PANAS-TDL2: A Psychrometric Deep Learning Model for Characterising post-COVID-19 Twitter Perceptions of Tourist Destinations

Alejandro Peña[1[0000-0001-9441-9280]] and Joao Vidal[2[0000-0001-9727-0557]]
and J.D. Gonzalez-Ruiz[3[0000-0003-4425-7687]] and Lina Sepulveda [1[0000-0003-1749-816X]]
[1] Escuela de Administración, Universidad EAFIT, Medellín, Colombia, 055422
[2] Polytechnic of Porto/CEOS.PP, S. Mamede de Infesta, Oporto, Portugal
[3] Departamento de Economía, Universidad Nacional de Colombia, Medellín, Colombia
japena@eafit.edu.co, cajvidal@iscap.ipp.pt,jdgonza3@unal.edu.co, lmsepulvec@eafit.edu.co

Abstract. Tourism is one of the main sectors in recent years that has significantly impacted the countries' economies worldwide. For this reason, many governments worldwide have begun a series of actions to improve travellers' perceptions of their tourist destinations, using the power of social networks. However, the characterization of these perceptions is not an easy task. For this fact, in this paper, we proposed a neural network model with a stacked deep learning structure, which integrates a Softmax function inspired by a PANAS-t scale (A Psychometric Positive and Negative Affect Scale) to characterize the traveller's perceptions of tourist destinations (PANAS-tDL2 model), from a series of comments posted by them in a social network as Twitter in the framework of COVID-19. For the classification of comments in each sentiment category that define the PANAS-t scale, each comment was subjected to preprocessing using NLP (Natural Language Processing) techniques. The results show the capacity of the proposed PANAS-tDL2 model to evaluate the COVID-19 impact from comments posted by travellers about a tourist destination worldwide. For its capacity, the PANAS-tDL2 model can be extended to create tourist packages, experiences and services tailored for travellers, or to create insurance products for reputation to protect the travellers in tourist destinations against climatic, political, or sanitary risks, as one generated by the COVID-19 pandemic.

Keywords: Deep Learning, PANAS- Scale, NLP (Natural Language Processing), COVID-19 Impact.

1 Introduction

One of the main sectors that have long driven countries' economies has been the tourist sector. Several researchers have shown the positive impacts that tourism has brought to countries worldwide, that are reflected in taxes, the protection of historical legacy, and in recent years, the protection of the environment. However, in the context of the COVID-19 pandemic, this industry harmed the country's economies due to massive lockdowns. To assess this impact, many researchers used different economic and financial indicators; however, these assessments left out significant impacts on tourism, such as those generated by travellers' perceptions about a country's health conditions based on comments posted by them on social networks such as Twitter or Facebook, or trip platforms like TripAdvisor and Booking.com.

To assess the impact of COVID-19 on the tourism sector in a holistic way, many researchers have turned to the characterization of travellers' perceptions in social networks using different models and methodologies. Several development trends can be identified in the scientific literature in this context. A first development trend discusses the impact of the COVID-19 pandemic on the global economy and international tourism. In this regard, Haryanto (Haryanto, 2020) highlighted the suspension of international flights due to the policy of seat distance between passengers, which fell by more than half, causing significant losses for tourist destinations. IATA (IATA, 2020) estimates that the COVID-19 crisis generated an approximate reduction of 55% in 2020 against the revenues reported by the airlines in 2019 (US 314 billion). This COVID-19. This airline crisis had a knock-on effect on a tourist destination like Indonesia. Indonesian government reported losses close to US 500 million per month (Dinarto, Wanto, & Sebastian, 2020). Tourist destinations such as Costa del Sol in Spain (Arbulu, Razumova, Rey-Maquieira, & Sastre, 2021), or the Riviera Maya in México (Rubio-Cisneros, Montero-Muños, Rubio-Cisneros, Morales-Ojeda , & Pech, 2022) suffered similar impacts. This development trend clearly shows the impacts that COVID-19 pandemic had on the tourist sector in general, significantly affecting countries where the economy revolves around tourism (Usui, Sheeran, Asbury , & Blackson, 2020).

A second development trend focuses on characterising travellers' perceptions against different tourist destinations within the framework of COVID-19 pandemic. Within this trend, Viñan-Ludueña et al. (Viñan-Ludeña & de Campos, 2021) analyse tourist content in Twitter about places, events, restaurants, hotels, etc. related to the Province of Granada (Spain). Kamata (2022) analyses the perception of residents in different tourist destinations using variables such as place attachment, distinctiveness, positive impact, negative impact, and attitude-related to fear of new outbreaks of COVID-19 as a result of welcoming tourists. Rahman et al. (2021) explore the impact of the COVID-19 pandemic on tourists' travel risk based on tourist perceptions and its effect on society using a sample of respondents. Finally, Kumar & Nafi (2021) measure the effect of the COVID-19 pandemic on the tourism industry in Bangladesh, using journals, historical records, newspaper articles, World Health Organization statistics, governmental data, and website materials on COVID-19 incidences in tourism. This trend highlights the importance of assessing the impact of COVID-19 by analysing news, social media materials and comments posted by travellers and tourist safety officers worldwide on social networks against tourist destinations.

A third development trend shows how web-scrapping techniques integrated with Machine & Deep Learning models have made it possible to assess travellers' perceptions in social networks against products and services offered by different tourist destinations worldwide. Han & Anderson (2021) show a growing interest in understanding the online travel marketplace using analytical models based on collet/scrape methodologies, experiments, or survey collections. For its part, Johnson et al. (2012) describe the use of automated methodologies to extract review data on the Canadian province of Nova Scotia from a significant travel review website, highlighting automated web harvesting as a web scraping methodology. In a third study, Darmawiguna et al. (2019) show how web-scraping methodologies can be hybridised with clustering methods to deliver accurate information on Bali, analysing information from tourist portals. Another group of studies within the web-scrapping process aims to generate relevant information for travellers before visiting a tourist place. In this way, Sherman et al. (2020) express the need to create platforms for travellers to evaluate a tourist destination based on a combination of reviews from different websites about negative and positive aspects. Finally, Peña et al. (2021) propose a novel PANAS-t Deep Learning model to characterize sentiments in comments posted by travellers against tourist destinations in a social network as Twitter using NPL techniques. In this development trend, Kalvet et al. (2020) indicate the complexity of the tourist market and the knowledge gap regarding suitable methods and data sources to measure the impacts of a risk as the COVID-19 pandemic generates on countrie`s economies. In general, this trend highlights how web-scraping, NLP, machine learning and artificial intelligence models could automate the evaluation of the benefits that the tourism sector brings to countries using elements such as communications in all its forms, including media (text, image, audio), perceptions, knowledge, planning and reasoning (Mich, 2022).

In line with the third trend identified in the scientific literature, in this study, we proposed a neural network model with Stacked Deep Learning structure to evaluate the travellers' perceptions of five tourist destinations (Colombia, Italy, México, Spain, USA) from a series of comments posted by them in a social network as Twitter in the framework of COVID-19 evolution (Fischer & Krauss, 2018). The proposed model integrates a novel *Softmax* function based on PANAS-t scale (Positive and Negative Affect Scale – *Softmax-PANAS-t function*) (Peña, Mesias, Patiño, & de Carvalho, 2021), which allows classifying the Twitter comments as Negative (NP) and Positive (PP) perceptions, and in an 11-Sentiment defined by the *PANAS-t scale*: Guilt, Fear, Sadness, Hostility, Shyness, Fatigue, Surprise, Joviality, Self-assurance, Attentiveness and Serenity. According to the structure of the activation function, this scale goes from (-1) *Guilt* to (1) *Serenity*, setting up a PANAS-TDL2 model.

For the analysis and validation of the proposed PANAS-tDL2 model 100,000 comments (*10-words*) were obtained from the Twitter social network using a web-scrapping methodology (Dewi & Alvin, 2019). Each comment was subjected to a pre-processing of Tokenization, StopWord Filter, and WordNet Lemmatizer using TextBlob (Mostafa, Ahmed, & Junayed, 2021). Comments were grouped in PP & NP, and the 11-Sentiment scale (Baseline Scenario). Subsequently, the comments were grouped in the five aforementioned tourist destinations before (2019-2020), during (2020-2021), and after COVID-19 worldwide pandemic (2021-2022), to analyse the evolution of travellers' perceptions. In this way, the comments posted by travels for a before stage (2019-2020) were used to set up the PANAS-tDL2 model (First Stage), while the remaining comments were used to validate the model in the absence of a learning process (Second Stage).

For the first stage, the proposed PANAS-tDL2 model reached compression rates close to 90% on average about the configuration of stacked layers structure using a sequential autoencoder strategy and taking as reference the number of neurons that make up the first layer (Peña, et al., 2021). For the characterization of comments based on the Softmax *PANAS-t* function, the proposed PANAS-tDL2 model reached IOAs close to 95% on average. In the Second Stage, the PANAS-tDL2 model was validated for comments obtained in 2020-2021 (during-COVID-19) and in 2021-2022 (post-COIVD-19) in the absence of a learning process. The results show the re-localisation of travellers' perceptions toward the Positive Affect region that define the PANAS-t scale as a result of massive vaccination worldwide. It is important to note that for 2019-2020 (*before-COVID-19*), the travellers' perceptions were located in the Positive Affect zone defined by PANAS-t scale, while for 2020-2021 these perceptions turned towards the Negative Affect zone of this scale, as a result of the impact of COVID-19. The latter shows the suitable behaviour exhibited by the PANAS-tDL2 model in the absence of a learning process, thanks to sensitivity to identify sentiments from comments posted by travellers in a social network such as Twitter against travel destinations in the framework of COVID-19 worldwide pandemic.

The paper has four sections describing the behaviour of the proposed PANAS-tDL2 model characterizing perceptions of travellers against tourist destinations in a social network such as Twitter. The first section (Introduction) shows the problem statement and the literature review. A second section (Methodology) describes the case of study, the structure of the proposed model, and the experimental validation. A third section shows the results achieved by the proposed model, and finally, the paper shows a series of conclusions and future work to create new tourist products and services based on the travellers' perceptions from any social network.

2 Methodology.

Social networks have become the reference data source for assessing concepts, products, and marketing campaigns. Regarding the tourism industry, Twitter is being more valuable for exploring the perceptions of travellers against tourist destinations, enabling decision-makers to create new products, services, and experiences. We propose the following methodology to assess the impact of COVID-19 on a tourism destination based on comments posted by travellers on a social network such as Twitter.

2.1 Case of Study.

For the analysis and validation of the PANAS-TDL2 model, a total of 100,000 comments (10-words) related to travellers' perceptions against five (5) tourist destinations (Colombia, Italy, Mexico, Spain, USA) were collected from the Twitter social network using a web-scrapping methodology (Dewi & Alvin, 2019). Each comment was subjected to a pre-processing of Tokenization, StopWord Filter, and WordNet Lemmatizer using the API TextBlob in Pyhton (Mostafa, Ahmed, & Junayed, 2021). After pre-processing, the comments were grouped in PP (Positive Perceptions) & NP

(Negative Perceptions), and the 11-Sentiment categories (Guilt, Fear, Sadness, Hostility, Shyness, Fatigue, Surprise, Jovialty, Selg-assurance, Attentiveness, Serenity) that define the *PANAS-t* scale as well (Baseline Scenario). In this way, each word is rated in the interval $[-1\ (Guilt), 1(Serenity)]$ according to the sentiment it represents, as well as a frequency value, representing the times that a word appears in the set of comments defined by this study.

To evaluate the impact of COVID-19 on different tourist destinations, the comments were grouped in the five destinations mentioned above (approximately 20,000 comments): Colombia, Italy, Spain, Mexico, and the United States of America (USA). In this way, about 33,500 comments corresponding to the period 2018-2019 (before COVID-19) were used to configure the proposed PANAS-tDL2 model (First Stage). The remaining comments (approximately 66,500 comments) for 2019-2020 (during COVID-19) and 2020-2021 (post-COVID-19) periods will be used to validate the PANAS-tDL2 model without a learning process.

2.2 Deep Learning Model with PANAS-t SoftMax Function (PANAS-tDL2).

For the psychometric characterization of perceptions against a tourist destination based on the comments posted by travellers on a social network such as Twitter, we propose a neural model with a deep learning structure (stacked layer), which integrates a novel *Softmax* activation function inspired on the PANAS-t scale, which shall be known as PANAS-tDL2 (PANAS-t Deep Learning). The PANAS-tDL2 model is denoted and defined:

$$y_{SL,k} = \sum_{j=1}^{n_{nc}} \ldots \sum_{j_1=1}^{n_1} \sum_{j_o=0}^{n_o} w_{j_{nc},j_{nc-1}} \ldots .. \left(w_{j_1,j_o}\left(w_{j_o,i}.x_{i,k}\right)\right) (1)$$

Where: $y_{SL,k}$: Stacked Deep Learning Output. $w_{j_{nc},j_{nc-1}}$: Stochastic connections between stacked layers. nc: Number of Stacked Layers. np: Number of words that make up a comment. k: Number of comments. According to the *PANAS-t scale*, the *Softmax* activation function can be expressed as follows:

$$yr_k = \frac{1}{1 + e^{-y_{SL,k}}} (2)$$

Where: $y_{SL,k}$: Represents the psychometric response of PANAS-tDL2 model in agreeing to an input comment ($x_{i,k}, i = 1,2,3, \ldots .., np$). It is important to note that this response can be characterized according to PANAS-t scale as follows:

$$S_s = \frac{1}{ns} : yr_k \in [0,1]$$

Where: S_S: Sentiment scale. ns: Number of sentiments (*11 sentiments*).

PANAS-t scale can be represented through a Radar Plot known as the multidimensional classification chart. In the context of deep learning, the PANAS-t scale can be expressed according to the *Softmax* function, or a logistic cumulative distribution function (CDF) (Fig. 1). The Radar Char and Softmax function have the PA zone (left side) and a NA zone (right side). The *x-axis* represents the positivity of a comment (−1 (*Guilt)*, 1 (*Serenity*)), while the y-axis defines the 11-sentiment scale.

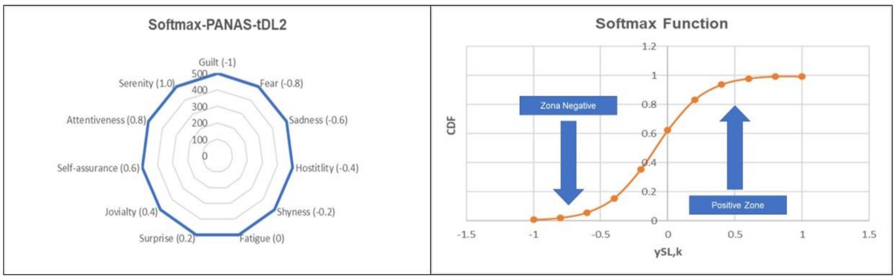

Fig. 1. Softmax PANAS-t activation function

For the configuration of the stacked layers, we proposed an *auto-encoder* sequential methodology based on the generalized Jacobi method:

$$w_{j_{nc},j_{nc-1},k} = w_{j_{nc},j_{nc-1},k-1} + \propto_{nc}. e_k. \left(w_{j_{nc},j_{nc-1}} \dots \dots w_{j_o,i}\right). x_{i,k} (3)$$

Where: $\propto_{nc}$: Learning factor for $nc - layer$. e_k: Mean square error.

$$e_k = (yd_k - yr_k)(4)$$

Where: yd_k: Represents the psychometric characterization for each $k - input$ comment. This comment is denoted and defined:

$$yd_k = \sum_{i=1}^{np} \frac{f_i. sp_i}{sf} : sf = \sum_{i=1}^{np} f_i \ (5)$$

Where: f_i: Frequency for $i - word$. sp_i: Psychometric value for the $i - word$. This psychometric value depends on a specific dictionary used to characterize the travellers' perceptions.

2.3 Metrics.

For the analysis and validation of the proposed PANAS-tDL2 model, the following metrics were used:

- **Relative Position**: Let *S* the set of tweets for a particular event of risk (e.g. COVID-19, natural disasters, political events, etc.) and S_s the subset of these tweets related to *s* sentiment. B_s represent the relative occurrence of sentiment *s* for event *S*. The B_s can be expressed as follows (Gonçalves, Benevenuto, & Cha, 2013):

$$\beta_s = \frac{|S_S|}{|S|} (6)$$

- **Score Function:** Let s categories that define the PANAS-t scale, the *Score Function* $P(s)$ can be expressed as follows (Gonçalves, Benevenuto, & Cha, 2013):

$$P(s) = \frac{(\propto_S - \beta_S)}{\alpha_S}, if\ \beta_S \leq \propto_S \ (7)$$

$$P(s) = -\frac{(\beta_S - \alpha_S)}{\beta_S}, if\ \beta_S \geq \propto_S \ (8)$$

The values for $P(s)$ are defined in the interval $[-1,1]$ for each s sentiment. $P(s) = 0$ means that the event has no increase or decrease for the s sentiment compared with the whole set of comments (S). A $P(s) \geq 0$ represents an increase for an s-sentiment, while a $P(s) \leq 0$ represent a decrease for an *s-sentiment*.

2.4 Experimental Validation.

For the analysis and validation of the PANAS-tDL2 model, two stages were proposed. In the first stage, the Stacked Layers Structure of the proposed model was configured using a sequential *autoencoder* methodology based on comments posted by travellers in 2019 (before COVID-19) for the five destinations selected by this study. Each Staked Layer will be configured using two parameters: the number of layers (n_{nc}), and the number of neurons (n_o) for the first layer. The relationship between the input and output comments for each stacked layer will be evaluated using the Index of Agreement (IOA-IC). It is essential to stand out that each comment was subjected a NLP process as aforementioned in the case of study section. In this stage, it is expected that the proposed PANAS-tDL2 model reached IOAs above 85% on average, according to the results achieved by Peña et al. (2021) and Goncalves et al (2013).

In this same stage, the overall performance of PANAS-tDL2 model will be done based on the structure of the *Softmax-PANAS-t* function for the period *before-COVID-19* for whole destinations. In this way, the comments (35,000) will be grouped into two sets of comments: a first set grouped approximately 70% of comments (*Configuration Set - LS*), while the second set grouped approximately 30% of comments (*Validation Set - VS*). In this stage, it is expected that the PANAS-tDL2 model reached Index of Agreements (IOAs-ICs) close to 100% and similar Skewness Indices as well, regarding the LS and VS sets.

In the second stage, the proposed PANAS-tDL2 model will be evaluated according to its sensibility (absence of learning). For this reason, the remaining comments (66,500) will be grouped during (2020-2021) and post-COVID (2021-2022). In this stage, the sensitivity of the proposed PANAS-tDL2 model is expected to shift the comments toward the negative zone of the PANAS-t scale (Radar Chart of Sentiments) during-COVID-19 pandemic for each tourist destination and shifts the comments toward to positive zone for *post-COVID-19* stage. In this way, the evolution of comments toward different zones of PANAS-t scale will be evaluated using the relative position (B_s) and the score function ($P(s)$). In this way, positive values for $P(S_k)$ represents the evolution of comments in the radar chart, while the difference of the absolute maximum values for two different periods ($dP(S) = P(S_k) - P(S_{k-1})$) represents the impact of COVID-19 on a tourist destination ($dP(S)$). Here, $dP(s)$ positive values represent the recovery of positive perceptions of travellers against a tourist destination, in the opposite, negative values show the negative perception remained by travellers despite COIVD-19 mitigation in the world because of massive vaccination.

3 Analysis of Results.

Figure 1 shows the behaviour exhibited by the model during the Stacked Layers Structure configuration. Here, the PANAS-tDL2 model shows an ascending behaviour about the Index of Agreement (IOA-IC), as the number of neurons for the first stacked layer (no) and the number of layers (n_{nc}) increased. This Figure shows that stabilisation in learning was achieved for an IOA-IC index close to 87% (IOA-IC:0.878825), for a seven layers (n_{nc}: 7), and for a total of 750 (no) neurons for the first layer. However, it is important to note that, despite the increased flexibility of the PANAS-tDL2 model, this didn`t significantly improve classifying comments by sentiment category.

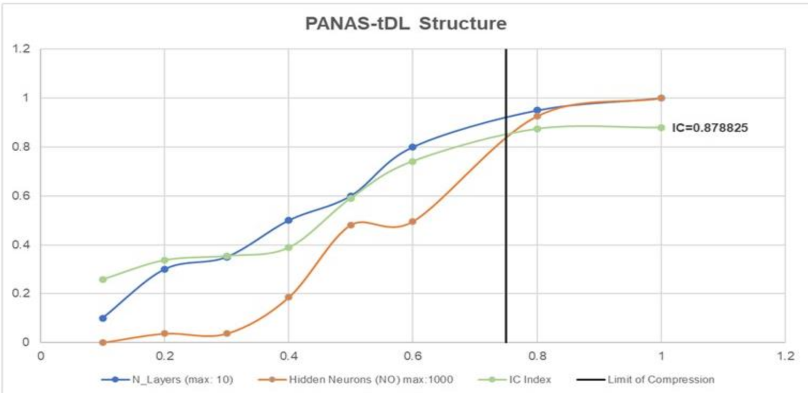

Fig. 2. PANAS-tDL2 Model Set-Up.

Figure 3 shows the results achieved by the PANAS-tDL2 model against the characterization of the *Softmax-PANAS-tDL* function. Here, the PANAS-tDL2 model achieved IOA-ICs close to 95% (LS: 0.954457) on average against the classification of comments by sentiment category (Configuration Stage). Notably, the PANAS-tDL2 model achieved IOA-ICs above 90% (VS:0.913662) on average during the validation stage, demonstrating the proposed model's autonomy in classifying comments posted by travellers on Twitter against the different tourist destinations in the framework of COVID-19 pandemic.

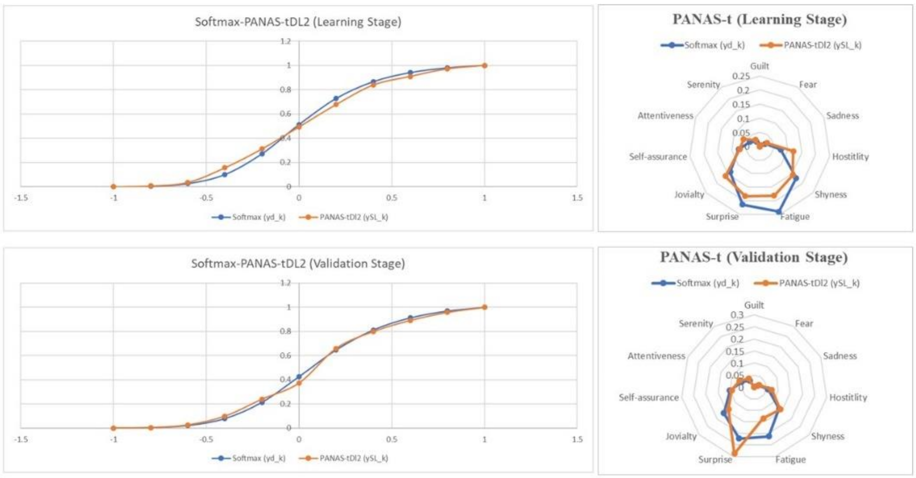

Fig. 3. Softmax-PANAS-tDL Characterization

Table 1 shows the evolution experienced by comments posted by travellers on Twitter against five (5) tourist destinations during the COVID-19 pandemic ($P(S_{during})$), and after the mitigation of COVID-19 pandemic ($P(S_{post})$) because of a massive vaccination worldwide. Here, the comments grouped as *during-COVID-19* were mainly located in the negative zone (NA-red color) of the *Softmax-PANAS-tDL* function ($P(S) \geq 0$), which demonstrates the impact of COVID-19 pandemic on each tourist destination. The negative signs evidenced a reduction in comments related to a negative and positive sentiment category defined by the PANAS-t scale. The difference between the absolute maximum values during ($P(S_{during,k-1})$) and after ($P(S_{post,k})$) of COVID-19 pandemic ($dP(S) = P(S_k) - P(S_{k-1})$), showed that the tourist destinations that had the greatest recovery in positive perceptions were Italy ($dP(S)$: 0.113712), followed by Colombia ($dP(S)$: 0.051164), while destinations such as USA, Mexico and Spain showed negative values due to a process of recovery of traveller's confidence in these countries as tourist destinations.

Table 1. Evolution of Comments – Sentimental Perceptions

	Colombia		Spain		Italy		USA		Mexico	
Sentiment	P(S-During)	P(S-Post)	P(S-During)	P(S-Post)	P(S-During)	P(S-Post)	P(S-During)	P(S-Post)	P(S-During)	P(S-Post)
Guilt	0.056465	-0.056643	0.019868	-0.019774	0.032112	-0.032047	0.013579	-0.013499	0.013579	-0.013564
Fear	0.305962	-0.307832	0.231538	-0.231965	0.245387	-0.245526	0.217612	-0.216544	0.217612	-0.216802
Sadness	0.366306	-0.370158	0.463026	-0.453920	0.406452	-0.403617	0.498287	-0.492604	0.498287	-0.496203
Hostitlity	0.091046	-0.120604	0.138995	-0.111364	0.136010	-0.115629	0.128682	-0.147555	0.128682	-0.140594
Shyness	-0.119119	0.103163	-0.109932	0.113432	-0.098867	0.150162	-0.120482	0.130993	-0.120482	0.122976
Fatigue	-0.224235	0.149241	-0.230815	0.173178	-0.216778	0.330490	-0.226517	0.195144	-0.226517	0.182257
Surprise	-0.208547	0.275399	-0.225028	0.229960	-0.213736	0.118980	-0.217673	0.163197	-0.217673	0.169188
Jovialty	-0.135251	0.158894	-0.147466	0.177231	-0.144398	0.092062	-0.146045	0.210143	-0.146045	0.212535
Self-assurance	-0.074025	0.089618	-0.079466	0.052789	-0.081023	0.057921	-0.081751	0.111266	-0.081751	0.114779
Attentiveness	-0.038491	0.051345	-0.040268	0.031717	-0.042677	0.030013	-0.043040	0.034610	-0.043040	0.039395
Serenity	-0.020112	0.027577	-0.020451	0.038717	-0.022481	0.017190	-0.022652	0.024848	-0.022652	0.023900
Min-Max	0.224235	0.275399	0.230815	0.229960	0.216778	0.330490	0.226517	0.210143	0.226517	0.212535
Difference	0.051164		-0.000855		0.113712		-0.016374		-0.013982	

It is important to note the relocation of traveller's perceptions towards the positive zone of the *Softmax-PANAS-tDL* function because of the recovery of the countries' tourism sector. Figure 4 shows how the comments *post-COVID (grey line)* have similar structures to the PANAS-t scale for comments grouped for *before-COVID-19* set *(blue line)*. This similarity between structures can be evidenced through the IOAs achieved by the PANAS-tDL2 against the characterization of the *Softmax-PANAS-t* function (Table 2) for these periods. These IOAs reached values 90% on average, showing the relocation of travellers' comments toward the positive zone of the PANAS-t scale to different tourist destinations because of massive COVID-19 vaccination.

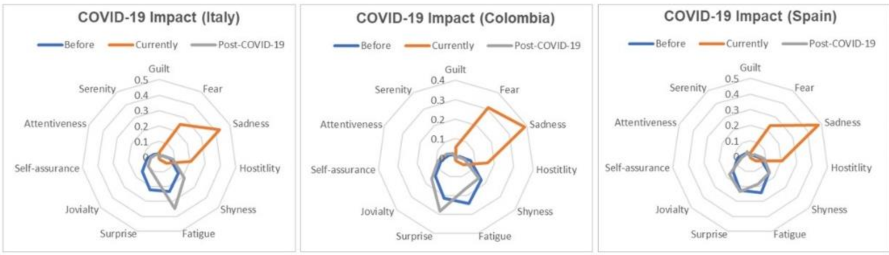

Fig. 4. Radar Chart Softmax Function – COVID-19 Impact

Table 2. Index of Agreement between the structure of Softmax function for before-COVID-19 and post-COVID-19.

	Colombia	**Spain**	**Italy**	**USA**	**México**
IOA	0.916567163	0.921887815	0.91721479	0.930779498	0.930779498

4 Conclusions and Future Work.

The proposed PANAS-TDL2 model made it possible to evaluate the impact of the COVID-19 pandemic on different tourist destinations worldwide, based on comments from travellers on a social network such as Twitter. To assess this impact, the proposed model carried out a psychrometric characterization of these comments, incorporating a novel activation function of the *Softmax* type, which was inspired in the PANAS-t scale. Thanks to its capacity for adaption and learning and its deep learning structure, the proposed model correctly identified the sentiment structure that defines the PANAS-t scale for different periods of COVID-19 pandemic evolution and its impact on different tourist destinations.

In the study carried out, it can be seen how the comments posted by travellers before the COIVD-19 pandemic were located in the positive zone of PANAS-t scale, while during COVID-19 pandemic, these comments evolved towards the negative zone of this scale, as a result of a deterioration of travellers' confidence against in visiting a tourist destination. It is essential to note that as a result of massive vaccination worldwide, these comments shifted towards the positive zone of PANAS-t scale, which shows the recovery of travellers' confidence on different tourist destinations. It is also important to note that this relocation of travellers' perceptions has been heterogeneus, mainly due to the recovery dynamics of the countries as tourist destinations related to vaccination and security.

According to the above, the proposed PANAS-tDL2 model can be extended to characterise plans, services and gastronomic or sensory experiences in an specific tourist destination. The latter will allow the creation of tourism products according to travellers' perceptions based on social networks. Likewise, the proposed PANAS-tDL2 model can be extended to create tourism insurances in agreement with the positive travellers' perceptions of a tourist destination using social networks.

Based on the results of the PANAS-tDL2 model, the authors propose as future work the application of this model to assess the risk of an organisation's business operations when it has a strong presence, or its activity depends to a large extent on social networks. This risk assessment will allow the creation of new assurance schemes by reputation, based on the people's perceptions in social networks would lead to new reputation assurance schemes in social networks, according to people's perceptions of a product or service based on comments posted on a social network such as Twitter. Likewise, due to its capacity for adaptation and learning, the model can be extended to evaluate these perceptions on other social networks such as Facebook and Instagram or specialised travel platforms such as TripAdvisor or Booking.

References

Arbulu, I., Razumova, M., Rey-Maquieira, J., & Sastre, F. (2021). Can domestic tourism relieve the COVID-19 tourist industry crisis? The case of Spain. *Journal of Destination Marketing & Management, 20*.

Darmawiguna, I., Pradnyana, G., & Santyadiputra, G. (2019). The Development of Integrated Bali Tourism Information Portal using Web Scrapping and Clustering Methods. *The 1st International Conference on Vocational Education and Technology. 1165(2019)*, págs. 1-10. IOP Publishing. doi:doi:10.1088/1742-6596/1165/1/012010

Dewi, L., & Alvin, M. (2019). Social Media Web Scraping using Social Media Developers API and Regex. *Procedia Computer Science 157*, 444-449. doi:https://10.1016/j.procs.2019.08.237

Dinarto, D., Wanto, A., & Sebastian, L. (2020). *COVID-19: Impact on Bintan's Tourism Sector.* Singapore: RSIS - Nanyang Technological University.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research 270(2)*, 654-669.

Han, S., & Anderson, C. (2021). Web Scraping for Hospitality Research: Overview, Opportunities, and Implications. *Cornell Hospitality Quarterly Vol. 62(1)*, 89-104. doi:https://doi.org/10.1177/1938965520973587

Haryanto, T. (2020). EDITORIAL: COVID-19 PANDEMIC AND INTERNATIONAL TOURISM DEMAND.

JDE(JournalofDevelopingEconomies) Col. 5(1). doi:http://dx.doi.org/10.20473/jde.v5i1.19767

IATA. (2020). *COVID-19 Impact on Asia-Pacific Aviation Worsens.* Singapore: IATA Press Room. Obtenido de https://www.iata.org/en/pressroom/pr/2020-04-24-01/

Johnson, P., Sieber, R., Magnien, N., & Ariwi, J. (2012). Automated web harvesting to collect and analyse user-generated content for tourism. *Current Issues in Tourism Vol 15(3)*, 293-299. doi:http://dx.doi.org/10.1080/13683500.2011.555528

Kalvet, T., Olesk, M., Tiits, M., & Raun, J. (2020). Innovative Tools for Tourism and Cultural Tourism Impact Assessment. *Sustainability 17(7470)*. doi:doi:10.3390/su12187470

Kamata, H. (2022). Tourist destination residents' attitudes towards tourism during and after the COVID-19 pandemic. *Current Issues in Tourism 25(1)*, 134-149. doi:https://doi.org/10.1080/13683500.2021.1881452

Kumar, D., & Nafi, S. (25 de 06 de 2021). *Impact of COVID-19 Pandemic on Tourism: Perceptions from Bangladesh*. Obtenido de SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3632798

Mich, L. (2022). AI and Big Data in Tourism. En Springer, *Applied Data Science in Tourism* (págs. 3-15). doi:https://10.1007/978-3-030-88389-8_1

Mostafa, G., Ahmed, I., & Junayed, M. (2021). Investigation of Different Machine Learning Algorithms to Determine Human Sentiment Using Twitter Data. *Information Technology and Computer Science 2*, 38-48. doi:https://10.5815/ijitcs.2021.02.04

Peña, A., Mesias, J., Patiño, A., & de Carvalho, J. (2021). PANAS-TDL: A Psychrometric Deep Learning Model for Characterizing Sentiments of Tourists Against the COVID-19 Pandemic on Twitter. En A. Abreu, D. Liberato, E. Gonzalez, & J. Garcia, *Smart Innovation, Systems and Technologies, 209.* Springer. doi:https://10.1007/978-981-33-4260-6_15

Peña, A., Patiño, A., Chiclana, F., Caraffini, F., Gongora, M., Gonzalez-Ruiz, J., & Duque-Grisales, E. (2021). Fuzzy convolutional deep-learning model to estimate the operational risk capital using multi-source risk events. *Applied Soft Computing 107*. doi:https://doi.org/10.1016/j.asoc.2021.107381

Rahman, M., Gazi, M., Bhuiyan, M., & Rahaman, A. (2021). Effect of Covid-19 pandemic on tourist travel risk and management perceptions. *PLOS ONE 16(9)*. doi:https://doi.org/10.1371/journal.pone.0256486

Rubio-Cisneros, N., Montero-Muños, J., Rubio-Cisneros, I., Morales-Ojeda , S., & Pech, M. (2022). Tourists perceptions of Caribbean islands facing environmental threats before the COVID-19 health crisis: Holbox Island and Archipelago of Bocas del Toro. *PLOS ONE 17(3)*. doi:https://doi.org/10.1371/journal.pone.0262014

Sherman, S., Joy, J., Dsouza, D., & Kaul, P. (2020). Dynamic Review Modelling and Recommendation of Tourism Data. *Proceedings of the International Conference on Recent Advances in Computational Techniques (IC-RACT) 2020.* SRRN. doi:http://dx.doi.org/10.2139/ssrn.3685669

Usui, R., Sheeran, L., Asbury , A., & Blackson, M. (2020). Impacts of the COVID-19 pandemic on mammals at tourism destinations: a systematic review. *Mammal Review*, 492-508. doi:10.1111/mam.12245

Viñan-Ludeña, M., & de Campos, L. (2021). Analyzing tourist data on Twitter: a case study in the province of Granada at Spain. *Journal of Hospitality and Tourism Insights*. Obtenido de https://www.emerald.com/insight/2514-9792.htm

1.