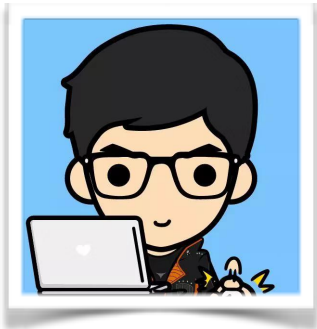




Introduce NLP framework **Kashgari**

Eliyar Eziz





Eliyar Eziz 艾力亚尔

Google Developers Expert @ Machine Learning

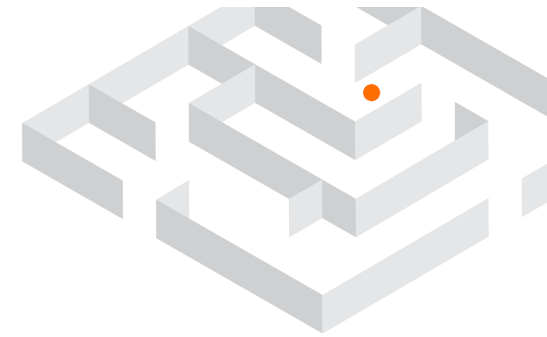
Yodo1 AI Specialist

Open-Source Project

- Kashgari : NLP Transfer learning framework | 1.3k Stars
- BMPlayer : iOS Swift Video Player | 1.4k Stars
- Blog <https://eliyar.biz>
- Github <https://github.com/BrikerMan>
- Email eliyar917@gmail.com



Topics



NLP

- What is NLP
- Two main tasks in NLP

Kashgari

- What problem kashgari solved

Open source experiences

- How I benefited from my Open-source project
- What need to prepare
- Which tools I use

帮我订一张到上海的
机票



```
{  
  "domain": "booking_flight",  
  "confidence": 0.9721,  
  "entities": [  
    {  
      "entity": "from_location",  
      "value": "北京",  
      "confidence": 0.98  
    },  
    {  
      "entity": "to_location",  
      "value": "上海",  
      "confidence": 0.94  
    }  
  ]  
}
```

Two Main Task in NLP

I want to watch a movie of Jackie Chan

- Text Classification → Watching movie
- Text Labeling → Actor: Jackie Chan

Adopt BERT takes Days



Adopt BERT takes Days

- Repeated Work
- Too Complex
- Not Scalable
- Not production ready





Kashgari

simple and powerful NLP Transfer learning framework

Use Cases



中海油

NLP 隐患分析 领域术语的识别与抽取
多标签文本分类



北京邮电大学



中科院信工所

细粒度情感分析



钛氩新媒体

智慧新闻领域的应用



思通聚宝

商业领域的情报能力



暴风 TV

语音助手语义理解



So.Car

UGC文本分类



妙健康

医学命名实体识别发现

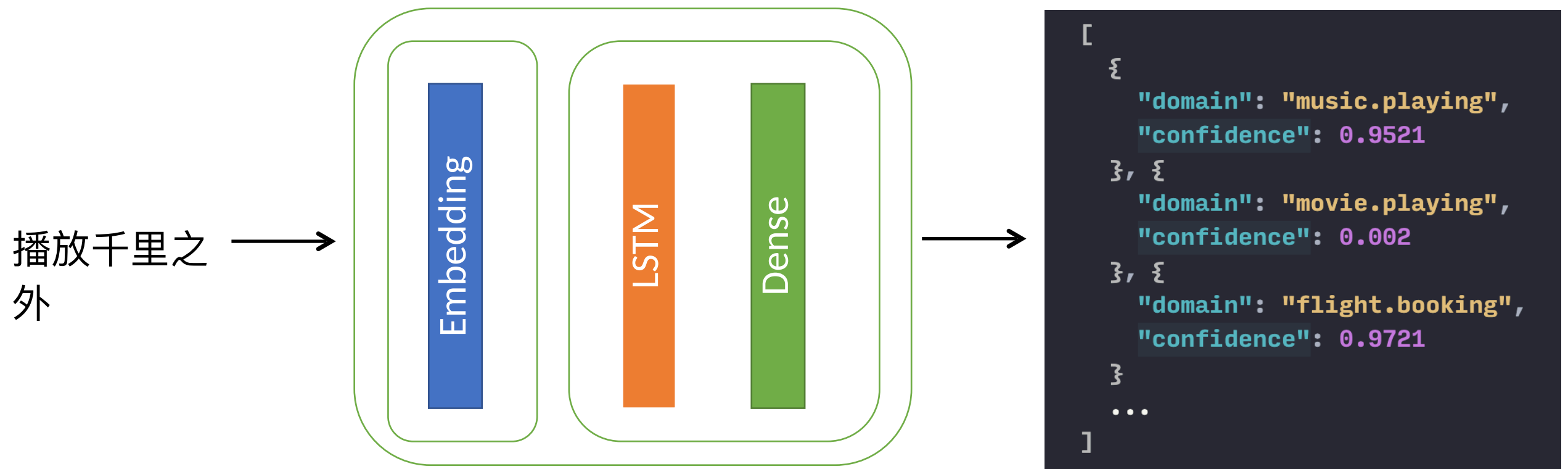


Ochoose

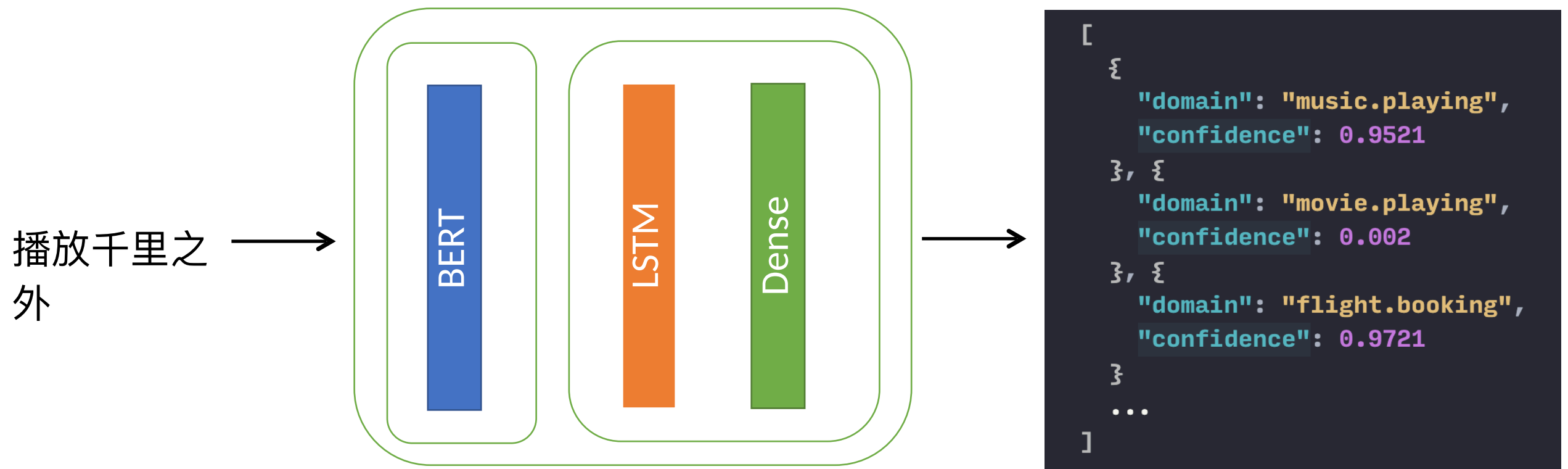
NLP 数字营销

And More .

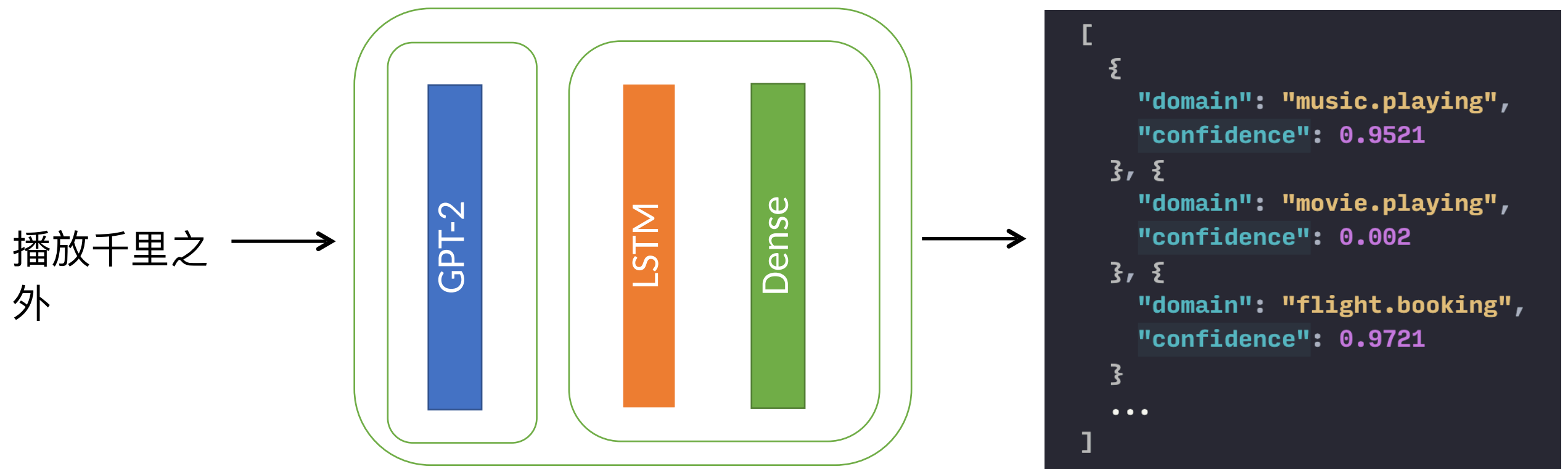
Kashgari Model Structure



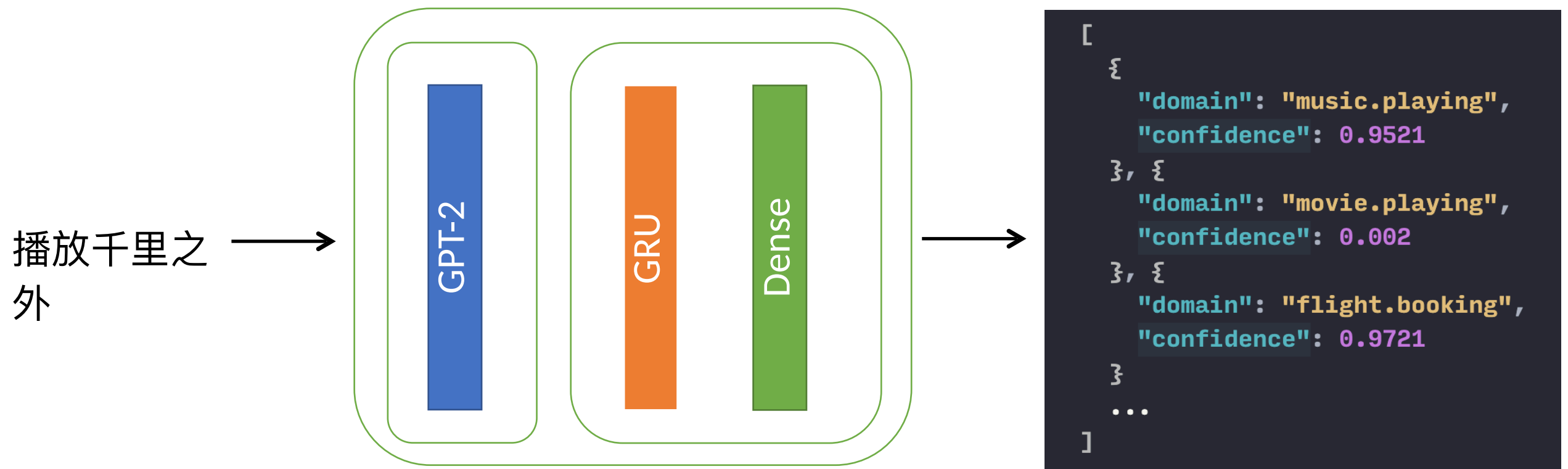
Kashgari Model Structure



Kashgari Model Structure



Kashgari Model Structure



Adopt BERT in 5 mins

- Human-friendly
- Built-in transfer learning
- Fully scalable
- Production Ready



```
import kashgari
from kashgari.embeddings import BERTEmbedding, GPT2Embedding
from kashgari.tasks.labeling import BiGRU_Model, BiLSTMModel

embedding = BERTEmbedding('<path-to-gpt-model-folder>',
                           task=kashgari.LABELING,
                           sequence_length=30)

model = BiLSTM_Model(embedding)
model.fit(train_x, train_y, valid_x, valid_y, epochs=50)
model.evaluate(test_x, test_y)
model.predict(test_x[:10])
```

```
import kashgari
from kashgari.embeddings import BERTEmbedding, GPT2Embedding
from kashgari.tasks.labeling import BiGRU_Model, BiLSTMModel

embedding = GPT2Embedding('<path-to-gpt-model-folder>',
                           task=kashgari.LABELING,
                           sequence_length=30)

model = BiLSTM_Model(embedding)
model.fit(train_x, train_y, valid_x, valid_y, epochs=50)
model.evaluate(test_x, test_y)
model.predict(test_x[:10])
```



```
import kashgari
from kashgari.embeddings import BERTEmbedding, GPT2Embedding
from kashgari.tasks.labeling import BiGRU_Model, BiLSTMModel

embedding = GPT2Embedding('<path-to-gpt-model-folder>',
                           task=kashgari.LABELING,
                           sequence_length=30)

model = BiGRU_Model(embedding)
model.fit(train_x, train_y, valid_x, valid_y, epochs=50)
model.evaluate(test_x, test_y)
model.predict(test_x[:10])
```

Kashgari

<https://github.com/BrikerMan/Kashgari>

Kashgari

IntroductionTutorialsEmbeddingsAdvancedAPIFAQAbout

Kashgari

Introduction

中文文档

Kashgari

license Apache-2.0 build passing coverage 94% downloads 25k pypi v1.0.0

🎉🎉🎉 We are proud to announce that we entirely rewritten Kashgari with tf.keras, now Kashgari comes with easier to understand API and is faster! 🎉🎉🎉

Overview

Kashgari is a simple and powerful NLP Transfer learning framework, build a state-of-art model in 5 minutes for named entity recognition (NER), part-of-speech tagging (PoS), and text classification tasks.

- **Human-friendly.** Kashgari's code is straightforward, well documented and tested, which makes it very easy to understand and modify.
- **Powerful and simple.** Kashgari allows you to apply state-of-the-art natural language processing (NLP) models to your text, such as named entity recognition (NER), part-of-speech tagging (PoS) and classification.
- **Built-in transfer learning.** Kashgari built-in pre-trained BERT and Word2vec embedding models, which makes it very simple to transfer learning to train your model.
- **Fully scalable.** Kashgari provide a simple, fast, and scalable environment for fast experimentation, train your models and experiment with new approaches using different embeddings and model structure.
- **Production Ready.** Kashgari could export model with `SavedModel` format for tensorflow

Table of contents

- Overview
- Our Goal
- Performance
- Tutorials
 - Quick start
 - Requirements and Installation
 - Example Usage
 - Run with GPT-2 Embedding
 - Run with Bert Embedding
- Contributing
- Reference

<https://pypi.org/project/kashgari/>



Open-Source

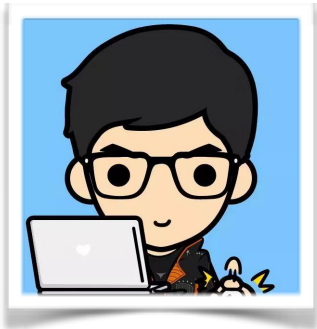
Why and How

Benefits

- Improve Coding Skills
- Social Network
- Google Developers Expert Title

Preparation

- Add test-cases / Linter
 - Add document
 - Add Issue Templates
- Travis CI / Github Action
 - mkdocs
 - Github Pages



Eliyar Eziz 艾力亚尔

Google Developers Expert @ Machine Learning

Yodo1 AI Specialist

Open-Source Project

- Kashgari : NLP Transfer learning framework | 1.3k Stars
- BMPlayer : iOS Swift Video Player | 1.4k Stars
- Blog <https://eliyar.biz>
- Github <https://github.com/BrikerMan>
- Email eliyar917@gmail.com