

“Best Dinner Ever!!”: Automatic Generation of Restaurant Reviews with LSTM-RNN

Alberto Bartoli Andrea De Lorenzo
Eric Medvet Dennis Morello Fabiano Tarlao

Department of Engineering and Architecture
University of Trieste
Italy



October 16th, 2016

<http://machinelearning.inginf.units.it>

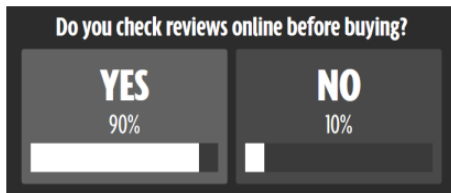
Table of Contents

- 1 Motivation
- 2 The tool
- 3 Experimental evaluation



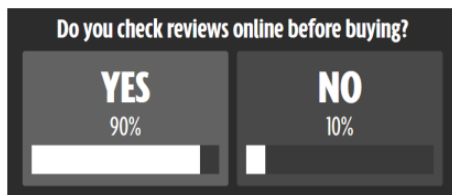
Product/service commerce

- People buy products/services online/offline
- When choosing seller, they trust other people's opinion (*reviews*)



Product/service commerce

- People buy products/services online/offline
- When choosing seller, they trust other people's opinion (*reviews*)



- A **malicious** seller might want to manipulate the choice (opinion spamming)
 - fabricating positive reviews for its products
 - fabricating negative reviews for competitors products



Review fabrication

Can be done “manually”:

Location : *Distance working*

Length of assignment : *2 months*

Budget : *\$100 – \$400*

Posted date : *04 March 2011*

Starting date : *ASAP*

Deadline for applications : *20 March 2011*

Required Skills : *Ability to write in English*

Submit application >

 Save to favourites >

 Subscribe to mailing list

Description

We are looking a “Ghost Writer” to write and post a total of 10 reviews and post them on Tripadvisor. The right candidate will write and post 5 good reviews about our hotel, and 5 very bad / bad reviews about another hotel (the full details will be given once an agreement has been made)

All reviews must be between 100 and a 1000 words. You must be able to create numerous email accounts and be able to change your IP address. We will pay \$30 for every good review that is posted, and \$40 for every negative review that is posted. We do not pay until the reviews are approved and posted so please do not apply for this position if you require any up-front payment. You can check our account history and see we have always paid as agreed.

“\$100–\$400” to “write and post a total or 10 reviews”, among which “5 good reviews about our hotel” and “5 very *bad* reviews about another hotel”!



Review fabrication: the next level

Can be done **automatically** by a tool?

- much cheaper (\approx free) for the single malicious seller
- much larger problem for the online retailer (Amazon, TripAdvisor, ...)
- (maybe) harder problem for opinion spamming researchers



Review fabrication: the next level

Can be done **automatically** by a tool?

- much cheaper (\approx free) for the single malicious seller
- much larger problem for the online retailer (Amazon, TripAdvisor, ...)
- (maybe) harder problem for opinion spamming researchers

Is that tool feasible?

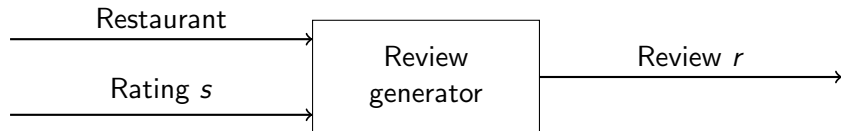


Table of Contents

- 1 Motivation
- 2 The tool
- 3 Experimental evaluation



Problem statement



r should:

- appear as generated by humans
- appear specific for restaurant
- express an overall rating s (from ★ to ★★★★★)

Input: what's a restaurant?

A set C of *categories*: e.g., Italian, Cafe, International, Mediterranean

Caffe San Marco

 400 Reviews | #9 of 666 Restaurants in Trieste |  Certificate of Excellence

\$\$ Italian, Cafe, International, Mediterranean 

Overview

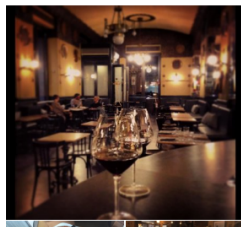
Reviews (400)

Details

Q&A

Location

 Save

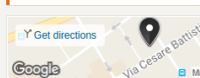


Reserve a Table Online

6:00 PM	6:30 PM	7:00 PM	7:30 PM
8:00 PM	8:30 PM	9:00 PM	9:30 PM
10:00 PM	10:30 PM	11:00 PM	11:30 PM

Is this an **European** restaurant?

Yes
 No
 Unsure

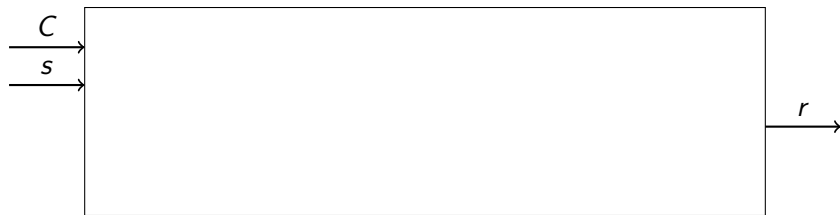


Via Cesare Battisti 18, 34125, Trieste, It
 +39 040 064 1724 Website

Today 8:30 am - 11:00 pm **Open**
 All hot

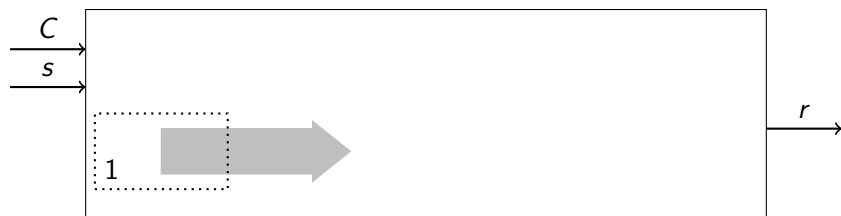


Method overview



Given an input C, s :

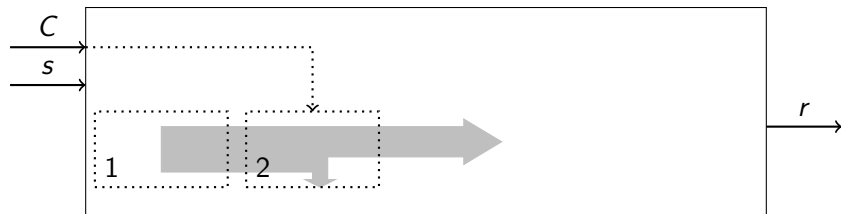
Method overview



Given an input C, s :

- 1 generate many “human-like” reviews (NLG w/ LSTM-RNN)

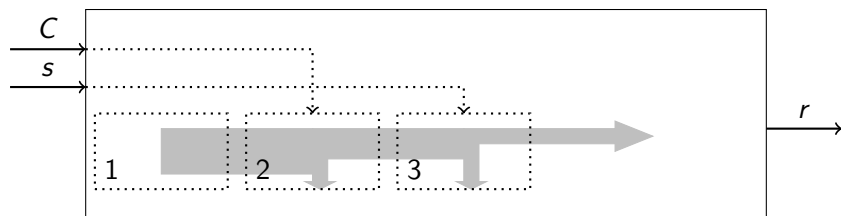
Method overview



Given an input C, s :

- ① generate many “human-like” reviews (NLG w/ LSTM-RNN)
- ② discard those not consistent with categories C (many binary classifiers)

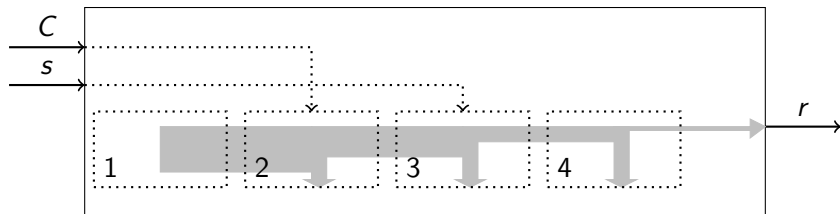
Method overview



Given an input C, s :

- 1 generate many “human-like” reviews (NLG w/ LSTM-RNN)
- 2 discard those not consistent with categories C (many binary classifiers)
- 3 discard those not consistent with rating s (one multiclass classifiers)

Method overview



Given an input C, s :

- 1 generate many “human-like” reviews (NLG w/ LSTM-RNN)
- 2 discard those not consistent with categories C (many binary classifiers)
- 3 discard those not consistent with rating s (one multiclass classifiers)
- 4 select randomly one review r among remaining reviews



Generating human-like reviews

Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN)

- works at the character level
- char-rnn library with default settings (3 layers of 700 neurons)
- trained on a corpus of 500000 reviews (\approx 1 month)
- when generating, seed is a random sentence of a real review
- first generated review is discarded (influence of the seed)

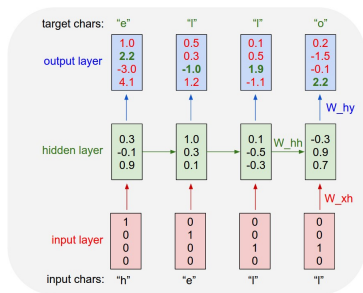


Table of Contents

- 1 Motivation
- 2 The tool
- 3 Experimental evaluation**



Aims

- Is an artificial review considered genuine? (*intrinsic evaluation*)
- Can an artificial review influence the human subject? (*extrinsic evaluation*)

Extrinsic performed first; 39 subjects involved, 3–4 forms each

Extrinsic evaluation

Simulates the restaurant choice by a user:

- each form with 3 reviews
- at least 1 artificial and 1 genuine

Uncle Sam's Meat & Wine

american, steakhouse

Review with ★★★★★

The atmosphere was very cozy. With small seating areas the noise is minimized. The service was good. [...]

Useful? Y N

Review with ★★

This place is dimly lit and reminded me of a bad prom decorating. The waitress was nice, but a little over [...]

Useful? Y N

Review with ★★★★★

Great food and even better atmosphere. It is a quiet darker setting with no windows. The service [...]

Useful? Y N

Would you go to this restaurant? Y N



Extrinsic evaluation

Simulates the restaurant choice by a user:

- each form with 3 reviews
- at least 1 artificial and 1 genuine

Genuine	Artificial	Going [%]	Not going [%]
\geq ★★★★	\geq ★★★★	47	53
\geq ★★★★	\leq ★★★	71	29
\leq ★★★	\geq ★★★★	24	76
\leq ★★★	\leq ★★★	23	77

Uncle Sam's Meat & Wine

american, steakhouse

Review with ★★★★★

The atmosphere was very cozy. With small seating areas the noise is minimized. The service was good. [...]

Useful? Y N

Review with ★★

This place is dimly lit and reminded me of a bad prom decorating. The waitress was nice, but a little over [...]

Useful? Y N

Review with ★★★★★

Great food and even better atmosphere. It is a quiet darker setting with no windows. The service [...]

Useful? Y N

Would you go to this restaurant? Y N



Extrinsic evaluation

Simulates the restaurant choice by a user:

- each form with 3 reviews
- at least 1 artificial and 1 genuine

Genuine	Artificial	Going [%]	Not going [%]
\geq ★★★★	\geq ★★★★	47	53
\geq ★★★★	\leq ★★★	71	29
\leq ★★★	\geq ★★★	24	76
\leq ★★★	\leq ★★★	23	77

Uncle Sam's Meat & Wine

american, steakhouse

Review with ★★★★★

The atmosphere was very cozy. With small seating areas the noise is minimized. The service was good. [...]

Useful? Y N

Review with ★★

This place is dimly lit and reminded me of a bad prom decorating. The waitress was nice, but a little over [...]

Useful? Y N

Review with ★★★★★

Great food and even better atmosphere. It is a quiet darker setting with no windows. The service [...]

Useful? Y N

Would you go to this restaurant? Y N



Extrinsic evaluation

Simulates the restaurant choice by a user:

- each form with 3 reviews
- at least 1 artificial and 1 genuine

Genuine	Artificial	Going [%]	Not going [%]
\geq ★★★★	\geq ★★★★	47	53
\geq ★★★★	\leq ★★★	71	29
\leq ★★★	\geq ★★★★	24	76
\leq ★★★	\leq ★★★	23	77

	Useful [%]	Not useful [%]
Genuine	80	20
Artificial	29	71

Uncle Sam's Meat & Wine
american, steakhouse

Review with ★★★★★

The atmosphere was very cozy. With small seating areas the noise is minimized. The service was good. [...]

Useful? Y N

Review with ★★★

This place is dimly lit and reminded me of a bad prom decorating. The waitress was nice, but a little over [...]

Useful? Y N

Review with ★★★★★

Great food and even better atmosphere. It is a quiet darker setting with no windows. The service [...]

Useful? Y N

Would you go to this restaurant? Y N



Intrinsic evaluation

Has this review been written by a human for this restaurant?

- forms with 5 reviews for each restaurant, 4 forms per user
- 4 kinds of reviews
 - R_{gs} **g**enuine for **s**pecific restaurant
 - R_{gd} **g**enuine for **d**ifferent restaurant
 - R_{ad} **a**rtificial for **s**pecific restaurant
 - R_{ad} **a**rtificial for **d**ifferent restaurant (no step 2)



Intrinsic evaluation

Has this review been written by a human for this restaurant?

- forms with 5 reviews for each restaurant, 4 forms per user
- 4 kinds of reviews
 - R_{gs} **g**enuine for **s**pecific restaurant
 - R_{gd} **g**enuine for **d**ifferent restaurant
 - R_{as} **a**rtificial for **s**pecific restaurant
 - R_{ad} **a**rtificial for **d**ifferent restaurant (no step 2)

	Yes [%]	No [%]
R_{gs}	81	19
R_{gd}	52	48
R_{as}	24	76
R_{ad}	24	76



Conclusions

Automatic Generation of Restaurant Review:

- “Is that tool feasible?” Yes! (we did it!)



Conclusions

Automatic Generation of Restaurant Review:

- “Is that tool feasible?” **Yes! (we did it!)**
- “Is an artificial review considered genuine?” ≈ 1 on 4
- “Can an artificial review influence the human subject?” **Unclear, deeper experiments needed**



Conclusions

Automatic Generation of Restaurant Review:

- “Is that tool feasible?” **Yes! (we did it!)**
- “Is an artificial review considered genuine?” ≈ 1 on 4
- “Can an artificial review influence the human subject?” **Unclear, deeper experiments needed**

Machine generated reviews might become a real threat for (e-)commerce!



Thanks!

