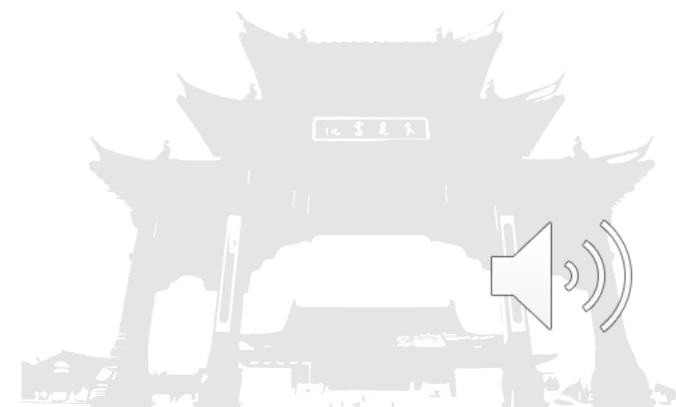


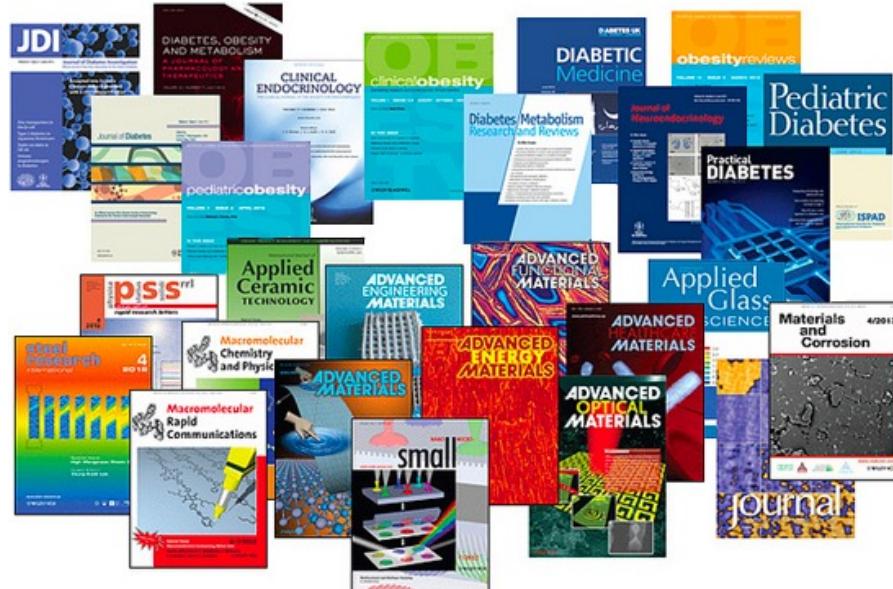


DeepKE: A Deep Learning Based Knowledge Extraction Toolkit for Knowledge Base Population

Ningyu Zhang, Xin Xu, Liankuan Tao, Haiyang Yu, Hongbin Ye, Shuofei Qiao, Xin Xie, Xiang Chen, Zhoubo Li, Lei Li, Xiaozhuan Liang, Yunzhi Yao, Shumin Deng, Peng Wang, Wen Zhang, Zhenru Zhang, Chuanqi Tan, Qiang Chen, Feiyu Xiong, Fei Huang, Guozhou Zheng, Huajun Chen



Knowledge Base Population



Text2Knowledge

Language Pretrain
BERT、GPT.....

属性 : 字

张之洞，晚清名臣、清代洋务派代表人物，祖籍直隶南皮，出生于贵州，字孝达。

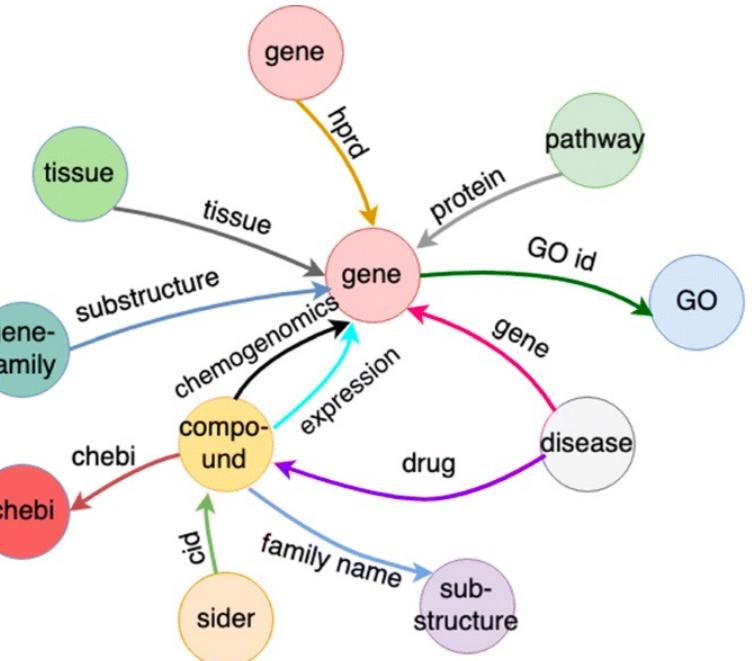
实体

属性值

isSpouseOf

Jack is married to the microbiologist known as Dr. Germ in the USA.

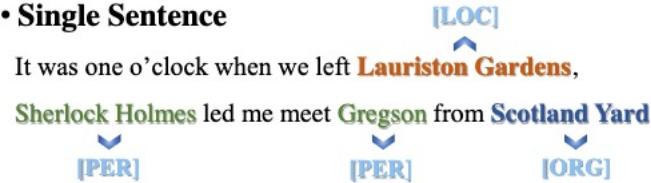
Entity Pair



DeepKE: Diverse Scenarios

- Single Sentence

It was one o'clock when we left **Lauriston Gardens**,
Sherlock Holmes led me meet **Gregson** from **Scotland Yard**.



The diagram illustrates the entity types for the entities mentioned. Lauriston Gardens is categorized as [LOC]. Sherlock Holmes and Gregson are categorized as [PER]. Scotland Yard is categorized as [ORG]. Arrows point from each entity to its respective category label.

- Document

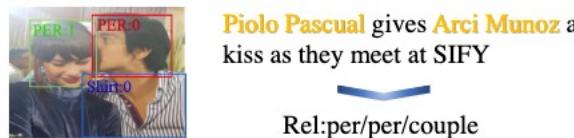
... Elias Brown (May 9, 1793 - July 7, 1857) was a **U.S.** Representative from **Maryland**. Born near **Baltimore**, **Maryland**, Brown attended the common schools ... He died near **Baltimore**, **Maryland**, and is interred in a private cemetery near **Eldersburg**, **Maryland** ...



Intra-sentence (**Maryland**, country, **U.S.**)
(**Baltimore**, located in, **Maryland**)
(**Eldersburg**, located in, **Maryland**)

Inter-sentence (**Baltimore**, country, **U.S.**)
(**Eldersburg**, country, **U.S.**)

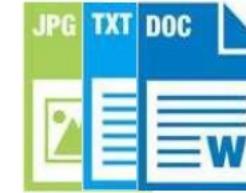
- MultiModal



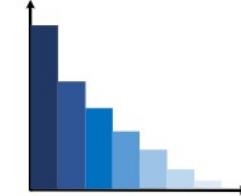
DeepKE

Knowledge Extraction Tool

Diverse Data



Low Resource



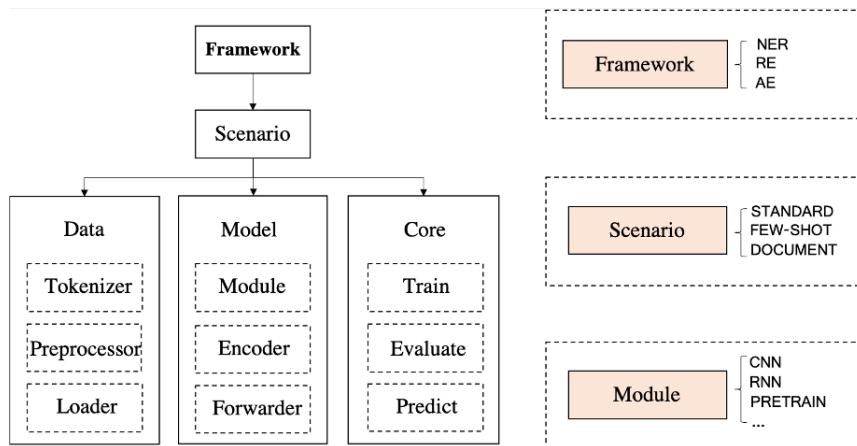
Off-the-shelf Usage Flexible Training



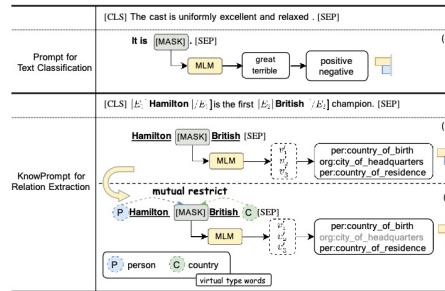
cnSchema
OpenKG 开放的中文知识图谱



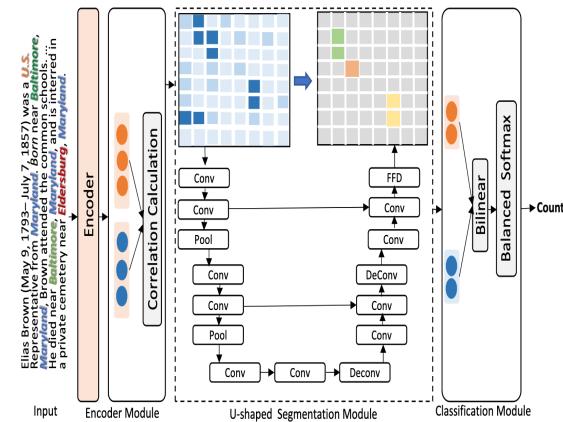
DeepKE: with Traditional and Cutting-edge Methods



```
# Import packages
from torch import optim
import deepke.relation_extraction.standard.models as models
from deepke.relation_extraction.standard.tools import preprocess, train, FocalLoss
# Preprocess data
preprocess(config)
# Choose the model, users can choose from CNN, RNN, Transformer, GCN, Capsule and BERT
__Model__ = {
    'cnn': models.PCNN,
    'rnn': models.BILSTM,
    'transformer': models.Transformer,
    'gcn': models.GCN,
    'capsule': models.Capsule,
    'lm': models.LM,
}
model = __Model__[config.model_name](config)
# Define the optimizer, scheduler and criterion
optimizer = optim.Adam(model.parameters(), lr=config.learning_rate, weight_decay=config.weight_decay)
scheduler = optim.lr_scheduler.ReduceLROnPlateau(optimizer, factor=config.lr_factor, patience=config.lr_patience)
criterion = FocalLoss()
# Train
train(epoch, model, train_dataloader, optimizer, criterion, device, writer, config)
# Save
model.save(epoch, config)
# Predict
model.load(config.saved_path, device=device)
model(preprocess(config.text))
```



KnowPrompt WWW'22

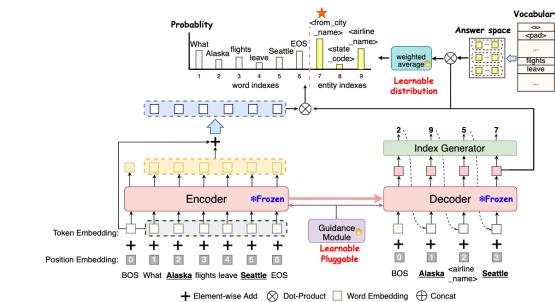


DocuNet IJCAI'21

KnowPrompt: Knowledge-aware Prompt-tuning with Synergistic Optimization for Relation Extraction WWW'22
LightNER: A Lightweight Generative Framework with Prompt-guided Attention for Low-resource NER Arxiv'21
Document-level Relation Extraction as Semantic Segmentation IJCAI'21

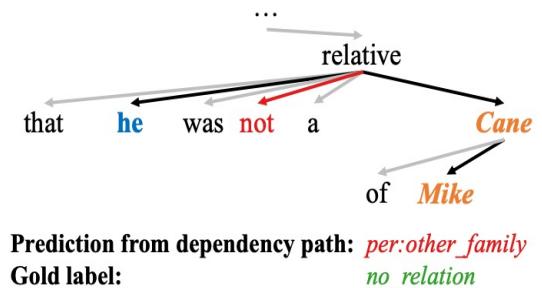
求是创新

Graph Convolution over Pruned Dependency Trees Improves Relation Extraction EMNLP'18



LightNER Arxiv' 21

I had an e-mail exchange with Benjamin Cane of Popular Mechanics which showed that **he** was not a relative of **Mike Cane**.



C-GCN EMNLP'18

Step1 : go to the task folder

```
cd example/ner/standard
```

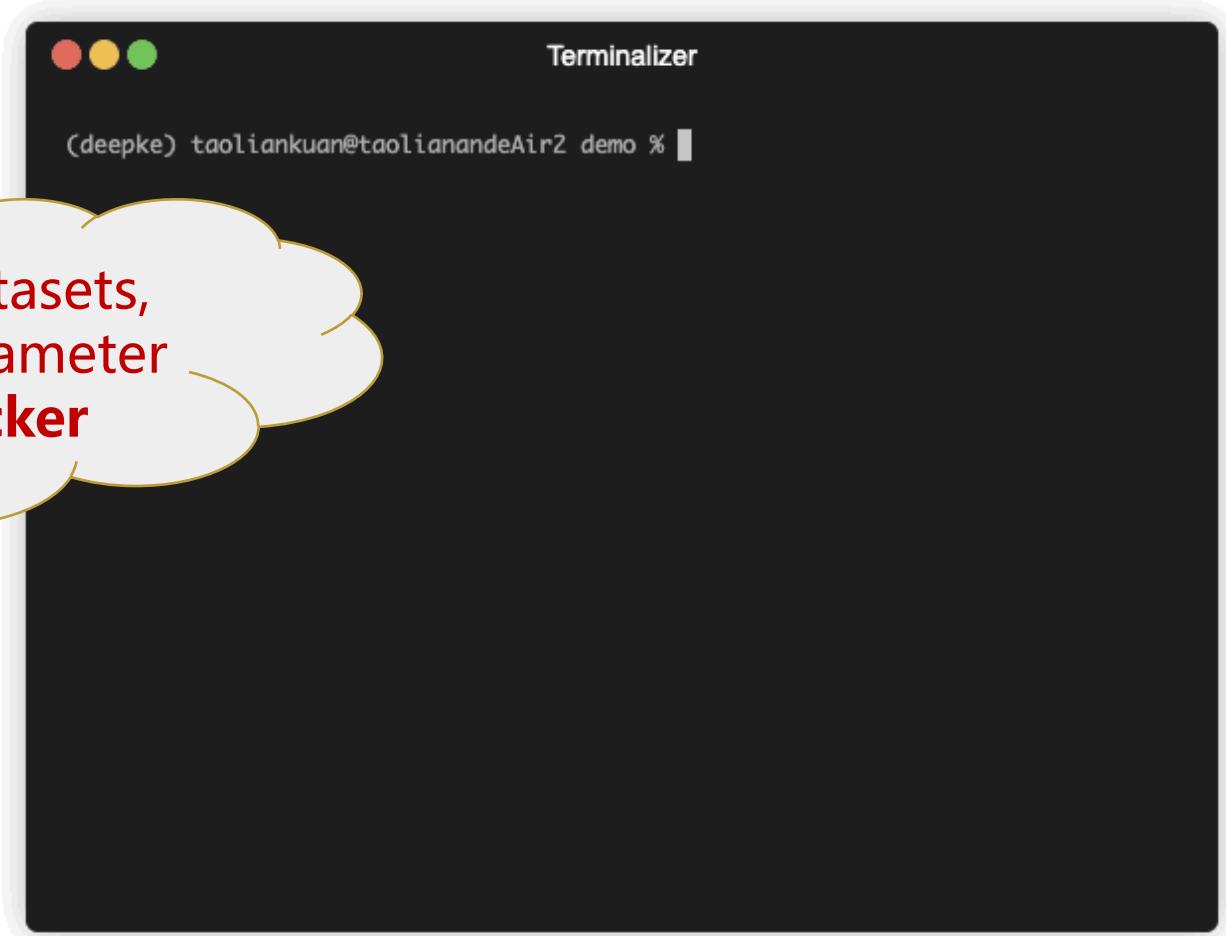
Step2: Train

```
python run.py
```

Step3 : Evaluation

```
python predict.py
```

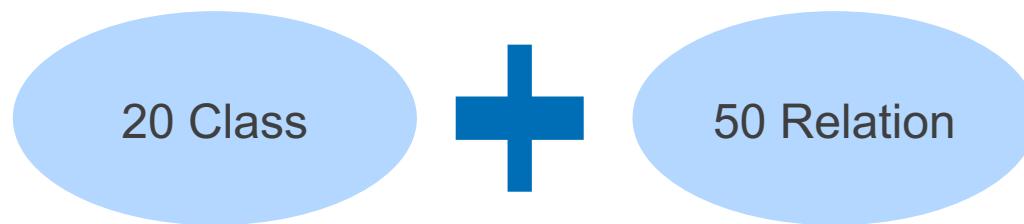
**Customize datasets,
auto-hyperparameter
tunning,docker**



DeepKE: cnSchema-based Off-the-shelf Models



DeepKE provides off-the-shelf extraction models based on cnSchema: supporting **28 entity types** and **50 relation classes**.



Entity Recognition

DeepKE(NER), Roberta-wwm-ext, Chinese

[Google dowload](#)

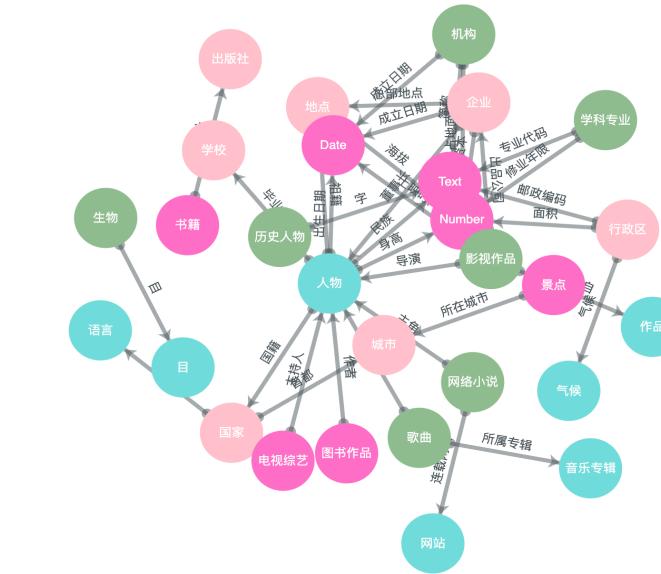
[Baidu Netdisk download](#)(password:u022)

Relation Extraction

DeepKE(RE), Roberta-wwm-ext, Chinese

[Google dowload](#)

[Baidu Netdisk download](#)(password:2goe)



```
1 {  
2   "text": "查尔斯·阿兰基斯 (Charles Aránguiz)，1989年4月17日出生于智利圣地亚哥，智利职业足球运动员",  
3   "spo_list": [  
4     {"predicate": "出生地",  
5      "object_type": "地点",  
6      "subject_type": "人物",  
7      "object": "圣地亚哥",  
8      "subject": "查尔斯·阿兰基斯"  
9    },  
10   {"predicate": "出生日期",  
11     "object_type": "Date",  
12     "subject_type": "人物",  
13     "object": "1989年4月17日",  
14     "subject": "查尔斯·阿兰基斯"  
15   }]  
16 }
```



DeepKE: Documents&Tutorials



DeepKE

Search docs

GETTING STARTED

Start

Install

Example

FAQ

PACKAGE

DeepKE

Attribution Extraction

Name Entity Recognition

Few Shot

Models

Module

Utils

Standard

Relation Extraction

Models

deepke.name_entity_recognition.few_shot.models.model module

```
class deepke.name_entity_recognition.few_shot.models.model.PromptBartEncoder(encoder) [source]
```

Bases: `torch.nn.modules.module.Module`

```
forward(src_tokens, attention_mask=None, past_key_values=None) [source]
```

Defines the computation performed at every call.

Should be overridden by all subclasses.



Although the recipe for forward pass needs to be defined within this function, one should call the `Module` instance afterwards instead of this since the former takes care of running the registered hooks while the latter silently ignores them.

```
class deepke.name_entity_recognition.few_shot.models.model.PromptBartDecoder(decoder, pad_token_id, label_ids, use_prompt=False, prompt_len=10, learn_weights=False) [source]
```

Bases: `torch.nn.modules.module.Module`

```
forward(tgt_tokens, prompt_state) [source]
```

Defines the computation performed at every call.

Should be overridden by all subclasses.



Although the recipe for forward pass needs to be defined within this function, one should call the `Module` instance afterwards instead of this since the former takes care of running the registered hooks while the latter silently ignores them.

```
decode(tokens, state) [source]
```

Comprehensive documents

求是创新

```
File Edit View Insert Runtime Tools Help Changes will not be saved
+ Code + Text Copy to Drive
Import package
[ ] import torch
from tqdm import tqdm
import numpy as np
from itertools import chain
from torch.utils.data import Dataset, DataLoader
from torch.nn.utils.rnn import pad_sequence
from transformers import BartTokenizer
import importlib
from deepke.name_entity_re.few_shot.models.model import PromptBartModel, PromptGeneratorModel
from deepke.name_entity_re.few_shot.module.datasets import ConllNERProcessor, ConllNERDataset
from deepke.name_entity_re.few_shot.module.train import Trainer
from deepke.name_entity_re.few_shot.module.metrics import Seq2SeqSpanMetric
from deepke.name_entity_re.few_shot.utils.util import *
from deepke.name_entity_re.few_shot.module.mapping_type import mit_movie_mapping, mit_restaurant_mapping, atis_mapping

wandb: You can find your API key in your browser here: https://wandb.ai/authorize
wandb: Paste an API key from your profile and hit enter, or press ctrl+c to quit: .....
wandb: [REDACTED] No API key specified.
wandb: You can find your API key in your browser here: https://wandb.ai/authorize
wandb: Paste an API key from your profile and hit enter, or press ctrl+c to quit: .....
wandb: Appending key for api.wandb.ai to your netrc file: /root/.netrc
wandb: wandb version 0.12.9 is available! To upgrade, please run:
wandb: $ pip install wandb --upgrade
Syncing run royal-bee-1 to Weights & Biases (docs).

!wget 120.27.214.45/Data/ner/few_shot/data.tar.gz
!tar -xzvf data.tar.gz

./data/
./data/atis/
./data/atis/test.txt
./data/atis/10-shot-train.txt
./data/atis/20-shot-train.txt
./data/atis/50-shot-train.txt
```

Google Colab Tutorials



DeepKE: Online System



DeepKE is a knowledge extraction toolkit supporting **low-resource, document-level** and **multimodal** scenarios.

PyTorch  Transformers



DeepKE
Knowledge Extraction Tool

DeepKE Public

An Open Toolkit for Knowledge Graph Extraction and Construction

Python ⭐ 1.2k 📈 355

<https://github.com/zjunlp/DeepKE>

OpenKG.CN 中文开放知识图谱

Homepage Demo Download Contact Us Doc OpenKG English 中文

Sentence

- Standard Named Entity Recognition
- Standard Attribution Extraction
- Standard Relation Extraction

Document

- Document Relation Extraction
 - Example
 - Input

MultiModal

- MultiModal Relation Extraction
 - Example
 - Input

Canton Memorial Civic Center

The {{0}|LOC|Canton Memorial Civic Center} is a multi-purpose arena located in {{1}|LOC|Canton}, {{2}|LOC|Ohio}, {{3}|LOC|United States}, and is currently the home arena for the {{4}|ORG|Canton Charge} of the {{5}|ORG|NBA G League}. Built in {{6}|TIME|1951}, previous sports teams that have played at the center include the {{7}|ORG|Canton Legends} indoor football team, {{8}|ORG|Canton Invaders} indoor soccer team, and {{9}|ORG|Ohio Aviators}; of the {{10}|ORG|American Basketball Association}. The building is owned by the {{11}|LOC|City of Canton} and operated by {{11}|ORG|SMG}. Capacity is {{12}|NUM|5,200} in the arena, and up to {{13}|NUM|600} in the {{14}|LOC|McKinley Room}. The facility has over years hosted concerts, professional wrestling cards, political rallies, family shows, and features a number of annual {{15}|MISC|Pro Football Hall of Fame} festival events.

Paris Belongs to Us

The {{0}|MISC|Paris Belongs to Us} (sometimes translated as {{1}|MISC|Paris Is Ours}) is a {{2}|TIME|1951} {{3}|LOC|French} mystery film directed by {{4}|PER|Jacques Rivette}. Set in {{5}|LOC|Paris} in {{6}|TIME|1957} and often referencing {{7}|PER|Shakespeare}'s play {{8}|MISC|Pericles}, the title is highly ironic because the characters are immigrants or alienated and do not feel that they belong at all. The story centres on an essentially innocent young university student called {{9}|PER|Anne} who through her older brother meets a group of friends haunted by mysterious tensions and fears that lead {{10}|NUM|two} of them to commit suicide. Among them is her opposite, a femme fatale called {{11}|PER|Terry} who has had affairs with all the men. The source of the malaise affecting the group is never explained, leaving viewers to ponder how far it might be an amalgam of individual imbalances, general existentialist anxiety, or the more specific paranoia of the {{12}|MISC|Cold War} as the world faced the possibility of nuclear annihilation.

Click

The Great Appeal

The {{0}|MISC|The Great Appeal} ({{1}|MISC|Italian}: {{0}|MISC|Il Grande appello)) is a {{2}|TIME|1936} {{3}|LOC|Italian} war film directed by {{4}|PER|Mario Camerini} and starring {{5}|PER|Camillo Pilotto}, {{6}|PER|Roberto Villa} and {{7}|PER|Lina d'Acosta}. It is sometimes known by its alternative title {{0}|MISC|The Last Roll-Call}. {{4}|PER|Camerini} was considered to have no sympathies with the {{8}|ORG|Fascist} regime of {{9}|LOC|Italy}, but he made this propaganda film that endorsed the colonial policies of the {{3}|LOC|Italian} government. It was one of a number of {{10}|LOC|African}-set films made during the {{8}|ORG|Fascist} era including {{11}|MISC|The White Squadron} ({{2}|TIME|1936}), {{12}|MISC|Sentinels of Bronze} ({{13}|TIME|1937}) and {{14}|MISC|Luciano Serra, Pilot} ({{15}|TIME|1938}). The film portrays the rediscovery of his patriotism of an {{3}|LOC|Italian}, who eventually dies for his country.

Click

Visitors' number: 15650
Copyright© 2019-2022 OpenKG & ZJUKG All Rights Reserved

http://deepke.zjukg.cn/CN/ner_st.html





Thank You!

A C C E P T M Y E N D L E S S G R A T I T U D E

