Advances in Intelligent Systems and Computing 1430

Daniel H. de la Iglesia
Juan F. de Paz Santana
Alfonso J. López Rivero *Editors*

New Trends in Disruptive Technologies, Tech Ethics and Artificial Intelligence

The DITTET 2022 Collection

Springer

An Advertising Real-Time Intelligent and Scalable Framework for Profiling Customers' Emotions

Leandro Alves[1], Pedro Oliveira[1], João Henriques[1,2,3](✉), Marco V. Bernardo[1,4], Cristina Wanzeller[1,3], and Filipe Caldeira[1,2,3]

[1] Polytechnic of Viseu, Viseu, Portugal
{pv23844,estgv9081}@alunos.estgv.ipv.pt,
{joaohenriques,mbernardo,cwanzeller,caldeira}@estgv.ipv.pt
[2] University of Coimbra, Coimbra, Portugal
[3] CISeD – Research Centre in Digital Services, Polytechnic of Viseu, Viseu, Portugal
[4] University of Beira Interior, Covilhã, Portugal

Abstract. The advertising industry is continuously looking up for effective ways to communicate to customers to impact their purchasing. Usually, profiling them is a time-consuming offline activity. Therefore, it becomes necessary to reduce costs and time to address consumers' needs. This work proposes a scalable framework enabled by a Machine Learning (ML) model to profile customers to identify their emotions to help to drive campaigns. A multi-platform mobile application continuously profiles consumers crossing the front stores. Profiling customers according to their age and hair color, the color of their eyes, and emotions (e.g. happiness, sadness, disgust, fear) will help companies to make the most suitable advertisement (e.g. to predict whether the advertising tones on the front store are the adequate ones). All that data are made available in web portal dashboards, wherein subscribers can take their analysis. Such results from the analysis data help them to identify tendencies regarding the culture and age, and drive companies to fit front stores accordingly (e.g. to discover the right tones for the season). This framework can help to develop new innovative cost-effective business models at scale by driving in real-time the advertisements to a huge number of consumers to maximize their impact and centralizing the data collected from a large number of stores to design future campaigns.

Keywords: Machine Learning · Cognitive services · IoT · Advertising · Scalability

1 Introduction

The advertising industry is continuously looking up for effective ways to communicate to customers to impact their purchasing. However, profiling customers takes a huge burden and it is a time-consuming offline activity. Therefore, it

D. H. de la Iglesia et al. (Eds.): DiTTEt 2022, AISC 1430, pp. 55–68, 2023.
https://doi.org/10.1007/978-3-031-14859-0_5

becomes necessary to understand consumers to reduce costs and impact their decisions on purchasing. In that regard, a large range of several known methods exist helping them to collect consumer satisfaction from surveys (e.g. in the paper, online, mobile). Unfortunately, such an approaches can become expensive and outdated (not reflecting the current opinion), not confidential, and time-consuming (questionnaires too long), that may result in skewed answers when consumers become upset. Subsequently, data should be analyzed, conclude and when we use phone calls or paper surveys, we still need to translate assessments into digital format to support further analysis. The advertisement industry is trying to reduce the burden on non creative work.

This paper proposes extracting genuine reactions and determining the age, gender, and consumers characteristics moving around shops helping driving campaigns. Video from customers was captured to predict emotions and extract features (e.g. age, gender, makeup, hair color, glasses). This work can leverage new innovative business models supported by the use of time slots on the public screen with advertising content to well-defined targets to guarantee the right audience for announcements and scale their application to a huge number of front stores and companies.

The remainder of this paper is organized as follows. Section 2 presents the related work. The Sect. 3 presents the proposed framework. Section 4 presents the experimental work achieved results and discussion. Finally, Sect. 5 concludes the paper.

2 Related Work

Nguyen et al. [1] conducted a comprehensive review on existing approaches in store layout design and modern AI techniques that can be utilized in the layout design task.

Chen et al. work [2] was focused on achieving a sustainable user experience with a smart system installed in an autonomous store to understand in-store customer behavior by continuously observing and refining the user experience.

Newman et al. [3] tried to understand the behaviour of customers by tracking them with in-store CCTV cameras to detect patterns of behaviour for optimising the space and store performance.

Mete Taspinar et al. [4] have worked on a similar solution aiming to capture the attention of consumers. They set a billboard with a video and when a consumer approaches, and then, a camera takes a photo to be delivered to a server, collects facial insights, substitutes the face on the video with the one collected, and with this dynamics, retrieves the consumer attention during the promotional video.

Taspinar et al. [5] put its focus on purchases from consumers by providing billboards in the store to stimulate consumer purchases. Their methods included a ML algorithm to identify the products fitting with the ones already purchased. The selection of the products de+ends on the inventory, wherein the selected products are old in stock to combine.

Advertising Technologies [6] already reported a similar solution. Unfortunately, the company suspended all activity due to the pandemic and detailed information was not found.

3 Proposed Framework

This section presents a reference architecture of the proposed framework. It presents the modules, implementation aspects and technologies. We will discuss the functional aspects of the proposed framework provided by each one of the modules as well the non-functional aspects to scale their functions and achieve good performance, resiliency, and availability.

3.1 Architecture

This section presents a reference architecture of the proposed framework, addressing the aforementioned requirements, as presented in Fig. 1. It includes the modules for Cognitive Inference, Face Detection, Message Broker, Core Engine, Database, and Reporting.

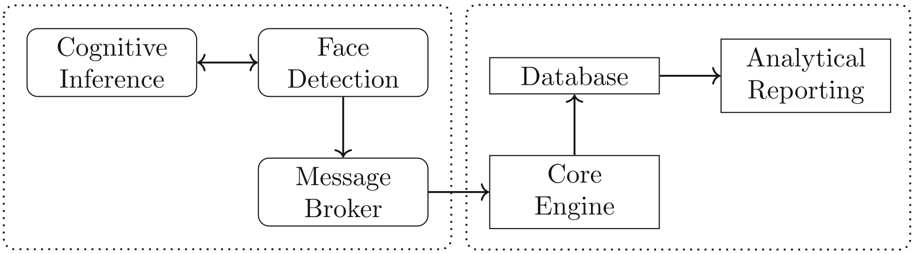


Fig. 1. Architecture scheme

The Face Detection Module takes the responsibility on detecting faces and capturing video frames to be delivered to the Cognitive Inference Module. After getting a response, this module returns the classified data to be moved to the Message Broker Module.

The Cognitive Inference Module interacts with the Face Detection Module to classify features from the received frame. These features include the consumer's characteristics (age and gender) emotions (anger, contempt, disgust, fear, happiness, sadness, and surprise), and the other ones (id, timestamp, store, campaign).

The Message Broker Module relies on the Publisher-Subscriber pattern and streaming capabilities to intermediate the communication in real-time. It receives the data from the Cognitive module and stores them temporally and made available to the Core Engine Module. This module decouples the gathering data stage from the analysis stage.

The Core Engine Module is decoupled from the Cognitive Inference Module and takes the responsibility of collecting the already classified data to store them in a database.

Finally, the Analytical Reporting module provides the capabilities to graphically present the data already available in the database and help to profile the consumers in future campaigns according to their emotions, characteristics, and locations.

3.2 Implementation

The adopted technologies in the implementation of the proposed framework can be depicted in Fig. 2.

The Face Detection Module of the proposed framework was implemented as a mobile application. It was developed in Xamarin [7] can be installed on mobile devices (e.g. smart devices, tablets, smartphones, and IoT devices) for Android, and IOS. It takes the video from cameras to collect data to detect human faces to be classified. The [8] ML Kit is a well-trained Google library providing the detection capabilities. After detecting a face, the frame is delivered to the Cognitive Inference Module for classification purposes. The extracted features include the user's emotions and physical characteristics. Then, the classified data will be moved to the Message Broker Module.

The Cognitive Inference Module was deployed with Microsoft Cognitive Services to extract emotional attributes. Figure 3 depicts the Azure Cognitive Services Architecture. One of the key points was the selection of the technology behind this implementation. While YOLOv5 [9] is much faster than Microsoft Cognitive Services [10,11], it requires more local computing resources. Hence the option was to use Microsoft Cognitive Services. Microsoft Cognitive Services [10,11] provides a solution for emotional recognition through the use of web service invocation. Two services have been configured in Azure: Event hub and Cognitive Services Face API.

The Message Broker Module was implemented as a service in Azure Event Hub [12], including a topic to collect the received data. It provides good performance for real-time data ingestion and micro batching on the same stream. Event Hubs is a fully managed, real-time data ingestion service that's simple, trusted, and scalable. Therefore, the platform is robust, reliable, scalable, works with [13] Advanced Message Queuing Protocol (AMQP), process 5000 concurrent AMQP connection, process 1 MB/s or 1000 events/s ingress and 2 MB/s or 2000 events/s egress, and make use of a shared access key authentication for secure proposes. Streaming millions of events per second from sources to build dynamic data pipelines and take immediate responses to business challenges. Integrate seamlessly with other Azure services to unlock valuable insights. At this stage, the data is available to subscribers of the topic. In the proposed framework, the collected data will be made available for further analytical reporting processing in real-time. With an Azure account, two services were set up: Event hub and Cognitive Services Face API. The retention time was set to one day.

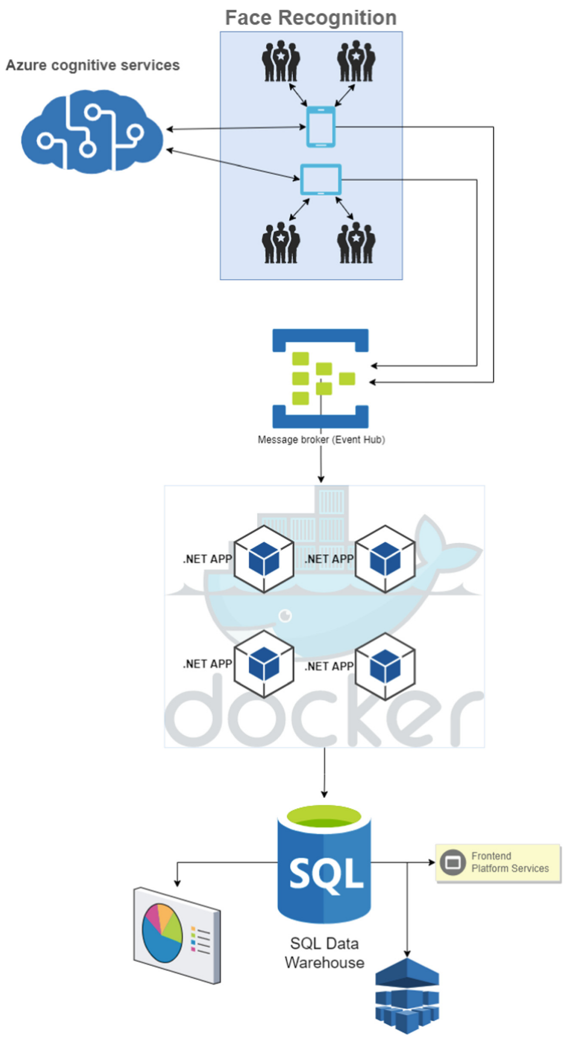


Fig. 2. Implementation framework

The Core Engine Module was implemented as a .NET Core service. .NET Core is a new Microsoft framework facilitating to have applications to Docker runtimes [14,15] on Linux or Windows. It was deployed as a Docker container leveraging the scalability of the framework. It retrieves the data from the Message Broker Module to be stored in the Database Module. The service subscribes to the topic of the Message Broker Module for listening to new events, collecting, and saving them in the Database. The application will listen for new events to be stored in an Azure SQL Database. The scaling process can be triggered as a result of the volume of data arriving at the event hub since they overcome the

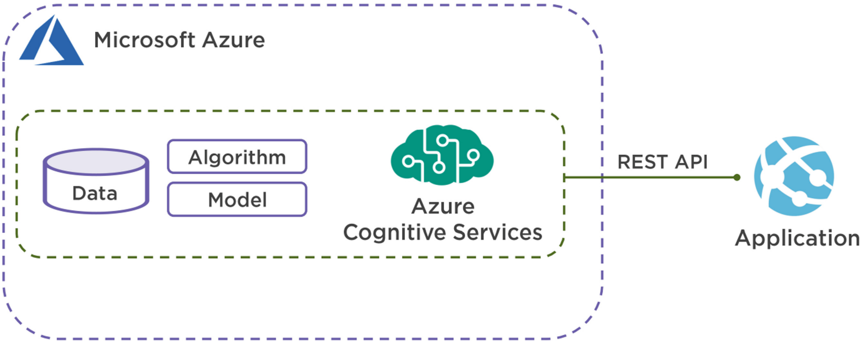


Fig. 3. Azure cognitive services architecture

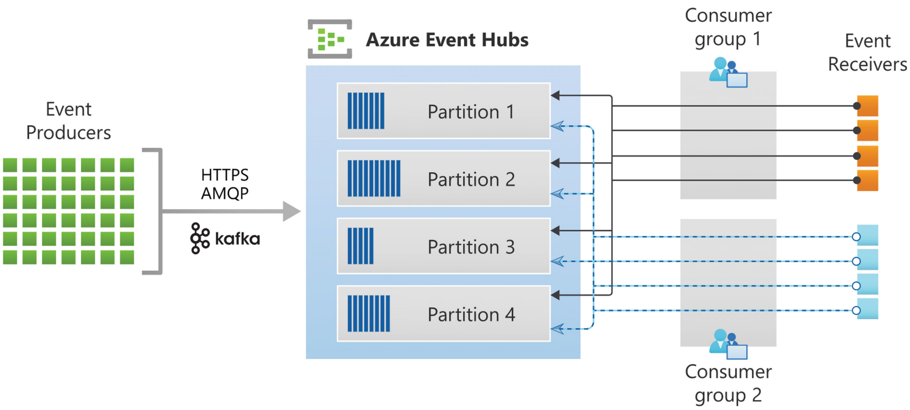


Fig. 4. Event hub architecture

given threshold. This can happen in case the number of IoT devices increases (Fig. 4).

The Database Module was deployed under SQL Azure. Its auto-scale mechanisms allow for an increase or decrease in the allocation of resources according to the required volume of data. The produced graphics include features for age average, gender, and emotional attributes by data interval, crossing campaigns with people's predicted reactions. and saved in Azure SQL Database to support further processes and analyses. The Database Module was implemented as a database on Azure will be deployed and configured. Therefore, the mobile application will monitor for new faces and once they are detected, all the information will be recorded into the database.

The Analytical Reporting Module provides the graphical capabilities to present the incoming data. This module was implemented in Grafana [16] to graphically present the collected data to operators. It summarizes the metrics from the user-defined queries against the data already stored in the Database

Module. The decision to select Grafana results from being an open-source tool, high scalable, lightweight, and boasting a responsive design. Figure 5 depicts a diagram of the architecture, wherein Grafana extracts the metrics by securely connecting to the Azure Database and takes SQL queries to collect the data by applying date intervals filters.

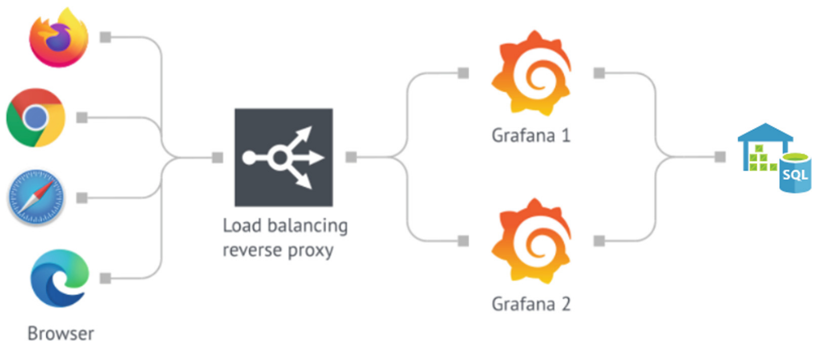


Fig. 5. Grafana—SQL AZURE architecture

Finally, a simulation tool was implemented in .NET producing the data feeding the framework. This data already includes the end features related with the customer profile, such as the gender and sentiment.

4 Experimental Work

Due to privacy concerns and the number of required physical stores involved, the option was to produce a dataset for simulation purposes. It will help to demonstrate the effectiveness of the framework and its scalability. In that aim, a significant amount of data of messages helped to evaluate the framework performance by the use of an application producing the events at a certain pace, according to several defined events per second, and specific ranges for age, campaign, and locals and 8 emotions (Anger, Contempt, Disgust, Fear, Happiness, Sadness, and Surprise).

The collected data in the dataset can help to extract insights aiming to fit the consumer's needs. The most valuable features include age, gender, and emotional status.

To take the experimental work of the proposed framework 27439 records were collected into a dataset with 13 features. These features to profile consumers include the ones related to their physical characteristics (age and gender) and emotions (Anger, Contempt, Disgust, Fear, Happiness, Sadness, and Surprise) and the other ones (Id, Timestamp, Store, Campaign) (Fig. 6).

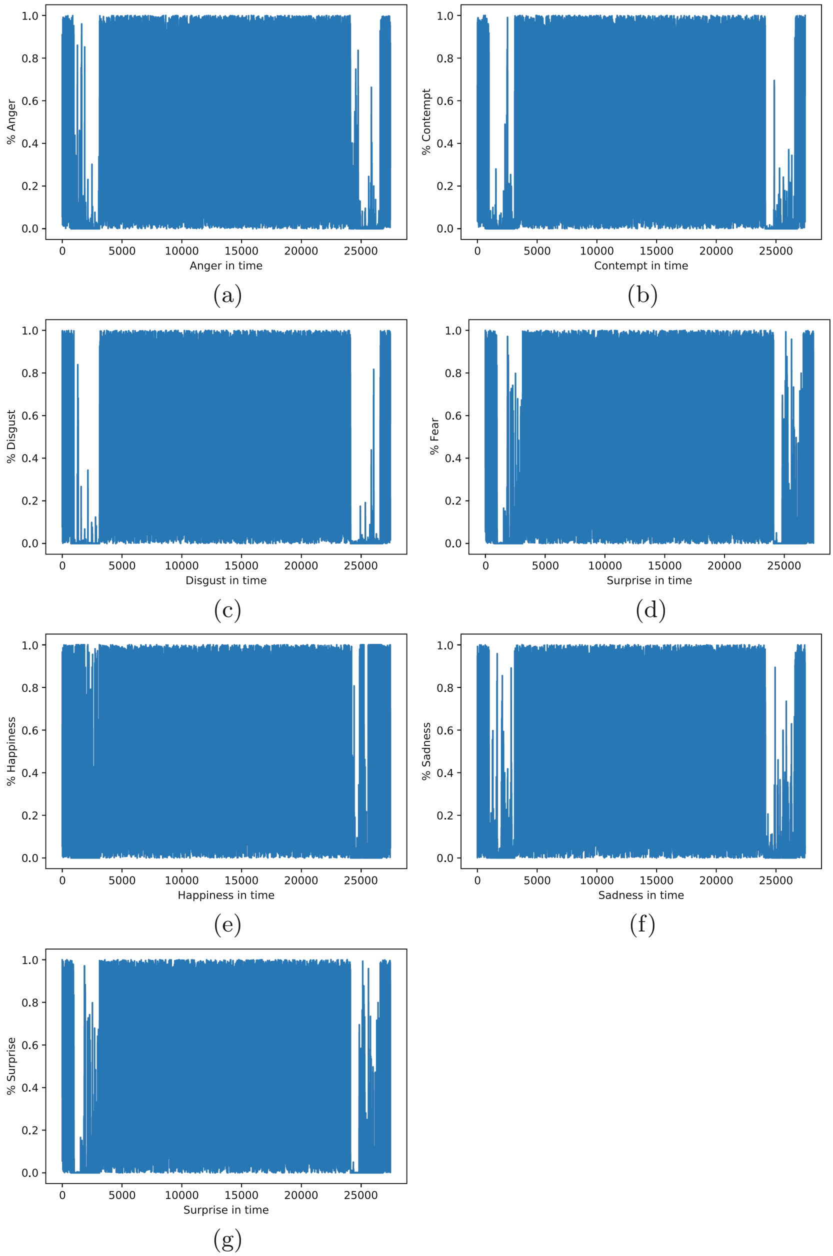


Fig. 6. Emotions in time: a) Anger, b) Contempt, c) Disgust, d) Fear, e) Happiness, f) Sadness and g) Surprise

Figure 7 depicts the frequency histograms for emotions.

Figure 8 represents the different boxplot values for emotions: anger, contempt, disgust, fear, Happiness, Sadness and Surprise.

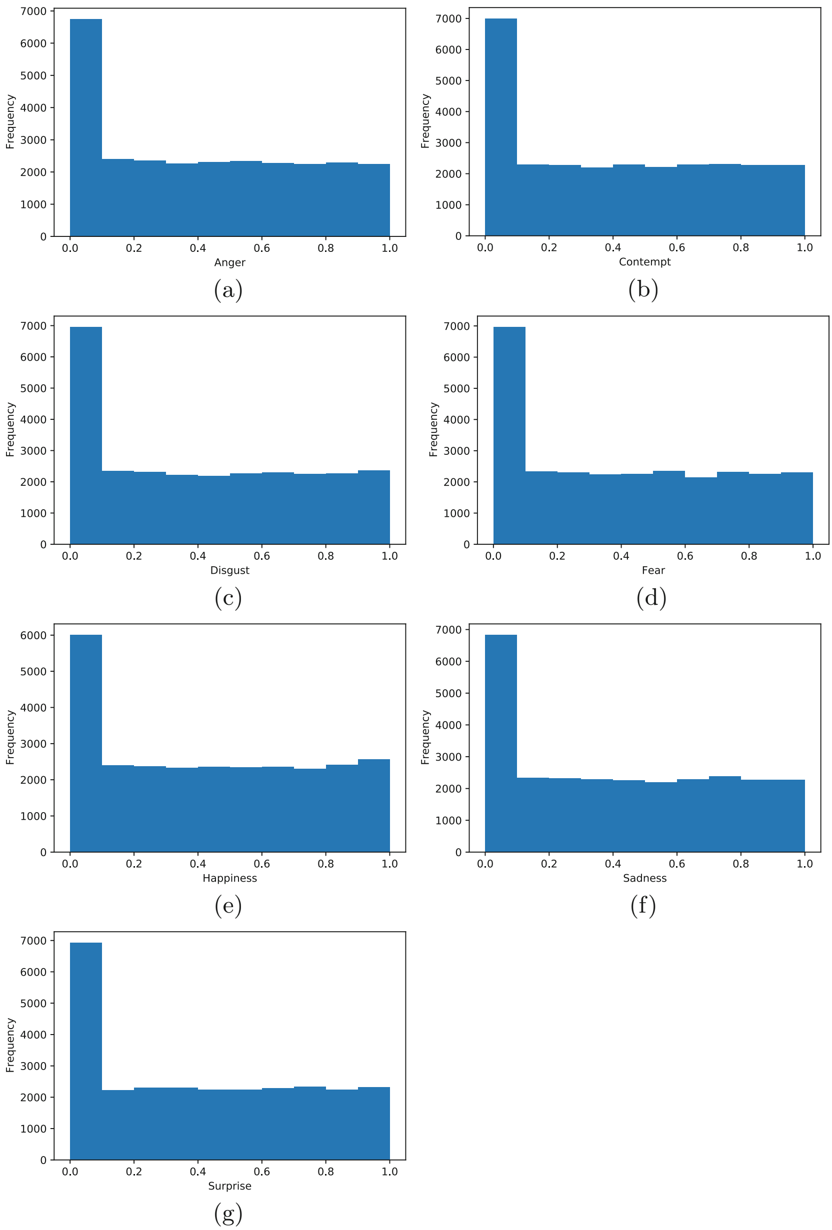


Fig. 7. Frequency histogram for: a) Anger, b) Contempt, c) Disgust, d) Fear, e) Happiness, f) Sadness and g) Surprise

Figure 9 summarizes the occurrences in the dataset by gender and store. Figure 10 provides the number of occurrences in the dataset by campaign. Figure 11 depicts the correlation between age in the Happiness.

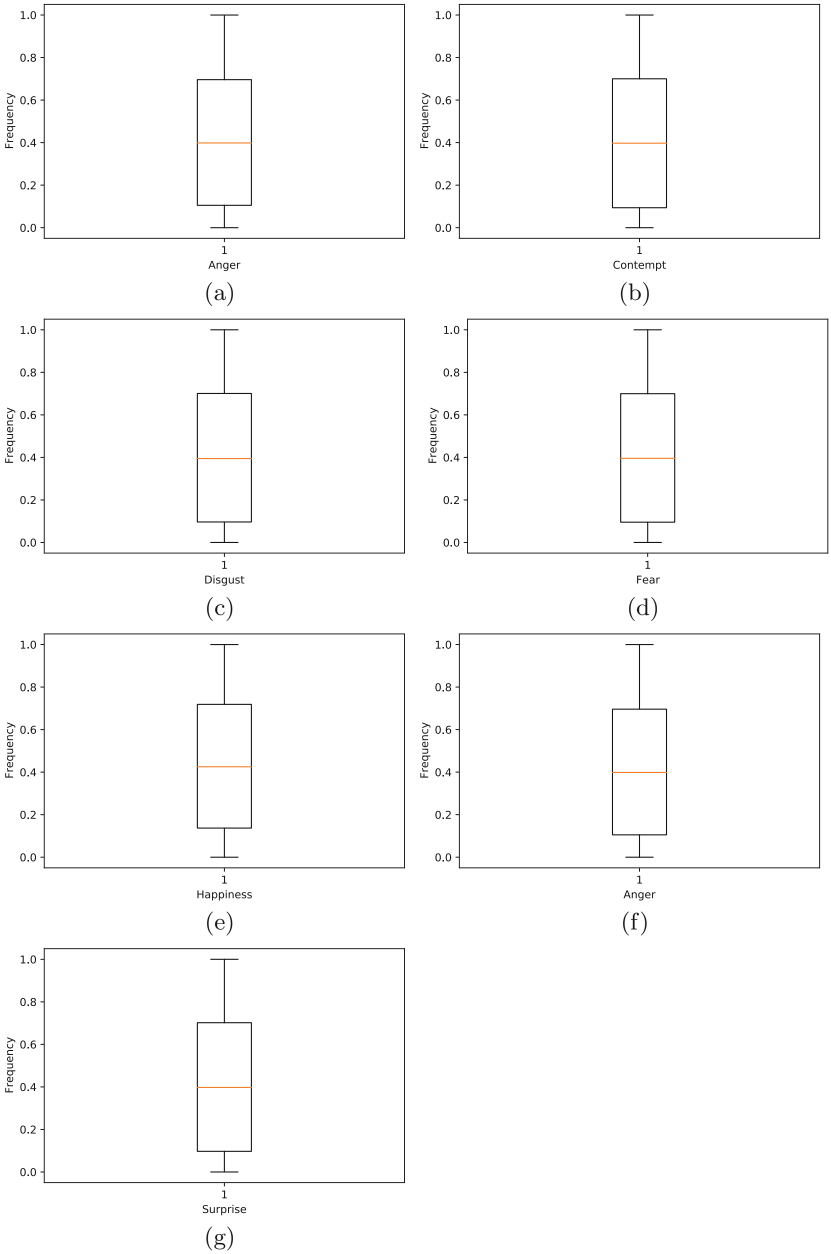
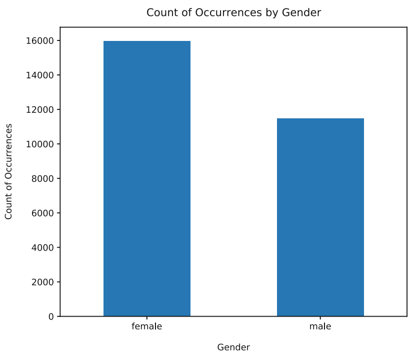
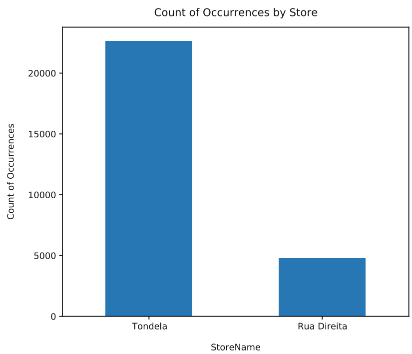


Fig. 8. BoxPlot values for a) Anger, b) Contempt, c) Disgust, d) Fear, e) Happiness, f) Sadness and g) Surprise



(a)



(b)

Fig. 9. Occurrences by a) Gender and b) Store

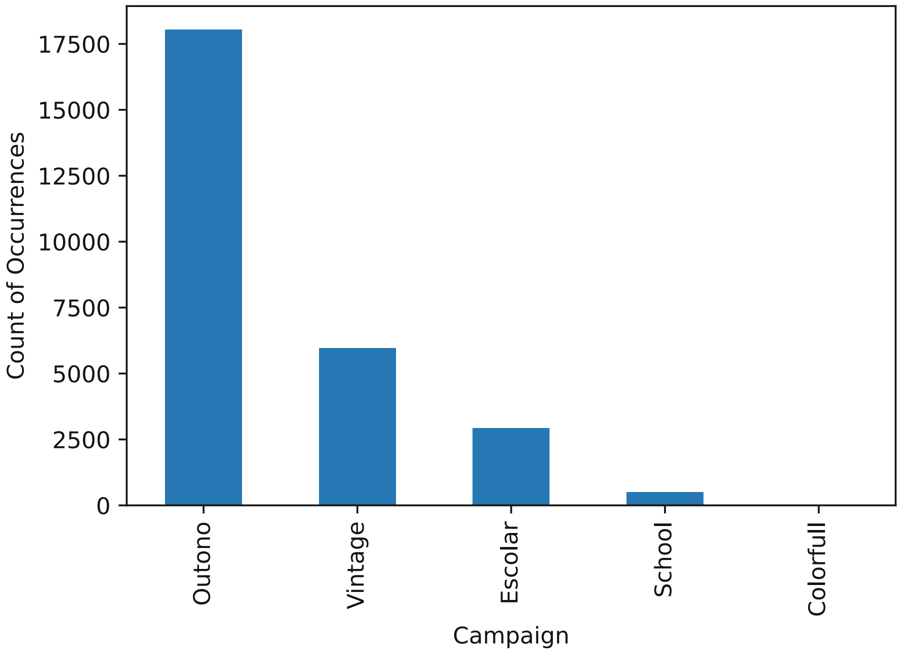


Fig. 10. Occurrences by Campaign

Figure 12 represents the correlation between all the numerical features in the dataset.

The datasets gathered the produced data from the simulator. Notwithstanding, the simulator did not considered key quality aspects in the generated data. For example, the dependency between parameters were not considered because there exist opposite emotions in the same instant, such as happiness and sad-

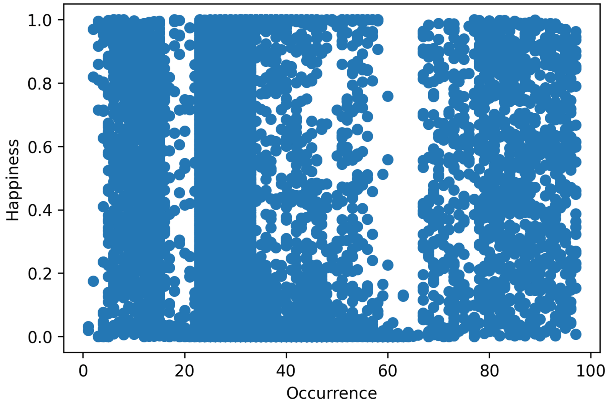


Fig. 11. Age vs Happiness

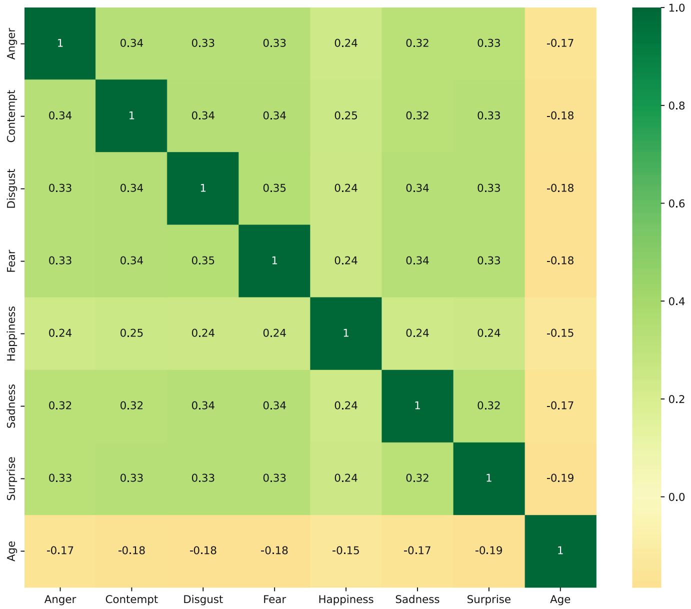


Fig. 12. Correlation of the dataset features (Emotions and Age)

ness. Also the simulation data depicts an approximate distributions for every sentiment, as denoted in Fig. 8.

Figures 9 and 10 reveal the variation stores and gender produced by the simulation tool feeding the framework.

The results demonstrate the effectiveness of the proposed framework on profiling customers and processing data. The achieved results demonstrate the proposed framework can help advertising companies reach the right target audience is to increase the ROI (Return Of Investment) of their projects. The development of a proper mobile application for multiple low-cost devices (e.g. tablets and smartphones), will help to scale new business through the use at the scale of the presented framework.

This framework can be extended to other use cases, e.g. means of transport like taxis or other public transportation, by installing a tablet in the back seat for consumer use, it would be possible to predict age and gender to decide the advertising contents (e.g. shaver products to male consumers over 30 years, video games to the age between 16 and 29 years, hair loss shampoos for consumers with baldness percentage greater than 40%, lady bags to female consumers between 25 and 45 years.

5 Conclusions

This work proposed a scalable framework enabled by a Machine Learning model to profile customers and focus advertisement. It includes the capabilities to perform continuous profiling by classifying consumers emotions while they cross the front of the stores. The main goal of this work is to profile customers. It adoptos automated emotion analysis mechanisms (e.g. gender, age, location, emotions) to select the adequate content for consumers and list ads allowing companies to reach the right audience at the right place and time. This approach can help companies increasing their sales and targeting the right consumers. In another way, the profile attributes prediction can help retailers and brands identifying the segment as the main consumers, helping to produce better publicity. Finally, using the platform to estimate the value of an exhibitor based on age is also an option. Knowing how many people spend their time on a particular areas can help to profile them. For example, spaces where most consumers ages are between 40 and 60, it will help to enable a campaign announcing a wines rather than toys. The experimental work included a prototype of the proposed framework capturing video to profile customers and also included a simulation tool producing the classified data from the front store to be ingested and processed. The achieved results demonstrate the effectiveness of the proposed framework to profile and analyse data at scale.

Therefore, the proposed framework can help to develop new innovative cost-effective business models at scale by driving in real-time the advertisements to a huge number of consumers to maximize their impact and centralizing the data collected from a large number of stores to design future campaigns. This way will be possible to drive advertisements to the right target of consumers at scale in an online and offline fashion.

In the future, aspects due to consumer privacy will be investigated. Moreover, the presented tsimulation tool requires improvement to overcome the quality issues in data revealed in the dataset. Finally, it is our aim to run experimental work in physical front stores as the way to obtain a real dataset.

Acknowledgements. This work is funded by National Funds through the FCT - Foundation for Science and Technology, I.P., within the scope of the project Refª UIDB/05583/2020. Furthermore, we would like to thank the Research Centre in Digital Services (CISeD) and the Polytechnic of Viseu for their support. This work was also supported by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020.

References

1. Nguyen, K., et al.: When AI meets store layout design: a review. Artif. Intell. Rev. 1–23 (2022)
2. Chen, S.-C., Shang, S.S.C.: Sustaining user experience in a smart system in the retail industry. Sustainability **13**(9), 5090 (2021)
3. Newman, A.J., Daniel, K.C., Oulton, D.P.: New insights into retail space and format planning from customer-tracking data. J. Retail. Consum. Serv. **9**(5), 253–258 (2002)
4. Taspinar, M., et al.: The importance of customized advertisement delivery using 3D tracking and facial recognition. In: 2012 Second International Conference on Digital Information and Communication Technology and it's Applications (DICTAP), pp. 520–524, May 2012. https://doi.org/10.1109/DICTAP.2012.6215365
5. Nigam, S., Asthana, S., Gupta, P.: IoT based intelligent billboard using data mining. In: 2016 International Conference on Innovation and Challenges in Cyber Security (ICICCS-INBUSH), pp. 107–110, February 2016. https://doi.org/10.1109/ICICCS.2016.7542329
6. AdMov Advertising Technologies (2017). http://www.facebook.com/admovtech.com/admovtech
7. What is Xamarin. http://dotnet.microsoft.com/en-us/learn/xamarin/what-is-xamarin
8. Google's machine learning team. Face detection (2021). http://developers.google.com/ml-kit/vision/face-detection
9. Nguy, L.H., Lam, K.N., Nguyen, K.-N.T., Kalita, J.: Facial expression recognition and image description generation in Vietnamese (2021). http://ebooks.iospress.nl/volumearticle/57836
10. Raman, A., Hoder, C.: Building Intelligent Apps with Cognitive APIs. O'Reilly Media, Sebastopol (2020)
11. Del Sole, A.: Introducing Microsoft cognitive services. In: Del Sole, A. (ed.) Microsoft Computer Vision APIs Distilled, pp. 1–4. Apress, Berkeley (2018). https://doi.org/10.1007/978-1-4842-3342-9_1
12. Klein, S.: IoT Solutions in Microsoft's Azure IoT Suite, pp. 273–289. Apress (2017)
13. AMQP. Advanced message queuing protocol. http://www.amqp.org/
14. Docker Inc., Docker. http://docs.docker.com/
15. Rousos, M., de la Torre, C., Wagner, B.: .NET Microservices: Architecture for Containerized .NET Applications. Redmond, Washington 98052-6399, Microsoft Developer Division, .NET and Visual Studio product teams, USA (2022)
16. Grafana. Grafana reference. http://grafana.com/