NLP in China Mobile

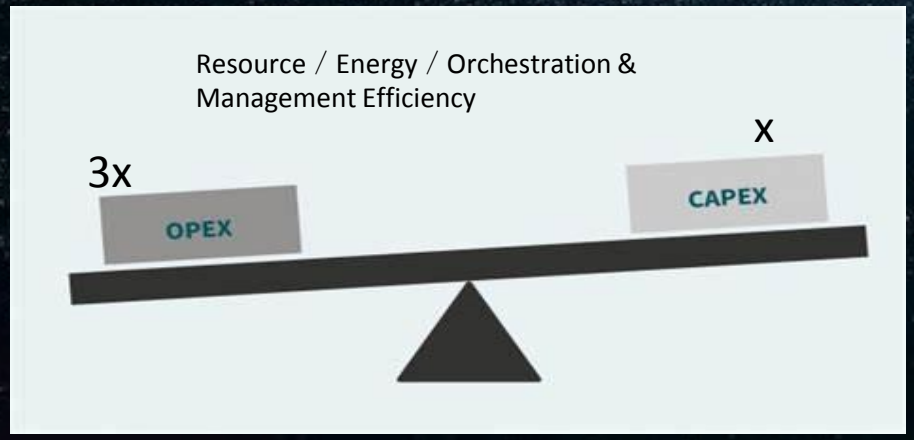
Junlan Feng, *China Mobile*

2019.10

Announcement: Lunch won’t be available until 15 mins later.

Sorry, It is not True!

Msg 1: Your work in this community has helped a lot to lower your Mobile Bill.



And your family's Mobile Bill

And many many others' Mobile Bill

Msg 2: China Mobile is tripling its investment on AI.

Not Interesting?

Way more as Planned!

中国移动 China Mobile

Msg 3: 5G will be intelligent .

If you tired of NLP, please come to work on 5G Intelligence.

It is fascinating, way beyond the Euclidean Space.

中国移动 China Mobile

Msg Summary:

We are hiring.

Please come to our booth to sense 5G .

China Mobile 5G Commercial Service Available in 50 Cities for 2019

中国移动 China Mobile

China Mobile 5G+ Plan

中国移动 China Mobile

JiuTian： China Mobile 5G + A Engine

Provide computing power、 data, environment, algorithms, solutions

Enable Industries

NLP in JiuTian： Intelligent Interaction

Context Understanding

Intelligent Q&A

Intelligent Recommendation

Intelligent Chat

Intelligent customer service

21,000,000+
Conversation / Month

Service Channels
228
Accuracy
92%+

Hospital Navigation

11,000,000
Patients
500+
Hopsitals

Accuracy
91%+

Spamming Phone Call Detection

Accuracy
98%

Smart OA

Assistant
OA
Agent

NLP in JiuTian： Network Complaint Service Robot

中国移动 China Mobile

Intent Analysis

Complaint Location and Demarcation

Troubleshooting Strategy Generation

Alarm Correlation and Compression

Fault Detection

Fault Location

Fault Handling

Network Complaint Service Robot

60,000,000+
Customers

Efficiency
90% ↑

Fault Location Accuracy
Top1 : 71%
Top3 : 86%
Top5 : 91%

Complaint Tickets
62% ↓

NLP in JiuTian： Video Recommendation on OTT TV

中国移动 China Mobile

Hybrid Candidate Generation

Rank Aggregation

Real-time Feedback

Comprehensive User Profiles

Video Recommendation on TV

200,000
Media Resources

8,000,000
Users

Play Rate
42% ↑

Income
12% ↑

NLP in JiuTian： Smart Contract Auditing

Recommendations for Contract Auditing

Smart Contract Auditing

1,000,000
Contracts / Year

300,000
Users

Recognition Accuracy
90%

Contract Auditing Accuracy
90%

Innovation1: Code-Switching Sentence Generation by Bert and Generative Adversarial Networks

■ Introduction

Table 1: Code-switching phenomena and examples

Code-switching phenomena	Examples
Vocabulary extension	• wifi, imax, ipad, hotdog, Gucci, latex, Hotmail, moto
Word translation	• 对我(To me) say bye
Phrase or sentence transformation	• 两个(Two) 宝贝(Babies) 一起(Together) cute 我(I) hold 不住(Can't) 了 • 江 南(Jiangnan) style
Spoken citation	• 那 书(The book) CoCo Chanel 特别 (So) 棒(Great)

■ Our Approach

➢ Bert-based Data Generation

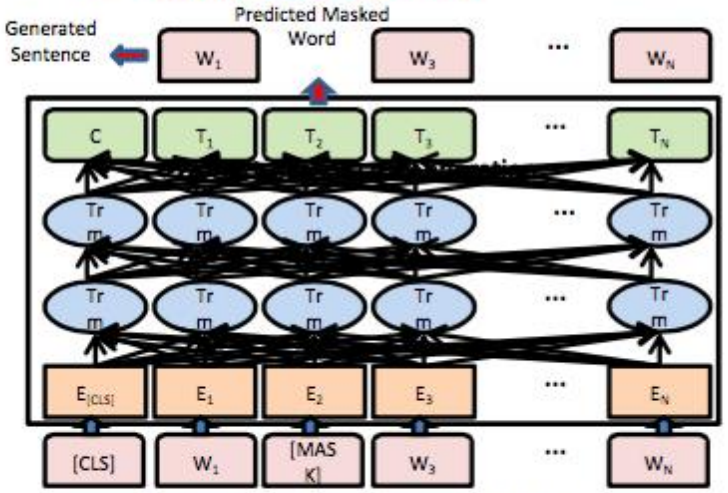


Figure 1: The process of generating a sentence by Bert.

– Use mix-lingual text to fine-tune: Mandarin-English code-switching spontaneous speech corpus OC16-CE80, train set

➢ GAN with Bert for Generation

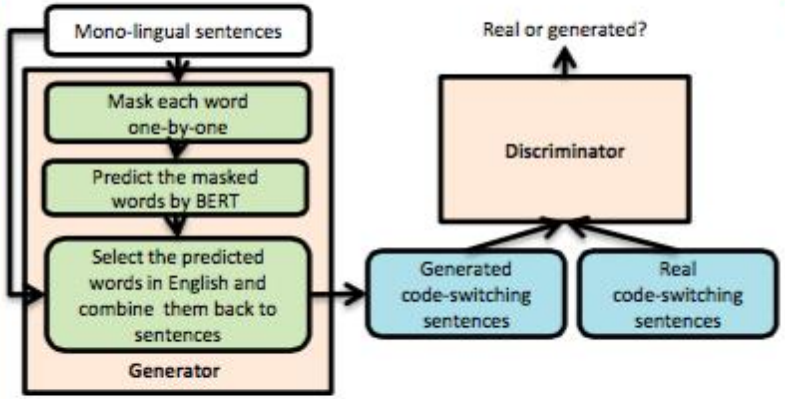


Figure 2: The diagram of the GAN method with Bert.

– Use mono-lingual text to generate code-switching data :
Mandarin speech corpus AISHELL-ASR0009-OS1

Generator Loss $L_G = L_{G_o} + L_{masked\_lm} \quad L_{G_o} = \frac{1}{M}\sum_j \log(1 - D(g_j^{Eng}))$

Discriminator Loss $L_D = \frac{1}{M}\sum_i \log(1 - D(x_i)) + \frac{1}{M}\sum_j \log(D(g_j^{Eng}))$

➢ English Word Prediction

Table 2: The prediction accuracy of the masked English words

Model	ACC
Bert-C	13.71%
+FineTuned-RandMask	42.69%
+FineTuned-EngMask	55.61%
+FineTuned-EngMask + GAN	58.67%
+FineTuned-EngMask + GAN+sel	**60.00%**

➢ Language Model Evaluation

Table 3: The perplexities of the generated data evaluated by original LM

Data	PPL
AISHELL mono	3591.66
Generated mix	1935.97
Generated mix $logP \geq -2$	1879.25
Generated mix $logP \geq -1$	**1755.70**

Table 4: The perplexities of the test set in OCC16-CE80 evaluated by different LMs

LM	PPL
Original LM	272.34
+AISHELL mono	449.27
+Generated mix	478.16
+Generated mix $logP \geq -2$	451.96
+Generated mix $logP \geq -1$	**383.78**

➢ Code-switching Speech Recognition

Table 5: The speech recognition WERs with different LMs

LM	Overall	Chi	Eng
Original LM	24.29	22.68	36.12
+AISHELL mono	**23.27**	**21.53**	36.09
+Generated mix	24.04	22.41	36.09
+Generated mix $logP \geq -2$	**23.92**	22.33	**35.57**
+Generated mix $logP \geq -1$	24.05	22.42	36.05

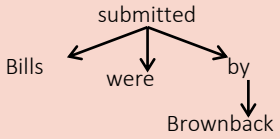
➢ Examples

Table 6: Several generated examples by the proposed mode

Original mono	Generated mixed
玩 得(Played) 相 当(Quite)开 快(Happy)	玩得 相当 **high**
如 今(Now) 三 星(Samsung) 正 在(Is) 进 一 步(Further) 推 动(Promoting) 自 己 的(Its) 虚 拟 现 实(Virtual reality) 业 务(Business)	如今 **Google** 正在 进 一步 推动 自己的 虚 拟现实 业务
保 监 会(CIRC) 五 部(Five measures) 解 决(Solve) 车 险(Vehicle insurance) 理 赔(Claim settlement) 难(Difficulty)	保监会 五部 解决 **p2p** 理赔 难
按照(According) 谁的(Whose) 孩 子(Child) 谁 抱 的(Who hold) 原 则 (Principle)	*Bert:* 按照 谁的 孩子 谁抱的 **baby** *Bert-GAN:* 按照 谁的 **baby** 谁抱的 原则
为 (For) 各 自 的(Each other's) 团 队(Team) 争 取(Fight) 胜 利(To win)	*Bert:* 为 各自的 团队 **solo** 胜利 *Bert-GAN:* 为 各自的 **team** 争取 胜利

Innovation2: A New Fine-Tuning Architecture Based on Bert for Word Relation Extraction

■ Introduction

Word-level relation classification	Examples
Dependency parsing	
Semantic relation classification	There were apples, **pears** and oranges in the **bowl**. (container, **pears**, **bowl**)

■ Contribution

• A new architecture to fine-tune BERT for word-level relation classification tasks.

• A probing task to test and visualize the extent to which BERT representations preserve the word relationship as BERT primarily built on Self-Attention Transformer mechanism.

■ Our Approach

➢ Attention map pair extraction

• As an intuition, noted that the learnt attention weights between words are good candidate features for word relation extraction.

• *Figure 1* illustrates the procedure how pair-wise attention is formed. Each cell corresponds to an attention weight.

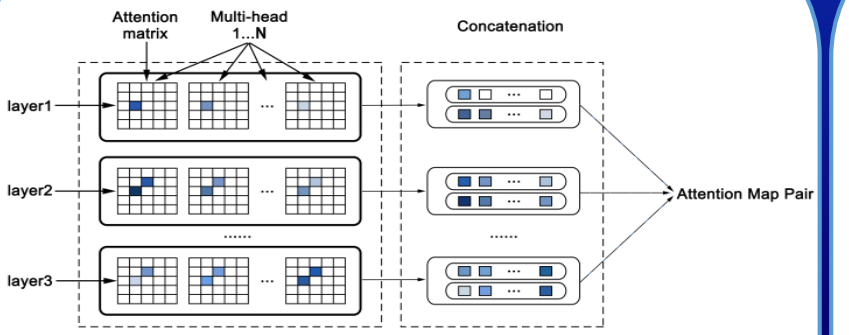


Fig.1. The process of attention map pair extraction

➢ Relation extraction based on Bert

• *Figure 2* indicates how to augment the basic BERT fine tuning architecture with pair-wise attention weights.

• Our argument is that these weights directly represent the rich relationship between words.

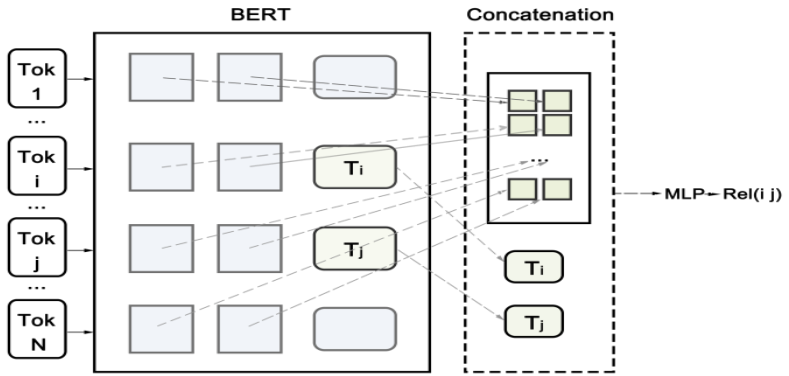


Fig.2. The process of relation extraction based on Bert

■ Result

➢ Model evaluation

Table 1. Accuary of UAS and LAS of on PTB

MODEL	UAS	LAS
fea:emb	62.1	55.5
fea:emb+att	77.2	68.3
fea:ptb:emb+att	76.9	67.4
fine-tune:emb	87.8	86.0
fine-tune:emb+att	89.3	87.8
fine-tune:emb+att(mst)	**90.9**	**88.9**

Table 2. F1 on semiEval-2010 task8

MODEL	F1
fea:emb	50.8
fea:emb+att	61.1
fea:ptb:emb+att	61.1
fine-tune:emb	76.5
fine-tune:emb+att	**78.4**

➢ Probing task for word relation extraction

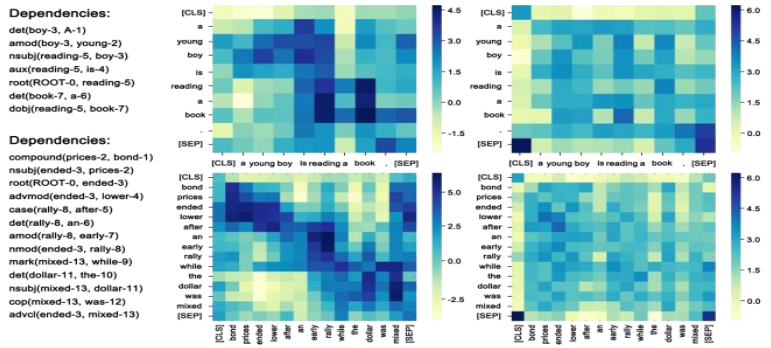
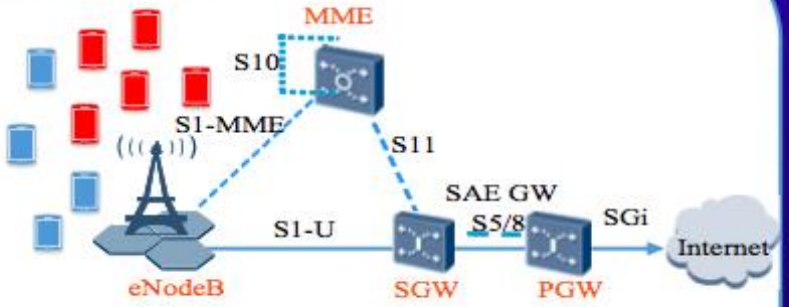


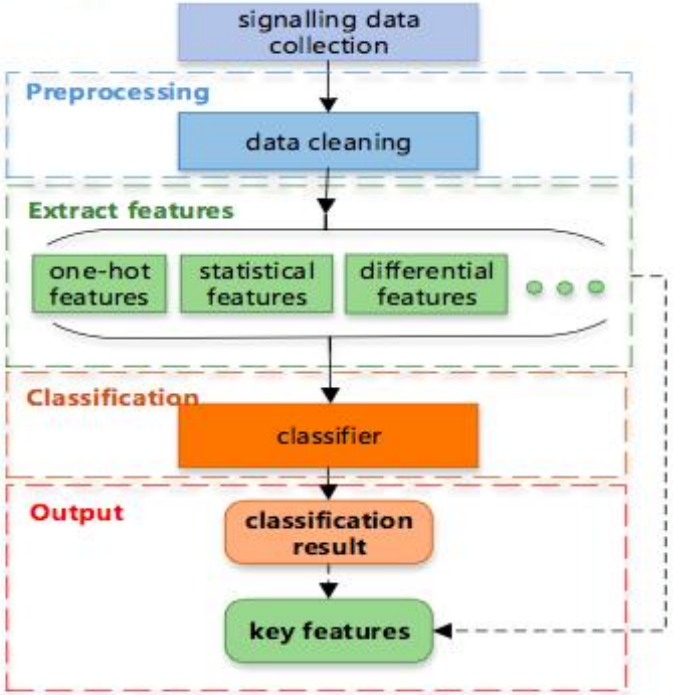
Fig.3. Attention map of top2 important features of different sentences for dependency relationship extraction

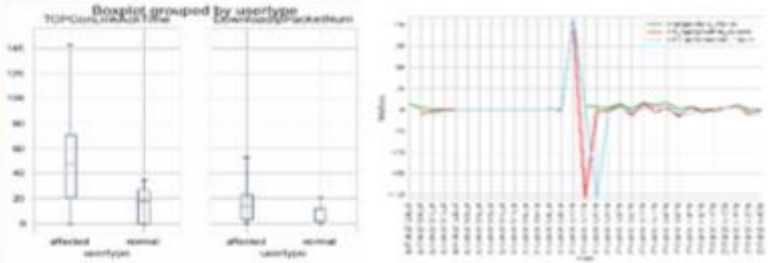
Innovation3: Smart Prediction of the Complaint Hotspot Problem in Mobile Network

Introduction



Our Approach

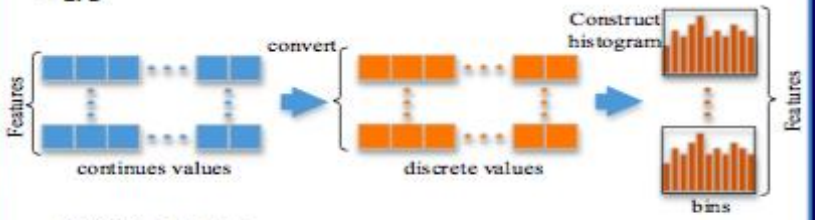




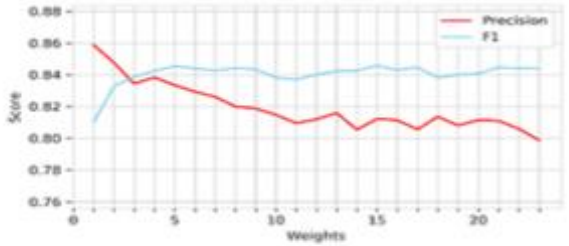
– GOSS

$$\bar{V}_j(d) = \frac{1}{n}\left(\frac{\left(\sum_{x_i \in A_l} g_i + \frac{1-a}{b}\sum_{x_i \in B_l} g_i\right)}{n_l^j(d)}\right) + \frac{1}{n}\left(\frac{\left(\sum_{x_i \in A_r} g_i + \frac{1-a}{b}\sum_{x_i \in B_r} g_i\right)}{n_r^j(d)}\right)$$

– EFB



– Weight selection

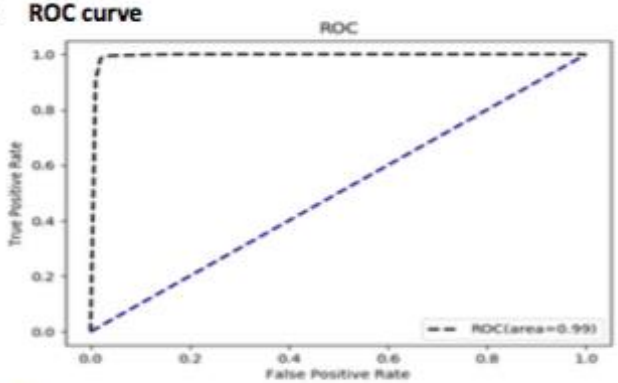


Method	Precision	F1 score
LightGBM	83.35%	**84.54%**
Decision Tree	75.71%	71.06%
Random Forest	83.25%	78.12%
GBDT	85.35%	80.90%
DNN	75.59%	78.81%
CNN	79.25%	76.83%

➢ **Features of TOP importance**

Control plane	User plane
Number of signalling data	Maximum of the spendtime
Sum of paging success number	Standard variance of windowSize
Total number of E-RAB release	Minimum of TCPSynNum
Total number of signalling data	Minimum of upload traffic
Total number of authentication failure	Sum of the spendtime

➢ **ROC curve**



R&D Concept

Openness, Cooperation, Mutual Benefit

Business Cooperation | University Cooperation | Industry Organization Cooperation

Joint Research | Ability Introduction | Joint Projects | Joint Labs | Industry Empowerment | Industrial Alliance

AIIA Tour of China Mobile (aiia.cmri.cn)

AI Self-Developed Competition of CMCC

5 of the Top10 teams used the JiuTian AI R&D Platform.

Smart Family Competition

91 Teams

Intelligent Hardware&
Intelligence Platform Related Topics.

AI Competition of the CMCC Network Department

JiuTian provided computing power, platform, technical trainings and guidance for participating teams.

Intelligent Network Competition

337 Teams, 40 Selected for Second Run

Network Data Flow Prediction,
Wireless Fault Root Cause Analytics.

中国移动5G+AI引擎： 九天

中国移动
China Mobile

望廬山瀑布
李白

日照香炉生紫烟，
遙看瀑布挂前川。
飛流直下三千尺，
疑是銀河落九天。

水調歌頭·重上井岡山
毛澤東

久有凌雲誌，
重上井岡山。
千裏來尋故地，
舊貌變新顏。
到處鶯歌燕舞，
更有潺潺流水，
高路入雲端。
過了黃洋界，
險處不須看。

風雷動，
旌旗奮，
是人寰。
三十八年過去，
彈指一揮間。
可上九天攬月，
可下五洋捉鱉，
談笑凱歌還。
世上無難事，
只要肯登攀。

中国移动
China Mobile

九天人工智能平台

jiutian.cmri.cn


中国移动
China Mobile
谢谢！