

数元灵个性化推荐服务使用手册

1. 数据准备

首先，需要将 MetaSpore 项目 clone 到本地。如果从 GitHub clone 较慢，也可以从 Gitee clone <https://gitee.com/meta-soul/MetaSpore>。

```
1 mkdir work && cd work
2
3 # github clone 下载较慢，也可以选择 git clone https://gitee.com/meta-soul/MetaSpore
4 git clone https://github.com/meta-soul/MetaSpore.git
5
6 cd MetaSpore/demo/ecommerce/
```

用户需要在 `init_data/script/` 目录中准备自己电商场景的数据集，包括：用户数据、商品数据、交互数据。

2. 场景配置

为了适配特定的业务场景，用户需要更改推荐系统的配置，包括：MetaSporeFlow 的总体配置、离线 DAG 调度配置、在线服务启动配置、数据库同步配置、样本生成配置、热门召回 / I2I 召回 / CTR 排序等模型配置。此外，在系统启动之后，用户还可以更改 AB 策略配置。下面会一一进行说明。

MetaSporeFlow 总体配置

```
demo > ecommerce > ecommerce_demo > ! metaspore-flow.yml
1  apiVersion: metaspore/v1
2  kind: MetaSporeFlow
3  metadata:
4    name: demo_metaspore_flow
5    uses:
6      - ./online_local_flow.yml
7      - ./offline-local-flow.yml
8      - ./crontab_scheduler.yml
9      - ./sync_data.yml
10     - ./join_data.yml
11     - ./train_model_pop.yml
12     - ./train_model_itemcf.yml
13     - ./train_model_deepctr.yml
14     - ./notify_load_model.yml
15
16  spec:
17    sharedVolumeInContainer: '/opt/volumes'
18
19
```

在 `./ecommerce_demo/metaspore-flow.yml` 文件中，引用所有在线和离线的调度配置，可以根据场景的需要进行调整。

离线 DAG 调度配置

```
demo > ecommerce > ecommerce_demo > ! crontab_scheduler.yml
1  apiVersion: metaspore/v1
2  kind: OfflineCrontabScheduler
3  metadata:
4    name: offline_crontab_scheduler
5  spec:
6    cronExpr: "* */1 * * *"
7    dag:
8      sync_data: [ "join_data" ]
9      join_data: [ "train_model_pop", "train_model_itemcf", "train_model_deepctr", "notify_load_model" ]
```

在 `./ecommerce_demo/crontab_scheduler.yml` 中设置，描述离线调度的依赖关系，可以根据场景的需要进行调整。

在线服务启动配置

```
demo > ecommerce > ecommerce_demo > ! online_local_flow.yml
1  apiVersion: metaspore/v1
2  kind: OnlineFlow
3  metadata:
4    name: online_local_flow
5  spec:
6    dockers:
7      recommend:
8        image: swr.cn-southwest-2.myhuaweicloud.com/dmetasoul-repo/recommend-service:1.0.0
9    services:
10     mongo:
11       host: 172.17.0.1
12       port: 27018
13       kind: mongodb
14       collection: [jpa]
15       options:
16         uri: mongodb://root:test_mongodb_123456@172.17.0.1:27018/jpa?authSource=admin
17     source:
18       user_key_name: user_id
19       item_key_name: item_id
20       user_item_ids_name: user_bhv_item_seq
21       user_item_ids_split: "\u0001"
22       user:
23         table: amazonfashion_user_feature
24         serviceName: mongo
25         collection: jpa
26         columns:
27           - user_id: str
28           - user_bhv_item_seq: str
29       item:
30         table: amazonfashion_item_feature
31         serviceName: mongo
32         collection: jpa
33         columns:
34           - item_id: str
35           - brand: str
36           - category: str
37     summary:
38       table: amazonfashion_item_summary
39       serviceName: mongo
40       collection: jpa
```

在 `./ecommerce_demo/online_local_flow.yml` 中设置，可以根据场景的需要进行调整。由于不同的场景数据特征会有较大差别，这里需要对数据集进行适配。

数据库同步配置

```
demo > ecommerce > ecommerce_demo > ! sync_data.yml
```

```
1  apiVersion: metaspore/v1
2  kind: OfflinePythonTask
3  metadata:
4    name: sync_data
5  spec:
6    scriptPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/setup/setup.py"
7    configPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/setup/conf/setup.yaml"
```

在 `./ecommerce_demo/sync_data.yml` 中设置，可以根据场景的需要进行调整。

样本生成配置

```
demo > ecommerce > ecommerce_demo > ! join_data.yml
```

```
1  apiVersion: metaspore/v1
2  kind: OfflinePythonTask
3  metadata:
4    name: join_data
5  spec:
6    scriptPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/dataset/gen_samples.py"
7    configPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/dataset/conf/gen_samples.yaml"
```

在 `./ecommerce_demo/join_data.yml` 中设置，可以根据场景的需要进行调整。如果数据特征列和 demo 数据集不一致，或者对特征预处理、样本采样等步骤有定制需求，可能需要对上述配置引用的 python 代码和配置进行更改。

热门召回模型配置

```
demo > ecommerce > ecommerce_demo > ! train_model_pop.yml
1  apiVersion: metaspore/v1
2  kind: OfflinePythonTask
3  metadata:
4    name: train_model_pop
5  spec:
6    scriptPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/pop/train.py"
7    configPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/pop/conf/popular_topk.yml"
```

在 `./ecommerce_demo/train_model_pop.yml` 中设置，可以根据场景的需要进行调整。如果特征列和 demo 数据集中不一致，可能需要对上述配置引用的 python 代码和配置进行更改。

I2I 召回模型配置

```
demo > ecommerce > ecommerce_demo > ! train_model_itemcf.yml
1  apiVersion: metaspore/v1
2  kind: OfflinePythonTask
3  metadata:
4    name: train_model_itemcf
5  spec:
6    scriptPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/itemcf/train.py"
7    configPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/itemcf/conf/swing_i2i.yaml"
```

在 `./ecommerce_demo/train_model_itemcf.yml` 中设置，可以根据场景的需要进行调整。如果特征列和 demo 数据集中不一致，可能需要对上述配置引用的 python 代码和配置进行更改。

排序模型配置

```
demo > ecommerce > ecommerce_demo > ! train_model_deepctr.yml
1  apiVersion: metaspore/v1
2  kind: OfflinePythonTask
3  metadata:
4    name: train_model_deepctr
5  spec:
6    scriptPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/deepctr/train.py"
7    configPath: "/opt/volumes/ecommerce_demo/MetaSpore/demo/ecommerce/deepctr/conf/deep_ctr.yaml"
```

在 `./ecommerce_demo/train_model_deepctr.yml` 中设置，可以根据场景的需要进行调整。如果特征列和 demo 数据集中不一致，可能需要对上述配置引用的 python 代码和配置进行更改。

系统启动和停止

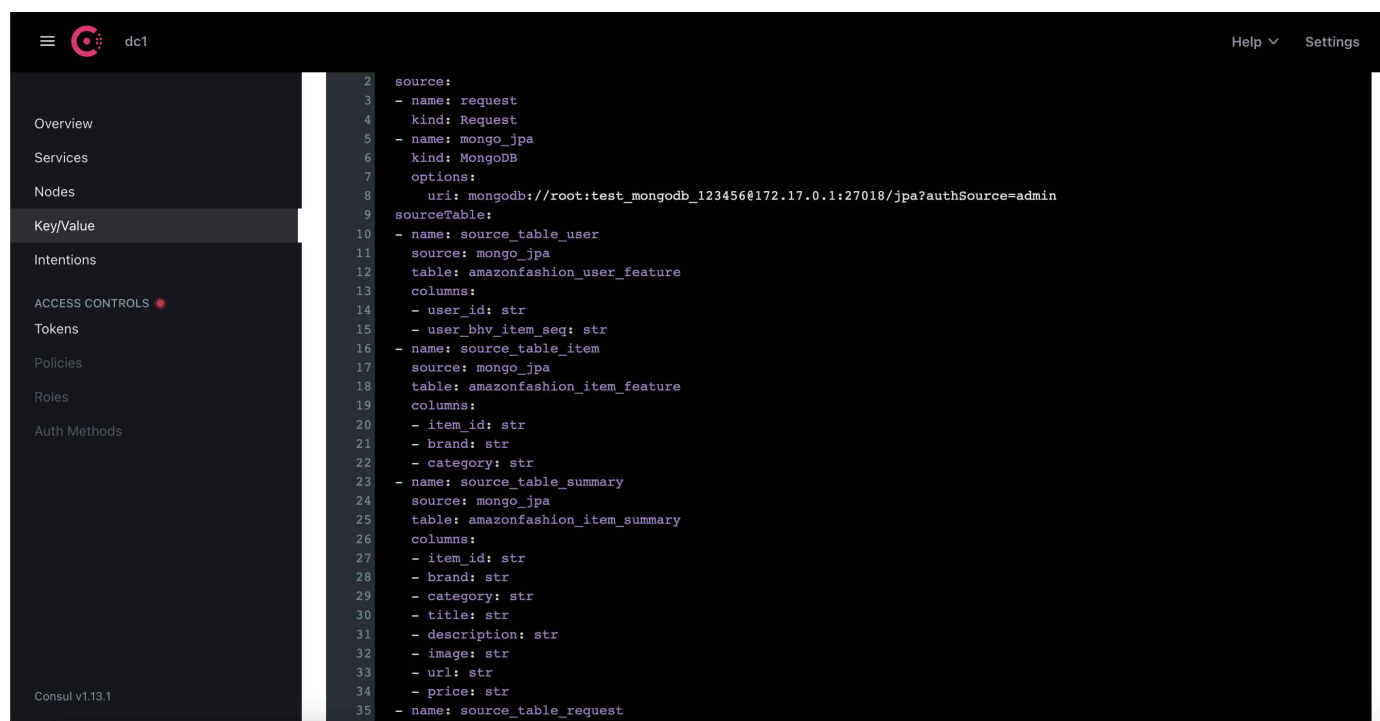
完成以上配置之后，可以通过运行 `./start.sh` 和 `./stop.sh` 进行一键启动、停止智能推荐系统。当启动完成后，如果在本机运行，可以使用 curl 命令进行验证：。

```
1 # 请使用真实用户 ID 替换 USER_ID 变量
2 curl -H "Accept: application/json" -H "Content-type: application/json" -X POST
   -d '{"user_id": "${USER_ID}"}' http://127.0.0.1:8081/service/recommend/guess-
   you-like
```

如果在远程服务器上运行，可以使用 SSH 隧道进行端口映射后，进行访问。

推荐策略实时调整

此外，当系统启动之后，我们还可以实时更新系统的 AB 策略。打开 <http://localhost:8500/ui/>，进入 Consul Web UI 界面，客户可以根据实际情况对推荐策略进行灵活调整：



3. Amazon 女装推荐 Demo

我们使用 [Amazon Review Dataset](#) 的女装子数据集进行演示。我们在 GitHub 项目中，加入了已经处理好的数据集：amazon_fashion_interaction.small.parquet、amazon_fashion_user.small.parquet、amazon_fashion_item.small.parquet。

假设当前的工作目录在 `~/work/MetaSpore` 中，可以一键启动智能推荐服务：

```
1 cd demo/ecommerce/  
2 ./start.sh
```

当启动完成后，如果在本机运行，可以打开网址：<http://localhost:41730> 查看推荐效果。如果在远程服务器上运行，可以使用 SSH 隧道进行端口映射后，进行访问。

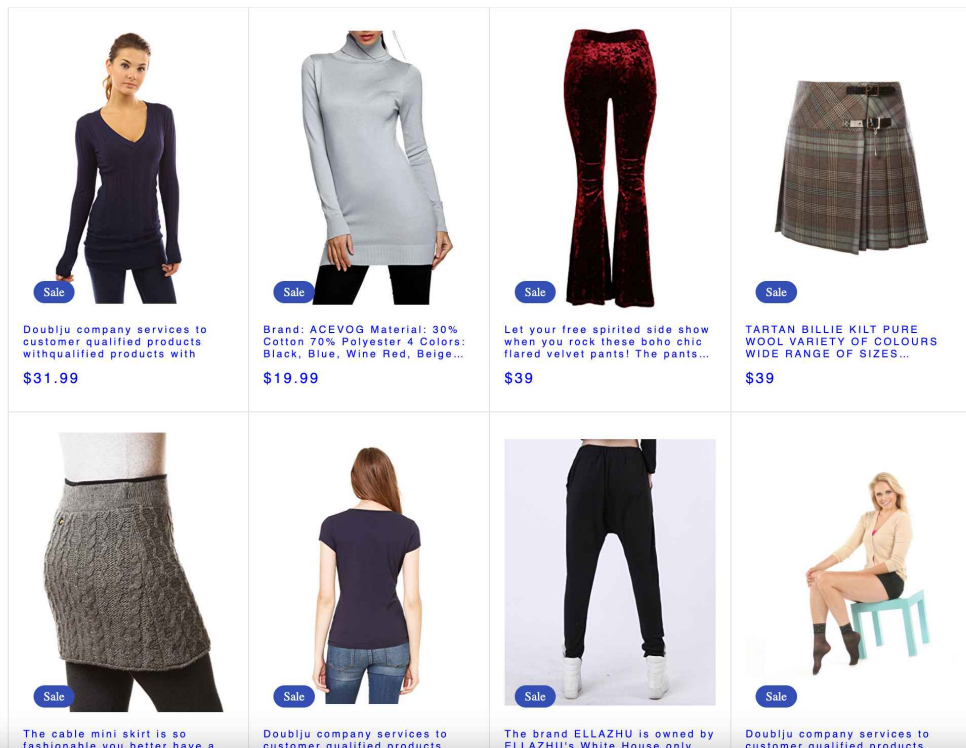
• 猜你喜欢推荐

Featured Products

A23P7HJBRQ0F7L

 Sale Cotton Blend Boat Neck Cap Sleeve Fitted Summer Casual Dress. \$28.99	 Sale Description: Material:Knitted blend Color:Red,Black,Grey Style:Sexy Club,Regular 100% Fashion sexy and... \$9.99	 Sale With embroidery work in the front. Suitable for all occasions. Unique patched products . \$32.99	 Sale "Amazing and Affordable!" "Great Quality, Comfortable and Cute!" "I Am In LOVE With This Dress!" "Very..." \$39
			

• 看了又看推荐



此外，用户也可以通过官网直接预览推荐效果：或 <http://recommend-demo.dmetasoul.com/>
如果想停用推荐服务，可以通过运行以下命令来完成。

```
1 ./stop.sh
```

如果有特殊的场景定制需求，欢迎您联系数元灵技术支持团队，感谢您的使用。