IIMAS-UNAM Team entry: Transformers Adapters for the Sentiment Analysis REST-MEX 2023

Eric Yaven Baez-Reyes[1,†], Isaac Barrón-Jiménez[1,†], Héctor Becerril-Pizarro[2,†], Ximena de la Luz Contreras-Mendoza[1,†] and Ivan Vladimir Meza-Ruiz[3,*,†]

[1]*Facultad de Ciencias, Universidad Nacional Autónoma de México, México*

[2]*Facultad de Estudios Superiores Zaragoza, Universidad Nacional Autónoma de México, México*

[3]*Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, México*

Abstract

This work describes and presents the methodology of the *IIMAS-UNAM* team followed in the sentiment analysis task in the REST-MEX 2023 competition on tourist reviews. It shows the use of the Transformers architecture together with Adapters achieves a sufficiently good performance in this type of task. The submitted model was ranked third place, reaching a 75.00 Sentiment Track Score.

Keywords

Adapters, Transformers, Sentiment Analysis

1. Introduction

This paper describes the procedures and results of the IIMAS-UNAM team in the REST-MEX 2023 competition within the sentiment analysis task, which aimed to predict three variables: *polarity*, *type of place*, and *country* that correspond to written reviews. These reviews were in Spanish from tourists in Mexico, Cuba and Colombia. The review consisted of a title and the review text. Table 1 shows an example of a review and the variables that have to be predicted.

An explanation of the relevance of this task in the tourism sector can be found in [1], where it is empirically established that the decisions made by tourists regarding the establishment they chose for their service are based on the reviews and comments that other people make about places of interest. Therefore, knowing options, destinations, and opinions is valuable for exploring routes and improving tourist destinations [2].

The data processing problem related to tourism has been extensively studied; a good summary

IberLEF 2023, September 2023, Jaén, Spain

*Corresponding author.

†These authors contributed equally.

eric315@ciencias.unam.mx (E. Y. Baez-Reyes); isaac.bj@ciencias.unam.com (I. Barrón-Jiménez); fzpsi.hp.unam@gmail.com (H. Becerril-Pizarro); xilucome@ciencias.unam.mx (X. d. l. L. Contreras-Mendoza); ivanvladimir@turing.iimass.unam.com (I. V. Meza-Ruiz)

0000-0002-7239-1480 (I. V. Meza-Ruiz)

Table 1
Review example and corresponding labels. In italics, translated text.

Title	Review	Polarity	Country	Type
LINDA EXPERIENCIA/*NICE EXPERIENCE*	Hermoso lugar para admirar las obras de Botero. Me encantó,porque el lugar es aseado,ordenado y cuenta con baño y tienda para el público que lo visita./*Beautiful place to admire the works of Botero. I loved it,because the place is neat,tidy and has a bathroom and a shop for the public that visits it.*	5	Colombia	Attractive

can be found in [3, 4, 5]. In those systematic reviews, there are different ways in which it has been evaluated, including text extraction, sentiment analysis, text grouping in the context of tourism text mining, application of tourism profiles, destination analysis and demand for tourism.

For our method, we propose using a pre-trained transformer [6] model and enhancing it by fine-tuning the adapters methodology[7]. We created three independent adapters models, one for each task: polarity, country and type classification. With this methodology, we reach the results presented in Table 2.

Table 2
Official scores of the presented model, Sentiment Track Score for the polarity variable, as well as the F1 score for each variable.

Task evaluation	Score
Sentiment Track Score	75.00%
Polarity Macro F1	59.32%
Type Macro F1	97.90%
Country Macro F1	90.24%

2. Previous work

In lasts years' contest edition, the sentiment analysis task consisted of predicting the polarity and type of place given an opinion issued by a tourist who travelled to the most representative places in Mexico [8, 9]. So the main difference with this year's contest is the addition of comments from tourist sites in two more countries: *Colombia* and *Cuba*. So now we have one more prediction task: predicting the country given a tourist opinion[10].

2.1. Transformers

The architecture *transformers* has shown excellent results for NLP tasks thanks to its multi-attention layers[11], which allow giving weights or importance to each word in the sentence, measuring its relation with the other words. All this through different models based on it, such as BERT [12], used in the last competition [13].

However, the problem we now face is fine-tuning the pre-trained model. Given the difficulty in fitting language models (LLM) for a specific task and dataset with customized domain, new training techniques such as *adapters* have been developed.

As sentiment classification tasks are normally worked on, a pre-trained model with millions of parameters is downloaded to make a fine adjustment on another specific task, which implies a computational cost in memory and time. This is why a relatively new alternative is the use of adapters. Adapters are part of the so-called additive fine-tuning methods, which add additional parameters or layers to the pre-trained model and train the newly added parameters [7].

Adapters are neural modules with a much smaller number of new parameters inserted between the original parameters of the pre-trained model. In this way, when training for a specific task, the only values that will be trained will be the new parameters, leaving the model parameters fixed. The advantages are that having fewer parameters to train means less training time and fewer data storage for the model. Another advantage is that the previously trained model is not modified so that information from the model that could be useful for future tasks is not lost [7].

The way they have been used in NLP, along with the transformers architecture, is to put a series of adapters parameters between each layer of the transformers architecture. The most common way to use adapters is through the *adapter-transformers* library, which is built on top of *Hugging Face* transformers library, so you don't need to call the transformers library again. So you only have to add a couple of lines of code to train. [7]

3. Preprocessing

The data provided by the competition are opinions of tourists written in Spanish who travelled to the most representative places, hotels, restaurants and attractions in Mexico, Colombia and Cuba on the page *TripAdvisor*, which had $251,702$ opinions labelled in polarity, type and country and $107,861$ opinions to test the data of this work in the competition.

For the pre-processign we use several *libraries*:

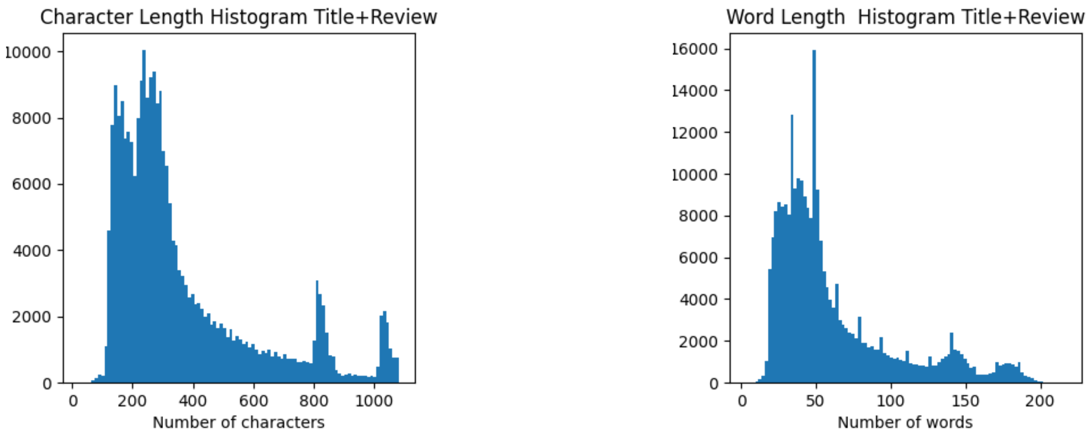
emoji is used in word processing to convert emojis to Unicode or vice versa.

Pandas to view and filter data.

Transformers provides an API that describes classes and functions such as configuration, model, tokenizer, and pipeline.

Among the modifications made to the data set during the data pre-processing stage are: having shortened the lengths of the title and body of the reviews to 75 and $2,500$ characters, respectively, based on the average number of characters for each column Since both the "Title" column and the "Review" column contained important review information, it was decided to concatenate both columns into a single one using the `[sep]` string to separate the content of both columns. Figure 3 present the character and word length histogram distribution of the concatenated Title+review.

Subsequently, data filtering continued: a "demojization" was performed (i.e. converting emojis to text strings), and repeated sentences were eliminated, which reduced the training data set from $251,702$ to $246,160$.



We take 95% of the data randomly for training, 2.5% for validation and 2.5% to test our results (with the same training data), in this way we use the vast majority of the data for training and because the Data set is very large, we got a good amount of data for validation and testing (around 6000 reviews for each set). The transformers model we used is the "xlm-roberta-base" model, which is pre-trained with 2.5 TB of data filtered by CommonCrawl. And as an adapter and fine-tuned an adapter. When training a pre-trained model of *Huggingface*, the first thing we did was to tokenize each of our strings, of which only the first 300 characters were used (max_lenght = 300). We proceeded to train the model; for each task, we performed 3 epochs with a batch size of 16.

During the training process, the model was tested on the evaluation set and the following results were obtained for Polarity task.

Table 3
Internal evaluation of the data set for the polarity variable.

Evaluation Parameter	Score
Evaluation loss	64.48%
Evaluation accuracy	72.68%
Evaluation F1	53.85%
Evaluation Precision	56.75%
Evaluation Recall	52.40%

The table 4 shows the accuracy value obtained for the polarity, country and type variables, respectively. As can be seen, the variable with the greatest difficulty to predict was polarity, obtaining a value of 73.98%. We believe that this low value, compared to the value obtained for the other variables, can be explained due to the difficulty of discerning between close ratings (e.g. see 5) given that the use of language can be similar in these cases. One proposal is to perform a linear regression with the “polarity” labels to improve its prediction.

On the other hand, for the country and type variables, the accuracy values 90.54% and 98.04% were obtained, respectively. This shows that these variables were easier to predict, a

Table 4
Official accuracy values for the different variables.

Variable	Accuracy
Polarity	73.97995606%
Country	90.54448699%
Type	98.04288774%

Table 5
Two reviews extracted from the trainig dataset showing different polarity values with similar language use.

Review	Polarity
El Maine fue capturado por películas de Edison en 1898. Me podría imaginar en mi mente. Las fotografías son fantásticas al atardecer. Muy bien mantenido. Buen sitio histórico!	4
experiencia unica, ver las piramides y toda la energia, plenos hermoso momento.. digno de ser visitado.	5

probable explanation for this is the independence between the values of these labels compared to the values for the polarity label.

Finally, 6 shows the general metrics obtained with the model in comparison with the rest of the participants. Despite the fact that in general an acceptable value was obtained in the sentiment analysis task, it can be observed that the metric related to the polarity variable had a negative impact on the final grade obtained in the task. However, it is necessary to note that this fact was common to the rest of the participants, showing once again the difficulty that exists in the task of predicting this type of variable from the body of a review. One way in which the results could be improved would be by refining or summarizing the tourist opinions, in this way the training time could be reduced and the most important part of the opinion could be considered mainly.

4. Conclusions

4.1. Technique and final result

This paper describes the tasks and procedures that the IIMAS-UNAM team followed to solve the REST-MEX 2023 competition in the sentiment task. A third place in ranking results was achieved after a simple pre-processing of the data and training that required the use of Transformers and Adapters.

Thus, it is shown that the use of Transformers together with Adapters, in particular, the instances *xlm-roberta* is effective in task related to the classification and analysis of text sentiment such as reviews in domains such as tourism.

Table 6
REST-MEX competition's ranking table.

Ranking	Run	Sentiment Track Score	Macro F1(Polarity)	Macro F1(Type)	Macro F1(Country)
1st	LKE-IIMAS-Team_RUN_2	0.7790	0.6217	0.9903	0.9420
2nd	javier_alonso-Team_sentiment_sub6	0.7662	0.6021	0.9885	0.9358
3rd	**IIMAS-UNAM-Team_resultados**	**0.7501**	**0.5932**	**0.9790**	**0.9024**
HM	INGEOTEC-Team	0.7376	0.5549	0.9805	0.9271
HM	UCT_UA-Team_run_01	0.7190	0.5153	0.9874	0.9384
HM	BUAA-Team_M1_M1_M2	0.7189	0.5265	0.9804	0.9236
HM	Dataverse-Team_Results_RestMex23	0.7174	0.5207	0.9760	0.9152
HM	SENA-Team_i-1	0.7009	0.5060	0.9727	0.8853
HM	UMSNH-Team_sentiment_results_xgb_ensemble	0.6992	0.5086	0.9733	0.8908
HM	Camed_CU-ES-Team_camed_camila_y_eduardo_results_output	0.6978	0.5104	0.9739	0.8536
HM	JL-Team-Added	0.6889	0.4818	0.9719	0.9081
HM	ITT-Team-Topics_k=1	0.6814	0.4866	0.9674	0.8763
HM	Algiedi-Team-5000	0.6692	0.4922	0.9542	0.8166
HM	Arandanito-Team_POS-ADJ+NOUN+VERB	0.6392	0.4427	0.9534	0.7810
HM	ABCD-Team_submission	0.4728	0.0960	0.8848	0.7864
HM	Olga-Team-dfTestfinal_LyS-SALSA	0.3144	0.3436	0.3324	0.3259
HM	The_Last-Team	0.2295	0.1831	0.3018	0.2501

References

[1] Q. Ye, R. Law, B. Gu, W. Chen, The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings, Computers in Human Behavior 27 (2011) 634–639. URL: https://linkinghub.elsevier.com/retrieve/pii/S0747563210000907. doi:10.1016/j.chb.2010.04.014.

[2] E. Olmos-Martínez, M. Á. Álvarez-Carmona, R. Aranda, A. Díaz-Pacheco, What does the media tell us about a destination? the cancun case, seen from the usa, canada, and mexico, International Journal of Tourism Cities (2023).

[3] Q. Li, S. Li, S. Zhang, J. Hu, J. Hu, A Review of Text Corpus-Based Tourism Big Data Mining, Applied Sciences 9 (2019) 3300. URL: https://www.mdpi.com/2076-3417/9/16/3300. doi:10.3390/app9163300.

[4] M. A. Álvarez Carmona, R. Aranda, A. Y. Rodríguez-Gonzalez, D. Fajardo-Delgado, M. G. Sánchez, H. Pérez-Espinosa, J. Martínez-Miranda, R. Guerrero-Rodríguez, L. Bustio-

Martínez, Ángel Díaz-Pacheco, Natural language processing applied to tourism research: A systematic review and future research directions, Journal of King Saud University - Computer and Information Sciences 34 (2022) 10125–10144. URL: https://www.sciencedirect.com/science/article/pii/S1319157822003615. doi:https://doi.org/10.1016/j.jksuci.2022.10.010.

[5] A. Diaz-Pacheco, M. A. Álvarez Carmona, R. Guerrero-Rodríguez, L. A. C. Chávez, A. Y. Rodríguez-González, J. P. Ramírez-Silva, R. Aranda, Artificial intelligence methods to support the research of destination image in tourism. a systematic review, Journal of Experimental & Theoretical Artificial Intelligence 0 (2022) 1–31. URL: https://doi.org/10.1080/0952813X.2022.2153276. doi:10.1080/0952813X.2022.2153276. arXiv:https://doi.org/10.1080/0952813X.2022.2153276.

[6] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, Roberta: A robustly optimized bert pretraining approach, arXiv preprint arXiv:1907.11692 (2019).

[7] J. Pfeiffer, A. Rücklé, C. Poth, A. Kamath, I. Vulić, S. Ruder, K. Cho, I. Gurevych, Adapterhub: A framework for adapting transformers, arXiv preprint arXiv:2007.07779 (2020).

[8] M. Á. Álvarez-Carmona, R. Aranda, S. Arce-Cardenas, D. Fajardo-Delgado, R. Guerrero-Rodríguez, A. P. López-Monroy, J. Martínez-Miranda, H. Pérez-Espinosa, A. Y. Rodríguez-González, Overview of rest-mex at iberlef 2021: Recommendation system for text mexican tourism (2021).

[9] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, A. Y. Rodríguez-González, D. Fajardo-Delgado, R. Guerrero-Rodríguez, L. Bustio-Martínez, Overview of rest-mex at iberlef 2022: Recommendation system, sentiment analysis and covid semaphore prediction for mexican tourist texts, Procesamiento del Lenguaje Natural 69 (2022) 289–299.

[10] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, A. Y. Rodríguez-González, L. Bustio-Martínez, V. Muñis-Sánchez, A. P. Pastor-López, F. Sánchez-Vega, Overview of rest-mex at iberlef 2023: Research on sentiment analysis task for mexican tourist texts, Procesamiento del Lenguaje Natural 71 (2023).

[11] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention Is All You Need, 2017. URL: http://arxiv.org/abs/1706.03762, arXiv:1706.03762 [cs].

[12] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, 2019. URL: http://arxiv.org/abs/1810.04805, arXiv:1810.04805 [cs].

[13] J. A. García-Díaz, M. Á. Rodríguez-García, F. García-Sánchez, R. Valencia-García, Umuteam at REST-MEX 2022: Polarity prediction using knowledge integration of linguistic features and sentence embeddings based on transformers, in: M. Montes-y-Gómez, J. Gonzalo, F. Rangel, M. Casavantes, M. Á. Á. Carmona, G. Bel-Enguix, H. J. Escalante, L. A. de Freitas, A. Miranda-Escalada, F. J. Rodríguez-Sanchez, A. Rosá, M. A. S. Cabezudo, M. Taulé, R. Valencia-García (Eds.), Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2022) co-located with the Conference of the Spanish Society for Natural Language Processing (SEPLN 2022), A Coruña, Spain, September 20, 2022, volume 3202 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2022. URL: https://ceur-ws.org/Vol-3202/restmex-paper13.pdf.

[14] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave,

M. Ott, L. Zettlemoyer, V. Stoyanov, Unsupervised Cross-lingual Representation Learning at Scale, 2020. URL: http://arxiv.org/abs/1911.02116, arXiv:1911.02116 [cs].