

Improving speaker consistency in neural conversational models with adversarial approach.

Любовь Михайловна Яронская

Московский физико-технический институт

23 мая 2017 г.

- $X = (x_1, x_2, \dots, x_t)$ - входная последовательность токенов,
 $Y = (y_1, y_2, \dots, y_{n_y})$ - выходная последовательность токенов.
- LSTM задает распределение $P(Y|X) = \prod_{t=1}^{n_y} p(y_t|x_1, \dots, x_t, y_1, \dots, y_{t-1}) = \prod_{t=1}^{n_y} \frac{\exp(f(h_{t-1}, e_{y_t}))}{\sum_{y'} \exp(f(h_{t-1}, e_{y'}))}$
- Общие ответы: 'I don't know.', 'Of course.'
Решение: $\max(\log P(R|M) + \log P(M|R))$.
- Пример несогласованности ответов:

message Where are you from?

response England, you?

message Where did you grow up?

response I grew up in Texas.

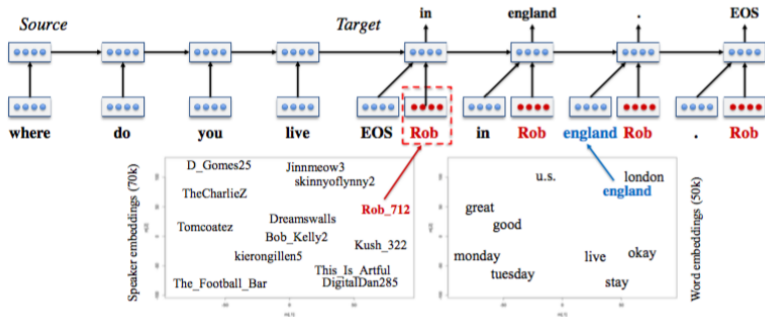
message How old are you?

response 16 and you?

message What's your age?

response 18.

A persona-based neural conversational model (Jiwei Li et al., 2016)



Несогласованность ответов в persona-based model

User2	
<i>message</i>	Where is your hometown?
<i>response</i>	I live in Indonesia.
<i>message</i>	Where are you from?
<i>response</i>	England, you?
<i>message</i>	In which city do you live now?
<i>response</i>	I live in Jakarta.
<i>message</i>	In which country do you live now?
<i>response</i>	I live in Indonesia.
<i>message</i>	What is your major?
<i>response</i>	Business, you?
<i>message</i>	What did you study in college?
<i>response</i>	Psychology, you?
<i>message</i>	How old are you?
<i>response</i>	I'm 18.
<i>message</i>	What is your age?
<i>response</i>	I'm 16.

Исследовать согласованность ответов диалоговой модели с применением соперничающих нейронных сетей.

Проблемы

Малый объем данных для обучения.

Неконсистентность ответов как следствие максимизации правдоподобия.

Предположения

RL дообучение.

Модели с латентными переменными лучше борются с двусмысленностью и неопределенностью.

GAN для генерации последовательностей

- Две модели - генератор G и дискриминатор D .
- D обучается различать реальные данные от сгенерированных, генератор обучается обманывать дискриминатор, порождая данные очень хорошего качества
- Обучение параметров дискриминатора D_ϕ :
$$\max_{\phi} (E_{Y \sim p_{data}} \log(D_\phi(Y)) + E_{Y \sim G_\theta} \log(1 - D_\phi(Y))).$$
- Обучение параметров генератора по policy gradient:
$$\nabla J = E_{X \sim p_{data}} E_{Y \sim G(Y|X)} \nabla \log G(Y|X) \cdot (D(Y) - b(X)),$$
 где в качестве бэйзлайна $b(X)$ скор от дискриминатора при жадном семплировании $D(Y_{greedy})$.

Добавление максимизации информации

В работе [Xi Chen et al., 2016] добавляют регуляризацию на информацию $I(c, G(x, c))$ между скрытыми переменными и ответами генератора.

Оценка сверху: $I(c, G(x, c)) \geq E_{c \sim P(c), x \sim G(z, c)} \log Q(c|x) + H(c)$.

Тогда если c - one-hot представление персонажа, то максимизация информации эквивалентна добавлению dense слоя с softmax нелинейностью от предпоследнего слоя дискриминатора и добавлении к функционалу дискриминатора минимизации кроссэнтропии.

Дискриминатор:

- LSTM + Attention
- Input: (Batchsize, $2 \cdot \text{EmbSize}$, QuestionLength, AnswerLength) + BatchNorm + Conv(3, 3) + MaxPool(2, 2) + Conv(3, 3) + MaxPool(2, 2)
- предыдущая модель с регуляризацией на информацию

- $Perplexity = e^{-\frac{1}{N_V} \sum_{n=1}^N \log P(U_1, \dots, U_k)}$, где N_V - мощность словаря, N - количество диалогов в тестовой выборке, k - размер диалога.
- Adversarial Success - доля фраз, не распознанных как сгенерированные.
- Количество правильно названных имен
- Human evaluation

- Сравнить результаты, получаемые использованием соперничающих нейронных сетей с предложенными ранее алгоритмами.

В качестве данных для предобучения дискриминаторов и дообучения генератора использовалась выборка субтитров "South Park" из 80714 фраз и 3194 персонажей. Валидационная выборка из 3200 фраз.

Сравнение качества полученных моделей

	AdvSucc	Perplexity
Baseline	-	7.96
LSTM+Att	0.699	7.65
Conv	0.977	13.86
Conv+Info	0.989	7.38

Визуализация внимания и примеры ревордов

0.646619



hello _EOS_ hello _EOS_ what is your name ?
GENERATED: my name is dr . biber .

Reward	Phrase
0.646619	my name is dr . biber .
0.386294	bye - bye , precious .
0.280729	my children , mr . jefferson says you already made a sex doll the school nurse gollum has UNK it
0.535254	trent boyett !
0.325994	very well .

Ответы моделей при контексте

["Hello" "Hello" "What is your name ?"]

Угаданные имена	Ответ
'president'	"it ' s the president of wall - mart . at south park elementary . the second was a hit"
'scott'	"my name . . . called scott , but my dad isn ' t even legal anymore ?"

Рис.: LSTM + Attention GAN

Угаданные имена	Ответ
'm'	"ooo stan marsh . a what are his name kyle gints ? i ' m sure as soon as filmore"
'man'	'hi fellas . i wanna go home . y - you gotta come meet the white man . get your'
'timmy'	'boooo !! show us what senator spears ? ! fudge said , timmy timmy timmy russell fan flying discipline'
'kenny'	'# and daddy my fat ugly little woman ! whoa whoa whoa whoa , whoa h . . . kenny'
'pip'	"eugh . uhuh , i ' m pip kids from my future ' s voice box salesman , travis harrisons"

Рис.: ConvGAN

Угаданные имена	Ответ	Прим.
'm'	"my name is <i>UNK</i> , and i ' m harry potter . i ' m the little boy . but"	-
'kyle'	'kyle .'	Ок
'snooki'	'snooki .'	Ок
'kenny'	'(a question as am , and is perhaps mr . havisham , but , his holiness , eric kenny'	-

Рис.: Conv + InfoApproximation GAN

- Модели на основе соперничающих нейронных сетей дают сопоставимые с базовой моделью результаты по метрике оценивания языковых моделей
- Дискриминаторы на основе CNN более сильные по AdvSucc и обучаются быстрее, так как меньшее число параметров.
- Обучение с максимизацией информации между one-hot вектором имени и ответом улучшает количество правильных ответов на вопрос об имени и качество ответа.

- 1 Iulian V. Serban, Alessandro Sordoni, Yoshua Bengio, Aaron Courville and Joelle Pineau. Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models, 2015.
- 2 Jiwei Li, Michel Galley, Chris Brockett, Georgios P. Spithourakis, Jianfeng Gao, Bill Dolan. A Persona-Based Neural Conversation Model, 2016.
- 3 Lantao Yu, Weinan Zhang, Jun Wang, Yong Yu. SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient, 2016.
- 4 Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, Pieter Abbeel. InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets, 2016.