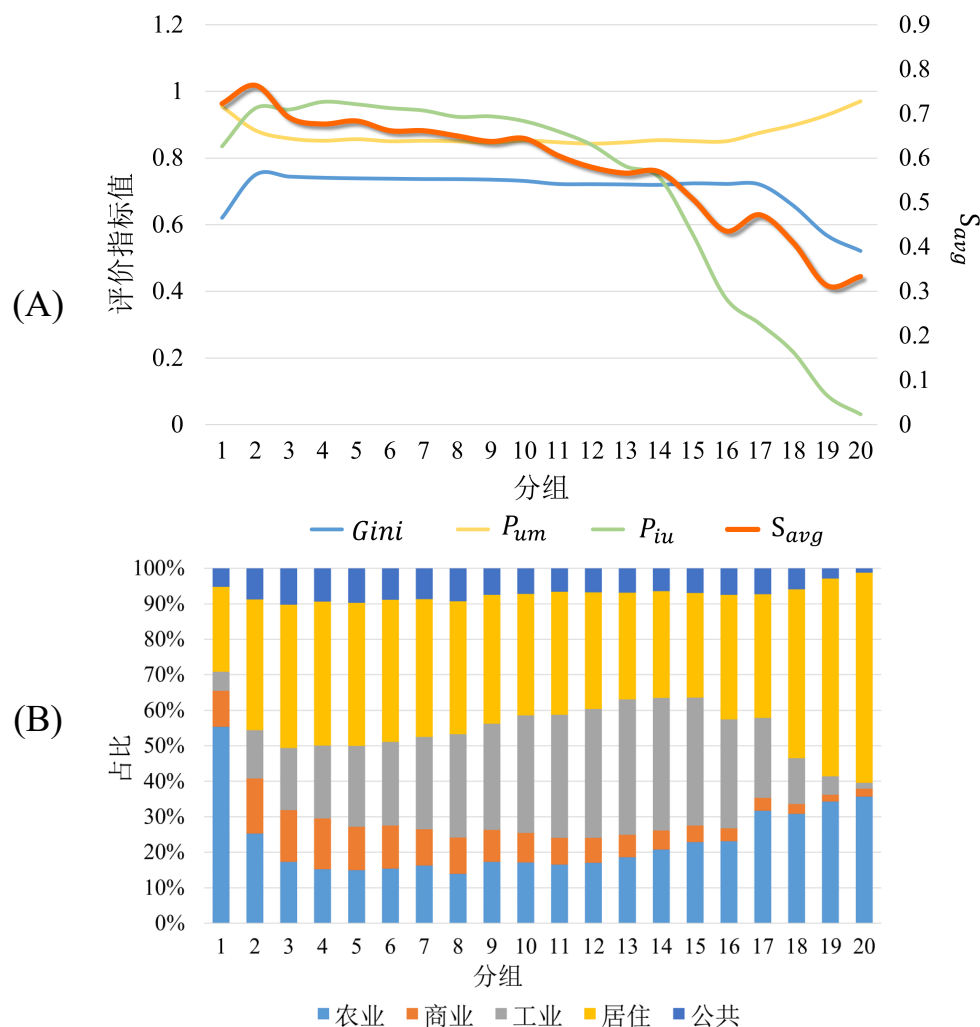
- **标注阶段**。志愿者在有预标注的情况下进行标注；
- **遗忘阶段**。志愿者在第一次标注后的一个月里，陆续完成少量其他的标注任务。
- **重标注阶段**。在不告知志愿者实验目的前提下，去除预标注标签，使志愿者重标注自己的样本。



4.1 DCAI-CLUD 地块属性数据筛选与预标注可行性研究



➤ 基于地块面积的数据筛选结果与评价



- 对9个城市内31960 块地块进行统计
- 将地块按面积从小到大平均分为 20 组。

- ▣ 面积 $\leq 38,931.315m^2$ (第1组)
- ▣ $38,931.315m^2 \leq \text{面积} \leq 676,818.47m^2$ (第2~16组)
- ▣ $676,818.47m^2 \leq \text{面积}$ (第17~20组)

当面积 $\leq 38,931.315m^2$ (第1组) or $676,818.47m^2 \leq \text{面积}$ (第 17~20 组) $Gini$ 和 P_{um} 的皮尔逊相关系数达到 **-0.914**, 呈现**极强的负相关性**。

(A) $Gini$ 、 P_{um} 、 P_{iu} 和 S_{avg} 随面积的变化; (B) 土地利用类别比例随面积的变化。

4.1 DCAI-CLUD 地块属性数据筛选与预标注可行性研究

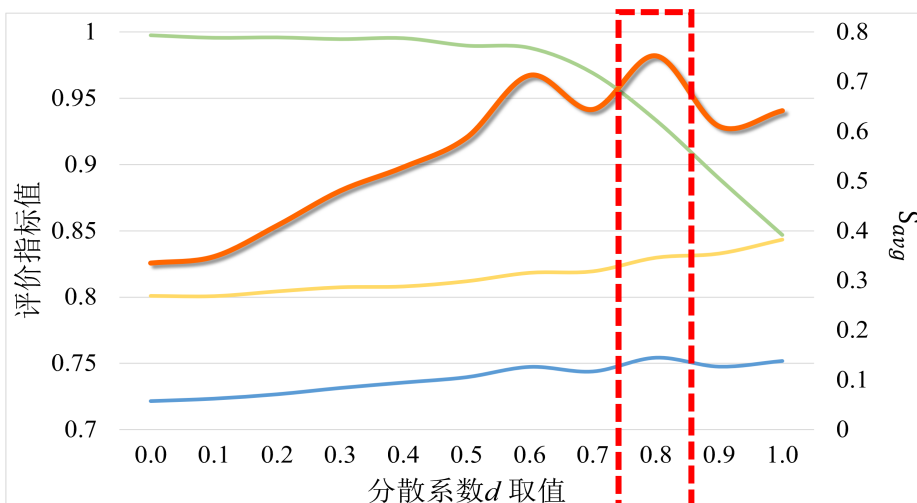


➤ 基于地块面积的数据筛选结果与评价

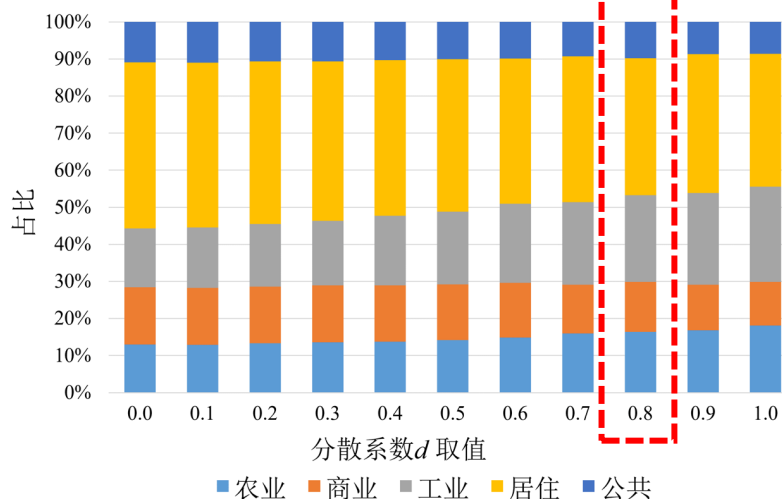
随着所选择地块逐渐远离城市中心 (d 逐渐增大)

- $Gini$ 和 P_{um} 波动上升, 而 P_{iu} 则加速下降
 - S_{avg} 呈波动上升趋势, 在 $d=0.8$ 时达到最大值 0.752。
-
- 农业自然用地和工业用地的比例分别增加了 39.88% 和 61.10%
 - 商业、居住和公用地的比例分别下降了 23.58%、20.04% 和 21.06%

(A)



(B)



(A) $Gini$ 、 P_{um} 、 P_{iu} 和 S_{avg} 随分散系数 d 的变化; (B) 土地利用类别比例随分散系数 d 的变化。

4.1 DCAI-CLUD 地块属性数据筛选与预标注可行性研究

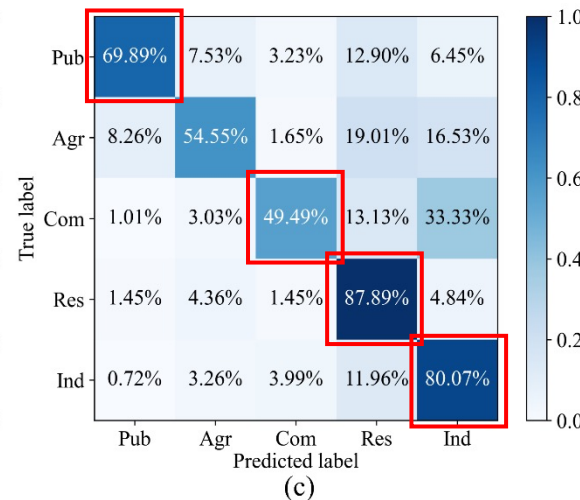
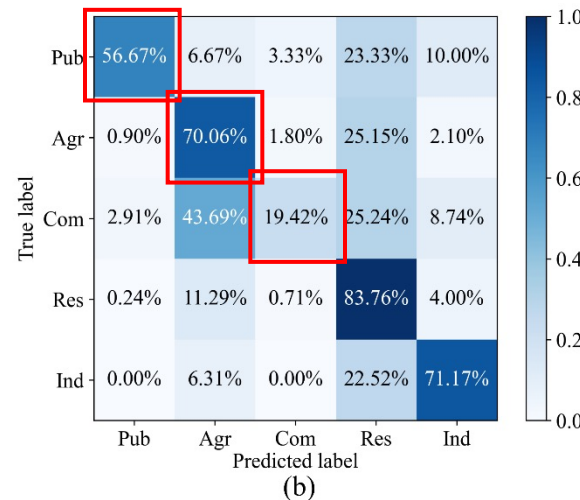
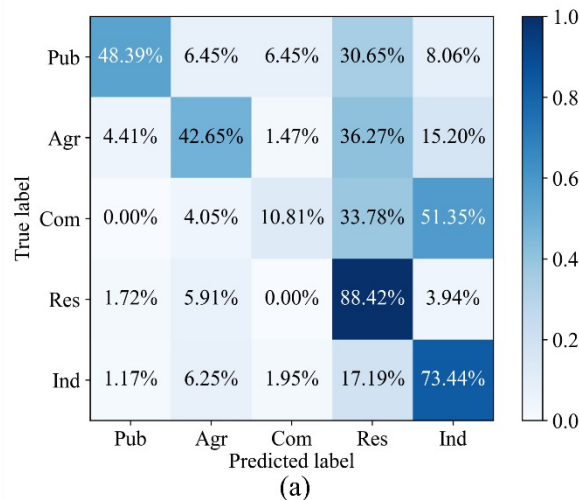


D_r : 随机抽样的数据集

D_s : 按面积筛选的数据集

D_{sl} : 按面积和位置筛选的数据集

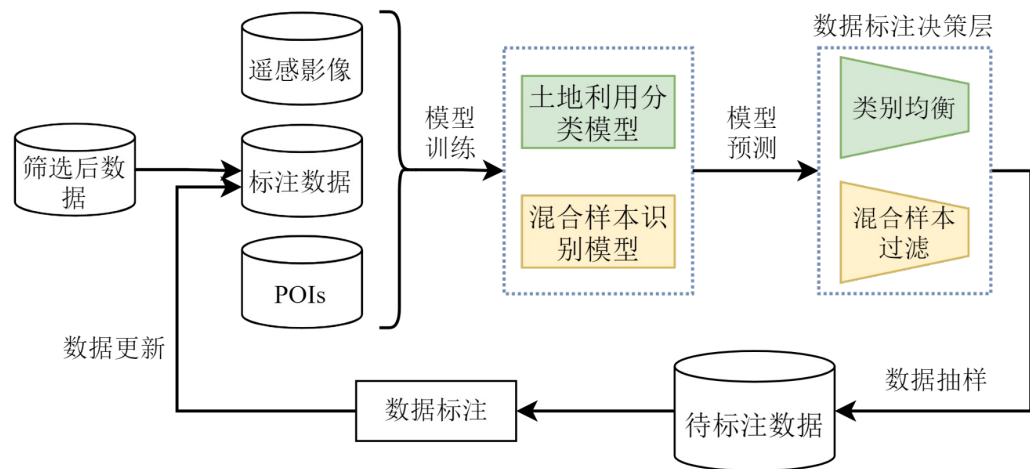
数据集	数据集质量评价指标			模型评价指标	
	$Gini$	P_{um}	P_{iu}	TA	Kappa
D_r	0.733	0.798	0.701	0.671	0.520
D_s	0.741 (+1.09%)	0.784 (-1.75%)	0.836 (+19.26%)	0.704 (+4.92%)	0.552 (+6.15%)
D_{sl}	0.763 (+2.97%)	0.787 (+0.38%)	0.925 (+10.65%)	0.762 (+8.24%)	0.664 (+20.29%)



4.1 DCAI-CLUD 基于人机协作的高质量土地利用数据集构建

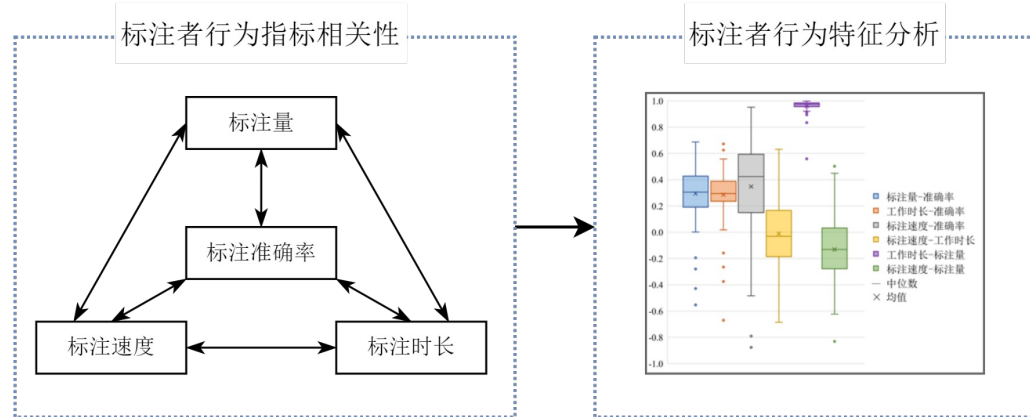


1. 基于双模态神经网络的标注决策辅助



- 构建了一种以数据为中心的土地利用数据构建框架。基于双模态土地利用分类模型的预测能力对数据进行类别均衡与混合样本过滤。在数据的不断积累过程中逐步优化模型质量来进一步提高模型辅助决策的水平

2. 基于皮尔逊相关系数的标注者指标相关性分析

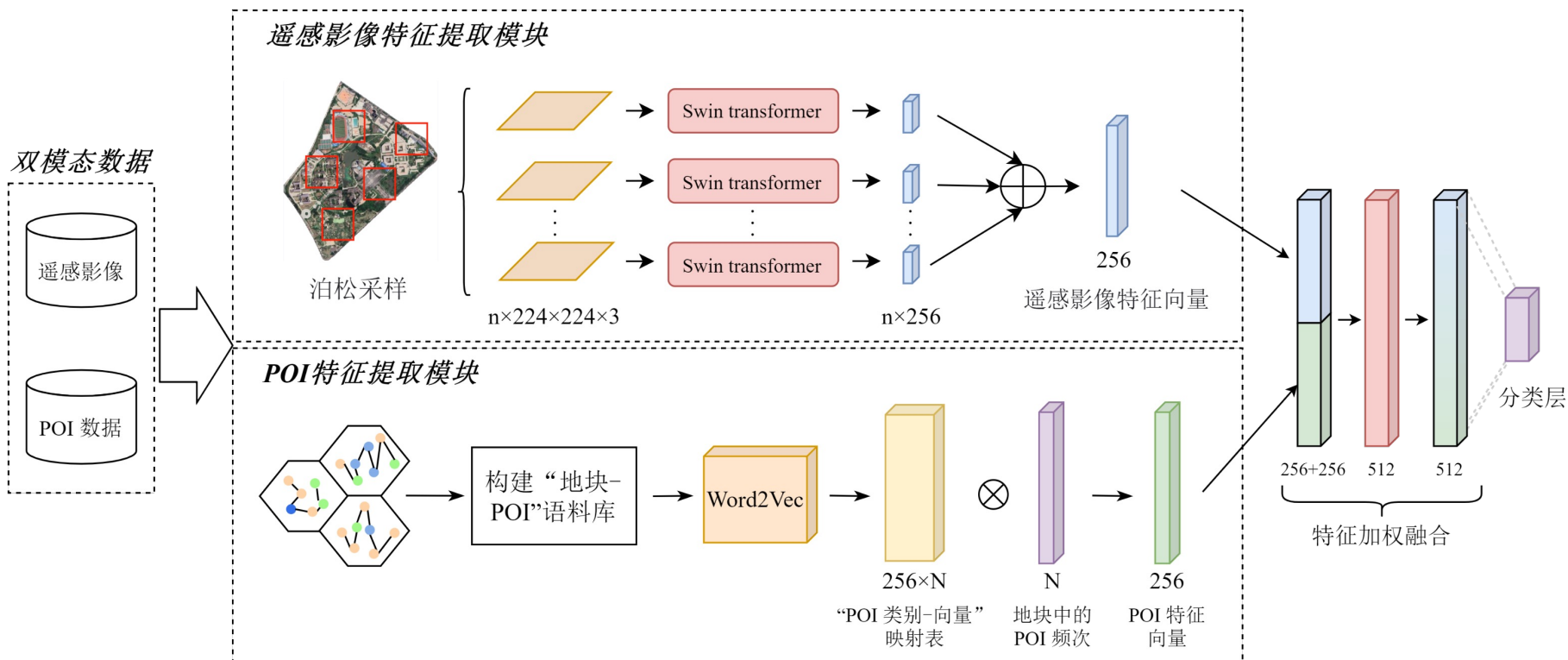


- 基于皮尔逊相关系数（Pearson correlation coefficient）分析标注者行为变量之间的相关性。

4.1 DCAI-CLUD 基于人机协作的高质量土地利用数据集构建



➤ 耦合遥感影像与POI的双模态土地利用分类模型



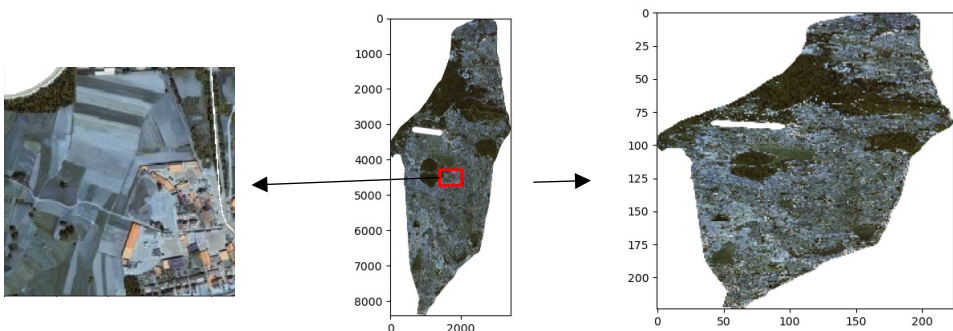
双模态土地利用分类模型BDF-Net结构

- 基于Swin Transformer (Liu et al., 2021) 提取遥感图像中的自然物理特征。
- 基于POI embedding (Yao et al., 2017) 提取POI数据中的社会经济特征。
- 通过自适应特征加权层 (Lu et al., 2022) 对两类特征进行加权融合

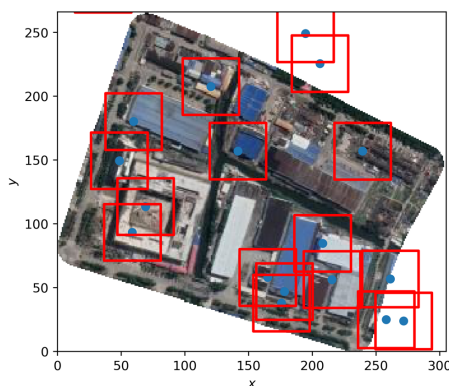
4.1 DCAI-CLUD 基于人机协作的高质量土地利用数据集构建



➤ 采样方式以及样本特征融合



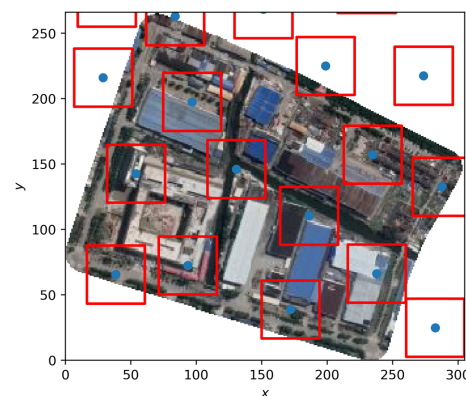
重采样



随机采样

泊松采样

- 均衡的在输入地块上采取多个样本，将不规则地块转化为若干个规则地块。
- 泊松采样相比于随机采样，获得的样本更均衡可以更好的反应地块的真实信息。
- 相比于重采样可以保留高信息密度的遥感图像局部细节信息。

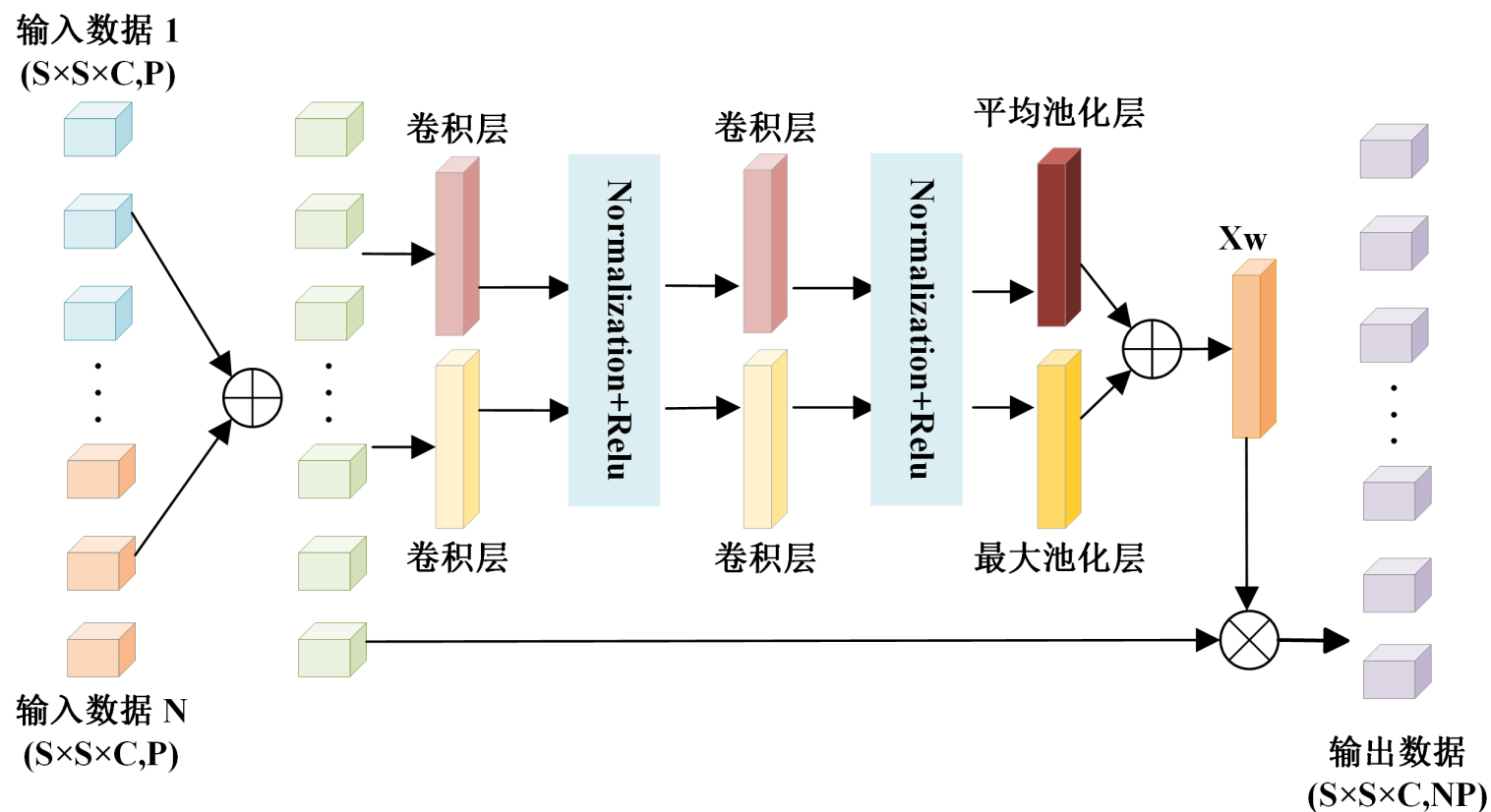


泊松采样

4.1 DCAI-CLUD 基于人机协作的高质量土地利用数据集构建



➤ 可学习权重模块 (LW module)



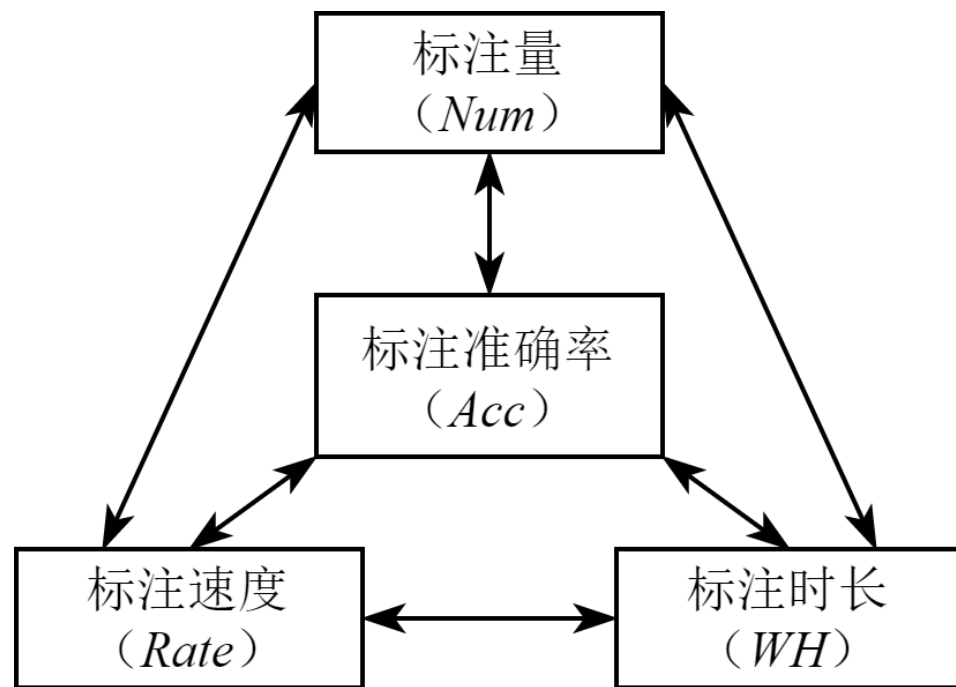
- 根据不同数据特征对土地利用分类结果的贡献程度自动学习并调整权重，以提高分类精度。
- 由两个卷积层、批量归一化层和激活函数组成，最后分别通过全局池化层或最大池化层并相加得到特征权重。

➤ 标注者的4项指标

- 每日标注数量（个）
- 每日标注时长（小时）
- 每日标注速度（个/小时）
- 每日标注准确率

➤ 计算4项指标之间的皮尔逊相关系数

$$\begin{aligned}\rho_{X,Y} &= \frac{\text{cov}(X,y)}{\sigma_X\sigma_Y} = \frac{E((X - \mu_x)(Y - \mu_Y))}{\sigma_X\sigma_Y} \\ &= \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}}\end{aligned}$$



4.1 DCAI-CLUD 基于人机协作的高质量土地利用数据集构建

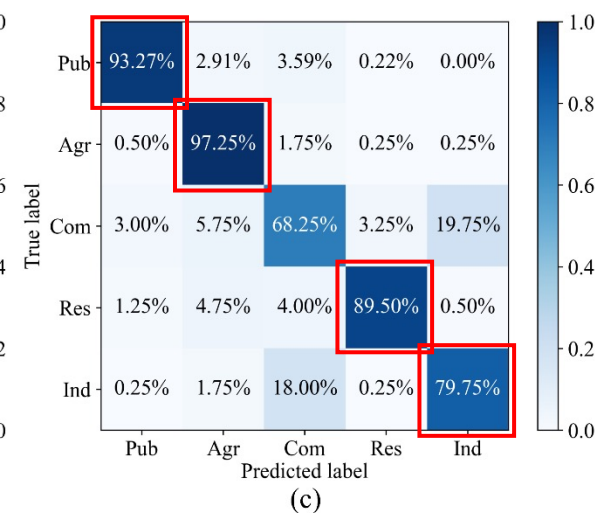
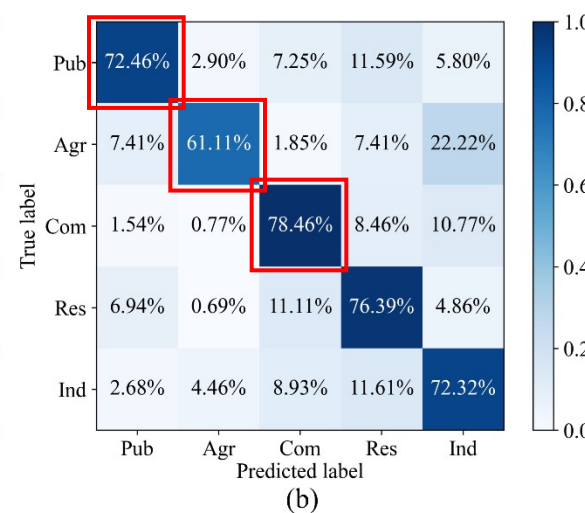
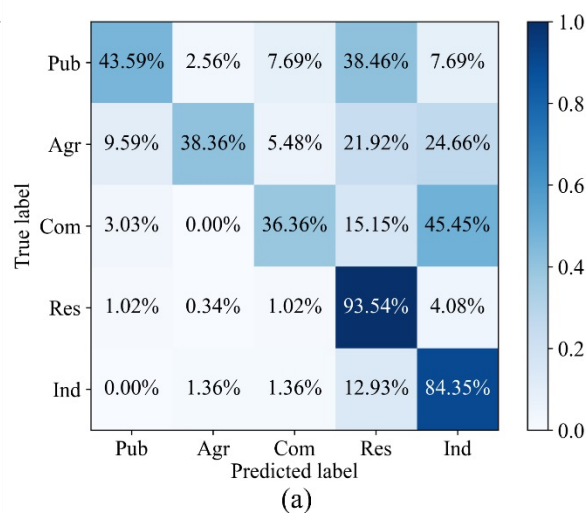


D_{np} : 没有进行预标注的数据集

D_p : 标注过程中使用了预标注
进行类别均衡的数据集

D_{pm} : 标注过程中使用了预标注
进行类别均衡和混合土地
利用样本过滤的数据集

数据集	数据标注效率评价指标		数据集质量评价指标		模型评价指标	
	Acc	$Rate$	$Gini$	P_{um}	TA	Kappa
D_{np}	89.92%	116.34	0.690	0.749	0.778	0.649
D_p	90.00% (+0.09%)	115.94 (-0.34%)	0.779 (+12.90%)	0.733 (-2.14%)	0.739 (-5.01%)	0.662 (+2.00%)
D_{pm}	95.25% (+5.83%)	150.04 (+29.41%)	0.757 (-2.82%)	0.866 (+18.14%)	0.858 (+16.10%)	0.822 (+24.17%)



模型的混淆矩阵: (a) M_{np} 、(b) M_p 、(c) M_{pm} 。

4.1 DCAI-CLUD 基于人机协作的高质量土地利用数据集构建



准确率和其他三项指标的皮尔逊相关系数在各区间人数占比

相关程度	皮尔逊相关系数取值范围	标注量	标注时长	标注速度		
		人数占比				
强正相关	0.8~1	0.00%	0.0%	6.89%	49.99%	熟能生巧
中等正相关	0.6~0.8	5.17%	6.89%	17.24%		
	0.4~0.6	25.86%	15.51%	32.75%		
弱正相关	0.2~0.4	43.10%	56.89%	15.51%	5.15%	慢工出细活
	0~0.2	18.96%	13.79%	15.51%		
	-0.2~0	1.72%	1.72%	1.72%		
	-0.4~-0.2	1.72%	3.44%	3.44%		
中等负相关	-0.6~-0.4	3.44%	0.00%	3.44%		
	-0.8~-0.6	0.00%	1.72%	1.72%		
强负相关	-1~-0.8	0.00%	0.00%	1.72%		

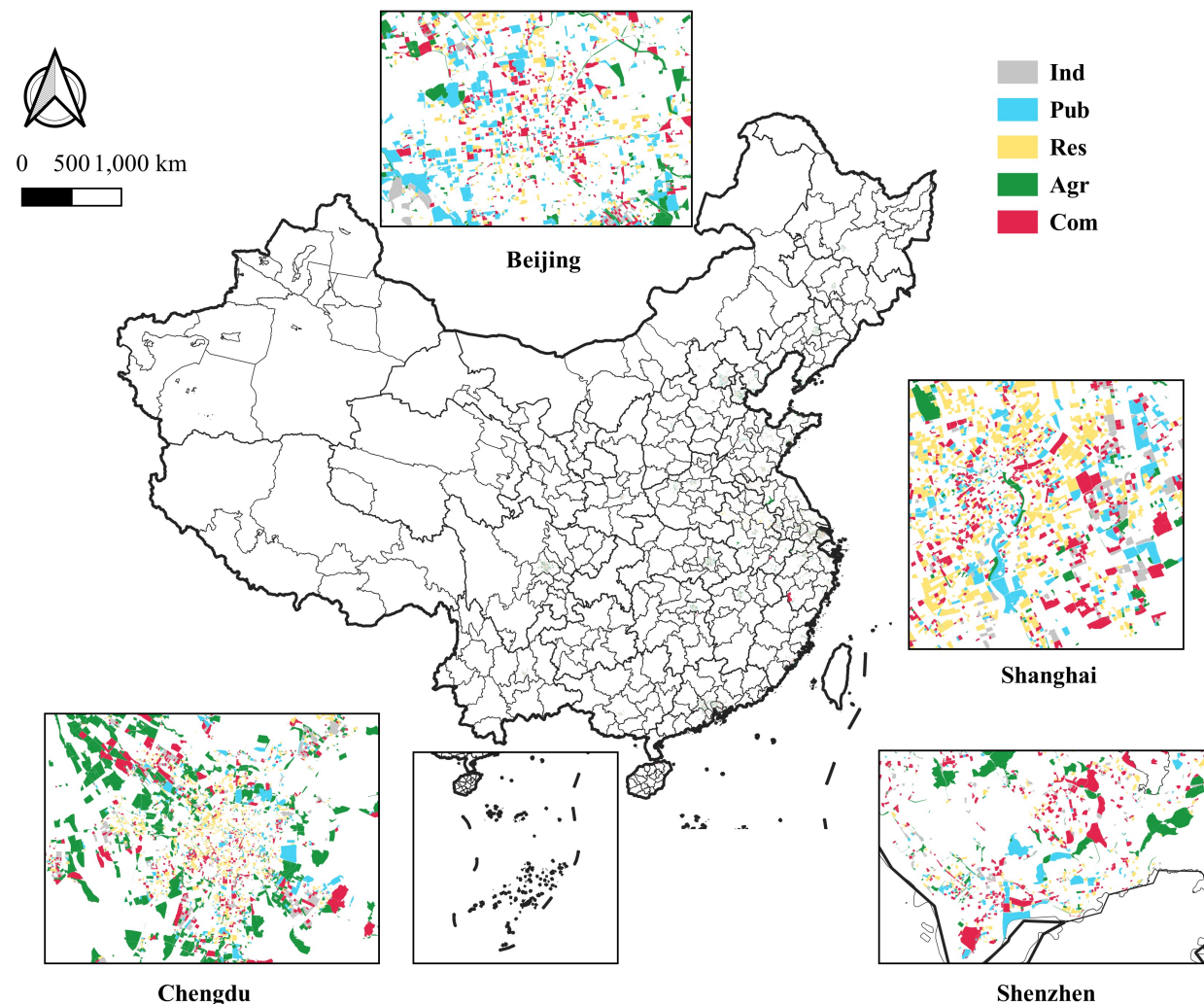
4.1 DCAI-CLUD MSLU-100K数据集



➤ 发布数据集MSLU-100K

MSLU-100K 是一个面向中国 **81** 个主要城市的多源土地利用遥感影像数据集，覆盖面积达 **983,215 平方千米**，具有极高的城市空间代表性。数据集通过结合人工智能和人工专家标注，采用**人机协作的标注框架**构建而成，全面体现了城市空间中居住、商业、公共服务、工业、农业等功能区域的空间分布与语义特征。

参考网址：<https://www.urbancomp.net/archives/cn-mslu-100k-land-use-classification-dataset-at-block-scale-for-multi-source-spatio-temporal-data-zh>

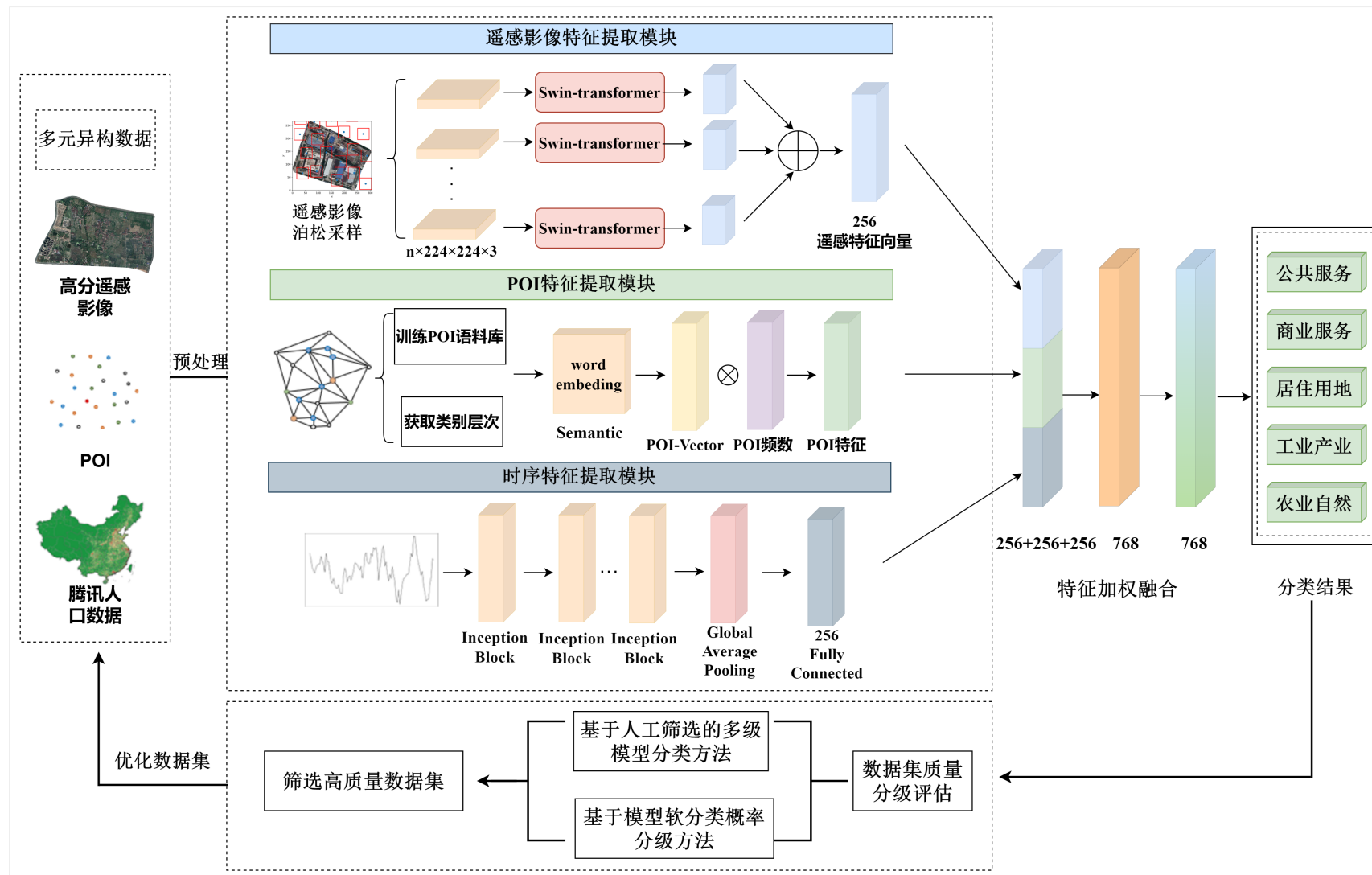


4.2 全国土地利用数据集质量评价方法



► 数据质量评估模型

- 来源于腾讯技术服务，每小时在固定范围内记录每个用户的地理位置。
- 获取的腾讯用户时序密度数据时间分辨率为1小时，空间分辨率约为1100米，时间为2019年5月7日。
- 腾讯技术覆盖终端10亿，日请求次数达1100亿次。已被证实能够在精细尺度下反映城市社会经济属性。

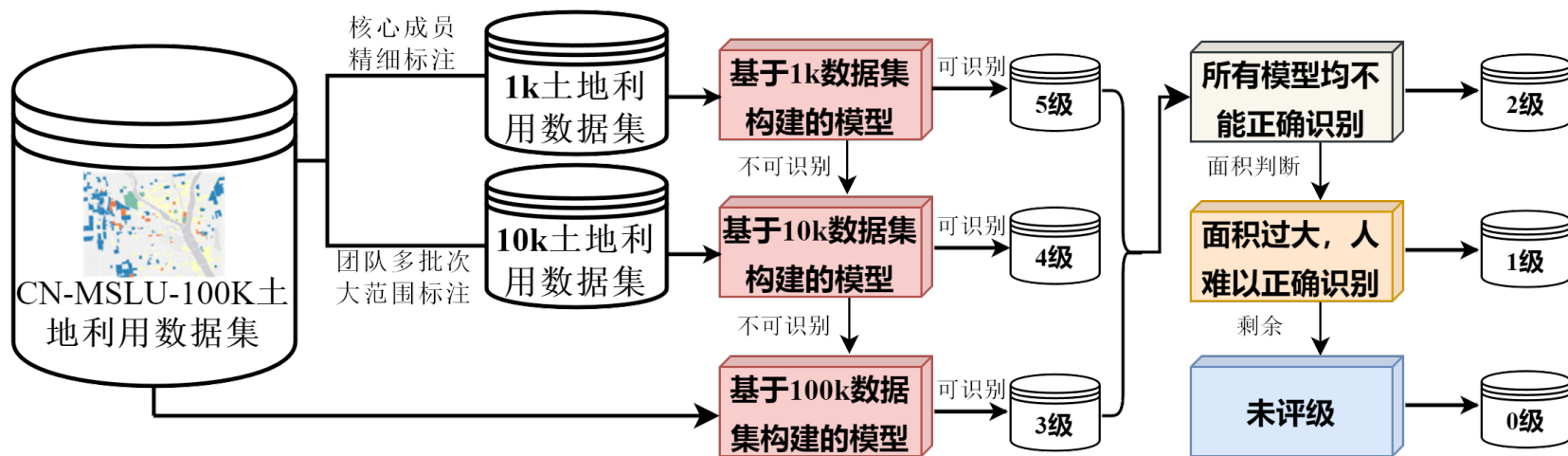


4.2 全国土地利用数据集质量评价方法



本研究提出两种土地利用数据集质量评价方法，对CN-MSLU-100K数据集进行明确等级划分，分为0~5，共6个等级。

➤ 基于人工筛选的多级模型分类方法

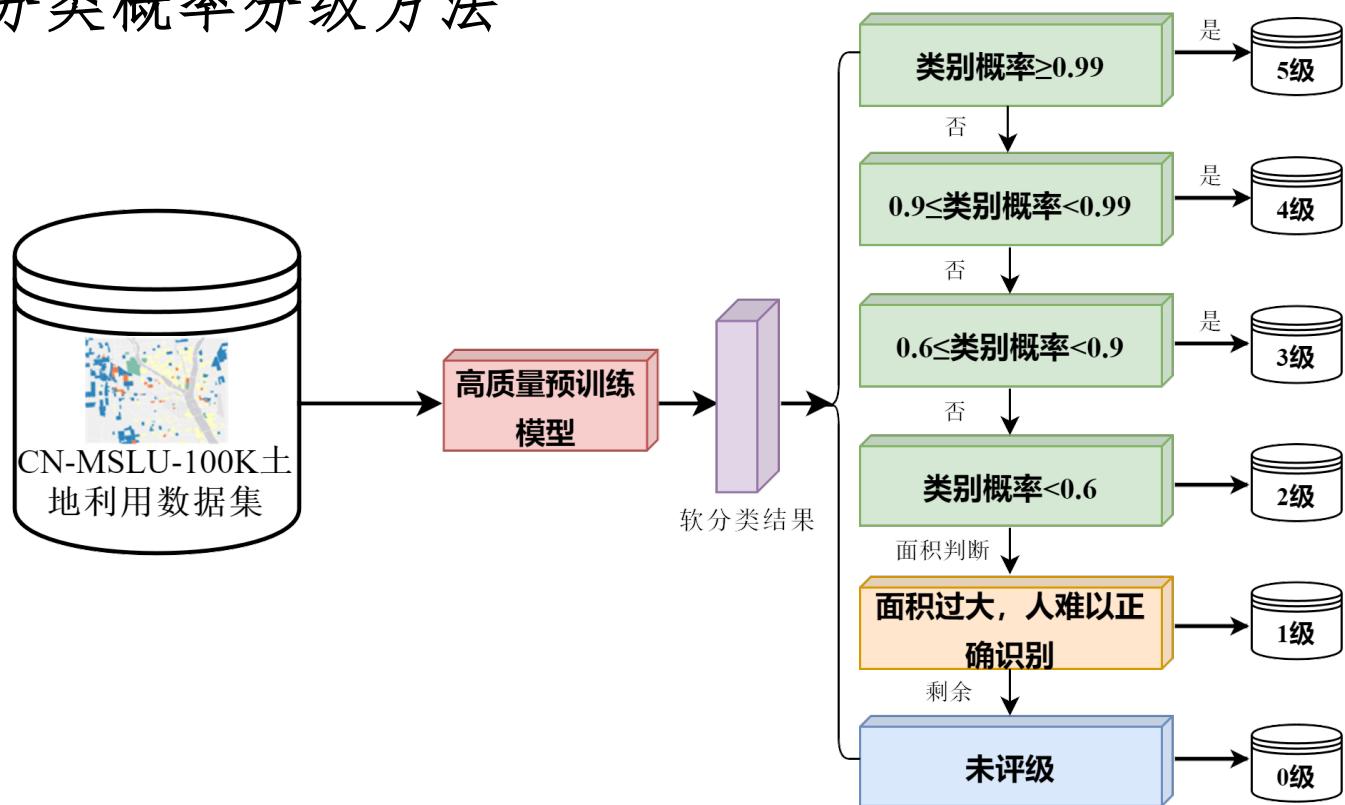


- 1k数据集：项目骨干成员精细标注，**准确无误**，作为后续标注的标准。
- 10k数据集：邀请外部专家多批次大量标注得到，**准确率高**。
- 根据输入特征的多少构建多级模型，利用**多级模型分类结果**，将数据分级，5级到0级质量递减。

4.2 全国土地利用数据集质量评价方法



➤ 基于模型软分类概率分级方法



- 选取已有模型中识别效果最好的模型得到软分类结果。
- 依据**概率分布结果**，对数据进行分级；概率越高代表模型认为该地块样本没有其他功能特征，具有较为纯净且明显的地块功能，数据质量好。
- 5级到0级质量递减。

4.2 全国土地利用数据集质量评价方法



一级类别	5.0	4.0	3.0	2.0	1.0	0.0
Res	12583(30.9%)	874(2.1%)	1133(2.8%)	14959(36.7%)	1399(3.4%)	9734(23.9%)
Com	754(11.30%)	246(3.7%)	168(2.5%)	4050(60.6%)	32(0.5%)	1434(21.4%)
Pub	1792(28.5%)	604(9.6%)	105(1.7%)	2618(41.6%)	17(0.2%)	1150(18.3%)
Ind	9810(40%)	1013(4.1%)	1283(5.2%)	4363(17.8%)	9(0.1%)	8020(32.7%)
Agr	10118(47.2%)	4701(22%)	749(3.5%)	428(2%)	5(0.1%)	5410(25.3%)
Tra	0(0%)	0(0%)	0(0%)	571(71.4%)	4(0.5%)	224(28%)
Unk	0(0%)	0(0%)	0(0%)	0(0%)	11(0.1%)	25058(99.9%)
合计	35057(27.9%)	7438(5.9%)	3438(2.7%)	26989(21.5%)	1477(1.1%)	51030(40.7%)

基于人工筛选的多级模型分类方法质量等级在各类别中的分布

一级类别	5.0	4.0	3.0	2.0	1.0	0.0
Res	15769(38.7%)	3713(9.1%)	2834(6.9%)	7233(17.7%)	1399(3.4%)	9734(23.9%)
Com	1088(16.3%)	420(6.3%)	602(9%)	3108(46.4%)	32(0.5%)	1434(21.4%)
Pub	637(10.1%)	1096(17.4%)	741(11.7%)	2645(42.1%)	17(0.2%)	1150(18.3%)
Ind	3699(15.1%)	7145(29.1%)	7283(29.7%)	3795(15.5%)	9(0.1%)	2567(10.5%)
Agr	11999(56%)	5116(23.8%)	2090(9.7%)	1258(4.2%)	5(0%)	943(4.4%)
Tra	0(0%)	0(0%)	0(0%)	571(71.4%)	4(0.5%)	224(28%)
Unk	0(0%)	0(0%)	0(0%)	0(0%)	11(0.1%)	25058(100%)
合计	33192(26.5%)	17490(13.9%)	13550(10.8%)	18610(14.8%)	1477(1.1%)	41110(32.8%)

基于模型软分类概率分级方法质量等级在各类别中的分布

- 整体上，两种方法中5级数据均占据较高的比重，说明数据集有约三分之一的数据质量很高。
- 基于人工筛选的多级模型分类方法以模型能否识别为评价标准，不能识别的2级数据占比较大。
- 基于模型软分类概率分级方法，以软分类概率为评价标准，各等级数量较为均衡。

Pub: 公共管理与公共服务设施用地; Com: 商业服务业用地;
Res: 居住用地; Ind: 工业产业用地; Agr: 农业自然; Tra:
交通设施用地; Unk: 跳过

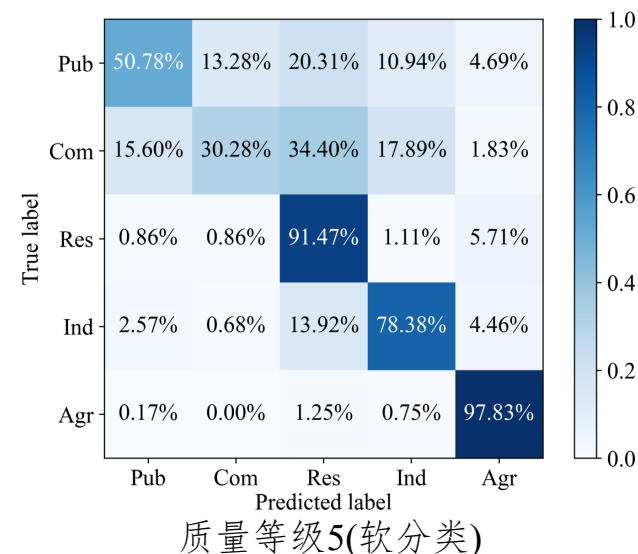
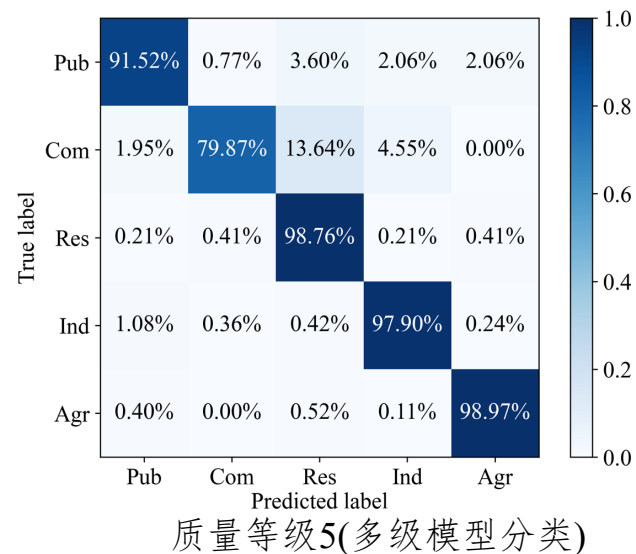
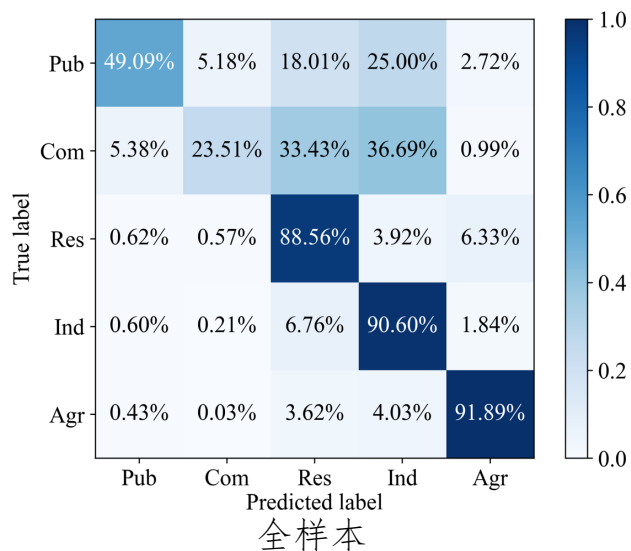
4.2 全国土地利用数据集质量评价方法



各quality等级中不同类别的典型样本

- quality等级越高，数据质量越好，地块形状越规整，POI类型越单一，分辨率相对越高，模型更容易识别。
- 混合类型的地块特征较为复杂，部分地块POI缺失，以及存在面积很大的地块，模型难以识别，quality等级偏低。
- 预训练模型的精度和标签的准确率也在很大程度上影响评价结果

4.2 全国土地利用数据集质量评价方法

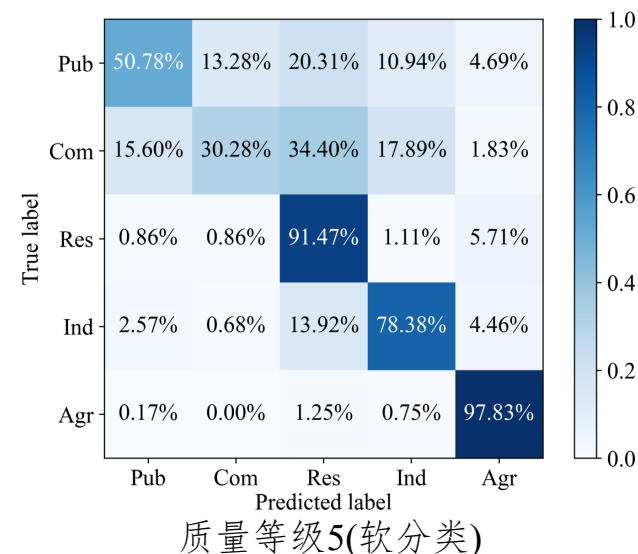
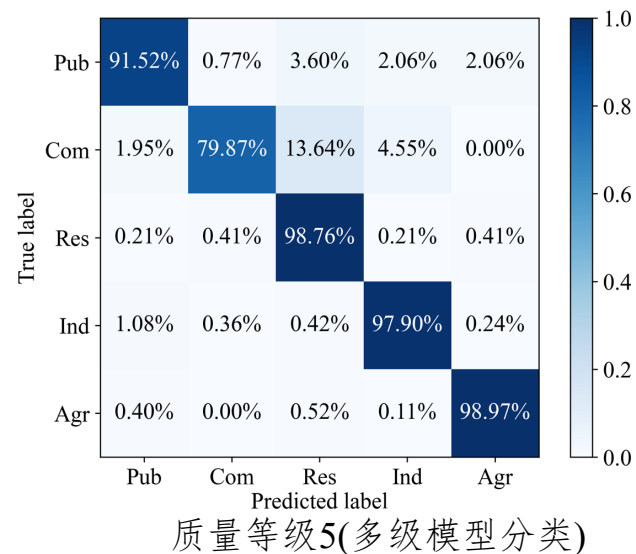
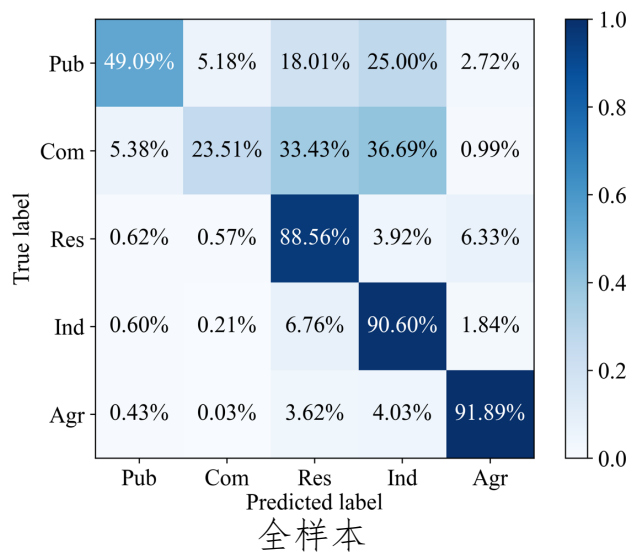


基于不同数据土地利用识别模型各指标精度对比

模型	TA	Kappa	F1 Score
全样本	0.860	0.804	0.855
5(多级模型分类)	0.975	0.965	0.962
5(软分类)	0.895	0.832	0.891

- 高质量的小数据集比低质量的大数据集训练得到的模型效果更好，土地利用功能识别能力更强。
- 基于人工筛选的多级模型分类方法在筛选高质量数据时优于基于模型软分类概率分级方法

4.2 全国土地利用数据集质量评价方法

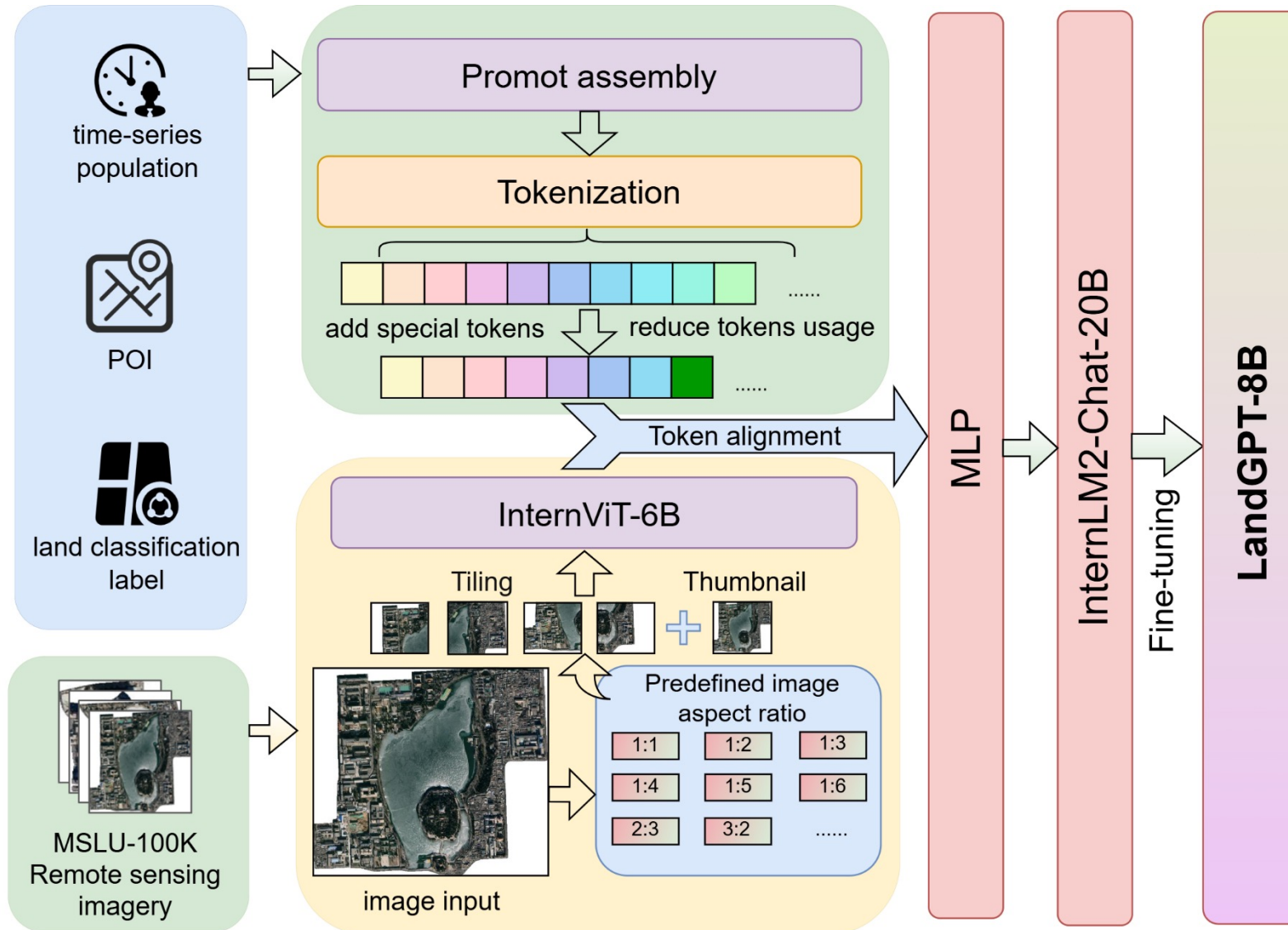


基于不同数据土地利用识别模型各指标精度对比

模型	TA	Kappa	F1 Score
全样本	0.860	0.804	0.855
5(多级模型分类)	0.975	0.965	0.962
5(软分类)	0.895	0.832	0.891

- 高质量的小数据集比低质量的大数据集训练得到的模型效果更好，土地利用功能识别能力更强。
- 基于人工筛选的多级模型分类方法在筛选高质量数据时优于基于模型软分类概率分级方法

4.3 LandGPT



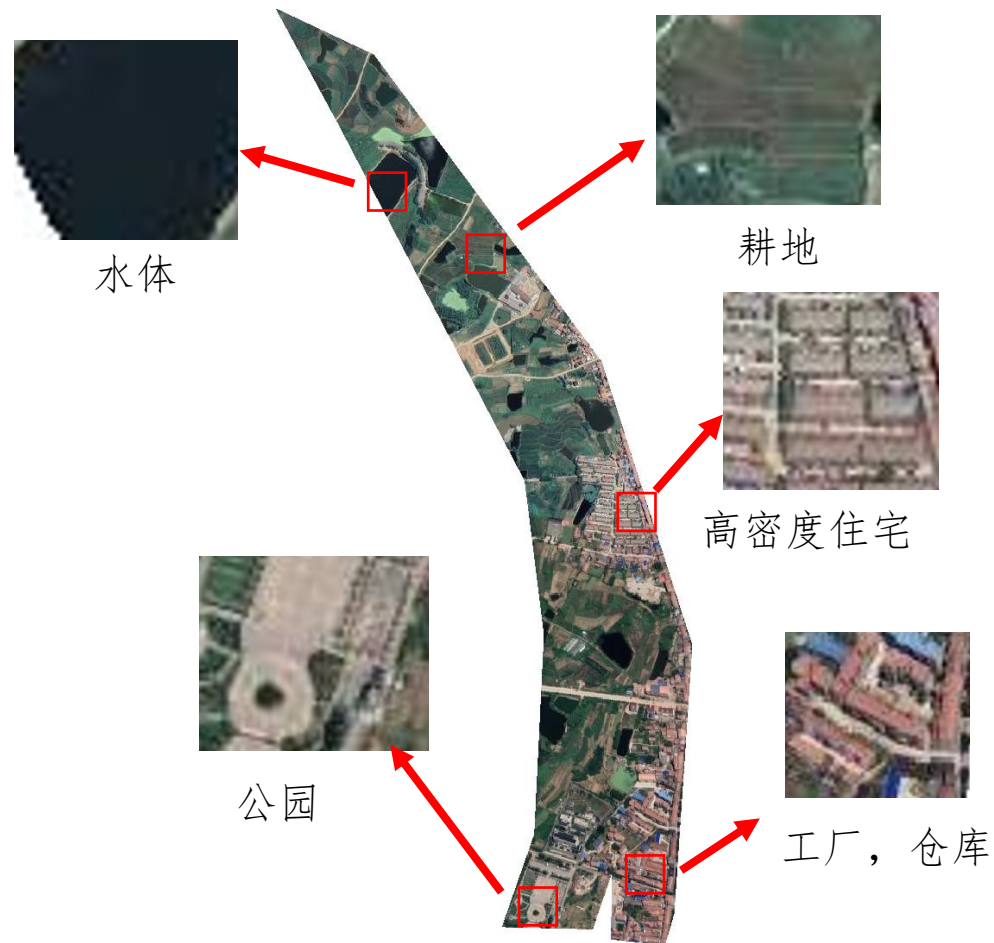
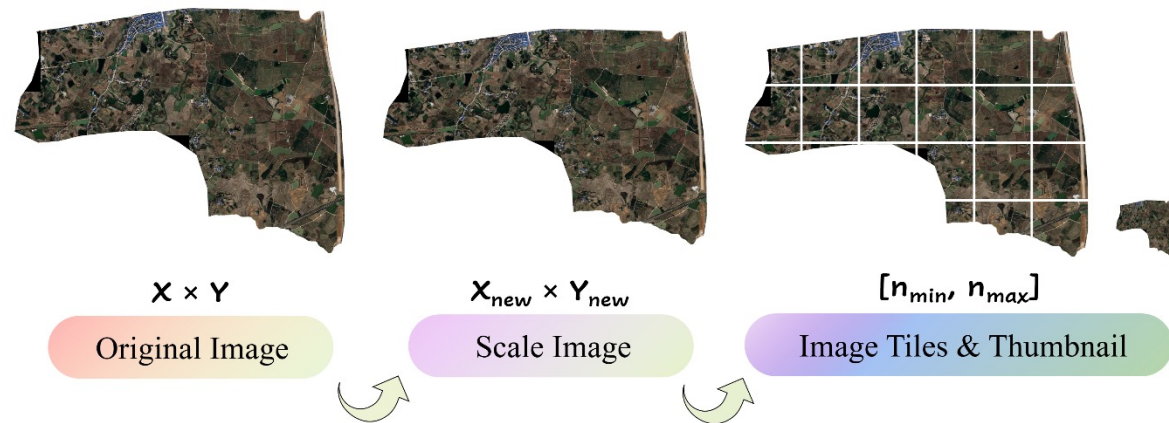
4.3 | LandGPT



➤ 图像处理

➤ 动态高分辨率

本研究要采用动态高分辨率进行高分辨图像的输入。根据原图进行比例匹配分割成不同的瓦片，以及一张被强制缩放的全局图，这种做法保证了模型结构只需要接收448*448大小的图像，也能获得高分辨率的全局视野：(Chen et al. 2024)



参考网址: [LandGPT: Multi-modal model for parcel land use classification using multi-source data - 城市之光 – City of Light](#)

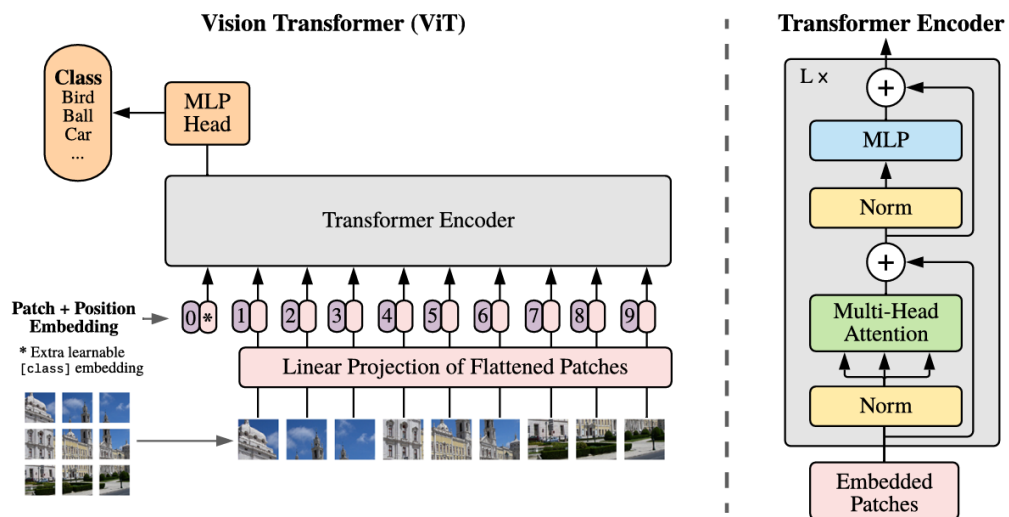
传统随机采样的方法会造成一定量的精度损失

4.3 LandGPT

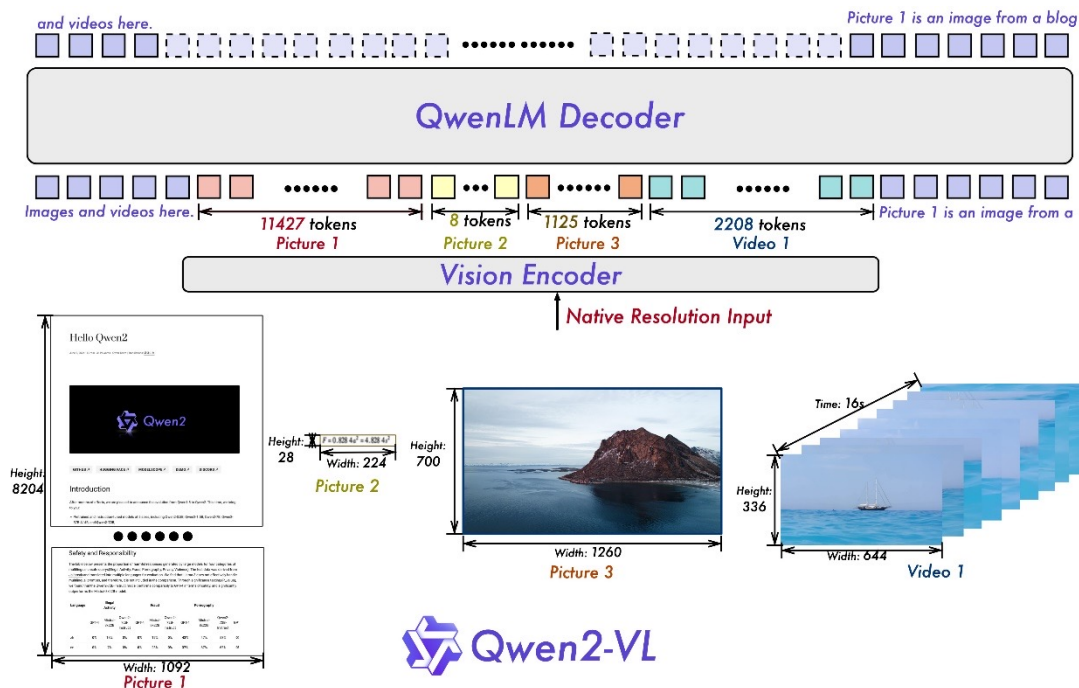


➤ 基于Vision Transformer的视觉编码器InternViT-6B

- 该编码器以**Vision Transformer**为基础框架(*Dosovitskiy et al. 2021*)，将视觉编码器的参数数量拓展到60亿个参数以同LLM的规模相匹配。
- 为了在精度、速度和稳定性之间做出平衡，模型最终选定了宽度为3200，深度为48层，MLP宽度为12800以及25个头数量的注意力(*Chen et al. 2024*)。



Vision Transformer结构图



Qwen2-VL 结构图

4.3 LandGPT



➤ 类GPT4-tokenizer的文本分词器

- 模型使用InternLM2 Tokenizer作为文本分词器，这是一种类GPT4-tokenizer的分词器，基于字节对编码（Byte Pair Encoding, BPE）技术（Philip et al. 1994）实现从字符串编码到令牌（tokens）的转换。
- 本研究在输入土地利用分类真值、对应地块POI信息、对应地块时序人口信息时使用了此分词器技术。



一级分类：Residential Districts
二级分类：High-rise Residential Buildings

洗浴推拿-2;中介-4;小区-9;培训机构-14;游泳馆-5;小学-3;
咖啡-4;彩票-1;农贸市场-1;洗衣-12;快餐小吃-11;建材店-3;
楼栋-105;其它医疗-4;图书馆-1;其它酒店住宿-2;其它美食-25;
其它便民-19;其它体育场所-5;文体办公用品-5;普通宾馆-2;.....

[14.0929545725776, 26.106366163715, 18.9010164997208,
33.7945215191649, 29.283344332327, 31.9617340926916,
32.5137835430916, 25.9815202789225,]

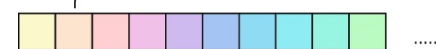
High-rise Residential Buildings

1266416604261741109560

Prompt组合:

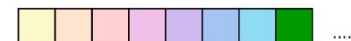
```
<image>  
The FirstLevel category of this image is{FirstLevel}.  
The SecondLevel category of this image is  
{SecondLevel}. This image contains some POI:{POI}.  
This image contains some pedestrian density data:  
{pd_data}
```

分词



添加特殊token

减少token用量



映射至唯一整数ID



分割长度与图像token对齐

➤ 微调数据准备

- 选取MSLU-100K数据集中5个大类下的18个小类作为二级分类指标

去除了原数据集中“居住用地”一级类别中语义模糊的“别墅;高档住宅”（与“多层和高层住宅”相冲突），“农村宅基地”（与“农村建筑与耕地”相冲突）两个二级分类类别

去除了原数据集中“工业产业用地”一级类别中数量过少的“建设用地”一个二级分类类别

- 选取与MSLU-100K数据集地块相对应的POI数据和人口时序数据配合不规则遥感地块影像数据

➤ 渐进式微调Prompt设置

- Prompt分为两步

第一步获取一级分类：

Please determine which of the following FirstLevel categories this remote sensing image belongs to: Residential Districts, Commercial Zones, Industrial Land, Public Services, or Agriculture and Nature. This image contains some POI (Point of Interest) information, which is now provided to you. You can refer to this POI information to make a judgment. The format of the POI information is: 'POI category'-'the number of occurrences of this category in the image'. {POI} This image includes the 24-hour pedestrian density for a certain day. A higher pedestrian density value indicates a larger number of people during that time period. You can refer to the pedestrian density of this area to make your judgment. The format of the pedestrian density data is: [density index from 0 to 1, density index from 1 to 2, ...,], covering a total of 24 hours. {Peodata}

➤ 微调Prompt设置

➤ Prompt分为两步

第二步根据一级分类的结果选取二级分类，并防止模型一级分类出错而添加独立的判断语句：

The FirstLevel category of this image is {predicted_first_level}.Please select the most likely SecondLevel among {'.'.join(second_levels)}. This image contains some POI (Point of Interest) information, which is now provided to you. You can refer to this POI information to make a judgment.If you believe that this image does not belong to any of the above-listed SecondLevel categories, you can also select the SecondLevel category you think this image should belong to based on the POI information below.The format of the POI information is: category'-the number of occurrences of this category in the image. {poi_info}

➤ 微调部分超参数

参数名	数值	参数名	数值
trainable params	37,748,736	freeze_llm	True
force_image_size	448	freeze_mlp	False
learning_rate	4e-5	freeze_backbone	True
max_seq_length	4096	max_seq_length	4096
use_llm_lora	16	optimizer	AdamW

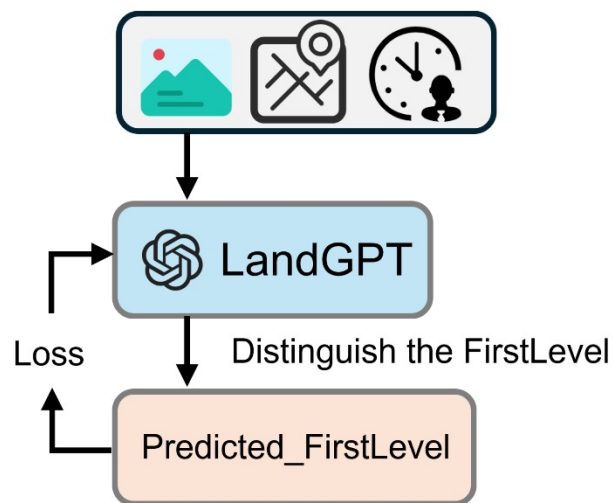
4.3 LandGPT



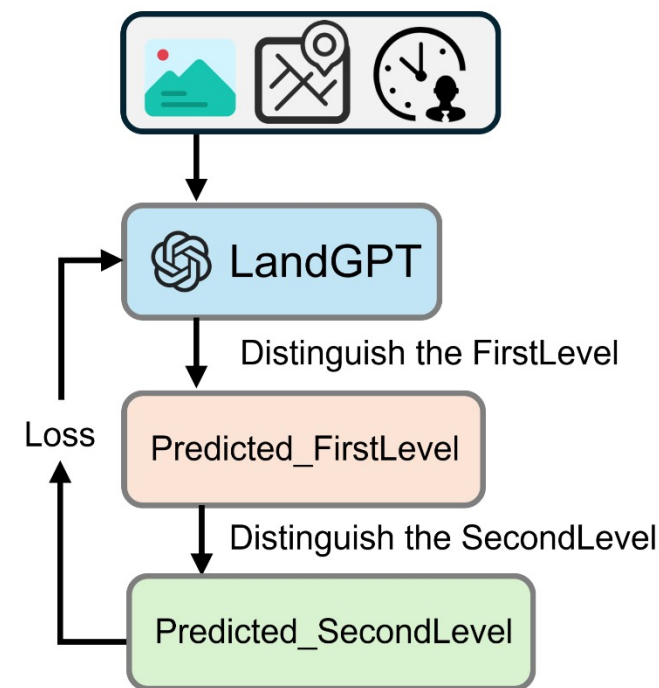
➤ “跨级判别” 框架

- 基于前文提出的渐进式微调策略，结合大语言模型的灵活指令机制，本研究进一步提出了一种“跨级判别”框架。
- 借助该框架，模型可以直接利用高精度的一级分类结果实现更为精准的二级分类，无需重复执行一级分类，从而降低了LandGPT在具体应用场景中的计算开销。

Train to distinguish the FirstLevel



Train to distinguish the SecondLevel

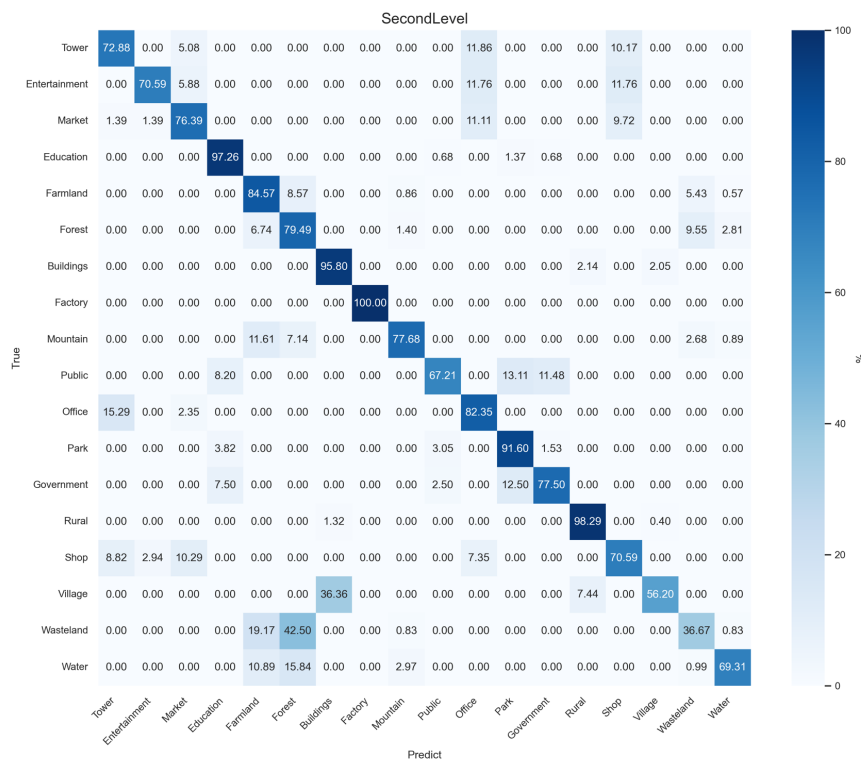


4.3 LandGPT



➤ 实验结果

- 分别在4张NVIDIA RTX 6000 Ada显卡和4张NVIDIA L20显卡进行了实验。实验一共进行了20批次约180个小时的训练。
- 实验使用约8万张训练集和约1万张测试集。在一万张测试集的测试下：LandGPT在跨级判别框架下进行的二级分类精度达89.7%，kappa系数为0.85。



LandGPT二级分类混淆矩阵图

4.3 | LandGPT



模型二级分类精度对照表 (%)
*表示采用了二级类别缩写。†表示采用了跨级判别框架。

SecondLevel	Swin	ResNet	Uknet	BDFnet	LandGPT (Ours)	LandGPT†
Tower*	2.99	2.49	0	6.97	46.6	72.88
Entertainment*	0	0	0	0	21.95	70.59
Market*	6.58	2.19	1.32	9.21	54.96	76.39
Education*	44.25	23.98	46.39	51.07	76.79	97.26
Farmland	23.25	23.98	29.66	25.64	45.44	84.57
Forest*	59.32	50.54	57.66	62.24	72.39	79.49
Buildings*	86.89	80.87	86.80	87.94	91.26	95.8
Factory*	80.58	84.25	79.54	81.28	93.77	100
Mountain	0	0	0	0	57.4	77.68
Public*	4.42	2.6	22.92	6.77	62.11	67.21
Office*	13.62	10.57	15.45	16.67	30.38	82.35
Park*	14.48	8.69	24.72	12.25	58.23	91.6
Government*	4.11	0	0.68	3.42	50.57	77.5
Rural*	84.5	87.73	83.34	85.85	79.39	98.29
Shop*	11.72	1.67	10.04	10.46	64.57	70.59
Village*	20.77	8.95	24.6	17.57	59.91	56.2
Wasteland	11.78	8.22	4.11	8.49	27.70	36.67
Water	24.26	34.55	18.54	21.28	65.28	69.31
Kappa	0.55	0.51	0.54	0.54	0.72	0.85

*Tower=Business Tower, Entertainment=Commercial Entertainment, Market=Commercial Market, Education=Educational and Research Institutions, Forest=Forestland and Grassland, Buildings=High-rise Residential Buildings, Factory=Industrial Park and Factory, Public=Non-profit Public Institutions (Museum; Stadium; Hospital), Office=Office Campus, Park=Parks and Squares, Government=Party and Government Institutions, Rural=Rural Architecture and Farmland, Shop=Shopping Center and Commercial Street, Village=Urban Villages,

4.3 LandGPT



➤ Prompt设计

本研究尝试让模型输出不同地块类别的概率，但是本研究发现模型对数字输出会有不稳定、理解具体计算不完全的问题。

询问不同类别概率

Model input

Q: Please provide its probabilities in these different FirstLevel categories: Residential Districts, Commercial Zones, Industrial Land, Public Services, and Agriculture and Nature (expressed in numerical values).

Model output

A: This image is a remote sensing image, and its probabilities in different FirstLevel categories are as follows: Residential Districts: 0.45, Commercial Zones: 0.3, Industrial Land: 0.2, Public Services: 0.1, Agriculture and Nature: 0.35.



询问不同类别概率并强调概率和为一

Model input

Q: The FirstLevel category of this image is Industrial Land. Please provide its probabilities in the following secondary categories: Industrial Park and Factory, Construction Site (expressed in numerical values, with the total sum of probabilities being 1).

Model output

A: The Industrial Park and Factory of this image is 0.9, and the Construction Site is 0.8.



切换人称尝试生成概率

Model input

Q: I need some insights on which FirstLevel category this image has the highest probability: Residential Districts, Commercial Zones, Industrial Land, Public Services, or Agriculture and Nature.

Model output

A1: This image belongs to the Residential Districts category with a probability of 0.734.

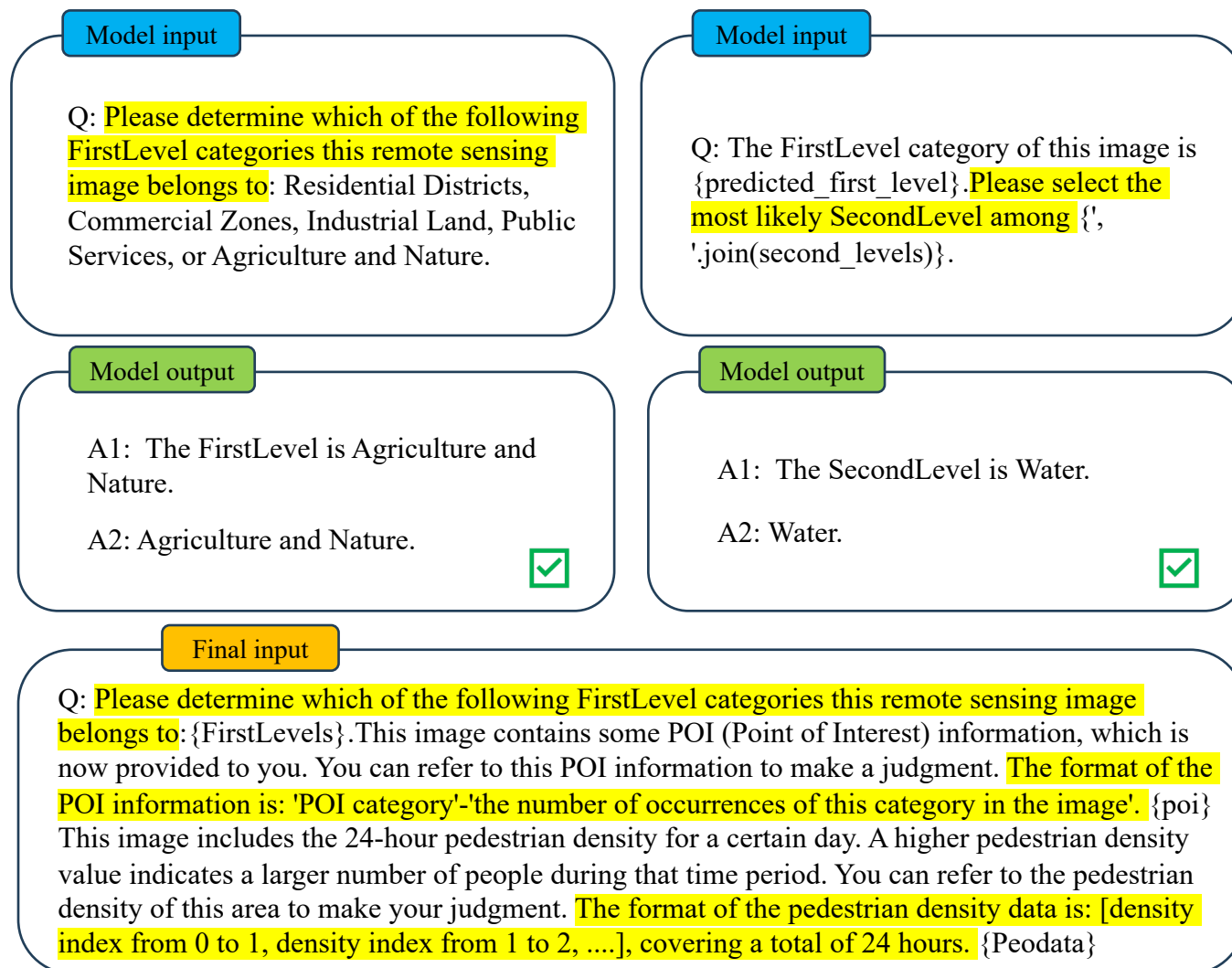
A2: This image belongs to the Residential Districts category with a probability of 0.652.



➤ Prompt设计

因此，本研究选取直接让模型输出结果的形式，进行prompt的设计和组装。最终，本研究选取了简短介绍任务要求，详细补充多源数据格式信息，明确要求给出唯一解的方式进行Prompt语句的组装。

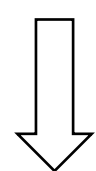
直接让模型输出结果，无需输出概率



➤ Prompt设计

本研究发现，在最后的二级推断中，加入让模型自动推断提供其的一级分类是否正确的语句，可以促使精度上升。此类型的自纠错语句有利于提升一级分类出错后，二级分类被纠正回正确类别的概率。

The FirstLevel category of this image is {predicted_first_level}. Please select the most likely SecondLevel among {'', '.join(second_levels)}. This image contains some POI (Point of Interest) information, which is now provided to you. You can refer to this POI information to make a judgment. The format of the POI information is: category'-the number of occurrences of this category in the image. {poi_info}

 精度上升2%，kappa系数上升0.1

The FirstLevel category of this image is {predicted_first_level}. Please select the most likely SecondLevel among {'', '.join(second_levels)}. This image contains some POI (Point of Interest) information, which is now provided to you. You can refer to this POI information to make a judgment. **If you believe that this image does not belong to any of the above-listed SecondLevel categories, you can also select the SecondLevel category you think this image should belong to based on the POI information below.** The format of the POI information is: category'-the number of occurrences of this category in the image. {poi_info}



Urbancomp

LandAI

Geographical Thinking Large Language Model Model

UrbanComp Lab

China University of Geosciences (Wuhan)

- 城市要素复杂交织，对城市发展至关重要。要深入挖掘其动态关联与潜在影响，需通过有效表征来精准洞察城市发展机制。
- 城市结构复杂且不断演变，对城市健康运行意义重大。深入解析其空间组织与功能关联，是优化城市布局、实现可持续发展的关键。
- 基于POI的表征方法能够有效挖掘其蕴含的语义信息与空间共现特征，广泛应用于土地利用分类等下游任务，为城市规划与管理提供有力支持。
- 多源大数据已经成为城市要素感知与分析的重要基础，为土地利用分类提供关键支撑，并能有效提高分类精度与智能化水平，为城市规划、管理和可持续发展提供有效支持。



谢谢!