



Engaging Content
Engaging People

The sentiment classification on indirect translation

Alberto Poncelas, Mohammad Aboomar,
Jan Buts, James Hadley, Andy Way



- Problem Statement
- Research Question
- Experiments
 - Classification on original sentences
 - Classification on direct translation
 - Classification on indirect translation



- Companies want to know what their client think.
- Analyze and classify their feedback (positive comments, complains, etc.)
- Categorize the feedback manually may be a laborious task.
- Alternatively, use an automatic classifier.



- Companies with international clients (hotel, restaurants, or software)
- Customer provide feedback in their native language.
- Most classifier (or data necessary for building it) are targeted for English language.



What are the solutions?

- 1) Build a classifier for each language?
 - How many classifier? What if we do not have data?
- 2) Translate into English, and classify.



Translate into English, and classify.

- Classifying the feedback in automatic translated sentences is less accurate:
 - Classifiers produce errors.
 - Translation models produce errors.
 - By combining both, we expect more errors.
- What if English is not my native language?
 - Translate the feedback twice?
 - Classify the feedback on my language.



RQ: How does influence the MT
in the classification?



- Languages: English, French, Spanish, Japanese.
- Train classifier in the native language.
- Compare the performance of the classifier in the native language and translated languages.
 - Direct Translation
 - Indirect Translation



Investigate the accuracy after performing translations. We perform the following experiments:

- Baseline: Classifier in the original language (L1)
- Direct: Classify on a MT-translated sentences (L1-EN).
- Indirect: Translate using a pivot language (L1-EN-L1).



- Languages: English, Spanish, Japanese, French.
- Domains: Hotels, restaurant and software.
- Classifier: LSTM classifier trained on:
 - English (3K feedbacks)
 - French (2K feedbacks)
 - Spanish (1.6K feedbacks)
 - Japanese (1.5K feedbacks)



Customer feedback classified as:

- Comment
 - *Software is good.*
- Complaint
 - *I hate it.*
- Request
 - *"where is notification icon in the notification bar?"*
- Bug
 - *"Got several confirmation emails"*
- Undetermined (not a feedback)
 - *Well...*



- Data is unbalanced
- Replicate less balanced for training.

	EN	FR	ES	JA
comment	1748	1232	1003	811
complaint	845	461	536	500
request	90	29	69	89
bug	53	39	13	80
undetermined	328	188	9	45

Accuracy in the original language:

LANGUAGE	ACCURACY
English	43%
French	47%
Spanish	62%
Japanese	41%



- Many errors between “undetermined” and “comment”.
- Example: “had to ask again .” is classified as undetermined (not a real review, and ambiguous to classify)
- Classified as comment by the automatic classifier.



- Languages: English, Spanish, Japanese, French.
- MT models: NMT model trained from English into:
 - French (51 M parallel sentences)
 - Spanish (39 M parallel sentences)
 - Japanese (8.7 M parallel sentences)

- EN → Language
- Evaluate FR, ES and JA model on the generated sentences.
- Different test sets.

LANGUAGE	ACCURACY (original sentences)	ACCURACY (translated)
French	47%	46%
Spanish	62%	37%
Japanese	41%	20%



- Mistakes between “undetermined” and “comment” are maintained.
- In French (and Spanish): same mistakes as the original.
- In Spanish: Classifier tend to use less “bug” and “undetermined” classes (as seen in training data).
- In Japanese: most of the sentences are classified as “undetermined”.



- Unlike Spanish and French, Japanese is more strict.
- Japanese MT-produced sentences are not strictly correct (but could be understandable by a native speaker).
 - Tend to keep unknown words as original (Example: “Facebook は instagram?”)

- Language → EN → Language
- Evaluate FR, ES and JA model on the generated sentences.

	LANG	EN	LANG
French	45%	43%	46%
Spanish	62%	63%	61%
Japanese	41%	43%	8%

- When translated into English the performance tend to maintain.
- Performance tend to decrease when translating into a more strict languages.



- How do they work in when combined.
- Use other languages as pivot.



END

