Models for Tourist Behavior Analysis Based on Neural Network

Sergei Mikhailov

ITMO University

St. Petersburg, Russia

samikhailov@itmo.ru

Abstract—**The tourism is one of the profitable spheres of the world economy. Tourists actively use mobile devices and e-tourism services during their trips. It is possible to use tourist behaviour analysis results for tourism services improvement. The paper shows the related work of the neural network based model for behaviour analysis.**

I. Introduction

Tourism industry is one of the drivers of the global economy and it grows rapidly over the last years. According to the UNWTO World Tourism Organization (https://www.e-unwto.org/doi/pdf/10.18111/9789284421152), the total number of international tourist trips and the international tourism income are constantly increasing. Around 1.5 billions international tourist arrivals were recorded in 2019 worldwide. The number of international travels are expected to increase by 3.3% per year between 2010 and 2030.

There is a substantial research development in IT and tourism over twenty years [1]. A huge increase in the use of the mobile technology among tourists have been happened over the past ten years [2]. Another example of the interaction of IT and tourism is the massive use of e-tourism services. E-tourism services provide a wide range of diverse functions, such as recommendations, trip planning, points of interests (POI) guide, that enhance the tourism experience.

It is possible to improve the quality of services by using the tourist behaviour analysis methods. The tourist-generated content such as photo/videos or POI reviews, spatial movement information, various sensors data can be used as behaviour data sources. The scientists use the Big Data methods and machine learning approach to construct models due to the big amount of behaviour information [3], [4]. Different approaches are used for the tourist behaviour models construction [5]. One of the possible solutions for behaviour analysis is working with neural networks. The neural networks models show more accurate results, but requires a lot of information for training purposes.

The presented paper aimed to show neural network usage for tourist behaviour analysis. The section II presents related works for different machine learning techniques for behaviour analysis and section III provides conclusion.

II. Related work

Different machine learning techniques can be used for the tourist behaviour model creation. The following approaches of machine learning which can be used in behaviour analysis is classification, clustering and time-series prediction. The chosen approaches is actively used in scientists researched and covers most analysis tasks.

A. Classification

The authors of paper [6] use deep neural network method for graph nodes classification. The deep neural network is composed of stacked sparse autoencoders and softmax layer, which could learn the node representation while encoding the rich nonlinear structural and semantic information. The positive pointwise mutual information matrix from the adjacency matrix is used ad source of neural network model. The proposed model has good performance and good classification rate.

The authors of paper [7] describe tourist behaviour mining from analysing photo content by using a computer deep learning model. 35356 Flickr tourists photos are identified into 103 scenes and analysed by ResNet-101 Deep learning model. Tourists cognitive maps with different perceptual themes are visualized by the authors according to photos geographical information. Statistical analysis and spatial analysis (by using hierarchical clustering analysis and ANOVA (analysis of variance)) are used for analysing tourist behaviour.

The work [8] proposes the sentiment analysis from tourist POI review. The analysis is based on bidirectional gated recurrent unit neural network model. This model consists of a topic model (lda2vec) and an attention mechanism. Lda2vec is used to discover the main topics of the review for the word vector. The attention mechanism is used to learn how to measure different weights of the words in the text to the overall meaning of it. The constructed model shows a good performance based on experiments with IMDB dataset.

The authors of paper [9] propose a convolution neural network for food labeling for tourists. Some foods from Indonesia have similarities, so food labeling assists to introduce meals guides and suggestions for tourists. Convolution neural network was chosen for reliable and fast process of classification of a complex and detail objects. The Indonesian food photos are used as main source of data.

B. Clustering

The paper [10] aims to identify the profiles of visitors to a heavily used tourist destinations using the self-organizing maps (SOM) analytical method. Around 2500 tourist interviews are used as main data source. Four heterogeneous tourist profiles

types are identified. The SOM method belongs to unsupervised learning algorithm and is widely used in clustering tasks.

The author [11] uses the SOM and Expectation Maximization (EM) techniques for data clustering in recommendation systems. The users ratings clusterization is achieved by using this model. The SOM model is applied on TripAdvisor datasets.

The work [12] clusters a textual conversational topics which shared through management tweets of upper class hotels in Philadelphia. The self-organizing map is used as analysing instrument. The Twitter and Google Trends data are mainly used as data sources.

C. Time series prediction

The work [13] presents the monthly tourist arrival forecasting framework. The framework consists of four steps: search engine platforms identification, data collection, deep learning model training and model result interpretation. The authors introduce the deep network architecture with historical time-series tourism demand data. The attention mechanism support has been added to the long short-term memory neural network model (LSTM). The attention mechanism makes the LSTM model more interpretive and allows the model to ignore irrelevant information.

The authors of paper [14] use the neural network autoregressive algorithm for monthly tourism demand forecasting of ten European countries. The denoising algorithms are used for prediction accuracy improvement. The singular spectrum analysis is used for denoising European tourist arrivals data. The proposed algorithm is compared with autoregressive integrated moving average and exponential smoothing models and result shows that proposed algorithm is more accurate.

The paper [15] describes a neural networks composition for a time series forecasting. Back-propagation neural network, dynamic architecture for artificial neural network, Elman artificial neural network and echo state network are selected as component forecasting models by authors. An input-hidden selection heuristic is designed to determine the input-hidden neuron combination for each component neural network. In-sample training–validation pair-based neural network weighting mechanism is studied to generate the associated combination weights.

The work described in [16] proposes to use LSTM methods to predict tourism flow. The authors state that LSTM methods perform better than auto regressive integrated moving average models and back propagation neural networks. The Adam optimizer with adaptive learning rates is used for the neural network optimization. The simplex and stacked LSTM networks are used for prediction and performance experiments.

The paper [17] uses sequential patterns of tourist activities and locations from social medias as main source of data. The convolutional long short-term deep learning method is used for prediction of the expected location. The proposed method combines convolutional neural network with LSTM. The paper authors state that their solution outperforms other neural network models when evaluating with the accuracy and loss metrics.

The work described in [18] presents a paired neural network model for tourist demand forecasting. The tourist arrival data is decomposed by two low-pass filters into long-term trend and short-term seasonal components, which are then modelled by a pair of autoregressive neural network models as a parallel structure. The model is evaluated by the tourists arrival data to United States from twelve source markets.

III. CONCLUSION

The classification of neural network models are used in different situation: for social media nodes classification, media content auto-labeling and sentiment analysis. For photo analysis usually the convolutional neural networks are used. The deep neural networks are more suited for a graph node and sentiment analysis.

The clustering analysis by neural networks not so widely represented in scope of tourist behaviour analysis. The main idea is usage of the self-organizing maps as unsupervised learning algorithm. The various textual data such as tourist questionnaires, ratings and reviews are used in this kind of tasks.

One of the most popular uses of neural networks for analyzing tourist behavior is time series prediction. The tourist demand is most popular target for behaviour model analysis. The LSTM models are widely used for this case because of the capability to "memorize" previous states. Some researches try to use composition of different neural networks as prediction solution.

REFERENCES

[1] Zheng Xiang, "From digitization to the age of acceleration: On information technology and tourism", *Tourism Management Perspectives*, vol. 25, 2018, pp. 147–150.

[2] Tan, Garry & Lee, Voon & Lin, Binshan & Ooi, Keng-Boon. "Mobile applications in tourism: the future of the tourism industry?", *Industrial Management & Data Systems*, vol 117, 2017, pp. 560–581.

[3] Miah, S. J., Vu, H. Q., Gammack, J., McGrath, M. "A Big Data Analytics Method for Tourist Behaviour Analysis", *Information & Management*, vol. 54, no. 6, 2017, pp. 771–785.

[4] Marine-Roig E., Clavée S.A., "Tourism analytics with massive user-generated content: A case study of barcelona", *Journal of Destination Marketing & Management*, vol. 4, no. 3, 2015, pp. 162–172.

[5] Jingjing Li, Lizhi Xu, Ling Tang, Shouyang Wang, Ling Li, "Big data in tourism research: A literature review", *Tourism Management*, vol. 68, 2018, pp. 301–323.

[6] Bentian Li, Dechang Pi, "Learning deep neural networks for node classification", *Expert Systems with Applications*, vol. 137, 2019, pp. 324–334.

[7] Kun Zhang, Ye Chen, Chunlin Li, "Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing", *Tourism Management*, vol. 75, 2019, pp. 595–608.

[8] Li, Qin, Li, Shaobo, Hu, Jie, Zhang, Sen, Hu, Jianjun, "Tourism Review Sentiment Classification Using a Bidirectional Recurrent Neural Network with an Attention Mechanism and Topic-Enriched Word Vectors", *Sustainability* , vol. 10, 2019, p. 9.

[9] R. P. Prasetya, F. A. Bachtiar, "Indonesian food items labeling for tourism information using Convolution Neural Network", *2017 International Conference on Sustainable Information Engineering and Technology (SIET)*, Malang, 2017, pp. 327–331.

[10] Taczanowska, Karolina, González, L.M., García-Massó, Xavier, Zieba, etc., "Nature-based Tourism or Mass Tourism in Nature? Segmentation of Mountain Protected Area Visitors Using Self-Organizing Maps (SOM)", "Sustainability", vol. 11, no. 1314, 2019.

[11] Nilashi, M., Bagherifard, K., Rahmani, M., Rafe, V. "A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques", *Computers & Industrial Engineering*, vol. 109, 2017, pp. 357–368.

[12] Le T., Pardo P., Claster W. "Application of artificial neural network in social media data analysis: a case of lodging business in Philadelphia", *Artificial Neural Network Modelling*, 2016, pp. 369–376.

[13] Rob Law, Gang Li, Davis Ka Chio Fong, Xin Han, "Tourism demand forecasting: A deep learning approach", *Annals of Tourism Research*, vol. 75, 2019, pp. 410–423.

[14] Emmanuel Sirimal Silva, Hossein Hassani, Saeed Heravi, Xu Huang, "Forecasting tourism demand with denoised neural networks", *Annals of Tourism Research*, vol. 74, 2019, pp. 134–154.

[15] Lin Wang, Zhigang Wang, Hui Qu, Shan Liu, "Optimal Forecast Combination Based on Neural Networks for Time Series Forecasting", *Applied Soft Computing*, vol. 66, 2018, pp. 1–17.

[16] YiFei Li, Han Cao, "Prediction for Tourism Flow based on LSTM Neural Network", *Procedia Computer Science*, vol. 129, 2018, pp. 277–283.

[17] J. Kanjanasupawan, Y. Chen, T. Thaipisutikul, T. K. Shih, A. Srivihok, "Prediction of tourist behaviour: Tourist visiting places by adapting convolutional long short-term deep learning", *2019 International Conference on System Science and Engineering (ICSSE)*, July 2019, pp. 12–17.

[18] Yuan Yao, Yi Cao, Xuemei Ding, Jia Zhai, Junxiu Liu, Yuling Luo, Shuai Ma, Kailin Zou, "A paired neural network model for tourist arrival forecasting", *Expert Systems with Applications*, vol. 114, 2018, pp. 588–614.