Use Intel Analytics Zoo to build an intelligent QA Bot for Microsoft Azure

3/11/2018

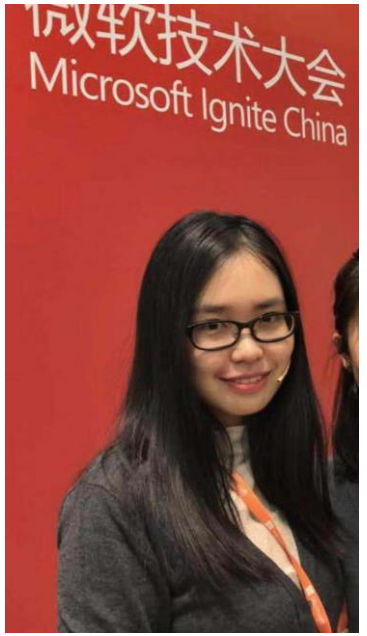
About Us

Kai Huang

Software Engineer from Intel Data Analytics Technology Team

Yuqing Wei

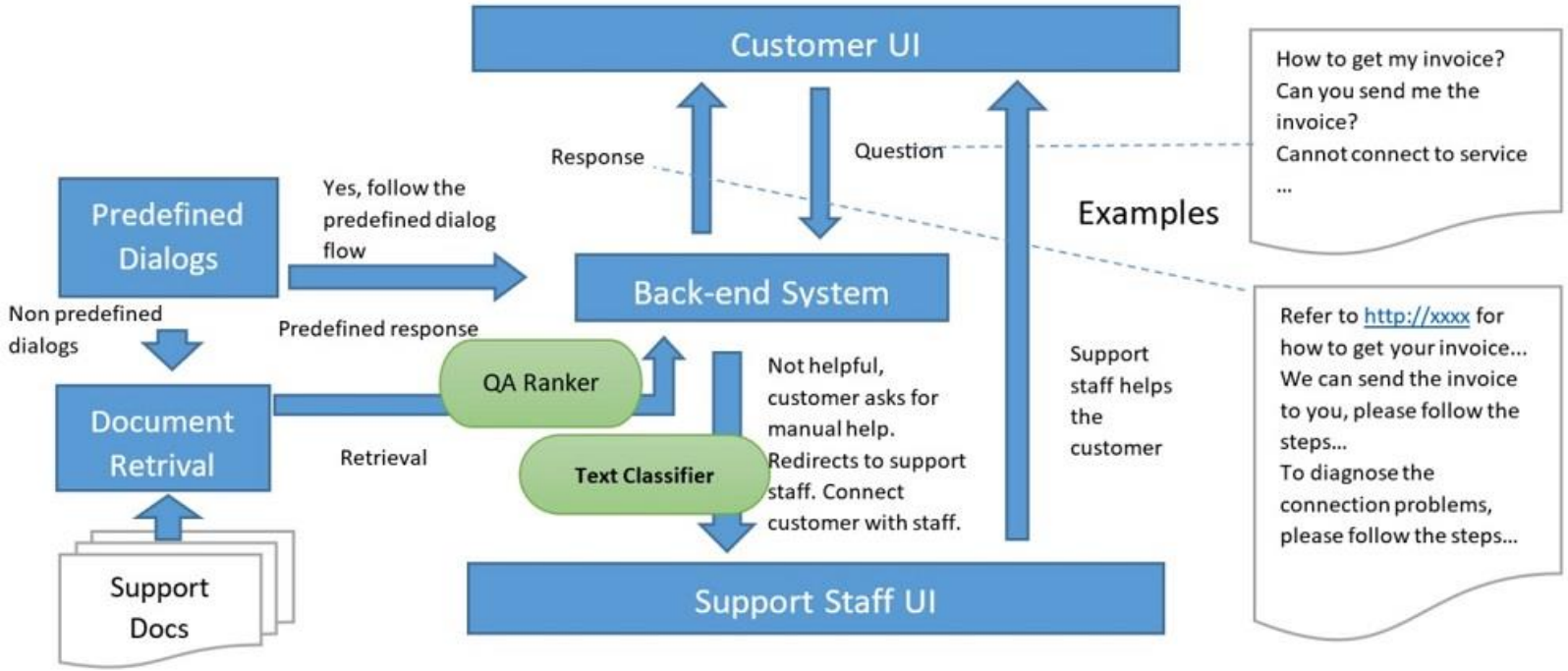
Software Engineer from Microsoft C+AI Team



Why customer support platform needs AI?

- Traditional vs recent intelligent platforms.
- Chat Bot is often one of the core intelligent components.
- To enhance user experience and relieve human workload.
- To provide technical support for Azure users effectively and efficiently.

Overall architecture



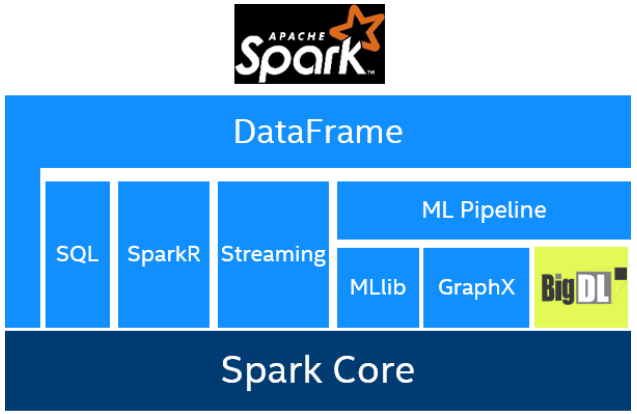
Overview of customer service platform
(basic modules in blue, intelligent modules in green)

Why neural networks?

- Neural networks are easier for feature extraction.
- TextClassifier module can be modified for sentiment analysis.
- Neural networks generally have better performance, especially on QA tasks and when we lack data.
- Common parts can share for different AI modules.

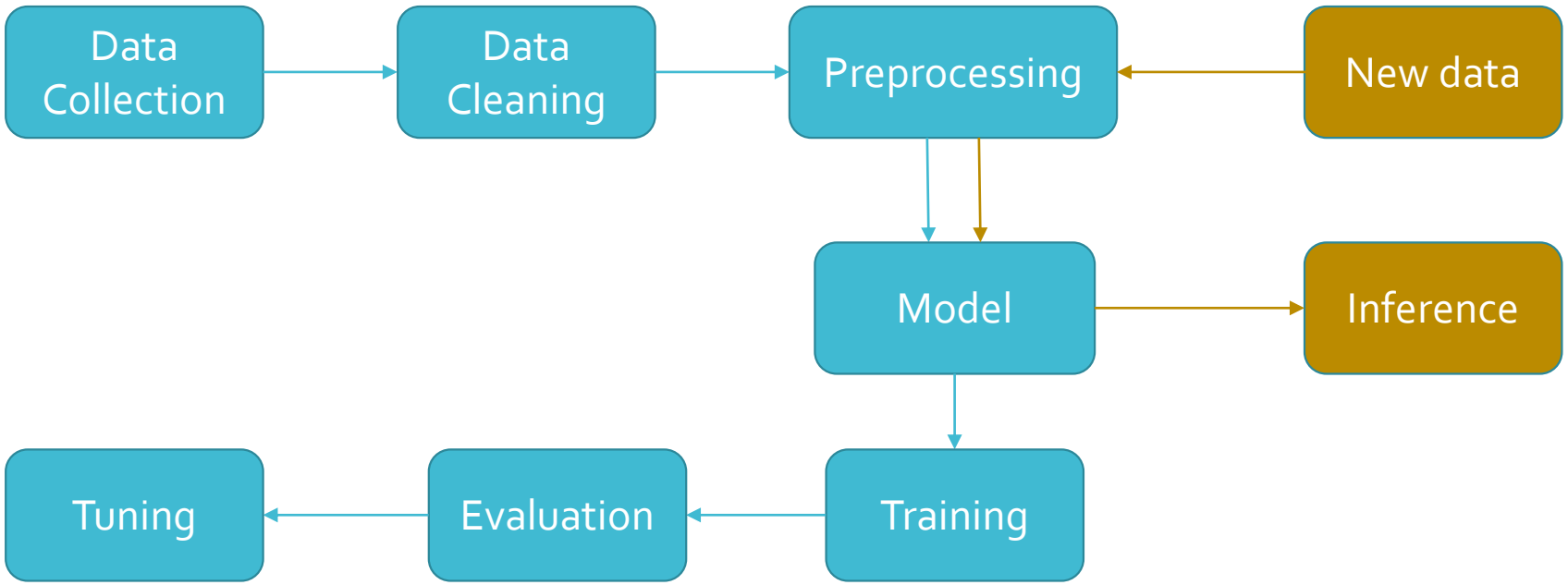
Why Analytics Zoo & BigDL?

- A unified distributed analytics + AI platform on Apache Spark.
- Provides pipeline APIs, prebuilt models and use cases for NLP tasks.
- Provide practical experience for Azure big data users to build AI applications.
- Preinstalled image on Azure Marketplace for easy deployment.



https://github.com/intel-analytics/analytics-zoo
https://analytics-zoo.github.io/

General steps for NLP tasks



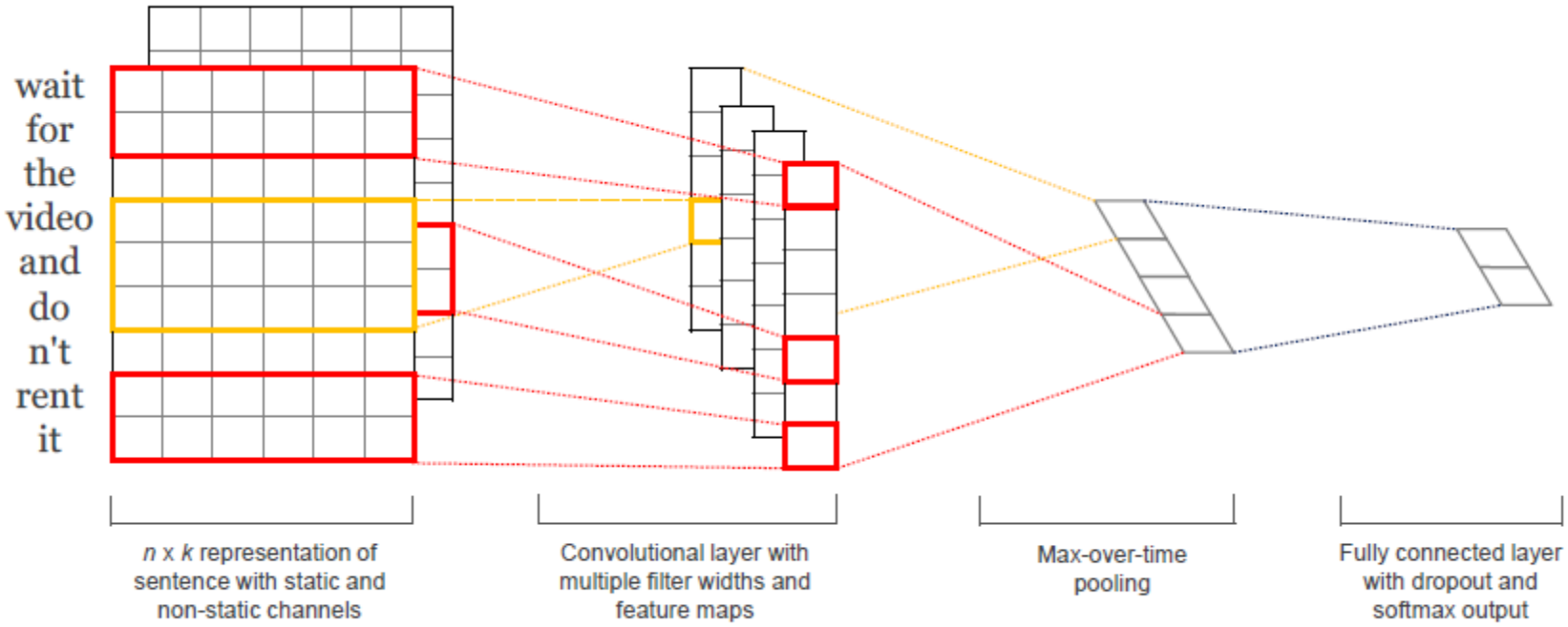
Data Preprocessing

- Read cleaned text data as RDD where each record contains two columns (text, label).
- Common Steps
  - Tokenization: https://github.com/fxsjy/jieba
  - Stopwords removal
  - Sequence aligning
  - Word2Vec: https://github.com/facebookresearch/fastText
  - Conversion to BigDL Sample -> RDD[Sample]

Define
TextC
mode

from zoo.models.textclassification **import** TextClassifier

_length=500,
n=256)



r 'gru'.
ler.

*Photo from: https://blog.csdn.net/littlely_ll/article/details/79151403

Train and evaluate model

- Analytics Zoo provides Keras-Style API for distributed training:

```
text_classifier.compile(optimizer=Adagrad(learning_rate, decay),
                        loss="sparse_categorical_crossentropy",
                        metrics=["accuracy"])

text_classifier.set_checkpoint(path)
text_classifier.set_tensorboard(log_dir, app_name)

text_classifier.fit(train_rdd, batch_size=..., nb_epoch=..., validation_data=val_rdd)

text_classifier.save_model(model_path)

text_classifier.predict(test_rdd)
text_classifier.predict_classes(test_rdd)
```

Ways for improvement

- Check your data first (quality, quantity, etc.).
- Use custom dictionary for tokenization if necessary.
- Train word2vec for unknown words if necessary.
- Hyper parameters tuning (learning rate, etc.).
- Add character embedding, etc.

Service Integration

- Prediction service implemented in Java
- POJO-like API for low-latency local inference

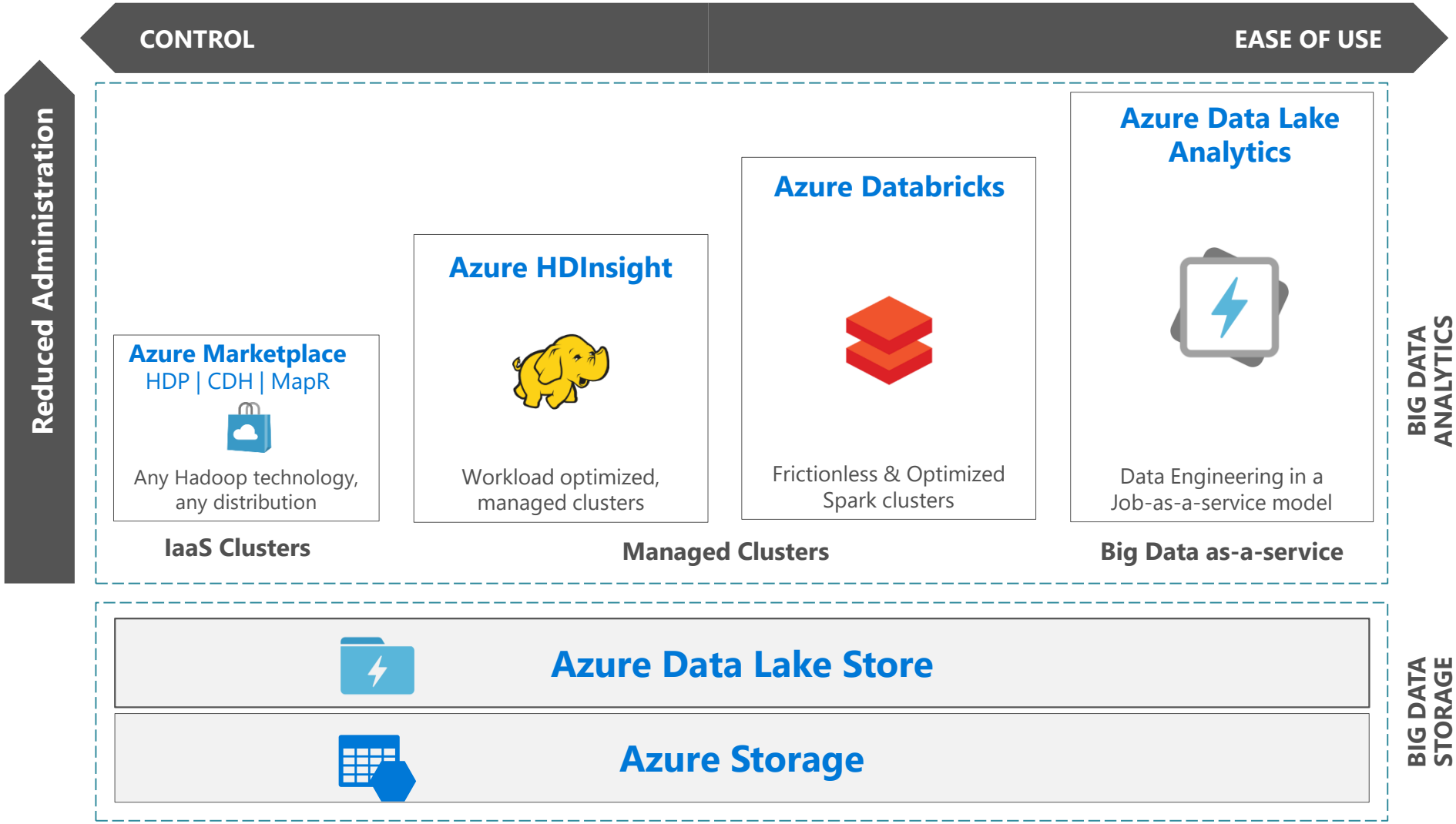
```
public class TextClassificationModel extends AbstractInferenceModel {
  public JTensor preProcess(String text) {
     //Re-implement the preprocessing using Java API
   }
}

TextClassificationModel model = new TextClassificationModel();
model.load(path);
String sampleText = "text content";
JTensor input = model.preProcess(sampleText);
List<JTensor> inputList = new ArrayList<>();
inputList.add(input);
List<List<JTensor>> result = model.predict(inputList);
```

WebService example: https://github.com/intel-analytics/analytics-zoo/tree/master/apps/web-service-sample

A glimpse of QA Ranker module

- Input: a query and a document pair.
- Similar preprocessing steps.
- Output: Relevance score or probability.

KNOWING THE VARIOUS BIG DATA SOLUTIONS



Spark Offerings in Azure

- Spark IaaS (Azure Marketplace)
  https://market.azure.cn/zh-cn
- Spark on Azure Batch using Docker
  https://azure.microsoft.com/en-us/blog/on-demand-spark-clusters-on-docker/
- HDInsight Spark
  https://docs.azure.cn/zh-cn/hdinsight/spark/apache-spark-overview
- Azure Databricks
  https://azure.microsoft.com/zh-cn/services/databricks/

Spark on HDInsight

- Provision cluster with a click of a mouse
- Fully supported by Microsoft and Hortonworks
- Supports Batch, ML, Streaming and SQL workloads

- Read data from **Azure Blob Storage**
- The Spark connector enables real-time analytics over globally distributed data in **Azure Cosmos DB**
- Powerful visualization of data in Spark with **Power BI**
- **VS Code** Integration

Bot Demo

- WeChat: Microsoft云科技
- Webchat: https://support.azure.cn/zh-cn/support/support-azure/

Blogs

- https://www.azure.cn/zh-cn/blog/2018/09/12/Using-Intel-Analytics-Zoo-to-inject-AI-into-customer-service-platform_PartI
- https://software.intel.com/en-us/articles/use-analytics-zoo-to-inject-ai-into-customer-service-platforms-on-microsoft-azure-part-1