

# GOTC

## 全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

# OPEN SOURCE , OPEN WORLD #

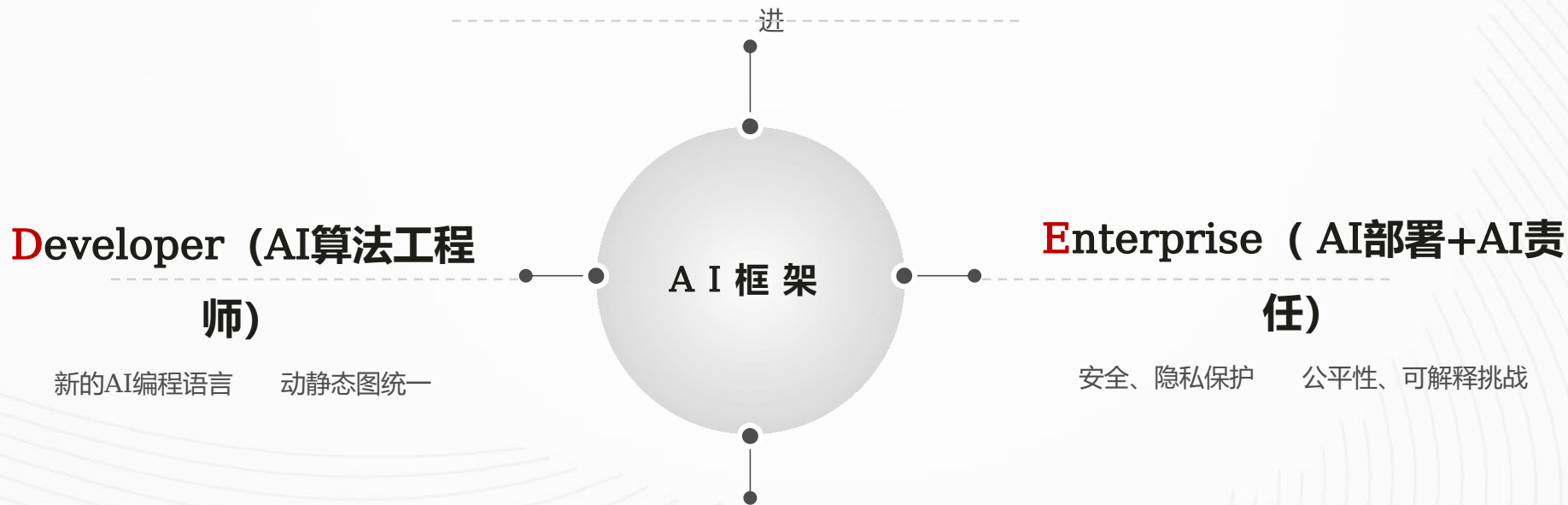
### 「AI、大数据与数字经济开源技术论坛」专场

MindSpore高阶API工具TinyMS架构设计与实践

胡晓曼 2021年7月10日

## Application+Bigdata ( AI模型和数据 )

模型规模和复杂度越来越高 (GPT-3参数量达1750亿)      从单一的NN向通用AI和科学计算演进



**Developer (AI算法工程师)**

新的AI编程语言      动静态图统一

**Enterprise ( AI部署+AI责任)**

安全、隐私保护      公平性、可解释挑战

## Chip (AI芯片和算力)

芯片/集群性能持续提升 (Atlas 900集群支持最大E级算力拓展)      CPU / GPU / NPU 多样化异构算力

# MindSpore开源社区现状

MindSpore  
累计下载量

41万+

贡献者 2200+人  
核心贡献者 900+人

MindSpore开  
发者数量

33万+

开发者积极贡献：  
Issue数1w+，PR数  
2.6w+

社区生态  
应用数量

120+

ModelZoo

2000+

上线应用

官网访问: 1500万+

国内第一热度AI开源社区

师资培训全国主要城市巡讲: 深圳、杭州、  
武汉、上海、成都、南京等场次,  
MindSpore 动手实践帮助教师迅速上手

国内举办了14场城市MSG, 在线  
举办了7场海外MSG(俄罗斯、印  
尼、新加坡、印度、中东), 以及6  
场校园行MSG

## 7大行业联合创新

医疗



“素问平台”

交通



智慧交通”

金融



“智慧金融”

“金融大脑”

互联网



好货推荐

内容安全

运营商



“智能运维”

制造



“工业质检”

能源

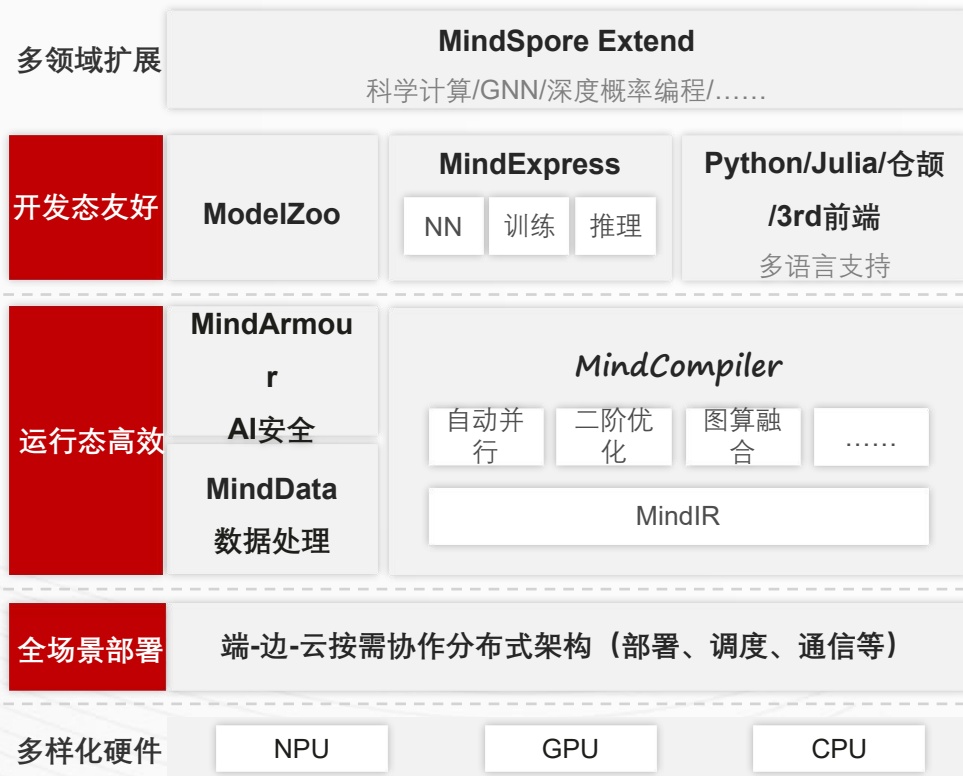


华为云

# 全球开源技术峰会

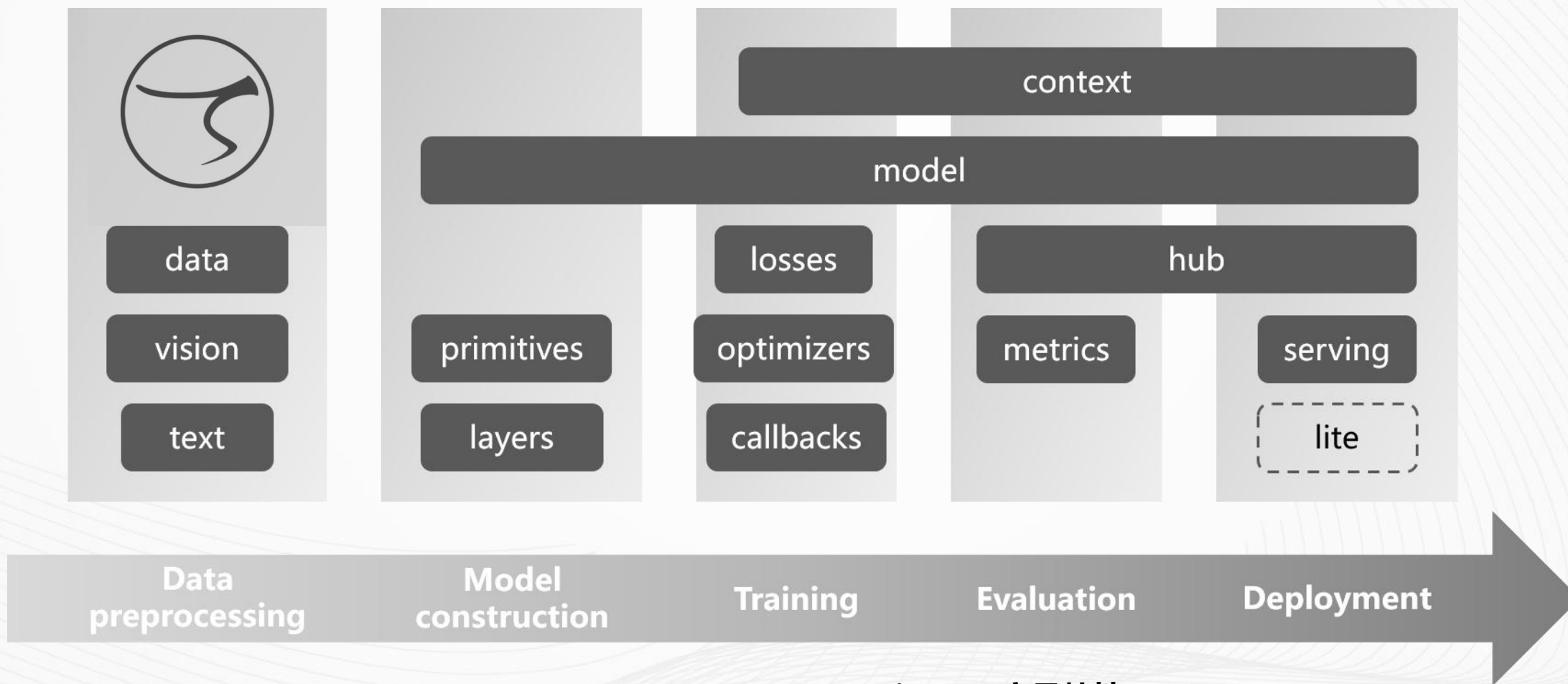
THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

## 整体架构



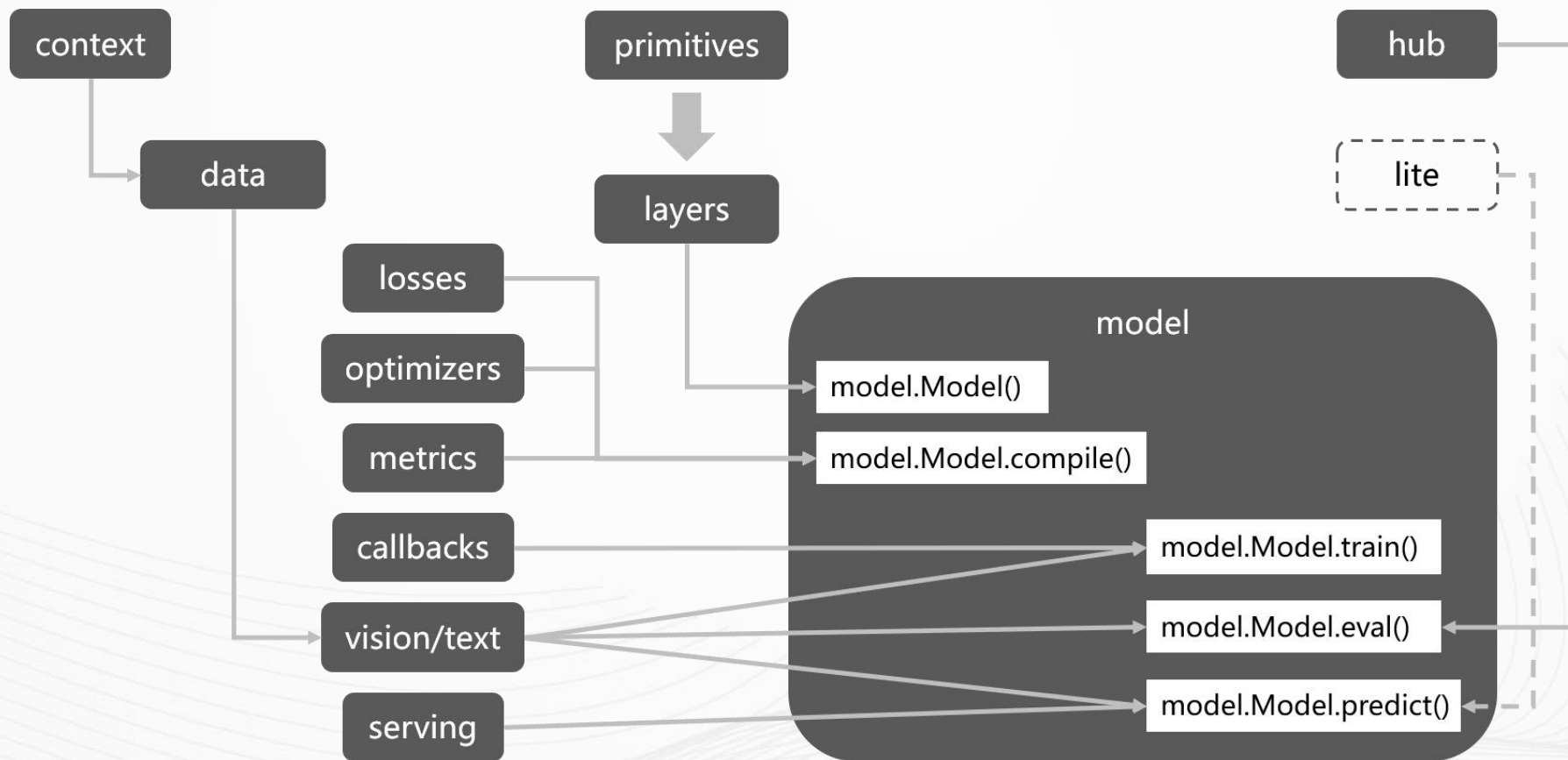
## 关键特性

- 1. 自动并行:** 动整图切分, 感知集群拓扑, 实现通信开销最小, 融合数据并行与模型并行;
- 2. 二阶优化:** 利用二阶计算修正梯度更新方向, 找到训练梯度最优下降路径, 从而加速训练收敛过程;
- 3. 动静态图结合:** 统一自动微分引擎支持动静态图, 一行代码完成模式切换, 兼顾模型开发和执行效率;
- 4. 全场景部署协同:** 超轻量技术, 模型压缩、自适应模型生成、IoT超轻量运行时, 基于异构的混合同行;
- 5. 全栈协同加速:** 通过多级并行、图算融合等深度图优化, 以及跨层内存复用技术, 发挥昇腾极致性能;
- 6. 迁移及调试工具:** 覆盖生态迁移、精度调优、性能调优全场景、可视化, 脚本迁移自动化, 降低开发门槛, 提高调试效率;
- 7. AI+科学计算, 场景应用创新, 拓展MindSpore的边界**
- 8. 安全可靠, 首个企业级AI可信计算框架 (Trusted AI)**



## TinyMS 产品特性

- 由MindSpore开源运营团队打造的全新MindSpore高阶API工具
- 让开发者1分钟上手体验全流程：训练、预测、部署一套走通
- 配备保姆级深度学习教程，真正对纯小白友好的零基础工具



# TinyMS模块介绍

## Data Preprocessing

### Step 1: data download

TinyMS supports users to complete the download, decompression and format adjustment operations of the data set with one click.

```
from tinynms.data import download_dataset

mnist_path = download_dataset('mnist', local_path='./')
```

```
from tinynms.data import MnistDataset

mnist_ds = MnistDataset(mnist_path, shuffle=True)
```

### Step 2: data loading

By completely inheriting MindSpore's native API, users can use the xxxDataset interface to instantiate different data sets very conveniently.

```
from PIL import Image
from tinynms.vision import mnist_transform

# Preprocessing a single one picture
img = mnist_transform(Image.open('picture.jpg'))
# Apply preprocessing to MnistDataset class instance
mnist_ds = mnist_transform.apply_ds(mnist_ds)
```

### Step 3: data processing

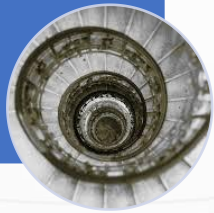
TinyMS directly corresponds to the processing of the dataset itself through the DatasetTransform interface, allowing users to utilize a single piece of data or the entire dataset with just one line of code.

# TinyMS模块介绍

## Training

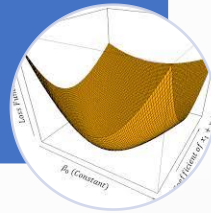
- TinyMS high-level API encapsulates the loss functions at the network level.

Loss  
function



- No need to know how optimizer works and use it with a single line of code.

Optimizer



- Automatically configure the training callback strategy.

Callback  
function



```
from tinymms.losses import SoftmaxCrossEntropyWithLogits
from tinymms.optimizers import Momentum

net_loss = SoftmaxCrossEntropyWithLogits(sparse=True, reduction='mean')
net_opt = Momentum(net.trainable_params(), 0.01, 0.9)
```

```
from tinymms.callbacks import mobilenetv2_cb

net_cb = mobilenetv2_cb()
```



## Evaluation

Metrics class	Introduction	Metrics class	Introduction
Accuracy	Calculates the accuracy for classification and multilabel data.	Metric	Base class of metric.
MAE	Calculates the mean absolute error.	Recall	Calculates recall for classification and multilabel data.
MSE	Measures the mean squared error.	Fbeta	Calculates the fbeta score.
Precision	Calculates precision for classification and multilabel data.	F1	Calculates the F1 score.
TopKCategoryical Accuracy	Calculates the top-k categorical accuracy.	Loss	Calculates the average of the loss.

### How to use?

```
from tinymms.model import Model
from tinymms.metrics import Accuracy

model = Model(net)
model.compile(metrics={"Accuracy": Accuracy()})
model.eval(ds_eval)
```

## Deployment

**Serving** is a module that supports performing model deploy and inference on both computers and mobile phones, with a **unified predict interface** to users.

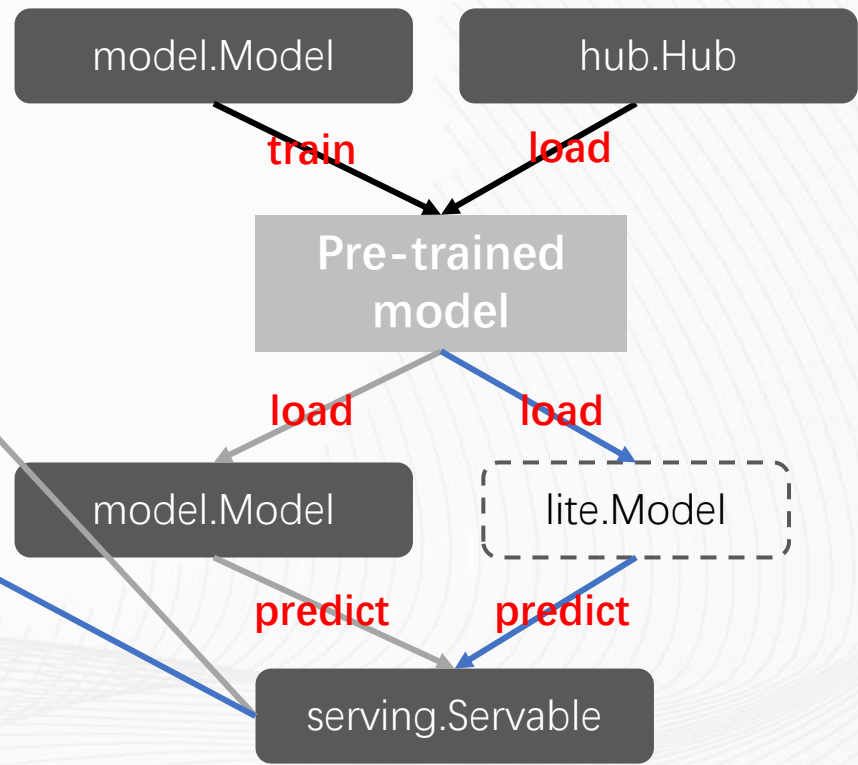


### TinyMS Serving Client

predict()

predict()

### TinyMS Serving Server



## Model construction

### Key features:

- Super easy to instantiate a network
- Provide a high-level Model API with the consistent development experience for TinyMS users

TinyMS	AI Frameworks (e.g. MindSpore)
<pre>from tinynms.model import lenet5  net = lenet5(class_num=10)</pre>	<pre>import mindspore.nn as nn  class LeNet5(nn.Cell):     def __init__(self, num_class=10, num_channel=1):         super(LeNet5, self).__init__()         self.conv1 = nn.Conv2d(num_channel, 6, 5, pad_mode='valid')         self.conv2 = nn.Conv2d(6, 16, 5, pad_mode='valid')         self.relu = nn.ReLU()         self.max_pool2d = nn.MaxPool2d(kernel_size=2, stride=2)         self.flatten = nn.Flatten()         self.fc1 = nn.Dense(16 * 5 * 5, 120, weight_init=Normal(0.02))         self.fc2 = nn.Dense(120, 84, weight_init=Normal(0.02))         self.fc3 = nn.Dense(84, num_class, weight_init=Normal(0.02))      def construct(self, x):         x = self.max_pool2d(self.relu(self.conv1(x)))         x = self.max_pool2d(self.relu(self.conv2(x)))         x = self.flatten(x)         x = self.relu(self.fc1(x))         x = self.relu(self.fc2(x))         x = self.fc3(x)         return x  net = LeNet5(class_num=10)</pre>

# TinyMS模块介绍

## Hub

TinyMS Hub是TinyMS生态的预训练模型应用工具，提供的预训练模型主要包括图像分类、目标检测、语义模型、推荐模型

- 向模型开发者提供方便快捷的模型发布、提交通道；
- 向应用开发者提供高质量的预训练模型，结合模型加载以及模型Fine-tune API快速完成模型的迁移到部署的工作。

## 加载预训练模型

```
from PIL import Image
from tinymms import hub
from tinymms.vision import mnist_transform
from tinymms.model import Model

img = Image.open(img_path)
img = mnist_transform(img)

# load LeNet5 pre-trained model
net = hub.load('tinymms/0.2/lenet5_v1_mnist', class_num=10)
model = Model(net)

res = model.predict(ts.expand_dims(ts.array(img), 0)).asnumpy()
print("The label is:", mnist_transform.postprocess(res))
```

## 加载模型ckpt文件

```
from tinymms import hub
from tinymms.model import lenet5
from tinymms.utils.train import load_checkpoint

ckpt_dist_file = '/tmp/lenet5.ckpt'
hub.load_checkpoint('tinymms/0.2/lenet5_v1_mnist', ckpt_dist_file)
net = lenet5()
load_checkpoint(ckpt_dist_file, net=net)
```

## 加载模型权重

```
from tinymms.utils.train import load_param_into_net

param_dict = hub.load_weights('tinymms/0.2/lenet5_v1_mnist')
net = lenet5()
load_param_into_net(net, param_dict)
```

配备保姆级AI课程，发布模型复现比赛，不到17h开发者成功复现

## TinyMS保姆级教程 4



【保姆级教程】EP04-30min  
速成Python指南（上） |  
896 4-23



【保姆级教程】EP03-30min  
速成Shell脚本命令  
7169 4-1



【保姆级教程】EP02-计算机  
是如何识别图像的？+TinyMS  
2483 3-31



【保姆级教程】EP01-最适合  
小白的深度学习入门课程来  
4628 3-30

领域	网络模型	难度	数据集
计算机视觉 (CV)	<a href="#">AlexNet</a>	初级	cifar10
计算机视觉 (CV)	<a href="#">DenseNet100</a>	初级	cifar10
自然语言处理 (NLP)	<a href="#">BERT</a>	中级	zhwiki
自然语言处理 (NLP)	<a href="#">SentimentNet</a>	中级	aclImdb_v1

**TinyMS**  
Tiny Modular Simple Development Kit for Deep Learning Frameworks  
<https://tinymms-ai.github.io>

### DenseNet100\_TinyMS\_XJY #54

zjuter0126 wants to merge 8 commits into tinymms-ai:main from zjuter0126:main

Conversation 3 | Commits 8 | Checks 4 | Files changed 7 | +468 -0

zjuter0126 commented on 31 Mar

What type of PR is this?

```
Uncomment only one /kind <> line, hit enter to put that in a new line, and remove leading whitespaces from that line:  
  
/kind bug  
  
/kind enhancement
```

8 Open | 79 Closed

- Amend the chinese version of hub contributing guideline ✓ #95 opened 24 days ago by leonwanghui • Review required
- finish AlexNet #92 opened on 1 Jun by zhuyifeng888 • Review required
- AlexNet\_TinyMS\_JeffDing ✓ #69 opened on 5 Apr by JeffDing • Review required
- DenseNet100\_TinyMs\_Qiu ✓ #61 opened on 1 Apr by yexijoe • Review required
- DenseNetBC100 ✓ #60 opened on 1 Apr by huicui1996 • Review required
- AlexNet\_TinyMS\_Qiu ✓ #56 opened on 31 Mar by yexijoe • Review required
- AlexNet\_TinyMS\_CuiHui\_ZJUT ✓ #55 opened on 31 Mar by huicui1996 • Review required

## MindSpore Study Group (MSG) 全球组织

区域型

高校型

企业型

**国内十三大城市：**北京、上海、深圳、广州、杭州、苏州、长沙、重庆、武汉、郑州、山东、南京、天津  
**海外七大城市：**莫斯科、印尼、新加坡、印度、中东、澳门、香港

**五场校园行：**哈工大深圳校区、北航、南方科技大学、天津大学、南京中医药大学等

联合**白玉兰开源、Julia社区、Graviti、Artefact、云启资本**等联合举办**初创企业行MSG、MSG-WomenInTech**等



聚焦三维用户矩阵，启动组织者自运转模式

14

# 全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

### MINDSPORE MSG企业行 南京站

革新未来 创造世界

2021.7.18 14:00-18:00

<b>符秋杰</b> 江苏昇腾生态创新中心CTO	<b>胡晓曼</b> 华为MindSpore运营总监
<b>刘占峰</b> 南京移嘉数据科技 产品研发中心总经理	<b>陈朕</b> AI在智能印章领域的应用及探索
<b>张晨</b> 创尔科技创始人、CEO	<b>李晓雪</b> 格物致市场总监
<b>朱光辉</b> 江苏鸿程大数据技术与应用 研究院副总经理	<b>胡小敏</b> 安元工业互联网领域的人工智能 解决方案经理
<b>杜振东</b> 南京云尚网络科技 区块链负责人	

嘉宾持续邀请中

主办方：Ascend 江苏人工智能生态创新中心（南京） MindSpore

协办方：DataFun

活动时间：2021年7月18日 13:30  
活动地址：江苏省南京市浦口区七里桥北路6号 南京集成电路培训基地

扫码报名

# 加入我们

GOTC

- Website: [www.mindspore.cn](http://www.mindspore.cn)
- Gitee: [gitee.com/mindspore](https://gitee.com/mindspore)
- GitHub: [github.com/mindspore-ai](https://github.com/mindspore-ai)
- iHub: <https://code.ihub.org.cn/companies/4vioxkz2>
- Mailing-list: <https://mailweb.mindspore.cn/postorius/lists/>



微信公众号



QQ交流群



Bilibili



抖音



官网

全球开源技术峰会

THE GLOBAL OPEN SOURCE TECHNOLOGY CONFERENCE

**GOTC**

**THANKS**

**全球开源技术峰会**

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE