

Transfert de ressources sémantiques pour l'analyse de sentiments au niveau des aspects

Caroline Brun

Naver Labs Europe, 38240 Meylan, France

caroline.brun@naverlabs.com

RÉSUMÉ

Dans cet article, nous abordons le problème de la détection de la polarité pour l'analyse de sentiments au niveau des aspects dans un contexte bilingue : nous proposons d'adapter le composant de détection de polarité d'un système préexistant d'analyse de sentiments au niveau des aspects, très performant pour la tâche, et reposant sur l'utilisation de ressources sémantiques riches pour une langue donnée, à une langue sémantiquement moins richement dotée. L'idée sous-jacente est de réduire le besoin de supervision nécessaire à la construction des ressources sémantiques essentielles à notre système. À cette fin, la langue source, peu dotée, est traduite vers la langue cible, et les traductions parallèles sont ensuite alignées mot à mot. Les informations sémantiques riches sont alors extraites de la langue cible par le système de détection de polarité, et ces informations sont ensuite alignées vers la langue source. Nous présentons les différentes étapes de cette expérience, ainsi que l'évaluation finale. Nous concluons par quelques perspectives.

ABSTRACT

In this paper, we address the problem of automatic polarity detection in the context of Aspect Based Sentiment Analysis (ABSA), in a bilingual setting : we propose to adapt the polarity detection component of an existing high performing ABSA system, designed with rich semantic knowledge of a given language to a more poorly endowed language regarding semantics. The underlying idea is to minimize the level of supervision needed to develop the source language polarity component. To achieve this, we leverage on automatic translation of the ABSA source language into the ABSA target language, word alignment of these parallel translations, extraction of semantically-rich features on the ABSA target language, and back mapping of these target features onto the ABSA source language. The different steps of the design of the system are presented together with its evaluation. We conclude with some perspectives.

MOTS-CLÉS : Analyse de sentiments au niveau des aspects, ressources sémantiques, traduction.

KEYWORDS: Aspect-Based Sentiment Analysis, Semantic Resources, Translation.

1 Introduction

Les réseaux sociaux sont une source de grands volumes de contenus générés par les utilisateurs, dans lesquels les millions d'opinions exprimées par ces utilisateurs sont accessibles. Initialement, les travaux à dédiés l'analyse de sentiments se concentraient sur une analyse globale de la polarité d'un document. Cependant, les opinions sont rarement monodimensionnelles, mais le plus souvent

multi-dimensionnelles. L'analyse automatique des sentiments au niveau des aspects (Aspect Based Sentiment Analysis en anglais, ABSA), (Liu, 2012), (Ganu *et al.*, 2009), s'attelle précisément à extraire et résumer les opinions décrivant l'avis des utilisateurs sur des entités spécifiques et sur leurs aspects, c'est-à-dire les différentes caractéristiques qui les qualifient. Cette thématique, particulièrement attractive mais aussi complexe a été introduite dans le cadre du challenge SemEval en 2014 (Pontiki *et al.*, 2014), 2015 (Pontiki *et al.*, 2015) et 2016 (Pontiki *et al.*, 2016). Une des tâches du challenge consiste à analyser des revues phrase à phrase, afin de détecter les termes cibles du domaine sur lesquels une opinion est exprimée et de leur associer leurs aspects sémantiques et la polarité correspondante, en différentes langues et pour plusieurs domaines. Nous avons participé l'édition 2016 avec un système développé pour le français et l'anglais, et dont les résultats étaient très satisfaisants (Brun *et al.*, 2016). Dans le présent article, nous proposons de décrire une expérience visant à capitaliser sur le système préexistant de l'anglais, pour développer un composant de détection de la polarité pour une langue additionnelle, l'espagnol. Après la présentation de travaux connexes, nous décrivons la tâche du challenge SemEval2016 et le système que nous avons développé. Nous détaillons par la suite les différentes étapes de son adaptation à l'espagnol, ainsi que l'évaluation finale. Nous concluons par quelques perspectives de poursuite de ce travail.

2 Travaux connexes

La plupart des systèmes de détection des sentiments associés aux aspects utilisent des algorithmes d'apprentissage automatique tels que les "Support Vector Machines" (SVM), (Wagner *et al.*, 2014; Kiritchenko *et al.*, 2014), ou des "Conditionals Random Field" (CRF) (Toh & Wang, 2014; Hamdan *et al.*, 2015), souvent combinées à de l'information sémantique lexicale, des n-grams, des parties du discours (POS), et quelquefois de l'information syntactico-sémantique plus fine. Par exemple, la méthode proposée par (Kumar *et al.*, 2016) s'est avérée très performante sur les différentes langues du challenge SemEval2016. Le système inclut des informations extraites de graphes de dépendances et de thesaurus distributionnels appris pour les différentes langues et domaines, en plus d'informations lexicales dérivées de corpus non annotés. Les méthodes basées sur l'apprentissage profond commencent également à être appliquées avec succès à la tâche : une méthode basée sur les réseaux neuronaux convolutifs pour la classification en aspects et en polarité a obtenu des résultats intéressants pour plusieurs langues et domaines (Ruder *et al.*, 2016).

Un autre champ d'étude, l'analyse multilingue de sentiments, est également pertinent pour le présent article. En effet, certains travaux se sont concentrés sur l'adaptation de ressources associées aux sentiments, comme des lexiques, à partir de langues richement dotées (typiquement l'anglais) vers des langues moins dotées. L'article de (Mihalcea *et al.*, 2007) présente une méthode détectant la subjectivité dans des contextes multilingues en utilisant des lexiques bilingues et des projections cross-lingues. La méthode proposée par (Banea *et al.*, 2008) a permis d'obtenir des lexiques de subjectivité pour des langues autres que l'anglais en traduisant automatiquement le corpus MPQA¹, pour finalement obtenir des phrases annotées en roumain. En se basant sur ces travaux, (Wan, 2009) a proposé une méthode basée sur un algorithme de "co-training" pour analyser les sentiments de textes chinois et anglais en mode cross-lingue. En outre, (Balahur & Turchi, 2012) et (Balahur *et al.*, 2014) ont montré que, en utilisant des données d'entraînement obtenues par traduction automatique, ils pouvaient développer des classifieurs statistiques catégorisant la polarité des tweets en plusieurs langues. Ces différents travaux montrent que la traduction automatique se montre utile pour la classification

1. http://mpqa.cs.pitt.edu/corpora/mpqa_corpus/

des sentiments pour des langues peu dotées en ressources lexico-sémantiques. Cependant, à notre connaissance, la traduction automatique n'a pas été utilisée dans le cadre de l'analyse de sentiments au niveau des aspects, ce qui est l'objet de l'expérience décrite dans le présent article.

3 Analyse de sentiments au niveau des aspects

L'analyse de sentiments au niveau des aspects est une tâche visant à capturer les sentiments exprimés sur différentes entités du monde, tels que des produits, des restaurants, des films, ou même des personnes, dans des contenus générés sur des médias sociaux. Les aspects sont des attributs de ces entités, par exemple le service dans un restaurant ou la qualité de l'écran d'un téléphone portable, généralement décrits sous forme d'une ontologie. Un système de détection des sentiments au niveau des aspects identifie l'entité sur laquelle une opinion s'exprime, qualifie la nature de l'aspect auquel il appartient et lui associe la polarité correspondante. La tâche d'analyse de sentiments au niveau des aspects de SemEval2016 (Pontiki *et al.*, 2016) était ainsi décomposée en trois sous-tâches : Phase A.1 : détection des termes sur lesquels une opinion s'exprime (terme cible) ; Phase A.2 : association de la catégorie sémantique de l'aspect associé au terme cible ; Phase B : association de la polarité (positif, négatif ou neutre) au couple <terme cible, catégorie de l'aspect>. Voici un exemple d'annotation pour l'anglais, dans le domaine des revues de restaurant :

```
<text>The food is fantastic and the waiting staff has been perfect.</text>
<Opinions>
<Opinion target="food" category="FOOD#QUALITY" polarity="positive" from="4" to="8"/>
<Opinion target="waiting staff" category="SERVICE#GENERAL" polarity="positive" from="31" to="44"/>
</Opinions>
```

Nous nous sommes intéressés au domaine des revues de restaurant, pour lequel les langues couvertes étaient l'anglais, le français, l'espagnol, le hollandais, le turque et le russe. Nous avons développé un système couvrant l'anglais et le français. Pour ce domaine, l'ontologie des aspects comprend 12 classes : *food#quality*, *food#style_options*, *food#prices*, *drink#quality*, *drink#style_options*, *drink#prices*, *service#general*, *ambience#general*, *location#general*, *restaurant#general*, *restaurant#prices* et *restaurant#misc* ; la polarité comprend 3 classes : positif, négatif, ou neutre. Nous nous intéressons ici au composant de détection de la polarité (phase B du challenge, pour laquelle l'annotation de référence en termes et en aspect est fournie), que nous transférons de l'anglais à l'espagnol. Le système initial que nous avons développé se base sur une méthode hybride combinant de l'information lexicale, syntaxique et sémantique à de la classification supervisée, (Brun *et al.*, 2016). Les traits linguistiques sont extraits par un analyseur syntaxique robuste et renseignent les différents algorithmes d'apprentissage. La détection des termes cibles des opinions est effectuée à l'aide d'un CRF (Lafferty *et al.*, 2001). La classification des aspects et des polarités associées aux termes sont réalisées par des modèles de régression logistique par ensemble. Une couche d'extraction de dépendances sémantiques dédiées à l'analyse de sentiments a été développée manuellement en aval de l'analyseur syntaxique, afin d'extraire des relations connectant prédicats de polarités et mots cibles des opinions, ce qui est illustré ci-dessous :

The food is fantastic and the waiting staff has been perfect.

Dépendances sémantiques : *sentim_posit(food,fantastic)*, *sentim_posit(waiting staff, perfect)*.

Ces dépendances sémantiques combinent les dépendances syntaxiques à des informations lexicales

concernant les schémas argumentaux, la polarité et les aspects. Elles sont utilisées comme traits pour la classification en polarité, et se montrent particulièrement discriminantes. Une validation croisée à 10 tours sur le corpus d’entraînement (environ 2000 phrases) a montré que leur introduction améliore d’environ 10 points l’exactitude de la classification en polarité. Lors de l’évaluation finale, le système s’est révélé très performant pour cette sous-tâche de SemEval2016, puisqu’il obtient les meilleurs résultats pour le français et l’anglais. Nous reportons les résultats obtenus par notre système (NLE) sur les différentes sous-tâches dans le tableau 1.

Tâches	LANGUE : EN				LANGUE : FR			
	A.1 (F1)	A.2 (F1)	A.12 (F1)	B (Acc)	A.1 (F1)	A.2 (F1)	A.12 (F1)	B (Acc)
Baseline	59.93	44.07	37.80	76.48	52.61	45.45	33.02	67.4
NLE	68.70	61.98	48.89	88.13	61.21	65.32	47.72	78.83
Meilleur système	73.03	72.34	52.61	88.13	61.21	66.67	47.72	78.83

TABLE 1 – Résultats de Semeval2016 (domaine des restaurants) : A.1=détection des aspects, A.2=détection des termes, A1.2=détection des couples <terme,aspect>, B=détection des polarités

Cependant, si ces dépendances sémantiques sont discriminantes pour la classification en polarité, ce sont des ressources manuelles coûteuses à développer, qui nécessitent un degré important de supervision. Leur adaptation à une autre langue s’avèrerait donc coûteuse également. Dans la suite de cet article, nous présentons une expérience visant à automatiquement transférer les dépendances sémantiques de l’anglais vers une autre langue, l’espagnol, pour laquelle nous disposons uniquement d’un analyseur syntaxique générique mais pas de ressources lexicales de polarité et d’aspects, ni de dépendances sémantiques de sentiment.

4 Expérience de transfert bilingue de traits sémantiques

Nous présentons ici notre expérience de transfert de traits issus des dépendances sémantiques de l’anglais ($L_{absa-target}$) vers l’espagnol ($L_{absa-source}$). Pour cela, nous utilisons la traduction automatique du corpus d’entraînement de l’espagnol vers l’anglais, le composant d’extraction de traits linguistiques de l’anglais présenté dans la section précédente, et une méthode d’appariement de ces traits vers la représentation vectorielle de l’espagnol pour la classification en polarité. La méthode d’appariement se base sur l’alignement de mots phrase à phrase. La figure 1 décrit ce processus.

4.1 Traduction des corpus

Afin de traduire automatiquement les corpus d’entraînement et de test de l’espagnol vers l’anglais, nous utilisons des outils génériques de traduction automatique disponibles sur internet. Nous avons ainsi utilisé « Google Translate »² ainsi que l’API du traducteur de Microsoft³, afin de comparer

2. <https://translate.google.com/>

3. <https://www.microsoft.com/en-us/translator/translatorapi.aspx>

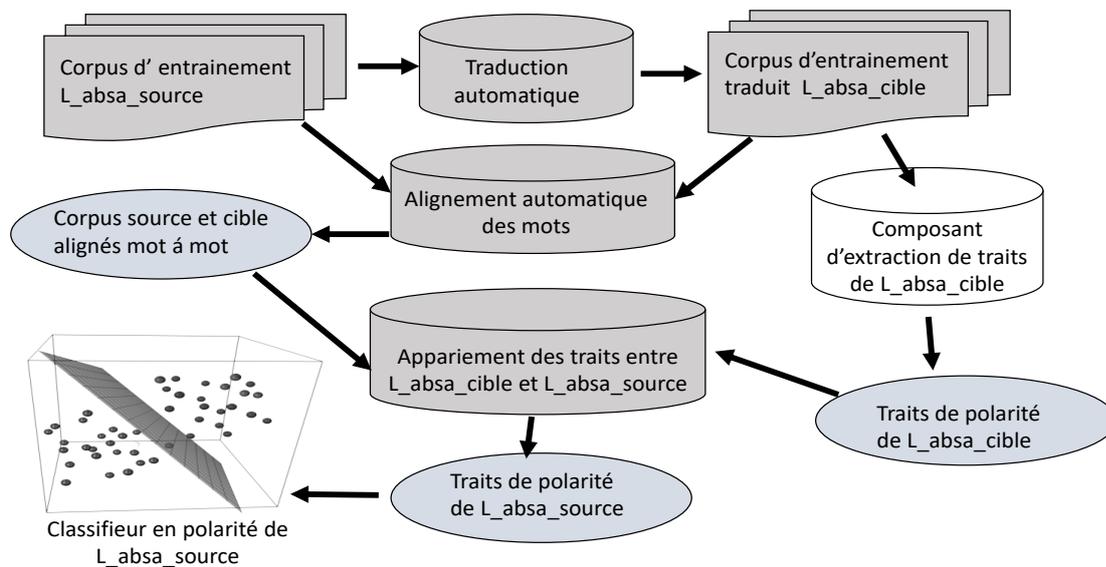


FIGURE 1 – Extraction des traits sémantiques pour $L_{absa-source}$

l'impact de la traduction sur notre méthode. Ces outils sont complexes à évaluer dans le contexte de notre expérience, puisque nous ne disposons pas d'une traduction de référence, cependant nous les avons choisis car leur performances sont généralement très compétitives, (Isabelle *et al.*, 2017), (Jimeno Yepes *et al.*, 2017). La liste des phrases de l'espagnol $L_{absa-source}$ alignées avec leur traductions automatiques vers l'anglais est le résultat de cette étape.

4.2 Alignement des mots

A ce stade, le corpus d'entraînement de l'espagnol est aligné phrase à phrase avec sa traduction en anglais. Afin de pouvoir apparier les traits sémantiques qui vont être extraits de la traduction anglaise par notre composant d'extraction linguistique développé pour l'anglais, nous devons aligner mot à mot chacune des phrases de ces deux corpus parallèles. Plusieurs outils « open-source » d'alignement de mots à partir de corpus parallèles sont disponibles, par exemple les outils décrits dans (Och & Ney, 2003), ou encore (Germann, 2008). Nous avons choisi *fast align* (Dyer *et al.*, 2013)⁴ pour sa simplicité d'utilisation. Cet outil prend en entrée des paires de phrases alignées et retourne les paires d'indices correspondant aux indices des mots alignés. Comme *fast align* est purement non supervisé, en plus des corpus parallèles dont nous disposons, nous lui avons également fourni des paires de mots du domaine (i.e. les revues de restaurants), c'est-à-dire des termes culinaires en espagnol et leur traductions anglaises, extraits de <http://www.linguasorb.com/Spanish/food-word-list>.

4.3 Appariement des traits sémantiques

Dans un premier temps, le système développé pour l'anglais est appliqué sur le corpus parallèle anglais traduit du corpus d'entraînement de l'espagnol : ce système annote les termes cibles des

4. https://github.com/clab/fast_align

opinions, leurs aspects et leurs polarités. Ce système utilise comme traits de décision les dépendances sémantiques que nous voulons transférer à l’espagnol. Ces traits peuvent être associés à des termes cibles des opinions ou à des phrases entières. D’autre part, le corpus d’entraînement de l’espagnol contient les annotations des termes cibles des opinions. L’alignement des mots présentés dans la section précédente nous permet d’aligner les termes cibles des opinions annotés dans le corpus d’entraînement espagnol avec les termes cibles des opinions extraites par notre système sur la traduction parallèle en anglais. L’appariement est réalisé en maximisant l’intersection des indices des mots contenus dans les termes source et cible. Suite à cet appariement des termes source et cible, les traits sémantiques associés aux termes anglais peuvent être directement transférés aux termes appariés en espagnol. Ils sont alors concaténés au vecteur de traits « baseline » dont nous disposons pour l’espagnol, lors de l’entraînement et de la phase de prédiction. Un exemple étape par étape est présenté sur le tableau 2.

espagnol	<i>El servicio es muy bueno y la calidad de la comida al mismo nivel</i>
traduction anglaise	<i>The service is very good and the quality of food at the same level</i>
alignement des mots	0-0 1-1 2-2 3-3 4-4 5-5 6-6 7-7 8-8 10-9 11-10 9-11 12-12 13-13 14-14
appariement des termes	<servicio, service>, <comida, food>
traits sémantiques anglais	sentim_posit(service, good) sentim_posit(food)
transfert vers l’espagnol	sentim_posit(servicio, bueno) sentim_posit(comida)

TABLE 2 – Exemple de transfert de traits

5 Evaluation

Le système développé pour l’espagnol et intégrant les traits transférés a été évalué selon le protocole et avec les données de l’espagnol de SemEval 2016. Le corpus d’entraînement comprend 2070 phrases et 2720 opinions exprimées (70.8% positives, 24.7% négatives et 4.5% neutres), le corpus de test comprend 881 phrases et 1072 opinions exprimées (69.9% positives, 25.7% négatives et 4.4% neutres). L’utilisation des traductions obtenues avec « Google Translate » (ABSA-GTR) et avec le traducteur de Microsoft (ABSA-MTR) permet d’évaluer l’impact de la traduction sur le système dans son ensemble. Un premier couple de « baselines » est également obtenu en appliquant le système complet développé pour l’anglais sur les traductions du jeu de test, puis en associant directement les polarités obtenues à la version espagnole via l’appariement des termes, pour les deux outils de traduction, ce qui donne les résultats de BL-GTR et BL-MTR. Une troisième résultat « baseline », ABSA-BL, applique directement le système minimal de l’espagnol en utilisant uniquement les traits standard (sac de mots, n-grams, dépendances syntaxiques générales). Les résultats en terme d’exactitude sont présentés sur le tableau 3, avec les résultats de référence de SemEval2016 (« baseline » : SE16-BL et meilleur système : SE16-BEST).

Système	BL-GTR	BL-MTR	ABSA-GTR	ABSA-MTR	ABSA-BL	SE16-BL	SE16-BEST
Exactitude	66.51	65.20	84.23	<u>84.04</u>	77.42	77.79	83.58

TABLE 3 – Exactitude des différents systèmes pour la détection de la polarité en espagnol

Les résultats basés sur les différents systèmes de traductions sont très proches, que ce soit pour les « baseline » ou les modèles utilisant le transfert de traits sémantique. Les résultats de ABSA-BL, système entraîné sur l’espagnol avec le jeu de traits standard, sont alignés avec la « baseline » de SemEval2016. Les « baselines » basées uniquement sur les modèle de l’anglais ont des résultats bien en deçà des autres systèmes, ce qui tend à justifier notre approche de transfert de traits. En effet, globalement le gain en performance obtenu par le transfert des traits sémantiques est important : les deux systèmes améliorent légèrement les résultats du meilleur système de SemEval2016 pour l’espagnol, (Kumar *et al.*, 2016), alors classé premier sur quatre participants. Les résultats détaillés par polarité pour ABSA-BL, ABSA-MTR et ABSA-GTR sont présentés sur le tableau 4.

Étiquette	ABSA-BL			ABSA-MTR			ABSA-GTR		
	P	R	F1	P	R	F1	P	R	F1
positive	0.78	0.94	0.85	0.85	0.94	0.89	0.86	0.94	0.90
négative	0.71	0.42	0.53	0.79	0.69	0.74	0.77	0.70	0.74
neutre	0.75	0.06	0.11	0.66	0.04	0.08	0.75	0.06	0.11

TABLE 4 – Résultats par polarité pour ABSA-BL, ABSA-MTR et ABSA-GTR

Les corpus étant fortement biaisés vers les opinions positives, le rappel et la précision sont importants pour cette étiquette. L’introduction via traduction de traits sémantiques relatifs à la polarité contribue donc significativement à l’amélioration des résultats, en particulier pour les polarités négatives. Nous prévoyons de réaliser une analyse des erreurs pour quantifier plus précisément, entre autres, celles qui sont dues à la qualité des traductions, à la qualité de l’alignement ou bien au système d’extraction de traits de l’anglais.

6 Conclusion et perspectives

Nous avons présenté dans cet article une expérience d’adaptation d’un système d’une langue (l’anglais), dotée de ressources sémantiques riches, vers une autre langue (l’espagnol), pour laquelle nous ne disposons a priori pas de ressources sémantiques dédiées à la tâche. À cette fin, nous avons développé une méthode de transfert de traits sémantiques de la langue richement dotée vers l’autre langue, méthode basée sur la traduction automatique du corpus annotés de la langue source vers la langue cible. Le système d’analyse de sentiments au niveau des aspects de la langue cible peut alors être appliqué afin d’extraire des traits sémantiques qui sont ensuite appariés par alignement puis transférés vers la langue source. Ces traits sont exploitables pour l’apprentissage du modèle d’annotation en polarité de la langue source par un pipeline similaire à celui de la langue initiale. L’évaluation des résultats est très prometteuse puisqu’elle montre une amélioration des meilleurs résultats de SemEval2016 pour l’espagnol. Par la suite, nous souhaitons appliquer une méthode similaire pour les autres langues de la tâche d’analyse de sentiments au niveau des aspects de SemEval16, à savoir le hollandais, le turque et le russe. Une autre perspective intéressante est l’investigation de méthodes de transfert d’apprentissage, comme celle proposée par (Zhou *et al.*, 2016) et de l’adapter à la tâche d’analyse de sentiments au niveau des aspects.

Références

- BALAHUR A. & TURCHI M. (2012). Multilingual sentiment analysis using machine translation ? In *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*, WASSA '12, p. 52–60, Stroudsburg, PA, USA : Association for Computational Linguistics.
- BALAHUR A., TURCHI M., STEINBERGER R., PEREA-ORTEGA J. M., JACQUET G., KUCUK D., ZAVARELLA V. & GHALI A. E. (2014). Resource creation and evaluation for multilingual sentiment analysis in social media texts. In N. C. C. CHAIR), K. CHOUKRI, T. DECLERCK, H. LOFTSSON, B. MAEGAARD, J. MARIANI, A. MORENO, J. ODIJK & S. PIPERIDIS, Eds., *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland : European Language Resources Association (ELRA).
- BANEA C., MIHALCEA R. & WIEBE J. (2008). A bootstrapping method for building subjectivity lexicons for languages with scarce resources. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC-08)*, Marrakech, Morocco : European Language Resources Association (ELRA). ACL Anthology Identifier : L08-1086.
- BRUN C., PEREZ J. & ROUX C. (2016). XRCE at semeval-2016 task 5 : Feedbacked ensemble modeling on syntactico-semantic knowledge for aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2016, San Diego, CA, USA, June 16-17, 2016*, p. 277–281.
- DYER C., CHAHUNEAU V. & SMITH N. A. (2013). A simple, fast, and effective reparameterization of ibm model 2. In *Proceedings of NAACL*, Atlanta : Association for Computational Linguistics.
- GANU G., ELHADAD N. & MARIAN A. (2009). Beyond the stars : Improving rating predictions using review text content. In *Proceedings of the 12th International Workshop on the Web and Databases*, Providence, Rhode Island.
- GERMANN U. (2008). Yawat : Yet another word alignment tool. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies : Demo Session, HLT-Demonstrations '08*, p. 20–23, Stroudsburg, PA, USA : Association for Computational Linguistics.
- HAMDAN H., BELLOT P. & BECHET F. (2015). Lsislif : Crf and logistic regression for opinion target extraction and sentiment polarity analysis. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, p. 753–758, Denver, Colorado : Association for Computational Linguistics.
- ISABELLE P., CHERRY C. & FOSTER G. F. (2017). A challenge set approach to evaluating machine translation. *CoRR*, **abs/1704.07431**.
- JIMENO YEPES A., NEVEOL A., NEVES M., VERSPOOR K., BOJAR O., BOYER A., GROZEA C., HADDOW B., KITTNER M., LICHTBLAU Y., PECINA P., ROLLER R., ROSA R., SIU A., THOMAS P. & TRESCHER S. (2017). Findings of the wmt 2017 biomedical translation shared task. In *Proceedings of the Second Conference on Machine Translation*, p. 234–247, Copenhagen, Denmark : Association for Computational Linguistics.
- KIRITCHENKO S., ZHU X., CHERRY C. & MOHAMMAD S. (2014). Nrc-canada-2014 : Detecting aspects and sentiment in customer reviews. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, p. 437–442, Dublin, Ireland : Association for Computational Linguistics and Dublin City University.
- KUMAR A., KOHAIL S., KUMAR A., EKBAL A. & BIEMANN C. (2016). Iit-tuda at semeval-2016 task 5 : Beyond sentiment lexicon : Combining domain dependency and distributional semantics

- features for aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, p. 1129–1135, San Diego, California : Association for Computational Linguistics.
- LAFFERTY J. D., MCCALLUM A. & PEREIRA F. C. N. (2001). Conditional random fields : Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning, ICML '01*, p. 282–289, San Francisco, CA, USA : Morgan Kaufmann Publishers Inc.
- LIU B. (2012). *Sentiment Analysis and Opinion Mining*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.
- MIHALCEA R., BANEA C. & WIEBE J. (2007). Learning multilingual subjective language via cross-lingual projections. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, p. 976–983, Prague, Czech Republic : Association for Computational Linguistics.
- OCH F. J. & NEY H. (2003). A systematic comparison of various statistical alignment models. *Computational Linguistics*, **29**(1), 19–51.
- PONTIKI M., GALANIS D., PAPAGEORGIOU H., ANDROUTSOPOULOS I., MANANDHAR S., AL-SMADI M., AL-AYYOUB M., ZHAO Y., QIN B., CLERCQ O. D., HOSTE V., APIDIANAKI M., TANNIER X., LOUKACHEVITCH N., KOTELNIKOV E., BEL N., JIMÉNEZ-ZAFRA S. M. & ERYİĞİT G. (2016). SemEval-2016 task 5 : Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval '16*, San Diego, California : Association for Computational Linguistics.
- PONTIKI M., GALANIS D., PAPAGEORGIOU H., MANANDHAR S. & ANDROUTSOPOULOS I. (2015). Semeval-2015 task 12 : Aspect based sentiment analysis. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, p. 486–495, Denver, Colorado : Association for Computational Linguistics.
- PONTIKI M., GALANIS D., PAVLOPOULOS J., PAPAGEORGIOU H., ANDROUTSOPOULOS I. & MANANDHAR S. (2014). Semeval-2014 task 4 : Aspect based sentiment analysis. In *International Workshop on Semantic Evaluation (SemEval)*.
- RUDER S., GHAFARI P. & BRESLIN J. G. (2016). INSIGHT-1 at SemEval-2016 Task 5 : Deep Learning for Multilingual Aspect-based Sentiment Analysis. *ArXiv e-prints*.
- TOH Z. & WANG W. (2014). Dlirec : Aspect term extraction and term polarity classification system. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, p. 235–240, Dublin, Ireland : Association for Computational Linguistics and Dublin City University.
- WAGNER J., ARORA P., CORTES S., BARMAN U., BOGDANOVA D., FOSTER J. & TOUNSI L. (2014). Dcu : Aspect-based polarity classification for semeval task 4. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, p. 392–397, Dublin, Ireland : Association for Computational Linguistics and Dublin City University.
- WAN X. (2009). Co-training for cross-lingual sentiment classification. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP : Volume 1 - Volume 1*, ACL '09, p. 235–243, Stroudsburg, PA, USA : Association for Computational Linguistics.
- ZHOU G., ZENG Z., HUANG J. X. & HE T. (2016). Transfer learning for cross-lingual sentiment classification with weakly shared deep neural networks. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '16*, p. 245–254, New York, NY, USA : ACM.

