



Java AI - 多语言文本特征向量提取

网址：www.aias.top

介绍

笔记简介:

- 面向对象: 深度学习初学者
- 依赖课程: **线性代数, 统计概率**, 优化理论, 图论, 离散数学, 微积分, 信息论

知乎专栏:

<https://zhuankan.zhihu.com/p/693738275>

Github & Gitee 地址:

https://github.com/mymagicpower/AIAS/tree/main/deep_learning

https://gitee.com/mymagicpower/AIAS/tree/main/deep_learning

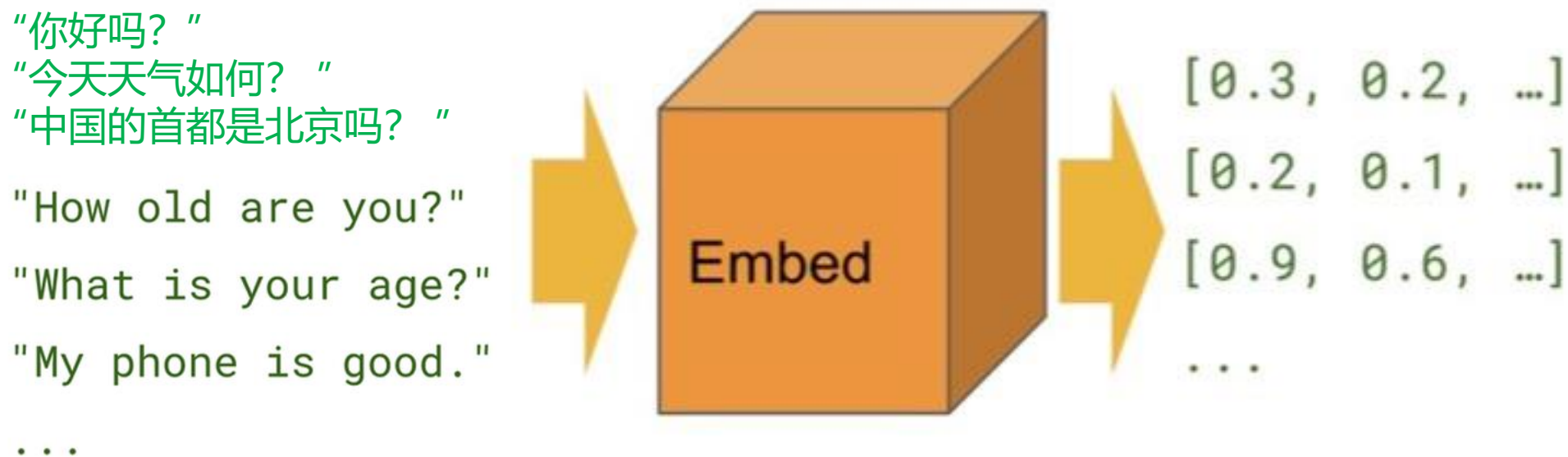
* 版权声明:

- 仅限用于个人学习
- 禁止用于任何商业用途

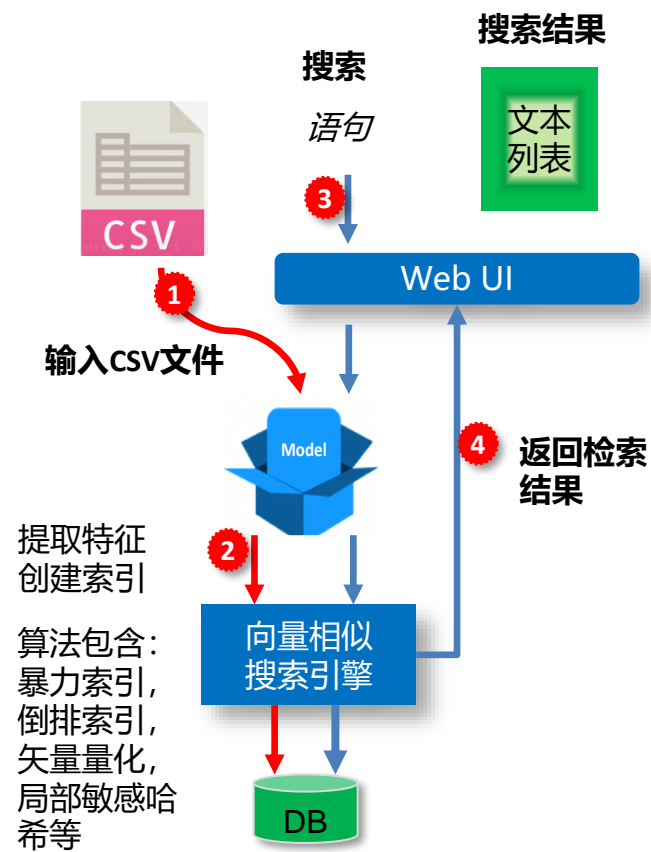
多语言文本语义特征提取工具箱

多语言文本语义特征提取工具箱提供 3 个 SDK:

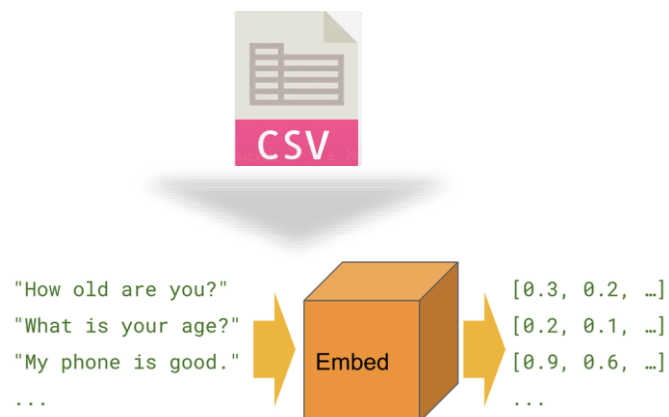
1. sentence_encoder_15_sdk (支持 15 种语言)
2. sentence_encoder_100_sdk (支持100种语言)
3. text2vec_base_multilingual_sdk (支持50+种语言)



应用场景 – 文本语义搜索原理简介

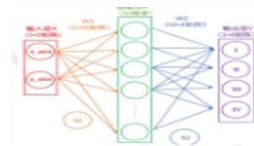


1. CSV文件



2. 特征向量

提取N维向量
(常用256/512维)



应用场景 – 用于RAG大模型领域

Dashboard / 文本搜索

5 [v] Karzai Promises Afghans Security for Election [Q]

```
▼ [  
  ▼ 0: {  
    id: 476,  
    score: 0.09515442,  
    text: ▼"Reuters - Afghanistan's President  
    Hamid Karzai promised Afghans greater  
    security when they go to vote in the  
    country's first ever democratic election  
    during an independence day speech on  
    Wednesday.",  
    title: ▼"Karzai Promises Afghans Security  
    for Election (Reuters)"  
  },  
  ▼ 1: {
```



网址

www.aias.top