UBO Team @ CLEF JOKER 2023 Track For Task 1, 2 and 3 - Applying AI Models In Regards To Pun Translation

Quentin Dubreuil[1]

[1]*Université de Bretagne Occidentale, Brest, France*

Abstract

Understanding and translating humorous wordplay poses challenges for both humans and computers due to the need for recognizing implicit cultural references, comprehending word formation processes, and interpreting multiple meanings. Therefore we will try to translate pun from English to French or Spanish within the JOKER track. This abstract will provide an overview of the subject and the different tasks involved. It will then delve into a description of each model employed, followed by an examination of the results obtained. Finally, a conclusion will be drawn based on the findings.

Keywords

CLEF, JOKER, SimpleT5, EasyNMT

1. Introduction

Humor is a universal form of communication that brings joy, laughter, and amusement to our lives. But when it comes to translating them, it becomes a delicate task of balancing the original humor with a new language and culture, whether it is humanly or automatically done. The objective of the JOKER [1, 2] track series at the Conference and Labs of the Evaluation Forum (CLEF)[3] is to unite linguists, translators, and computer scientists. The final objective of the JOKER project is to be able to translate jokes from a language to another. For JOKER-2023, Spanish has been added to the set of languages in the tasks, and is divided in 3 tasks:

- Task 1 : Detection of puns in English, French, and Spanish
- Task 2 : Location and Interpretation of puns in English, French, and Spanish
- Task 3 : Translation of puns from English to French and Spanish

For the task 1 and 2, we use the SimpleT5[4] model, because it was our first experience with training a model and it is an easy model to use, furthermore, many people already use ChatGPT, so we wanted to have a different approach. Whereas for the task 3, we used the 3 models of the EasyNMT [5] package :

- Opus-mt [6]
- mBart50_m2m [7]
- M2M_100 [8]

2. Approach

In the case of the SimpleT5 model, our process involves training the model using the training data and subsequently performing inference on the test data. However, for the three models of EasyNMT, they are pre-trained, so we only need to utilize these models directly on the test data without any additional training steps.

2.1. Use of SimpleT5 on the first two tasks

The training part was divided into two parts, reading the data and formating it, and then the training of the model. For both tasks, we had 3 files :

- An input file with "id" and "text".
- A qrels file with "text" and "wordplay" for the first task and "text", "location" and "interpretation" for the second task.
- A test file with "id" and "text".

Table 1: Example of the 2 columns needed for the SimpleT5 model

We had to format the data to make two fields in a dataframe for our model, the "source_text" (train) and the "target_text" (test). The "source_text" in both tasks was the "text" field, and the "target_text" field was the "wordplay" field for the task 1 while for task 2.1, it was the "location" field and for the task 2.2, it was the "interpretation" field.

"source_text"	"target_text"
At a flower show the first prize is often a bloom ribbon.	*yes*

Once we have our dataframe with the formated data, we can start the training of the SimpleT5 model, we first load the model, then we give this dataframe to our model. Once the training is finished, we do the same thing that we did before for the "source_text" and "target_text" but this time we only put the "text" in the "source_text", and the model will infer the result in the "target_text". Then we get what was generated in the "target_text" field and we format it with the required fields

2.2. EasyNMT package

The main difference between the SimpleT5 model and the 3 models of the EasyNMT package is, that they are already trained, we just need to load the model and use it on our test data.

For this task, we have only done it for the English -> French part, we have 2 files :

- An input file with "id_en" and "text_en"
- A qrels file with "id_en" and "text_fr"

Instead of training our models, we utilized the "apply" method on our dataframe to feed the "text_en" field to each of the models individually, retrieving the corresponding translations. Once we collected all the translations, we organized them according to the required fields, similar to how we formatted the data for the first two tasks.

3. Results

3.1. Task 1

In terms of the English corpus, our method achieved an F1 score of 0.51, indicating a moderate performance that neither falls into the category of being terrible nor exceptionally good. Despite not being the best approach, we were able to obtain some favorable results, occasionally discerning whether a sentence contained a wordplay. Moving on to the Spanish corpus, we obtained a slightly improved F1 score of 0.56 compared to English. Although the improvement was not significant, it suggests that we still achieved some noteworthy results in identifying wordplay, though not the optimal performance. The French corpus yielded the highest F1 score of 0.63 among all three languages. In this case, we were relatively successful in accurately determining whether a sentence contained wordplay on a frequent basis.

3.2. Task 2

3.2.1. Task 2.1

The results obtained from the English corpus were relatively satisfactory, with an accuracy of 0.77. We were able to successfully identify the intended location of the wordplay in most cases. In contrast, the results from the Spanish corpus were not as promising compared to the English corpus, yielding an accuracy of 0.57. However, the results from the French corpus were particularly inconclusive and the least successful among the three languages, exhibiting an accuracy of only 0.40.

3.2.2. Task 2.2

Regarding the English corpus, we achieved somewhat satisfactory results, with a majority of cases yielding both meanings of the word. Nevertheless, even when we managed to retrieve two possible meanings for the pun, they rarely aligned with the intended or expected meanings, indicating a significant deficiency in overall performance. On the other hand, when applying the same methodology to the French corpus, the results were significantly unfavorable, prompting us to disregard them altogether.

3.3. Task 3

The translation models, namely Opus-MT, mBART50_m2m, and M2M_100, exhibit comparable results when translating pun words from English to French, managing to maintain the meaning to a satisfactory extent (approximately 50% of the time). However, when it comes to preserving the pun itself, the success rate drops significantly to only 7%. This indicates that the employed models were not proficient in accurately translating pun words from English to French.

4. Conclusion

The SimpleT5 model demonstrated its effectiveness for task 1. However, when applied to task 2.1, it exhibited satisfactory performance only on the English corpus, while encountering limitations

when working with the Spanish and French corpora. As for task 2.2, the model's performance was notably poor, failing to provide accurate meanings for the puns.

However, in the case of task 3, the three models employed proved to be highly ineffective for English to French translation. This can primarily be attributed to the fact that these models are designed to primarily translate text between languages and are not specifically made to handle puns or prioritize maintaining their intended meaning.

To summarize, it is crucial to have high-quality training data created manually by linguists for the JOKER track in order to effectively train the models. Without such data, training the models effectively becomes challenging, as evidenced by the performance of the three models utilized in task 3.

5. Acknowledgments

References

[1] L. Ermakova, T. Miller, A.-G. Bosser, V. M. Palma Preciado, G. Sidorov, A. Jatowt, Science for Fun: The CLEF 2023 JOKER Track on Automatic Wordplay Analysis, in: J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, A. Caputo (Eds.), Advances in Information Retrieval, volume 13982, Springer Nature Switzerland, Cham, 2023, pp. 546–556. URL: https://link.springer.com/10.1007/978-3-031-28241-6_63. doi:10.1007/978-3-031-28241-6_63, series Title: Lecture Notes in Computer Science.

[2] L. Ermakova, T. Miller, F. Regattin, A.-G. Bosser, C. Borg, Mathurin, G. Le Corre, S. Araújo, R. Hannachi, J. Boccou, A. Digue, A. Damoy, B. Jeanjean, Overview of JOKER@CLEF 2022: Automatic Wordplay and Humour Translation Workshop, in: A. Barrón-Cedeño, G. Da San Martino, M. Degli Esposti, F. Sebastiani, C. Macdonald, G. Pasi, A. Hanbury, M. Potthast, G. Faggioli, N. Ferro (Eds.), Experimental IR Meets Multilinguality, Multimodality, and Interaction, volume 13390, Springer International Publishing, Cham, 2022, pp. 447–469. URL: https://link.springer.com/10.1007/978-3-031-13643-6_27. doi:10.1007/978-3-031-13643-6_27, series Title: Lecture Notes in Computer Science.

[3] T. T. S. V. A. G. D. L. M. A. M. V. G. F. N. F. Avi Arampatzis, Evangelos Kanoulas (Ed.), Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the Fourteenth International Conference of the CLEF Association (CLEF 2023), 2023.

[4] simpleT5, 2021. URL: https://github.com/Shivanandroy/simpleT5.

[5] EasyNMT, 2020. URL: https://github.com/UKPLab/EasyNMT.

[6] J. Tiedemann, S. Thottingal, OPUS-MT – Building open translation services for the World, in: Proceedings of the 22nd Annual Conference of the European Association for Machine Translation, European Association for Machine Translation, Lisboa, Portugal, 2020, pp. 479–480. URL: https://aclanthology.org/2020.eamt-1.61.

[7] Y. Tang, C. Tran, X. Li, P.-J. Chen, N. Goyal, V. Chaudhary, J. Gu, A. Fan, Multilingual Translation with Extensible Multilingual Pretraining and Finetuning, 2020. URL: http://arxiv.org/abs/2008.00401, arXiv:2008.00401 [cs].

[8] A. Fan, S. Bhosale, H. Schwenk, Z. Ma, A. El-Kishky, S. Goyal, M. Baines, O. Celebi, G. Wenzek, V. Chaudhary, N. Goyal, T. Birch, V. Liptchinsky, S. Edunov, E. Grave, M. Auli, A. Joulin, Beyond English-Centric Multilingual Machine Translation, 2020. URL: http://arxiv.org/abs/2010.11125, arXiv:2010.11125 [cs].

Appendices

A. Outputs example for Task 1

id	text	wordplay
en_4512	What is a lumberjack's favorite month? October !	no
en_2280	The Hong Kong businessman left a huge estate upon his death. This was China's great desire.	no
en_170	Old seamstresses never die, they just hide.	yes
en_3789	I have written several articles about sound effects. I have them all saved in the cloud.	yes
en_7608	When they bought a water bed, the couple started to drift apart.	yes
en_661	OLD BREAD MEN never die, they just lose their dough.	yes
en_5607	She was only a Fruit vendor's daughter, but, my, she had big melons.	yes
en_5205	Light at end of tunnel temporarily out of order.	no
en_6157	He crashed through several windows, but felt no pane.	yes
en_2572	Corn is so versatile that it is an a-maize-ing grain.	yes
en_1719	A dentist has to tell a patient the whole truth.	no
en_4362	Doctor, Doctor, I feel like a pair of curtains - Well, get a grip. Next.	yes
en_286	Please mix me a martini, said Tom, thirstily.	no
en_7889	Those who get up at sunrise have many ideas dawn on them.	yes
en_7051	Today is the first day of the rest of your life.	no
en_1311	Are giraffes good pets or are they too high maintenance?	no
en_3503	'Angel dust? Me? Never touch it!" Tom snorted.	yes
en_7020	I drove a shark crazy. He was titillated.	yes
en_1265	Yesterday's high was 78 and the low was 28, but those are not records.	no
en_7080	'I work at a bank," said Tom quietly.	no
en_2869	Kill them all! .... Let God sort them out.	no

B. Outputs example for Task 2.1

id	text	location
en_6889	Soft drink inventors saw a big popportunity.	popportunity
en_2277	I never found sending signals from ships challenging. I always had a flare for it.	flare
en_8363	OLD TELEPHONE OPERATORS never die, they just become disconnected.	disconnected
en_2354	When his wife asked for wooden walls in the basement, they had a panel discussion.	panel
en_6197	The sailor called his girlfriend's belly ring a naval destroyer.	naval
en_4290	'I prefer trout to salmon," Tom said officiously.	officiously
en_5207	There are many judges who would like to acquit smoking.	acquit
en_5162	I went to a buffet dinner with my neighbor, who is a taxidermist. After such a big meal, I was stuffed.	stuffed
en_1771	The unveiling of the statue was a monumental occasion.	monumental
en_5448	The rodent catcher was always trying to weasel his way out of work.	weasel
en_5781	OLD CREDIT CARD USERS never die, they just expire.	expire
en_3999	Wanting to buy his girlfriend a gift that she was sure to appreciate, Bugs Bunny decided that something on the order of twenty - four carats would be about right.	carats
en_6414	We welcome you with open psalms.	psalms
en_7818	The woman who rode her broom to a croquet game was a wicket witch.	wicket
en_4385	Did you hear about the vampire who used to torture his victims with music? His Bach was worse than his bite.	Bach
en_7889	Those who get up at sunrise have many ideas dawn on them.	dawn
en_3918	OLD GEOMETRY TEACHERS never die, they just go off on a tangent.	tangent
en_3035	She was only a Poet's daughter, but I've seen verse.	verse
en_8380	I phoned the zoo but the lion was busy.	lion

C. Outputs example for Task 2.2

id	location	interpretation
en_6889	popportunity	pop culture / pop;opportunity
en_2277	flare	flare up;blow up / flare
en_8363	disconnected	disconnect;disconnect / disconnect
en_2354	panel	panel of glass;panel of glass / piece;part
en_6197	naval	navire;ship / naval
en_4290	officiously	officiously / officiously
en_5207	cquit	culpable;cquitted of all charges;conviction;cquittal / clear
en_5162	stuffed	stuffed / fill
en_1771	monumental	monumental / monumental
en_5448	weasel	animal / weasel
en_5781	expire	expire / end;end
en_3999	carats	carat / cart
en_6414	psalms	Psalm / praise
en_7818	wicket	cricket equipment / wicket
en_4385	Bach	music / proper name
en_7889	dawn	saturday morning;dawn / dawn
en_3918	tangent	tangent / tangent
en_3035	verse	poesy;poetry / verse
en_8380	lion	Lion of Africa / light
en_2418	cooked	prepare;cook;make / cook

D. Outputs example for Task 3

id_en	text_en	text_fr
en_1007	Save the whales, spouted Tom.	"Il faut sauver les baleines", jeta Tom avant de se tasser.
en_1015	If a town's people have low IQs is the population dense?	Si un citadin a un faible QI, est-ce pour autant que la population niais pas pour quelque chose ?
en_102	A skier retired because he was going downhill.	Le skieur est parti à la retraite. Il n'arrivait pas à remonter la pente.
en_103	My wife uses a kitchen implement to shred garlic and parmesan cheese, which I hate. It really is the grater of two evils.	Ma femme écoute du hip hop quand elle cuisine des carottes ou du gruyère. Je n'aime pas ça mais elle me dit que ça l'aide à râper.
en_1031	Staying at the trendy, new hotel was the inn thing to do.	Je rêvais de dormir dans cet hôtel.
en_1033	While baking, I dropped a stick of margarine on the wooden tile floor, and when my neighbor slipped and fell, I said "It must have been the parkay."	En cuisinant, j'ai fait tomber un bâton de margarine sur le carrelage en bois, et quand ma voisine a glissé et est tombée, j'ai dit : "Ça doit être le parkay".
en_1047	Pablo and Jose waited for a signal from Pilar to join her in the Church Tower, but neither responded to the chiming because they didn't know for whom the belle was tolling.	Pablo et José attendaient que Pilar allume le grand Feu avant de la rejoindre à l'église, mais d'aucun ne réagit à la chaleur ne sachant à qui était destiné la femme
en_1051	Someone once accused me of stealing an old, rare, valuable stamp, and I philately denied it.	Un jour, quelqu'un m'a accusé d'avoir volé un timbre de valeur, rare et ancien. Je lui ai dit que, pour porter de telles accusations, il devait être complètement timbré.
en_1059	The whistling fisherman was always out of tuna.	Le pêcheur sifflant était toujours hors de thon.
en_1072	If you wear a blindfold at the shooting range, you won't know what you're missing.	Si vous portez un bandeau aux yeux, vous ne saurez pas ce que vous manquez.
en_108	OLD BAKERS never die, they just stop making lots of dough.	Les vieux boulangers ne meurent jamais, ils ne gagnent juste plus aussi bien leur pain.
en_1084	The fireplaces of oriental doctors have an Asian flue.	Je trouve que dans un intérieur rustique, un poêle à bois, ça a vraiment du chien
en_1085	After he bought a mirror he became very reflective.	Il réfléchit mieux maintenant qu'il a un miroir.
en_110	'This food tastes of plutonium," said Tom glowingly.	"La nourriture a goût de plutonium", dit Tom, rayonnant.
en_1106	OLD TYPISTS never die, they just lose their touch.	Les vieux dactylographes ne meurent jamais, ils perdent juste leur touche spéciale.
en_1107	'This salad dressing has too much vinegar," said Tom acidly.	Cette sauce salade contient trop de vinaigre, dit Tom d'un ton acide.
en_111	The local baker was paying his staff based on a flourly rate.	Tu sais que le boulanger du coin paye ses employés au blé noir ?
en_1120	OLD POLITICIANS never die, they just harangue in there.	Les vieux politiciens ne meurent jamais, ils ne sont juste plus trop dans leur état normal.
en_1121	Some people find fire drills quite Alarming.	Certaines personnes pensent que les exercices d'incendie peuvent amener les gens à se blesser dans le feu de l'action.
en_114	OLD FLORISTS never die, they just make alternative arrangements.	Un fleuriste ne meurt jamais, seulement, il perd de sa fraîcheur.