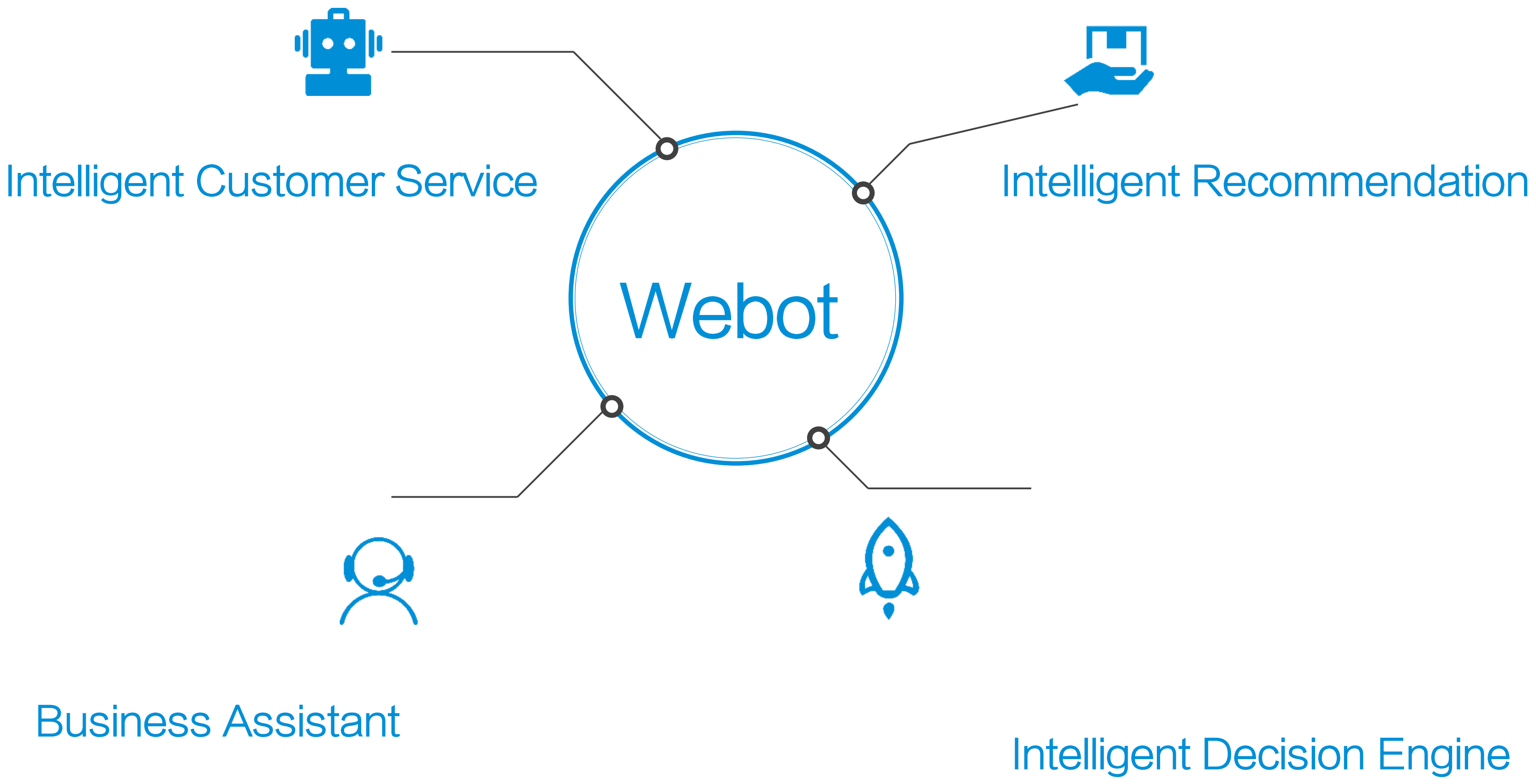
智言科技

Leo ZHOU@Webot.ai

Boost your business with conversational user experience!

You are competing with the best experience a consumer has ever had.

About Webot


Intelligent Customer Service
Intelligent Recommendation
Webot
Business Assistant
Intelligent Decision Engine

Research Team

Headquartered in Shenzhen
50% with Master degree
20% with PhD degree

Graduated from reputable universities in US, HK, UK, China

Industries

Application industry scenario

Finance

×

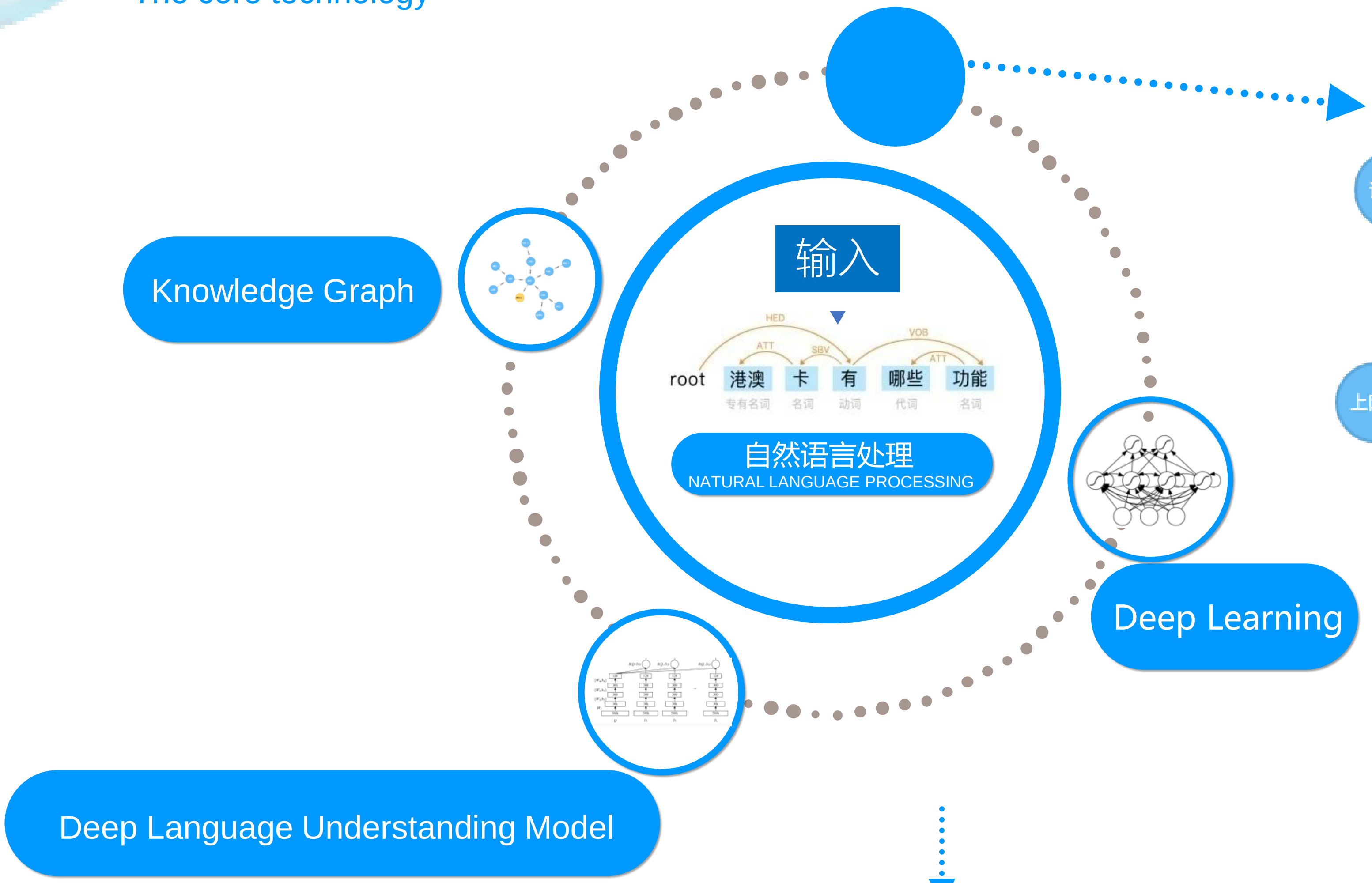
Insurance

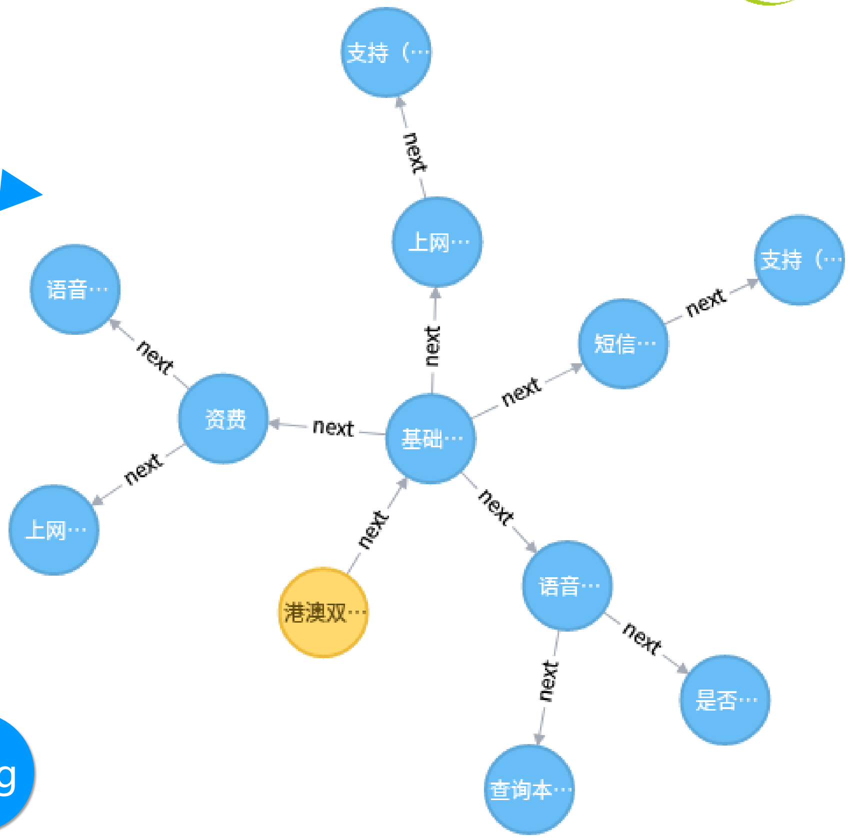
×

Tourism

Core Technology

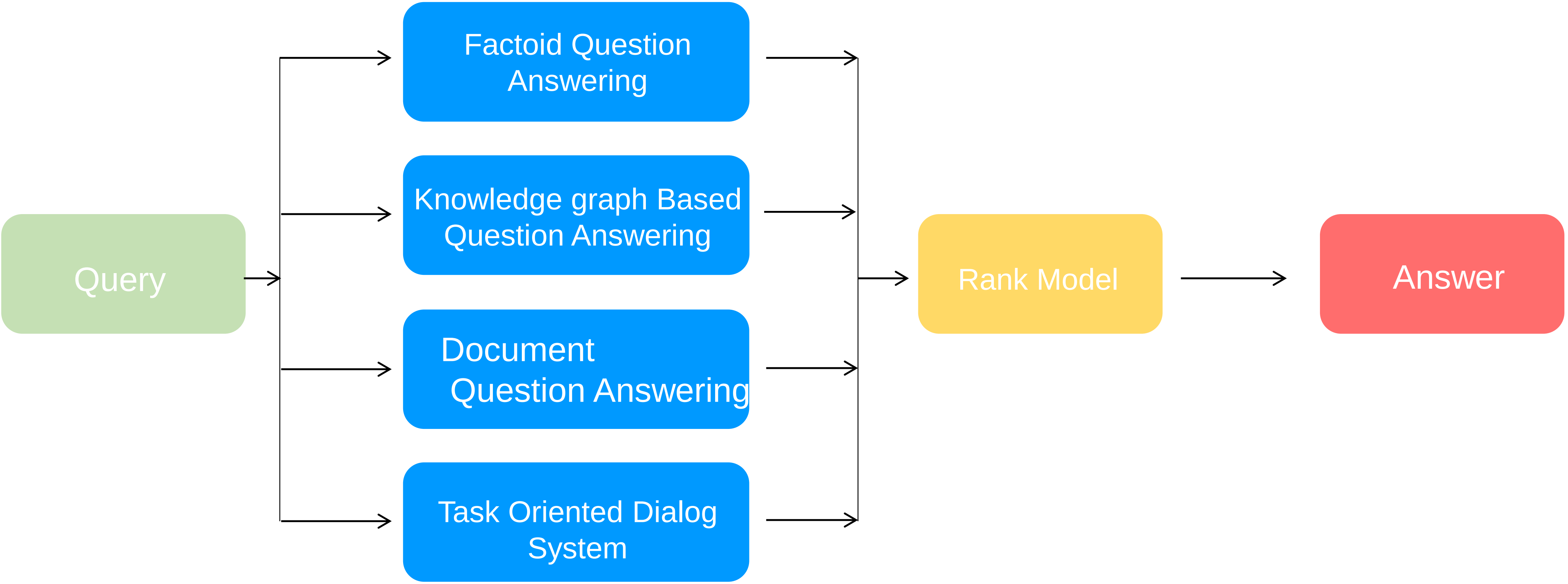
The core technology



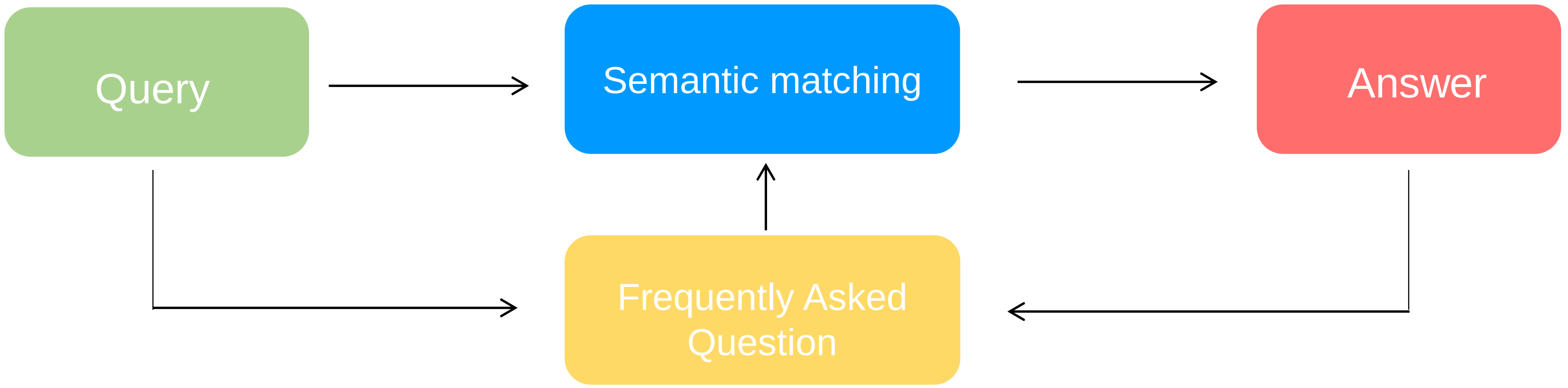


您好，港澳卡能够上网、打电话和发短信。

System Architecture



Factoid Question Answering



Techniques:

Cold start

1. **String fuzzy match**
2. **Key word extraction(text rank)**
3. **Topic model(Familia[1])**
4. **Word Mover's Distance[2] with Chinese character word embedding**
5. **Unsupervised sentence embedding. E.g FastText[3]**

With labeled data:

1. **Siamese network with RNN sentence encoder**
2. **GBDT**

[1] https://github.com/baidu/Familia

[2] Matt Kusner, .etl From word embeddings to document distances. Proceedings of the 32nd International Conference on Machine Learning, volume 37 of Proceedings of Machine Learning Research, pages 957–966

[3] Armand Joulin, etl. Bag of Tricks for Efficient Text Classification

Factoid Question Answering——Cold Start

Techniques:

Topic model

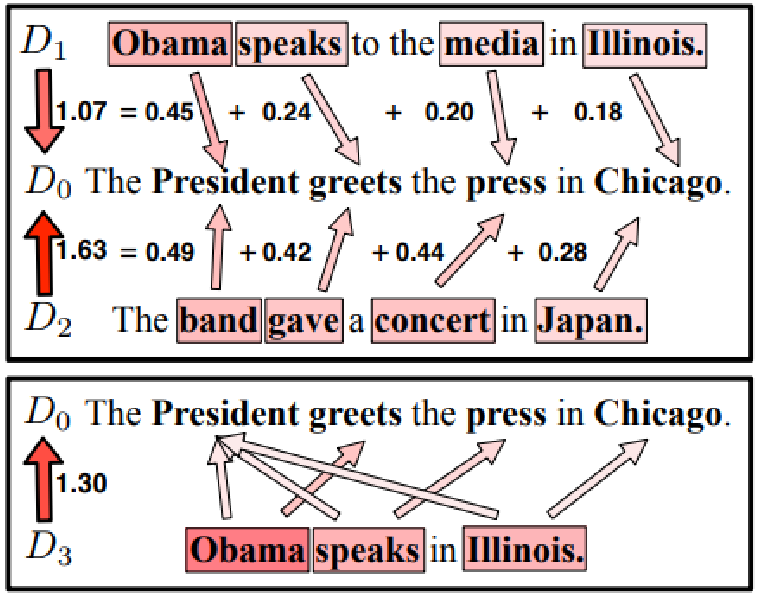
百度又一次展示了自动驾驶领域领导者的大气风范，发布了一项名为“Apollo(阿波罗)”的新计划，向汽车行业及自动驾驶领域的合作伙伴提供一个开放、完整、安全的软件平台，帮助他们结合车辆和硬件系统，快速搭建一套属于自己的完整的自动驾驶系统。

$$Similarity(q,c) = \prod_{w \in q} \sum_{k} P(w|z_k)P(z_k|c),$$

Inferred topic

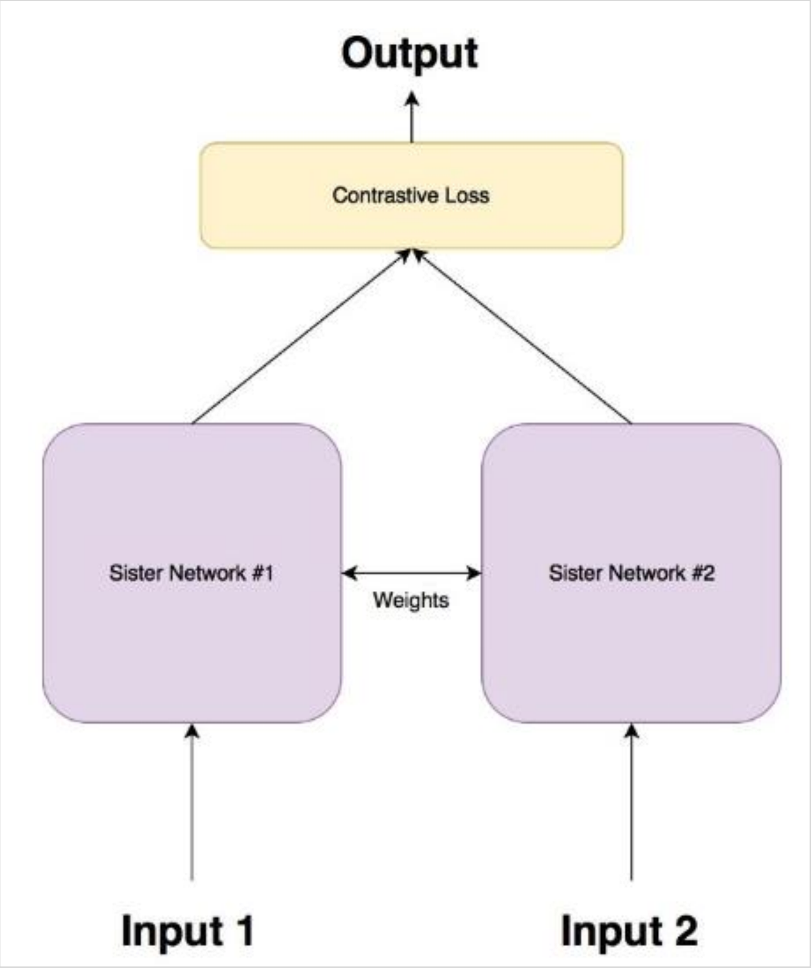
百度宣布阿波罗计划 开放自动驾驶技术有望改变汽车产业

Word Mover's Distance



[1] https://github.com/baidu/Familia
[2] Matt Kusner, .etl From word embeddings to document distances. Proceedings of the 32nd International Conference on Machine Learning, volume 37 of Proceedings of Machine Learning Research, pages 957–966
[3] Armand Joulin, etl. Bag of Tricks for Efficient Text Classification

Factoid Question Answering———with labeled data



For Siamese network with labeled data.

Contrastive loss:

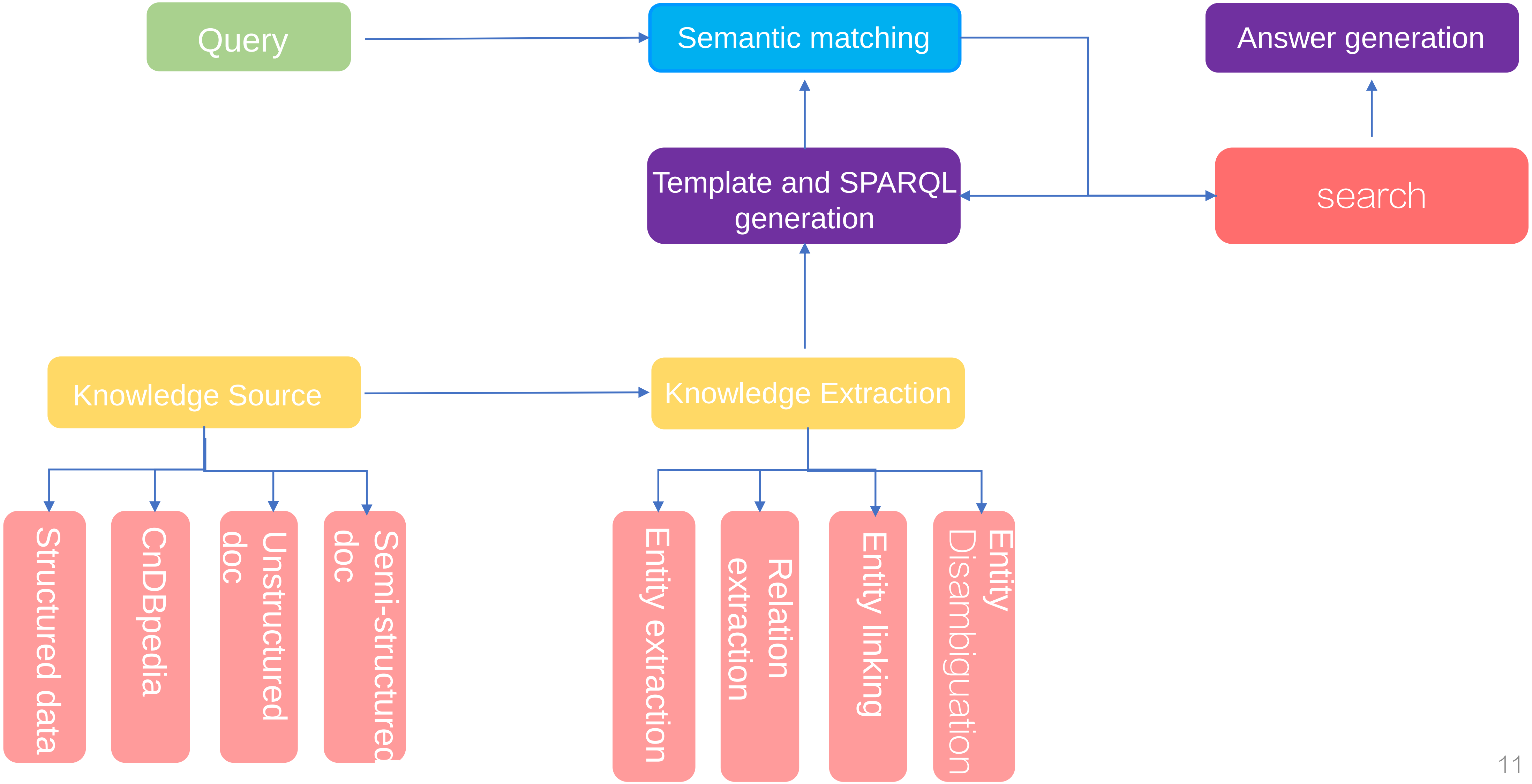
$$(1-Y)\frac{1}{2}(D_W)^2 + (Y)\frac{1}{2}\{max(0, m - D_W)\}^2$$

$$Dw = \sqrt{\{G_W(X_1) - G_W(X_2)\}^2}$$

Siamese network[1] using RNN sentence Encoder and wmd features with contrastive loss

[1] https://hackernoon.com/one-shot-learning-with-siamese-networks-in-pytorch-8ddaab10340e

Knowledge Graph based Question Answering



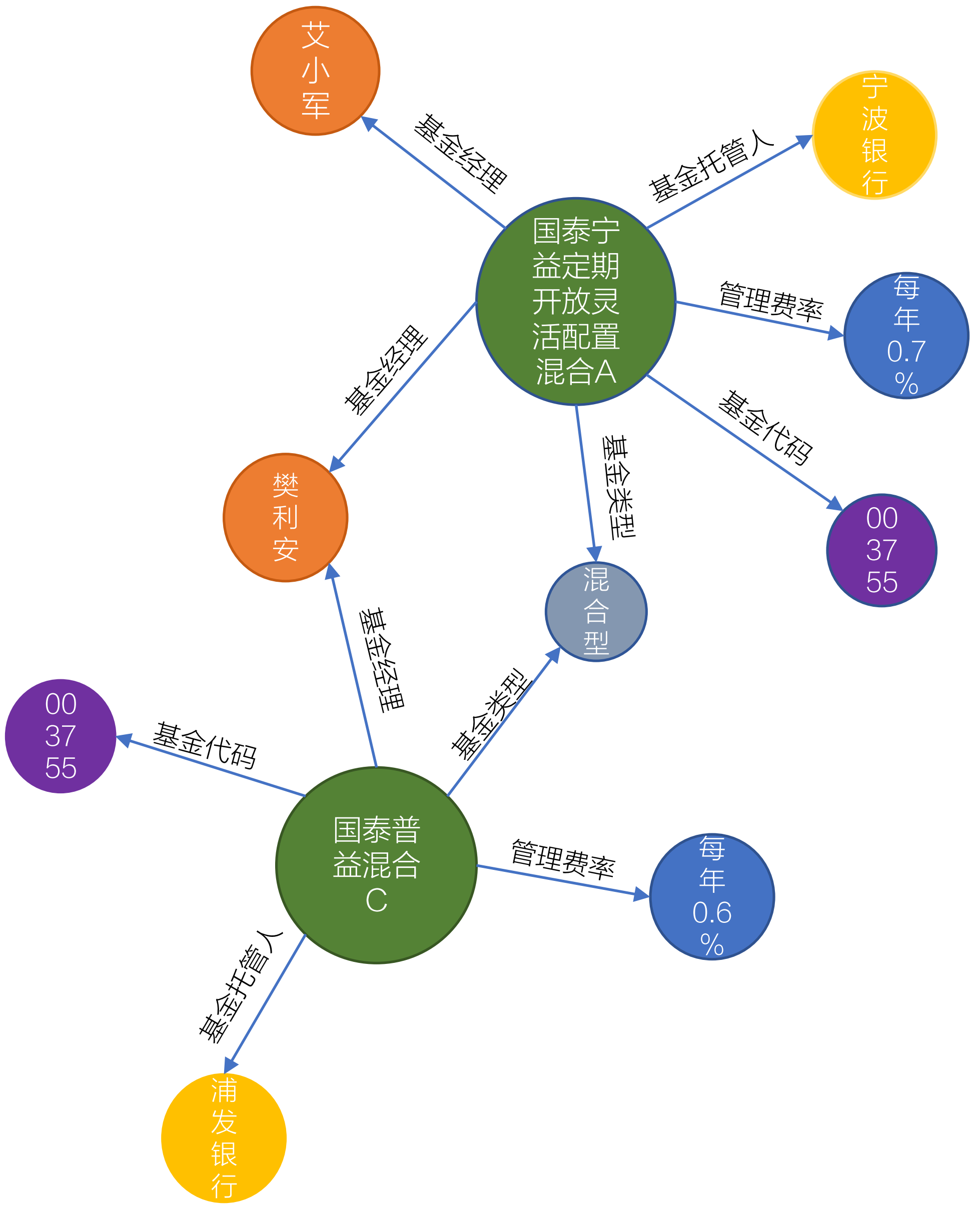
Knowledge Extraction

基本概况 其他基金基本概况查询： 003760 - 国泰宁益定期开放灵

基金全称	国泰宁益定期开放灵活配置混合型证券投资基金	基金简称	国泰宁益定期开放灵活配置混合A
基金代码	003760（前端）	基金类型	混合型
发行日期	2017年04月28日	成立日期/规模	2017年08月01日 / 2.054亿份
资产规模	1.91亿元（截止至：2017年10月27日）	份额规模	1.8268亿份（截止至：2017年08月01日）
基金管理人	国泰基金	基金托管人	宁波银行
基金经理人	樊利安、艾小军	成立来分红	每份累计0.00元（0次）
管理费率	0.70%（每年）	托管费率	0.10%（每年）
销售服务费率	0.00%（每年）	最高认购费率	0.60%（前端）
最高申购费率	0.60%（前端）	最高赎回费率	1.50%（前端）
业绩比较基准	沪深300指数收益率×50%+中债综合指数收益率×50%	跟踪标的	该基金无跟踪标的

基本概况 其他基金基本概况查询： 003755 - 国泰普益混合C

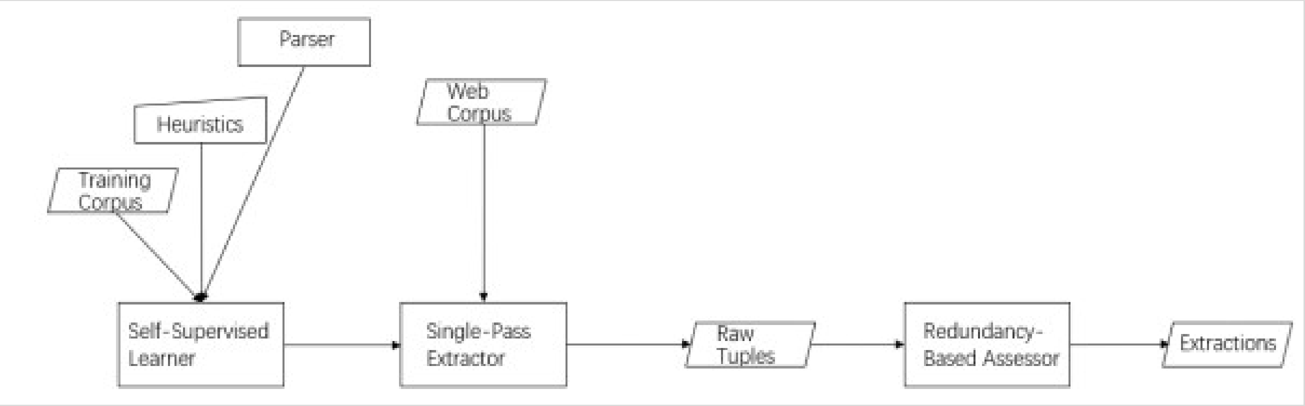
基金全称	国泰普益灵活配置混合型证券投资基金	基金简称	国泰普益混合C
基金代码	003755（前端）	基金类型	混合型
发行日期	2016年12月09日	成立日期/规模	2016年12月23日 / 2.002亿份
资产规模	2.22亿元（截止至：2017年09月30日）	份额规模	2.0037亿份（截止至：2017年09月30日）
基金管理人	国泰基金	基金托管人	浦发银行
基金经理人	樊利安	成立来分红	每份累计0.00元（0次）
管理费率	0.60%（每年）	托管费率	0.10%（每年）
销售服务费率	0.10%（每年）	最高认购费率	0.00%（前端）
最高申购费率	0.00%（前端）	最高赎回费率	0.50%（前端）
业绩比较基准	沪深300指数收益率×50%+中债综合指数收益率×50%	跟踪标的	该基金无跟踪标的



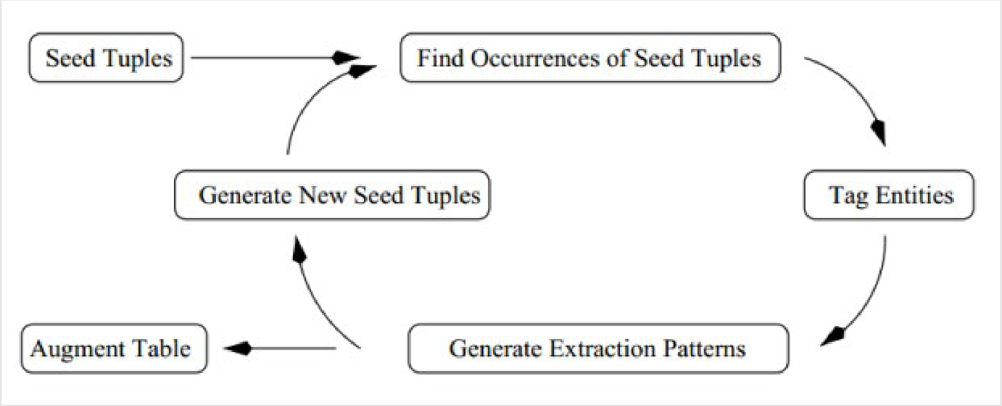
Knowledge Extraction———Cold start

Unstructured and semi-structured data

Open Information Extraction



Bootstrapping



Examples:

星展集团是亚洲最大的金融服务集团之一, 拥有约3千5百亿美元资产和超过280间分行, 业务遍及18个市场。

e1:星展集团, e2:亚洲最大的金融服务集团之一, r:是
e1:星展集团, e2:约3千5百亿美元资产, r:拥有
e1:业务, e2:18个市场, r:遍及

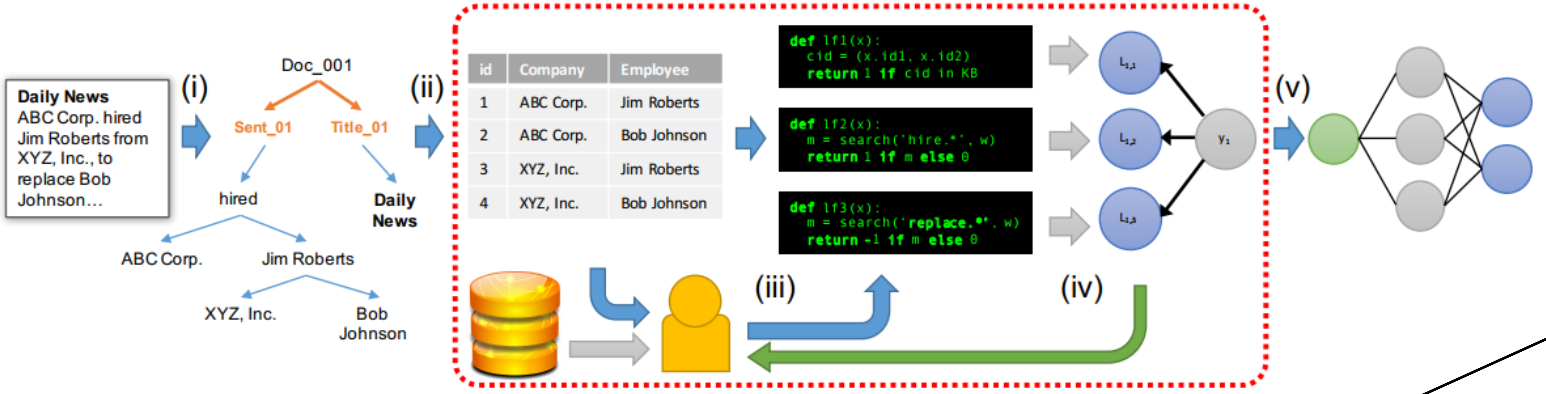
Data Programming in Snorkel

- The user
  - ***Loads in*** unlabeled data
  - ***Writes*** labeling functions (LFs)
  - ***Chooses*** a discriminative model, e.g., LSTMs

snorkel

- Snorkel
  - ***Creates*** a noisy training set- *by applying the LFs to the data*
  - ***Learns*** a model of this noise- *i.e. learns the LFs' accuracies*
  - ***Trains*** a *noise-aware* discriminative model

Data programming[1] for noisy labeled data construction using Snorkel



Label Function:

```
def LF_in_lexicion_1(c):
    t=c.text()
    if t[-1] in insurance_dict:
        return 1  else:    return 0

def LF_in_lexicion_2(c):
    t=c.text()
    after=t.next_word()
    if after in insurance_deny_tailword:
       return -1  else :    return 0

def LF_in_pos(c):
    t=c.text()
    before = t.forword_word()
    if before.pos in list_pos:
        return 1   else:    return 0
```

Examples:
公司冒用意外险，对企业有什么损害？
万能保险可靠吗？
重大疾病保险可以重复投保吗

Label function matrix:

[[0,1,0], [0,0,0], [0,0,0]]   [[0,1,0]]   [[0,0,1], [0,1,0]]

Logistic Regression:

$$\hat{\theta} = \operatorname{argmin}_{\theta} \sum_{x} \mathbf{E}_{\mu_{\alpha}} \left[l(x, y; \theta) \mid \lambda \right]$$

Candidates:
“公司冒用意外险”、“冒用意外险”、“意外险”
“万能保险” “重大疾病保险”

Select candidates with highest probabilities
意外险　万能保险　重大疾病保险

Alexander J Ratner .etl Data programming: Creating large training sets, quickly. NIPS 2016 pages 3567–3575.
https://hazyresearch.github.io/snorkel/

Knowledge Extraction

With labeled data-----Neural Entity Recognition & Neural Relation Extraction

NER extract name entity from sentence
NRE extract relation about two entities in sentence

Input sentence → NER → NRE → Triple (subject,relation_type,object)

“tesla”
“Palo Alto , California”
“2003”

→

<tesla,Palo Alto,California>
location
<tesla,2003>
Founded_time

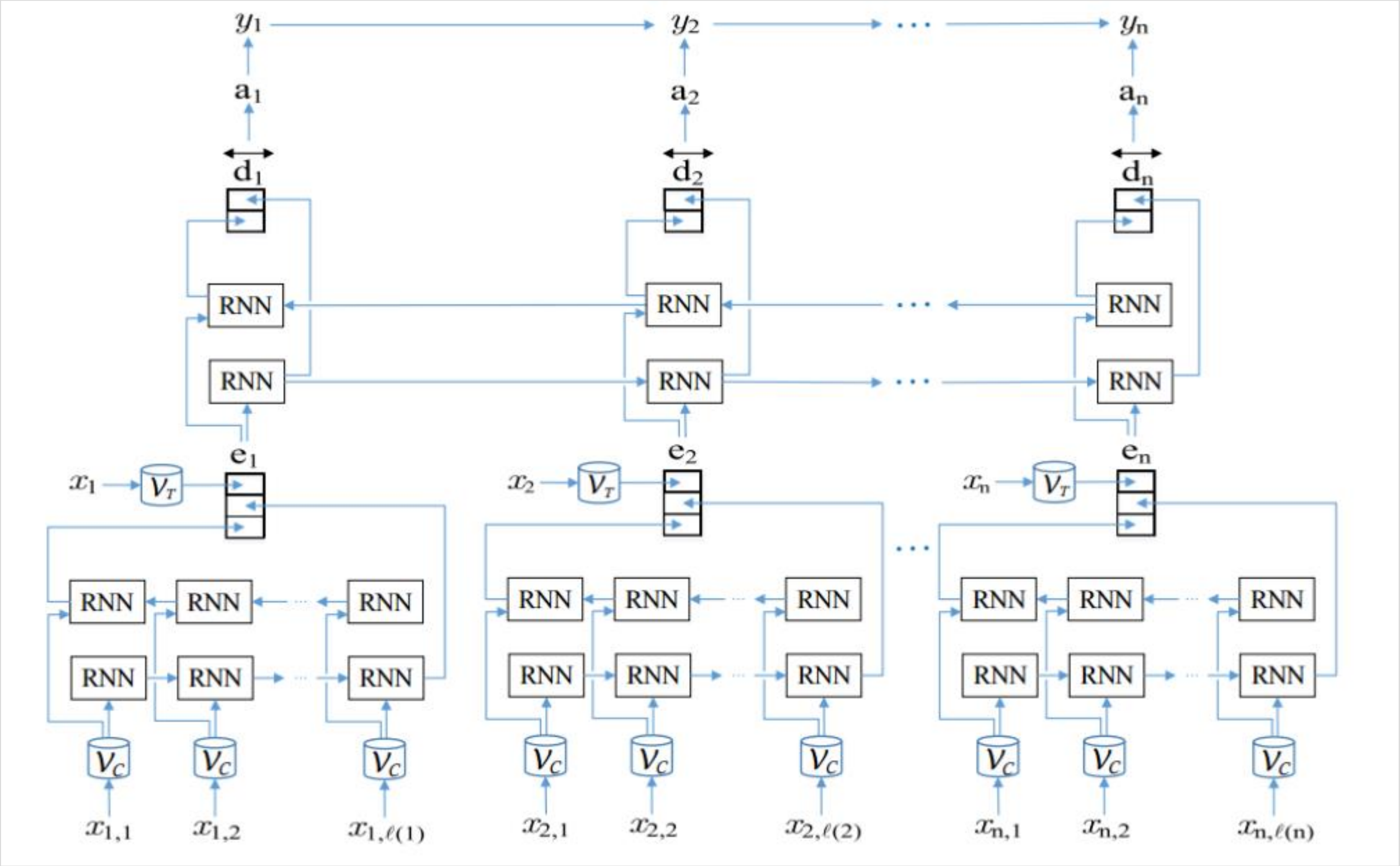
→

(“tesla”,”location”,” Palo Alto , California”)
(“tesla”,”founded_time”,”2003 ”)

Examples:

Tesla, Inc. (formerly named **Tesla Motors**) is an American automaker, energy storage company, and solar panel manufacturer based in Palo Alto, California.[6] Founded in 2003, the company specializes in electric cars, lithium-ion battery energy storage, and through their SolarCity subsidiary, residential photovoltaic panels.

NER Bi-LSTM+CNN+CRF Sequence labelling



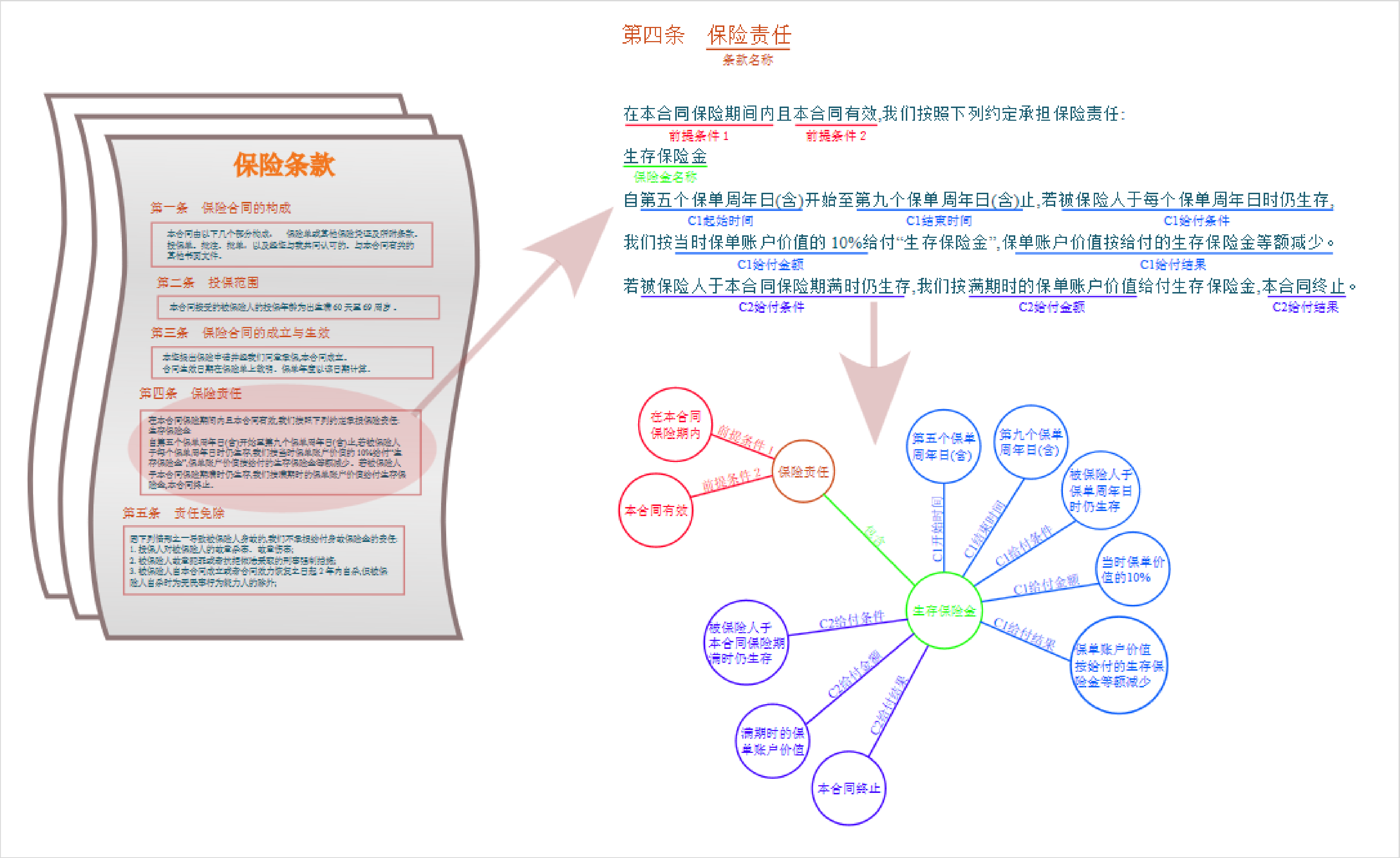
Applications:
Knowledge extraction
Slot extraction

留学生创业大厦附近有什么好吃的呢?

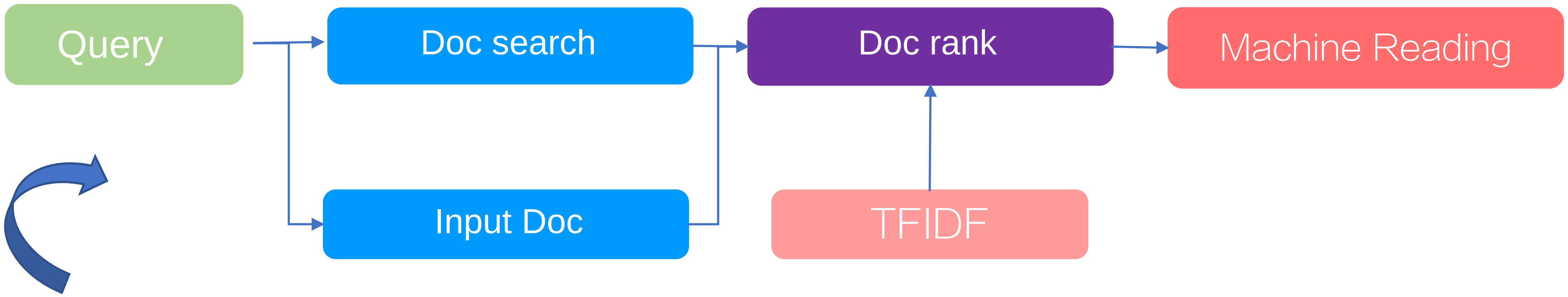
我在五道口地铁站,想去天安门,该怎么走呢

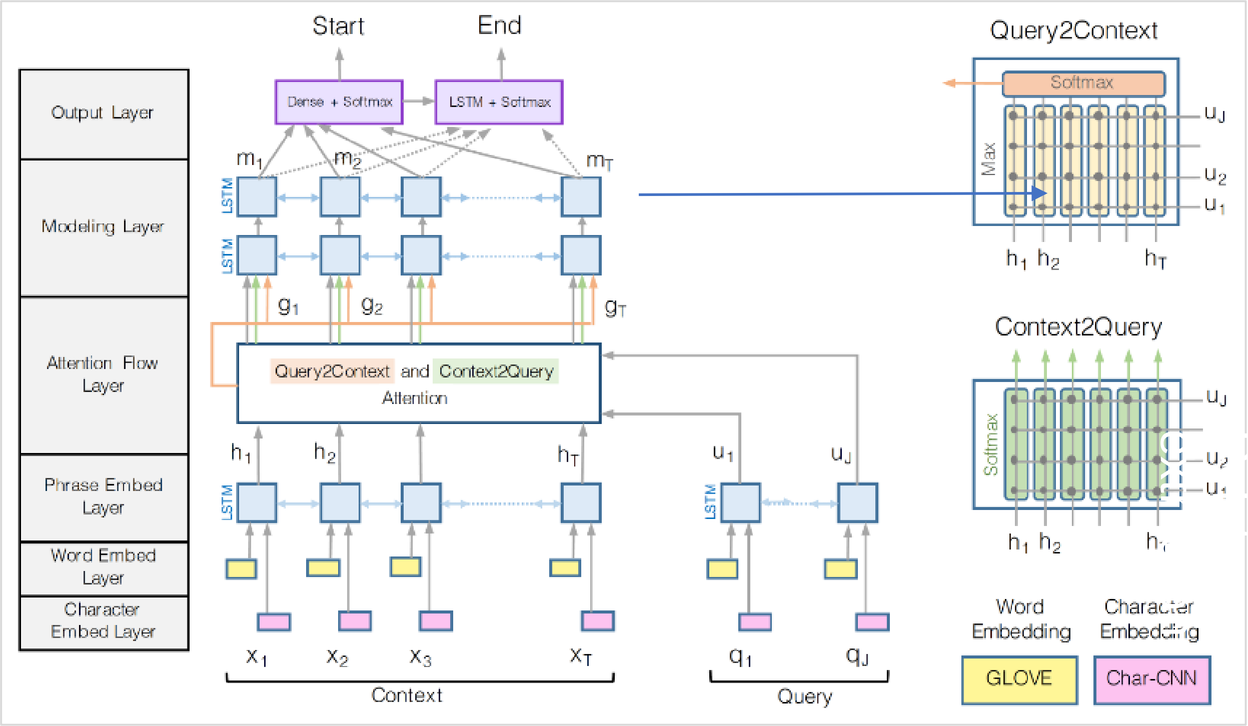
http://neuroner.com/

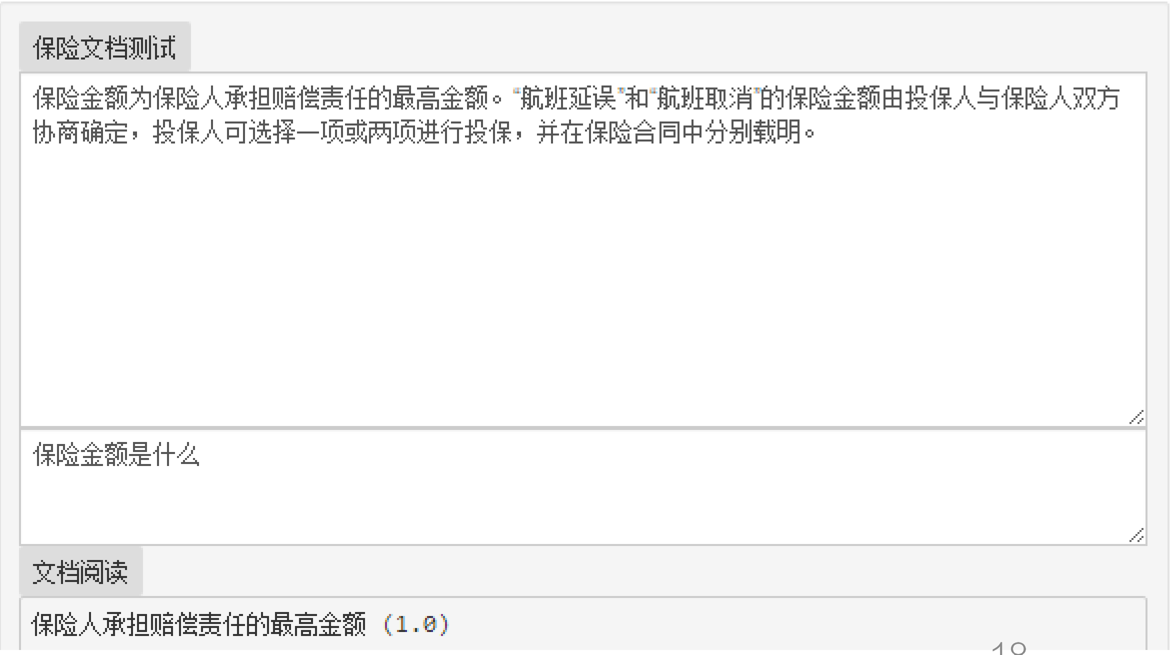
Neuro Relation Extraction as short text classification



DocQA for un-structured document based QA



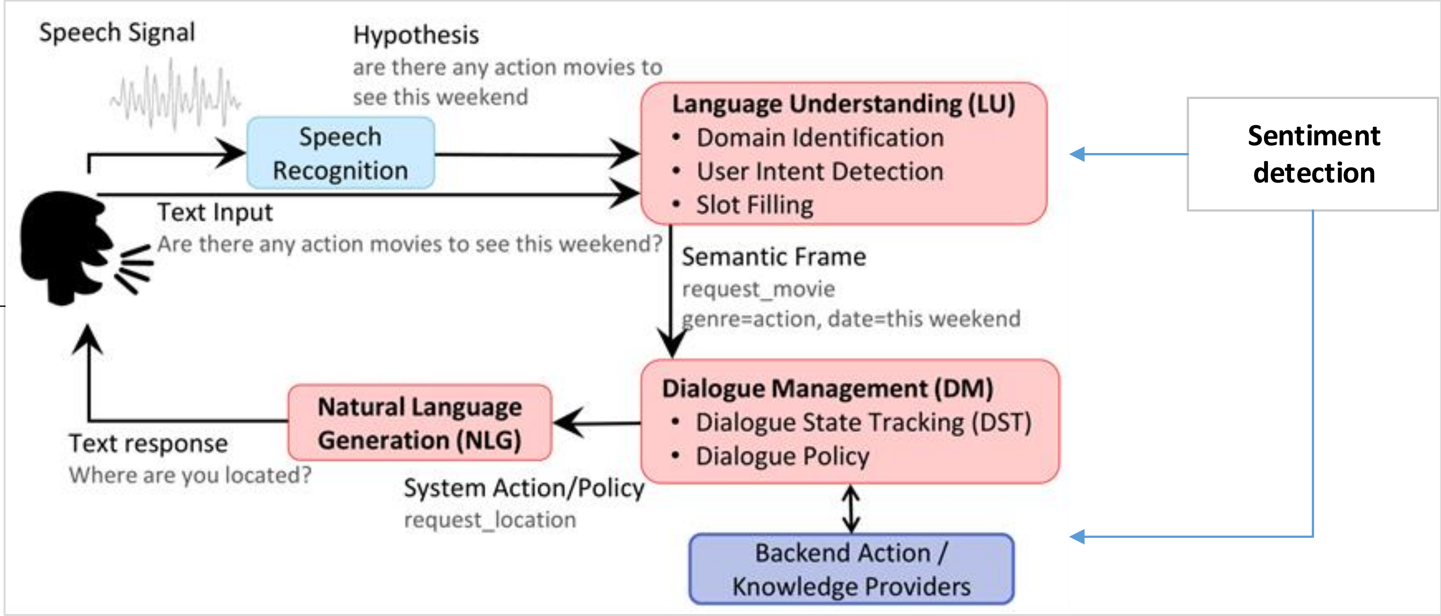




Minjoon Seo, et al. BI-DIRECTIONAL ATTENTION FLOW FOR MACHINE COMPREHENSION. ICLR17

Task Completion Dialog System

Add sentiment detection for better customer experience



Sentiment detection

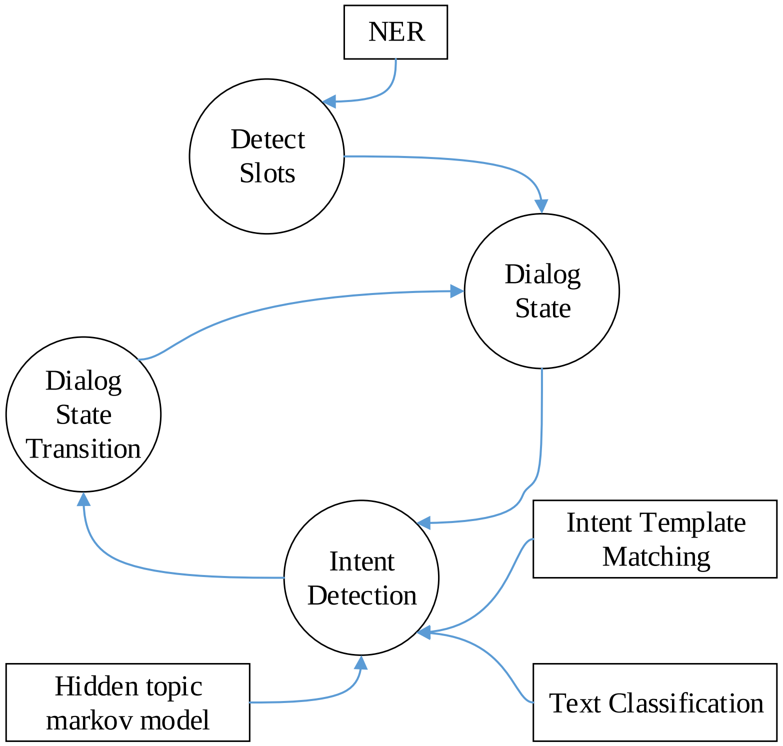
Domain identification,
Intent detection,
Sentiment detection
Those are typical
short text classification tasks

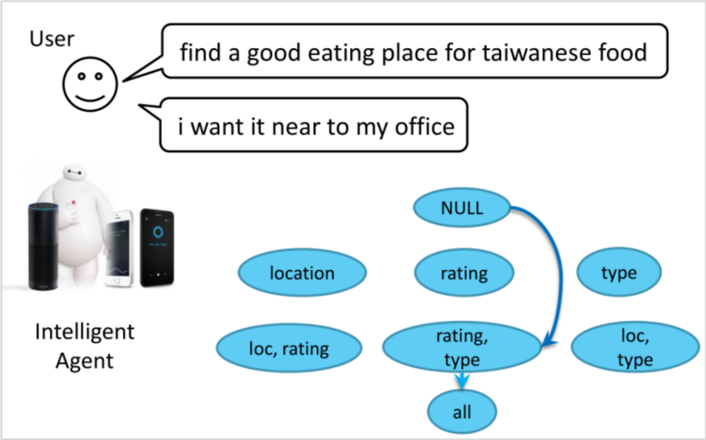
Output

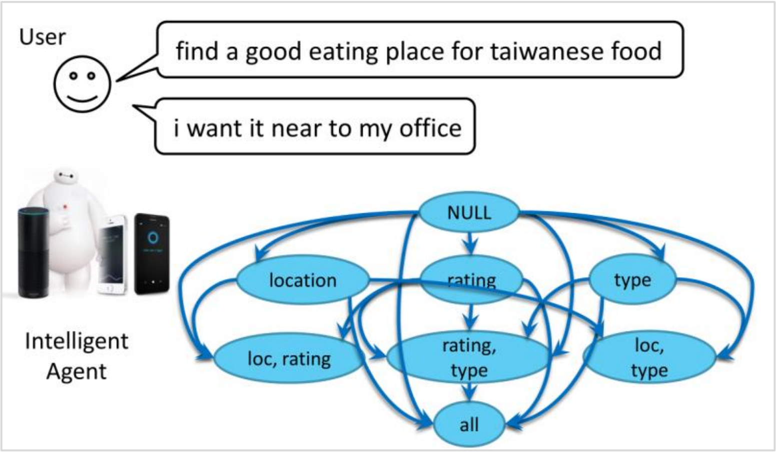
Chitchat

Cold Start No labeled data and dialog log

Hand Crafted State and state transition





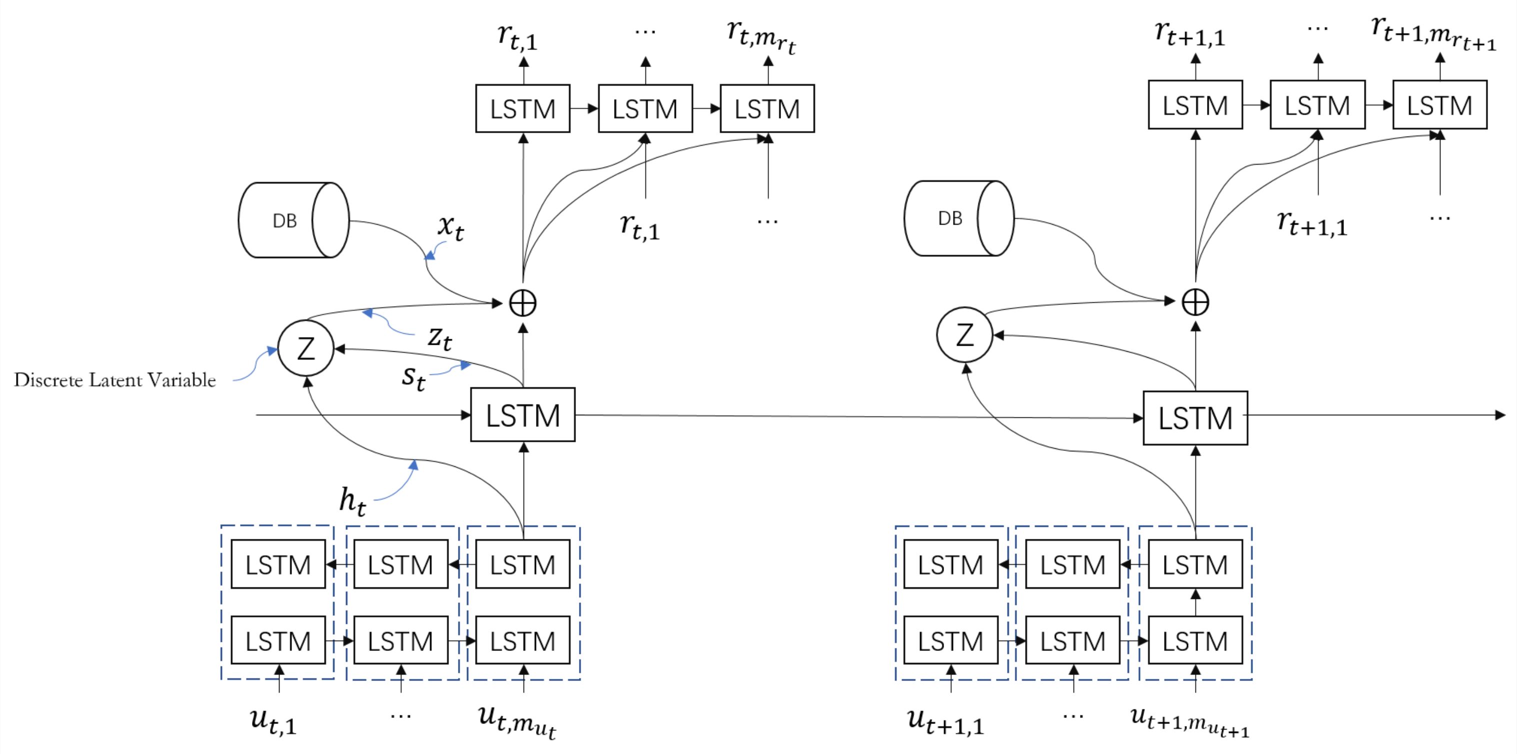


Dialog state transition

Task oriented dialog System with massive dialog log

End-to-end model that can be applied to unsupervised, semi-supervised and reinforcement learning

With dialogue data （no need for data labelling）: Submitted to top conference



1. Slot detection via NER or string fuzzy match
2. hierarchical encoder-decoder augmented with discrete latent intention for dialog modeling

HDLIED support supervised, unsupervised and reinforcement learning.

Finite and countable discrete latent can be Trained with neural variational inference learning(NVIL)[1] or Maximum log-likelihood Estimation. It can also tuned via REINFORCE[2] algorithms.

Models	Success rate 10	Success rate 30	Success rate 50	BLEU 10	BLEU 30	BLEU 50
HRED	75.6%			**0.339**		
HRED,+db_vector	73.1%			0.330		
HDLIED	73.8%	76.3%	77.5%	0.290	0.283	0.292
HDLIED, +db_vector	71.9%	76.9%	81.9%	0.296	0.294	0.279
HDLIED, +RL	74.4%	76.3%	78.1%	0.287	0.283	0.288
HDLIED, +db_vector, +RL	73.8%	75.6%	**82.5%**	0.294	0.292	0.279

Table 1: Offline results of our HDLIED via **exact MLE** varying latent intention dimension from 10 to 50.

Models	Success rate 30	Success rate 50	Success rate 70	Success rate 100	BLEU 30	BLEU 50	BLEU 70	BLEU 100
VHRED	71.9%	73.8%	74.4%	72.5%	0.329	0.325	**0.337**	0.297
HDLIED	72.5%	72.5%	75.0%	73.8%	0.310	0.317	0.315	0.312
HDLIED, +db_vector	75.6%	75.6%	73.8%	78.1%	0.317	0.305	0.331	0.301
HDLIED, +RL	**84.5%**	73.1%	76.9%	80.0%	0.237	0.318	0.318	0.321
HDLIED, +db_vector, +RL	74.4%	75.6%	74.4%	78.1%	0.316	0.305	0.327	0.306

Table 2: Offline results of our HDLIED via **NVIL** (Mnih and Gregor 2014) varying latent intention dimensions from 30 to 100.

1. No need for massive data labeling
2. End-to-end trainable with different learning paradigms: unsupervised, semi-supervised and reinforcement learning

Published Models	Success rate	BLEU
NDM+	76.1%	0.212
NDM+Att	79.0%	0.224
NDM+Att+SS	81.8%	0.240
LIDM*, I=50	66.9%	0.238
LIDM, I=70	61.0%	**0.246**
LIDM, I=100	63.2%	0.242
LIDM, I=50, +RL	82.4%	0.231
LIDM, I=70, +RL	81.6%	0.230
LIDM, I=100, +RL	**84.6%**	0.240

Needs pre-trained DST with labeled data for constraint and request slots

[1] Andriy Mnih and Karol Gregor. Neural Variational Inference and Learning in Belief Networks. Proceedings of the 34th International Conference on Machine Learning (ICML)
[2] Williams, Ronald J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. Machine learning 1992

Thanks

leo@webot.ai

www.webot.ai