

GLOBAL INTERNET ARCHITECTURE CONFERENCE

机器学习平台与开源技术融合

陈迪豪



## **GIAC**

#### 全球互联网架构大会

GLOBAL INTERNET ARCHITECTURE CONFERENCE



关注msup 公众号获得 更多案例实践 GIAC 是中国互联网技术领域行业盛事,组委会从互联网架构最热门领域甄选前沿的有典型代表的技术创新及研发实践的架构案例,分享他们在本年度最值得总结、盘点的实践启示。

2018年11月 | 上海国际会议中心



高可用架构 改变互联网 的构建方式



## Agenda

- ❖ 机器学习平台功能模块
- ❖ 机器学习工作流引擎
- ❖ 机器学习模型管理实践
- ❖ 智能机器学习平台与未来



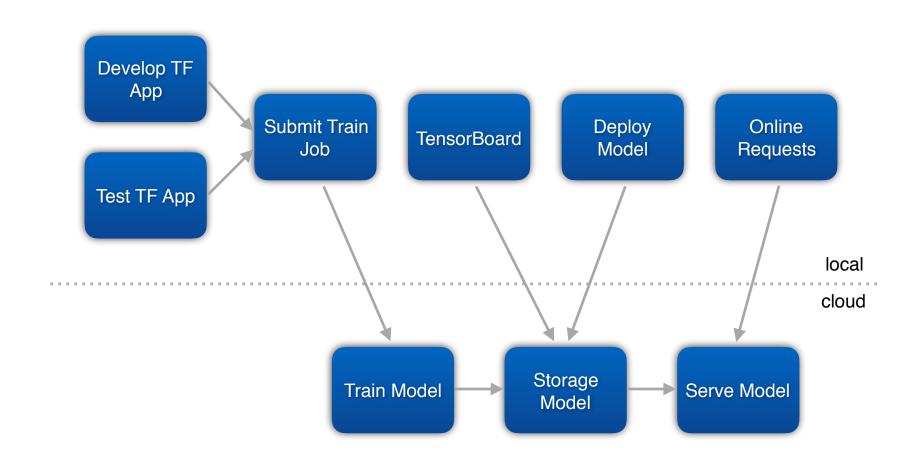
## GAC 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE

## Agenda

- ❖ 机器学习平台功能模块
- ❖ 机器学习工作流引擎
- ❖ 机器学习模型管理实践
- ❖ 智能机器学习平台与未来



## 机器学习平台功能模块





## 机器学习平台功能模块





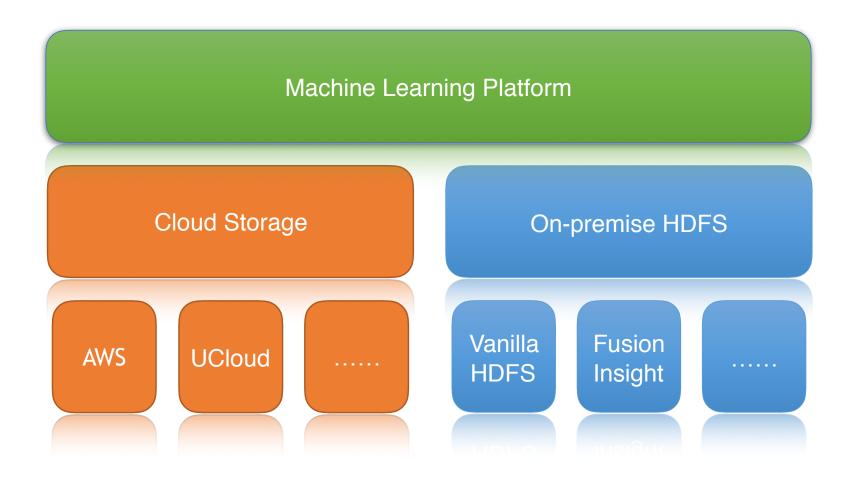
### 机器学习平台架构 - 元素系统

### **Element System**

- Table(DataFrame, Parquet...)
- Instance(GDBT format, TFRecords...)
- Model(GDBT format, TF SavedModel...)
- Report(Model evaluation report...)
- Log(Model training log, TF event files...)
- Script(FE script, Keras model...)



## 机器学习平台架构 - 元素系统





## 机器学习平台架构 - 算法集成

### **GDBT Algorithms**

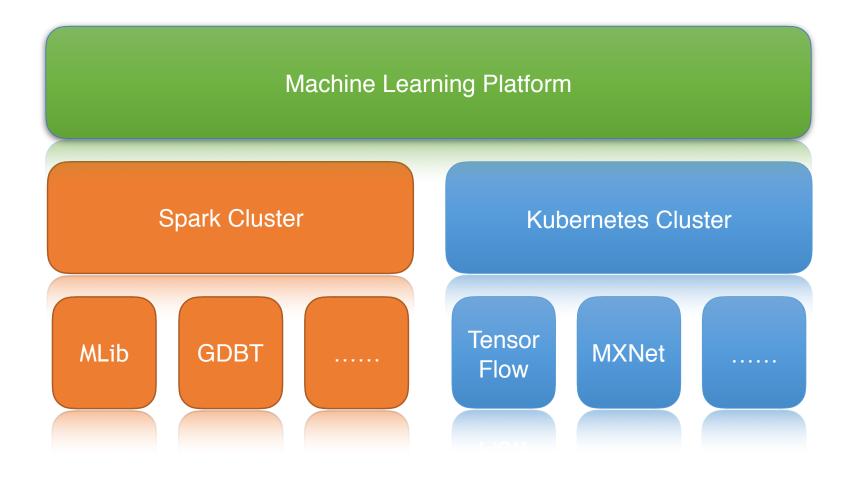
- High CV(LR, GBDT, SVM, HE-TreeNet, DNN)
- Auto ML(Auto-Feature, HP-Auto-Tune operators)

### **Open-source Algorithms**

- TensorFlow(Wide-and-deep, TextCNN, CTC)
- MXNet(Faster-RCNN, DeepText)

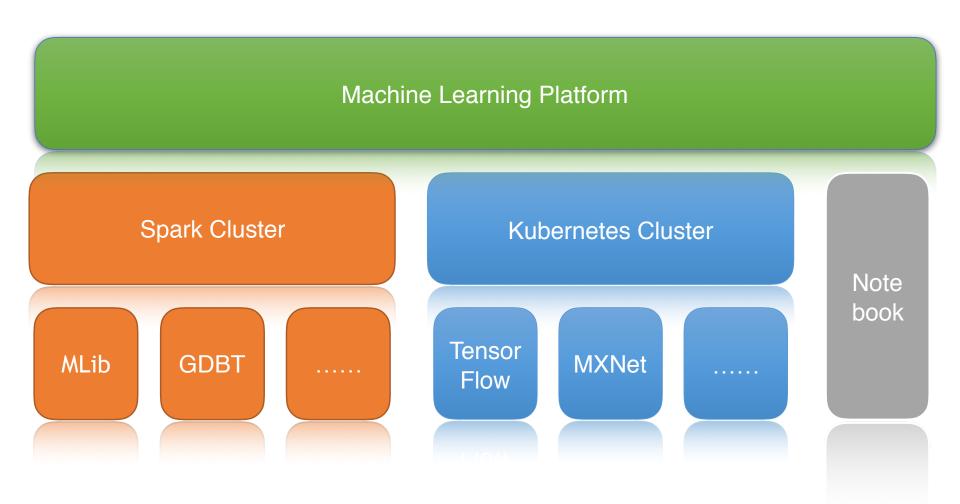


## 机器学习平台架构 - 算法集成





## 机器学习平台架构 - 算法集成





## GAC 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE

## 机器学习平台架构 - 在线服务

### **ML In Production**

- TSDB for online features(RtiDB)
- Workflow management
- Models and services
- Self-learning
- AB Testing



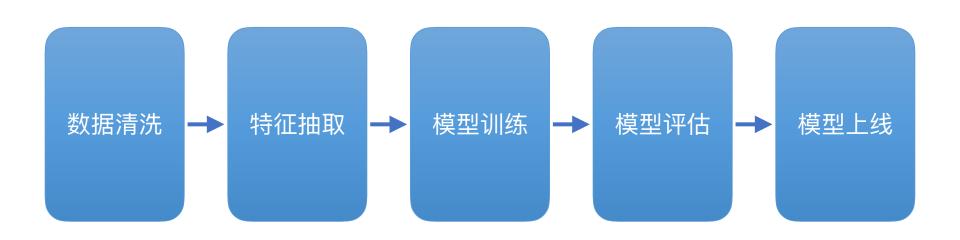
## GAC 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE

## Agenda

- ❖ 机器学习平台功能模块
- ❖ 机器学习工作流引擎
- ❖ 机器学习模型管理实践
- ❖ 智能机器学习平台与未来

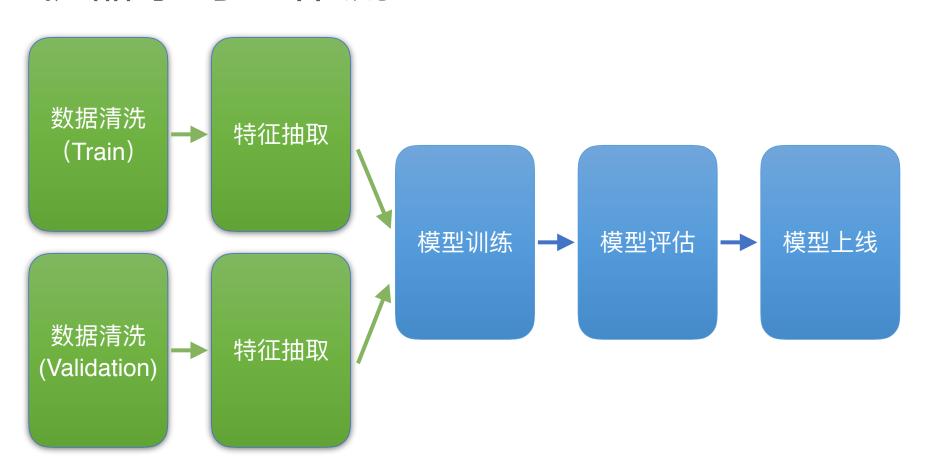


## 机器学习工作流





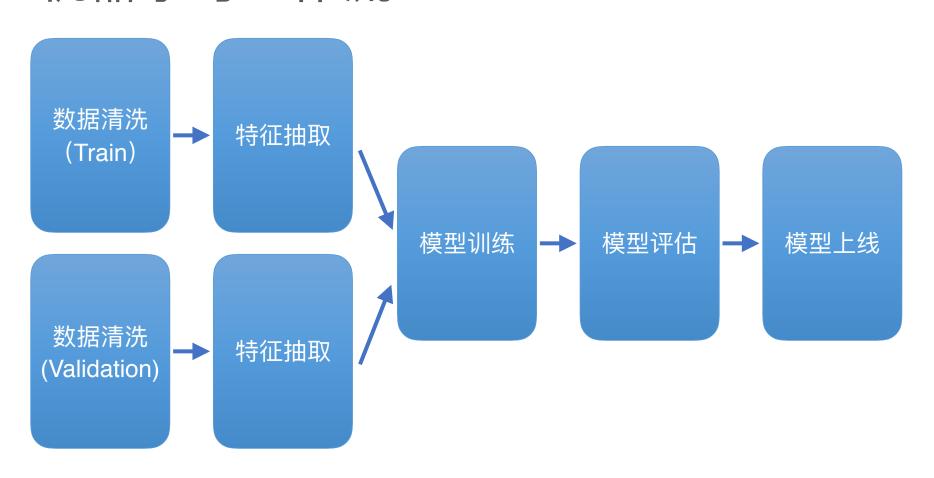
## 机器学习工作流





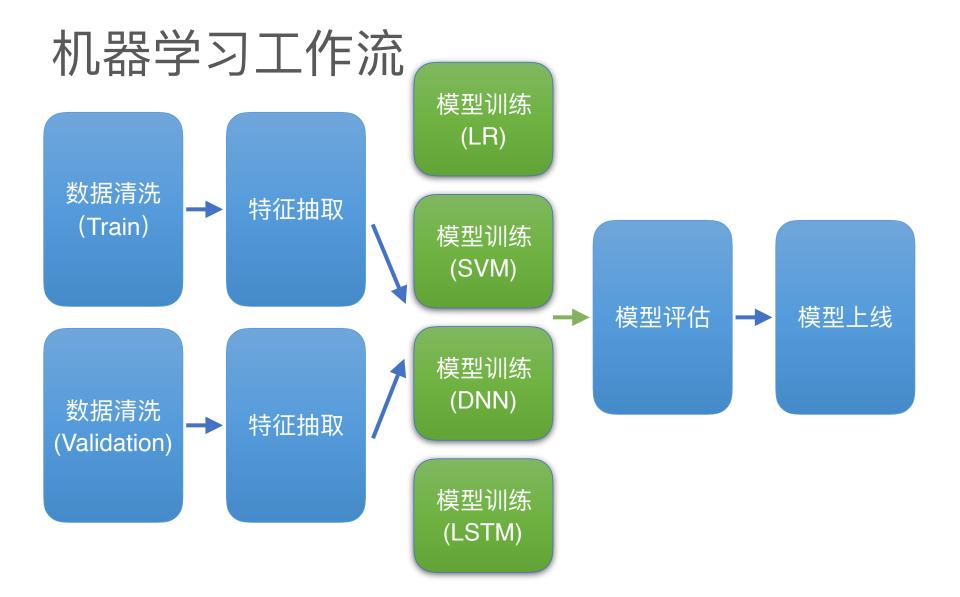


## 机器学习工作流



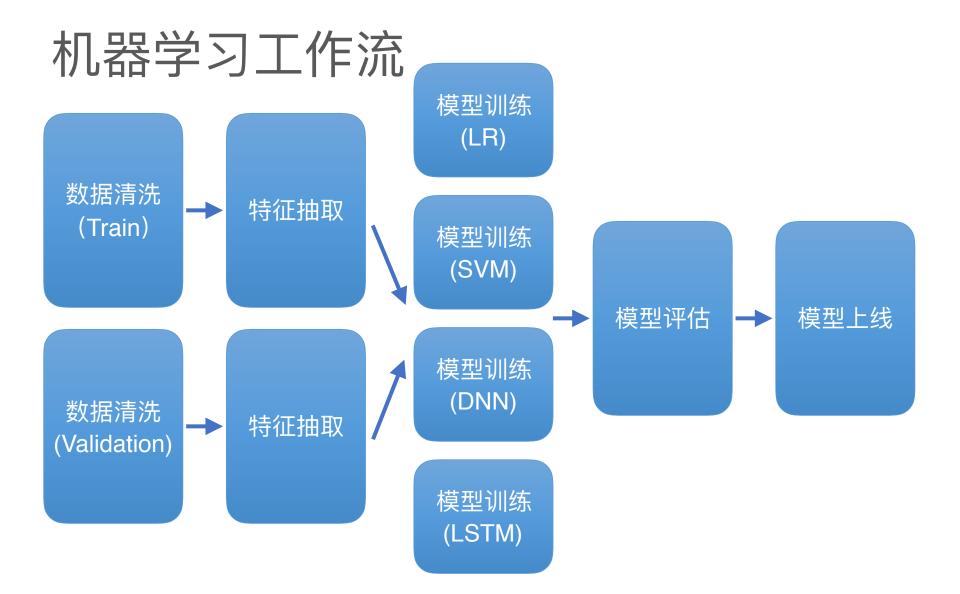


## 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE



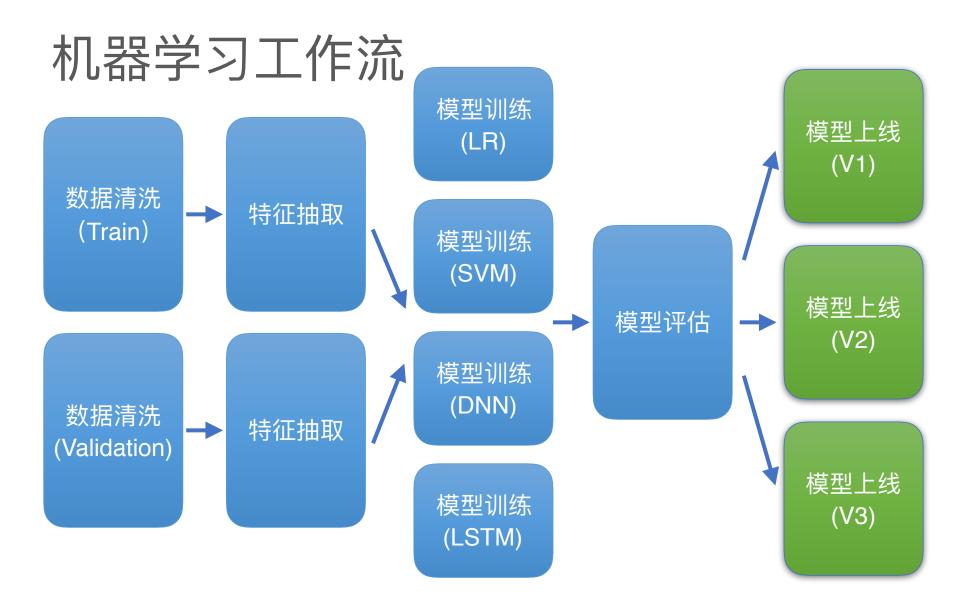


## G A C 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE





## GAC 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE





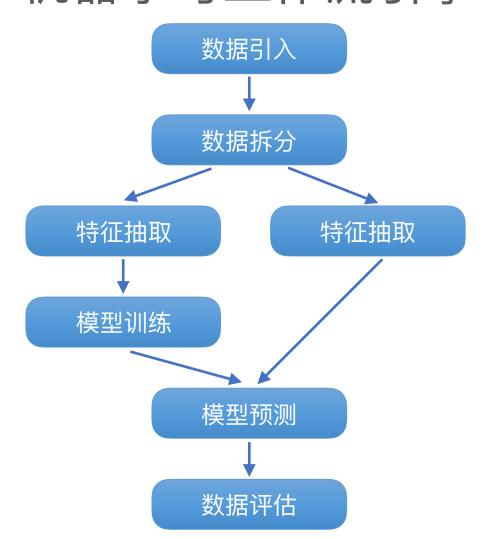
## 机器学习工作流引擎

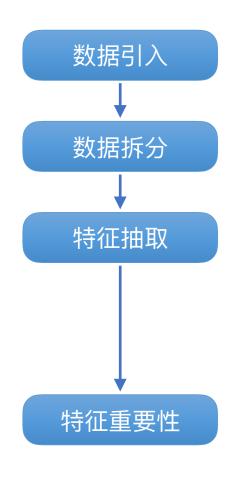
- ❖ Spark: 数据血缘关系
- ❖ TensorFlow: Op计算图
- ❖ DagChain(BlockChain): 侧链管理





## 机器学习工作流引擎







## 机器学习工作流引擎

- ❖ Operator算子开发
- ❖ Workflow持久化
- ❖ DAG搜索算法
  - · 运行当前节点
  - · 运行至当前节点
  - · 从当前节点运行



## Agenda

- ❖ 机器学习平台功能模块
- ❖ 机器学习工作流引擎
- ❖ 机器学习模型管理实践
- ❖ 智能机器学习平台与未来

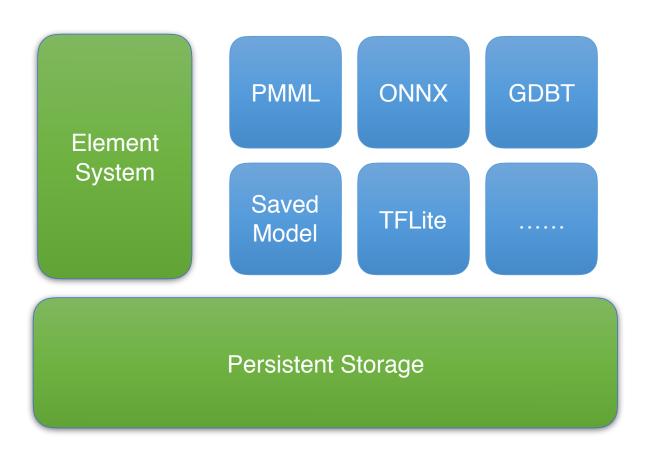


## GAC 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE

## 机器学习模型格式

- ♣ PMML
- **\*** ONNX
- GDBT format
- TensorFlow SavedModel
- TensorFlow TFLite
- **\***







- ♣ PMML
  - XML
  - Java server
  - Algorithms

```
<PMNL xnlns="http://www.dng.org/PMNL-4_1" version="4.1">
 <Header copyright="KNIHE">
  <Application name="KNIME" version="2.8.0"/>
</Header>
  <DataDictionary numberOfFields="5">
      <DataField dataType="double" name="sepal_length" optype="continuous">
    <Interval closure="closedClosed" leftMargin="4.3" rightMargin="7.9"/>
      <DataField dataType="double" name="sepal_width" optype="continuous">
      <Interval closure="closedClosed" leftMargin="2.0" rightMargin="4.4"/>
    <DataField dataType="double" name="petal_length" optype="continuous">
      <Interval closure="closedClosed" leftMargin="1.0" rightMargin="6.9"/>
    </DataField>
    <DataField dataType="double" name="petal_width" optype="continuous">
      <Interval closure="closedClosed" leftMargin="0.1" rightMargin="2.5"/>
    </DataField>
    <DataField dataType="string" name="class" optype="categorical">
      <Value value="Iris-setosa"/>
      <Value value="Iris-versicolor"/>
      <Value value="Iris-virginica"/>
    </DataField>
  </DataDictionary>
  <TransformationDictionary/>
    <SupportVectorMachineModel modelName="SVM" functionName="classification"</p>
algorithnName="Sequential Minimal Optimization (SMO)" synRepresentation="SupportVectors">
  <MiningSchena>
    <MiningField name="sepal_length" invalidValueTreatment="asIs"/>
    <MiningField name="sepal_width" invalidValueTreatment="asIs"/>
    <MiningField name="petal_length" invalidValueTreatment="asIs"/>
 <MiningField name="petal_width" invalidValueTreatment="asIs"/>
<MiningField name="class" invalidValueTreatment="asIs" usageType='predicted"/>
</MiningSchema>
```



#### **\*** ONNX

- MXNet/PyTorch/CNTK
- Protobuf
- Mxnet-model-server
- Simple TensorFlow Serving

```
BARAGoy tonch*ZACO.3:064E
WARRED TABLES AND
African Tabasas Africa VAVCANA
 Bacilla PARE ARE MOUNTED MANER MR
       611ations84A84A 4A4GFML
 ALARABLE ACTUATION
       Broadcast/@/A *A*E**$
PARA ANNA LIKAD
R12ABAR13"AGAG Proc14AB
Electric LishapodASSAE AAAGAAN
 Gstrides#A8848_4446294E/usr/local/tib/python2.7/site-packages/tondt/nn/functional
Aran/Look I/Lith/python2. Z/s ite-pockages/from/minuteles/mobile.py(325): __outl_
Apr/Local/Lib/python2.7/site-pockages/terch/jit/_init__py(2%): traced_inter-
/usr/lecel/lib/pythen2.7/site-packages/tarch/fit/_init__py(259): whasper
usr/local/lib/python2.7/site-packages/tarch/nn/mpdules/module.py(325): ....call.
/usr/loce]/lib/python2.7/site-pockages/torch/jit/_init_.py(217): trace
Assr/Local/Lib/python2.7/site-pockages/torch/arrs/_init__.py(116): _export
Asr/Local/Lib/python2.7/site-packages/torch/arrs/_init__.py(75): export
 /mnist.gy(96): matn
 /mnist.gy(108): ∠modules
```





- TensorFlow SavedModel
  - TensorFlow Serving(gRPC)
  - Simple TensorFlow Serving(HTTP)

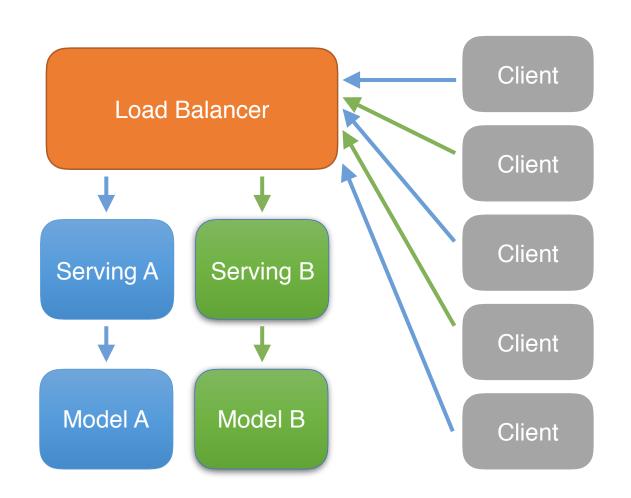
```
model_signature = signature_def_utils.build_signature_def(
    inputs={
        "keys": utils.build_tensor_info(keys_placeholder),
        "features": utils.build_tensor_info(inference_features)
    },
    outputs={
        "keys": utils.build_tensor_info(keys_identity),
        "prediction": utils.build_tensor_info(inference_prediction_op),
        "softmax": utils.build_tensor_info(inference_softmax_op),
    },
    method_name=signature_constants.PREDICT_METHOD_NAME)
```



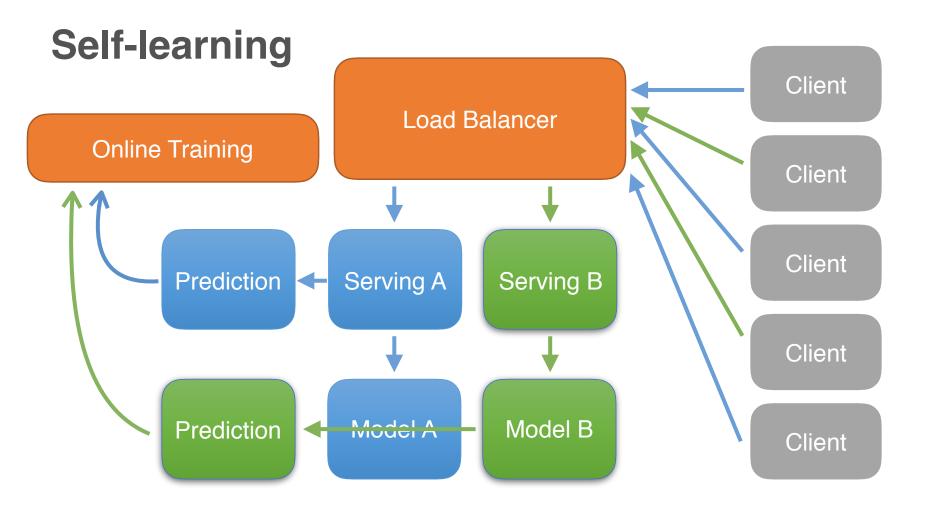




### **AB Testing**









## Agenda

- ❖ 机器学习平台功能模块
- ❖ 机器学习工作流引擎
- ❖ 机器学习模型管理实践
- ❖ 智能机器学习平台与未来



## GAC 全球互联网架构大会

- ◆ 大数据时代: Human -> Application
- ❖ AI时代: Data -> Application
  - Data -> (Data cleaning -> Feature engineering -> Hyper-parameter tuning) -> Model
  - · 基于数据建模 > 人类专家规则
  - · 基于数据调参 > 人类建模规则

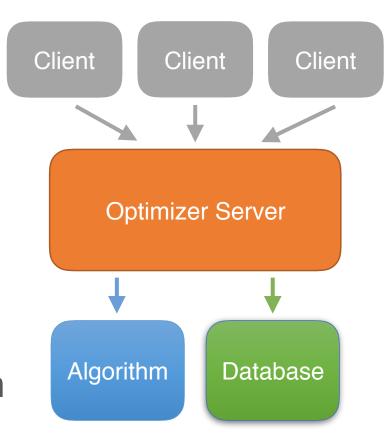


## GAC 全球互联网架构大会

- ❖ 自动特征组合系统(Auto Feature)
- ◆ 超参调优算子(GS LR, RS LR)
- ◆ 智能参数推荐(Hadoop, Spark, K8S)
- ❖ 超参调优服务(Google Vizior, Advisor)
- ❖ 自动神经网络结构生成(RL, BO, ENAS)



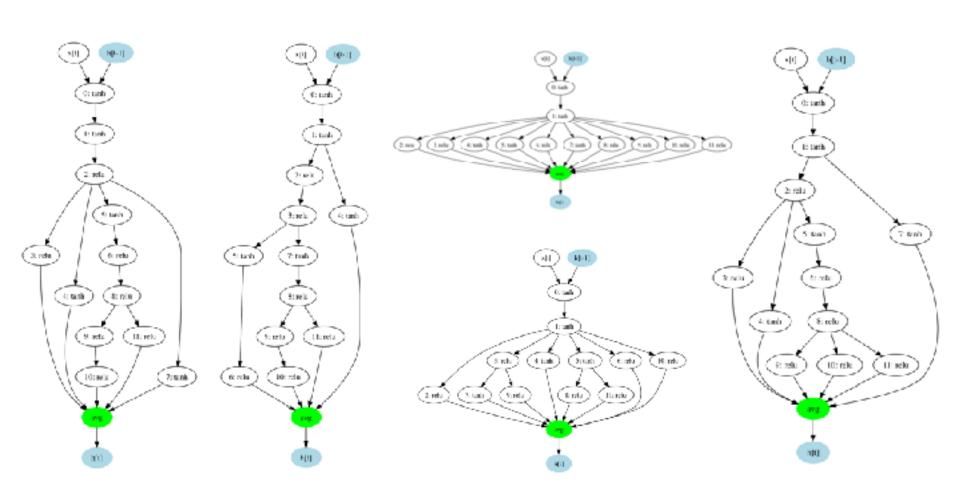
- ❖ 超参调优算法
  - Grid search
  - Random search
  - Bayesian optimization
  - Reinforcement learning
  - Particle swarm optimisation





- ❖ 神经网络调优算法
  - ► RL
  - ENAS(1GPU\*16H)







## 智能机器学习平台

### ENAS Model

- Understandable model description with JSON
- Graph visualization for any generated model
- Easy to extend the framework for other cells
- Model generation in pure Python statements
- Support black-box optimization algorithms



## GAC 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE

## Agenda

- ❖ 机器学习平台功能模块
- ❖ 机器学习工作流引擎
- ❖ 机器学习模型管理实践
- ❖ 智能机器学习平台与未来



## GAC 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE

### **Thanks**





chendihao@4paradigm.com talent@4paradigm.com

# G A C 全球互联网架构大会 GLOBAL INTERNET ARCHITECTURE CONFERENCE





关注公众号获得 更多案例实践