



Sentiment Analysis Through Machine Learning for the Support on Decision-Making in Job Interviews

Julio Martínez Zárate^(✉) and Sandra Mateus Santiago^(✉)

Politécnico Colombiano Jaime Isaza Cadavid, Medellín, Colombia
{julio_martinez54172, spmateus}@elpoli.edu.co

Abstract. In this paper, we propose a sentiment analysis model using machine learning for the support on decision-making in the process of job interviews. To do this, a characterization of the analysis of sentiments, job interviews and machine learning algorithms is first performed. Then, supervised machine learning with artificial neural networks is implemented in a prototype, due to the non-linear behavior described in the variables taken in the study and applying the Eye tracking technique. Finally, tests are carried out with people, in which, by asking questions of these, the involuntary movements of the pupil of the eye are analyzed, through the processing of a volume of data and the results of the ocular patterns are interpreted. Correlated with the questions of the test and with it, a final judgment is presented for the support of the decision making.

Keywords: Sentiment analysis · Machine learning · Job interview · Eye tracking · Neural networks

1 Introduction

The sentiment analysis is a process to extract the feelings and emotions of the users [1]. Liu [2] defines it as the field of study that analyzes the opinions, feelings, evaluations, validations, attitudes and emotions of people towards entities such as products, services, organizations, individuals, themes, events and its attributes. Chen [3] speaks about the different forms or computational approaches in which the sentiment analysis can be performed, such as: text-based, voice-based, visual and multimodal. One of the techniques or approaches used for this discipline is machine learning and Mitchel [4] defines it as a sub-area of computer science that studies methods to construct predictive computational models from observational data. The sentiment analysis can be applied to a myriad of disciplines and areas: in economics, in medicine, psychology, state security, politics; for the case of this work, it is applied in psychology and more exactly in the job interviews within organizations. The job interview is the most important process in recruitment and is used for various purposes: measurement of cognitive qualities, personality, motor and physical skills [5].

Job interviews are a popular selection technique from many points of view. In organizations around the world, job interviews are still one of the most widely used methods to evaluate candidates for employment. Among organizational decision

makers, interviews have been found to be the most preferred assessment method by supervisors and human resources professionals. In addition, applicants perceive that interviews are fair compared to other selection procedures and applicants expect interviews as part of a selection process. In fact, from the perspective of the applicant, obtaining a job interview is essential for success in the search for a job [6].

For the Sentiment Analysis, there are some instruments and techniques that also require a specialist person to interpret that analysis, as well as the costs of some of these devices are relatively high [7–9]. In pre-employment interviews, for example, a person is the one who analyzes behavior, gestures and certain key patterns such as the look, tone of voice and other expressions of the inquired or interviewed. On the other hand, it is worth mentioning one of the most used devices, the polygraph, which measures physiological alterations of people [10]; Furthermore, Chica [11] mentions some disadvantages of this device, where she states that, on the other hand, there are also several “tricks” that can alter the test. Another device is also the Magnetic Resonance Scanner, which uses one of the technologies considered to be the best in the detection of lies, however, it only focuses on this, it is very expensive and requires a rigorous process [12].

Because certain shortcomings and deficiencies in the techniques and devices among others, a low-cost model is proposed that can accurately interpret the feelings of people with eye tracking techniques and provide support to the current decision-making techniques for the personnel in charge of conducting job interviews in organizations, thanks to the information systems, mathematical theories and psychology.

This paper is organized as follows: in Sect. 2, the related works are presented. In Sect. 3, the sentiment analysis model using machine learning. In Sect. 4, the experiments and results. The conclusion is finally presented.

2 Related Works

There are some related works on sentiment analysis using machine learning:

Borth [13] proposed an approach based on understanding of the visual concepts that are strongly related to sentiments. Present a method built upon psychological theories and web mining to automatically construct a large-scale Visual Sentiment Ontology, using a detector library for visual sentiment analysis. Wang [14] proposed a visual sentiment analysis approach with coupled deep adjective and noun neural networks, considers visual sentiment analysis as a binary prediction problem that is to classify an image as positive or negative from its visual content. Baecchi [15] did a study that uses a multimodal feature learning approach, using neural network-based models, to address sentiment analysis of micro-blogging content, such as Twitter short messages, that are composed by a short text and, possibly, an image.

Zadeh [16] presented a model termed “tensor fusion” network (neural network-based) for sentiment analysis, highlighting the growth of research in this area through multiple modes and the use of machine learning. Similarly, Chinsatit [17] used neural networks-based pupil center detection method for a wearable gaze estimation, mentioning the applications that can have this in several disciplines of knowledge, including psychology. For its part, Poria [18] proposed a multimodal affective data

analysis framework to extract user opinion and emotions from video content and combines text, audio and video; the paper also proposes an extensive study on decision-level fusion.

On the other hand, Chen [3] uses a convolutional neuronal network for the prediction of sentiments through the joint learning of textual and visual sentiments from training examples. George [19] uses a convolutional neuronal network too, propose a real-time framework for the classification of eye gaze direction and estimation of eye accessing cues.

3 Sentiment Analysis Model Using Machine Learning

In this research, the characterization on sentiment analysis was initially carried out, among which are: textual, voice, visual and multimodal. After this characterization and given the final validation designed to apply it in job interviews, this work focuses on the visual approach. To do this, we use the register of positions through coordinates of the center of the pupil of the eye, following an algorithm called Eye-Tracking [20], which is the process of measuring the movement of an eye in relation to the head or the point where the gaze is fixed (Fig. 1).

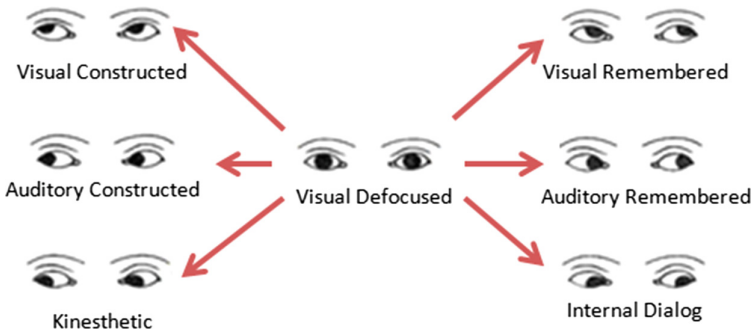


Fig. 1. Eye Accessing cues [21]

The model shown in Fig. 1 establishes that the specific characteristics of the thinking mechanism relate to a non-visual orientation of the gaze [21]. A person is related to the way he moves his eyes [22]. “According to a popular proverb, the eyes are the window of the soul. And, in fact, people have wondered for a long time if there is something in our eyes indicative of character” [23].

For this study, in addition, the exploration of the different automatic learning algorithms used for the sentiment analysis was carried out (Fig. 2).

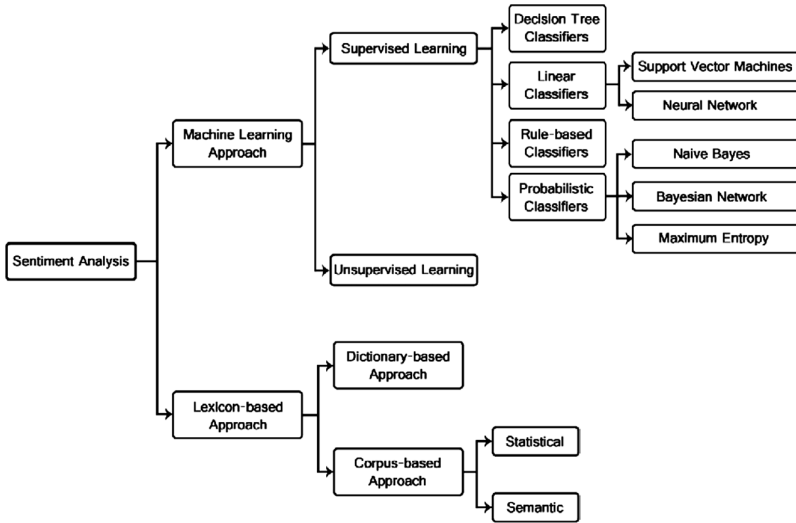


Fig. 2. Approaches and algorithms for sentiment analysis [24].

After making comparisons between the algorithms, the supervised machine learning approach was selected using artificial neural network techniques, using as a criterion that people have different behaviors in their eyes when asked about something and the time of permanence looking towards certain coordinates it is not the same for everyone, nor are the coordinates usually the same, so the relationship between the variables can be considered with a non-linear trend.

3.1 Architecture of the Artificial Neural Network

The prototype used in the sentiment analysis model proposed in the present article has enabled six of the seven patterns (visual defocused, visual created, visual remembered, auditory created, remembered auditory, kinesthetic sensations, internal dialog), excluding the defocused visual pattern.

Next, the architecture of the neural network is proposed to validate the coordinates of the model. The variables of inputs are related to the coordinates of the Cartesian plane (x, y) and the stay time of the gaze. When a question is asked during the job interview at a certain time t , the person fixe his gaze towards certain places, which, when captured, counts the number of times in all the coordinates:

- Number of looks towards visual defocused (#VD)
- Number of looks towards remembered visual (#VR)
- Number of looks towards visual created (#VC)
- Number of looks towards remembered auditory (#AR)
- Number of looks towards auditory created (#AC)
- Number of looks towards internal dialogue (#DI)
- Number of looks towards kinesthetic sensations (#KI)

The input data are normalized by Eq. (1):

$$x' = d1 + \frac{(x - xmin)(d2 - d1)}{xmax - xmin} \tag{1}$$

Where,

x : value to normalize

$[xmax, xmin]$: value range x

$[d1, d2]$: range to which the value of x will be reduced

The outputs are multiple and correspond to the value of the coordinates, since it is a multiclass classification layer (0, 1, 2, 3, 4, 5, 6 correspond to: VD, VR, VC, AR, AC, DI, KI, respectively).

After the characterization of the machine learning algorithms and the design of several neural network architectures, making combinations with different numbers of hidden layers, with different amounts of nodes in these layers, with different activation functions and different algorithms of learning, a multilayer perceptron is selected with two hidden layers of “processing of the sentiments” with 2 inputs (and Bias) and 7 neurons in each hidden layer, and the number of the output is 7 as well, as it was mentioned is multiple (classification). They are activated with the sigmoid activation function, as an error function, categorical cross entropy with Adam optimizer was chosen (Fig. 3).

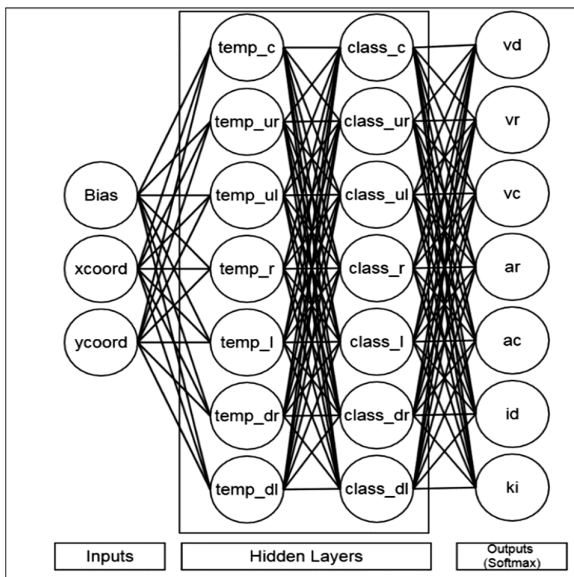


Fig. 3. Architecture of the multilayer perceptron neural network used to validate the patterns [Own development].

The model proposed in this paper can perform the visual sentiment analysis of people with machine learning, performing the interpretation of the eye Accessing cues with the help of the eye tracking technique. It is applied in job interviews now in which an interviewee is asked questions of a personal nature, and this fixes his gaze towards certain coordinates that have a meaning according to the areas of study of psychology. After the results obtained with the variables of the fixations of the gaze, the personnel in charge (usually of human resources), analyze said results and make their own decision about the candidate for the job.

3.2 Interview

The interview is based on the work done in Costa [25] on the five (5) personality dimensions (Big Five): Extraversion, kindness, responsibility, emotional stability, openness to experience. In Rauthmann [23] a study is carried out with the ocular tracking demonstrating with mixed linear models that the personality predicts the number of fixations, the duration of the fixation and the time of permanence in two different abstract animations. Hooft [26] investigate whether ocular tracking technology can improve people's understanding of the response process when it is said. For his part, Broz [27] collects the data from the look of human conversational pairs in order to understand which characteristics of the interlocutors of the conversation influence this behavior. These three (3) works made use of the Costa test [25].

Open questions are carried out based on the Costa test [25] carried out in order to evaluate each of the previous dimensions, so that the person responding projects their gaze towards certain coordinates involuntarily according to their feelings. The human being is capable of manifesting three (3) feelings at the same time with his eye cues.

4 Experimentos y Resultados

Next, the experiments and results made for the sentiment analysis model are presented. For the experiment, it should be suggested to the person not to wear glasses, not to turn the face to the sides, so that the webcam can constantly monitor the movement of the retina.

4.1 Prototype

The prototype consists of a client-server system with graphical user interface (GUI) web client, which captures the coordinates of the gaze and the time spent in these coordinates, using a conventional web camera. The detection of the retina is done with the help of the webcam, this is captured with a javascript library called Webgazer [28], which uses internal machine learning. First the calibration of the "library" is performed, the more it is calibrated, the greater the precision of the detection of the coordinates of the gaze.

In Fig. 4, the user is in front of the computer that has a conventional web camera and captures the movement of the retina, this is reflected in the graphical user interface web on the screen of said computer, capturing the needed information. The collection of the information of the variables is done in real time through the prototype eye-tracking and stored in a database for further processing.

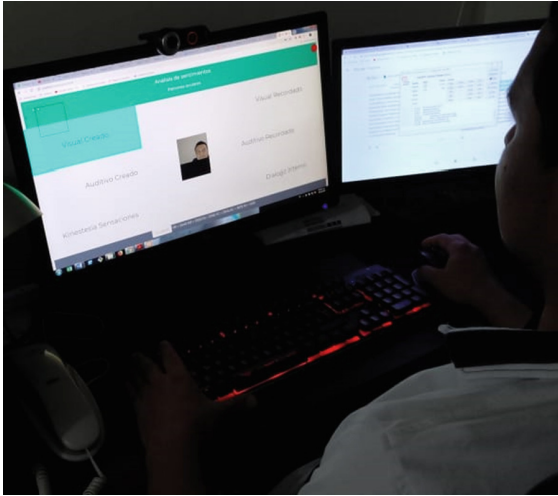


Fig. 4. GUI - prototype for eye tracking [Own development].

In each coordinate where the person interviewed fixes the gaze, the absolute frequency or number of fixations to that coordinate is calculated; then, a point (x, y) closest to the interpolation point described by the data set $(\min(\|X - IP\|))$, if it belongs to the first three (3) major absolute frequencies. It is making a scale of the point (x, y) captured in the width and height (in pixels) of the browser window viewport with respect to the square of the size of the rectangular image of the eye area. A translation of Cartesian coordinates is carried out and then the normalization of the data with respect to the pixel size of the rectangle of the eye region [29].

4.2 Evaluation

For training the neural network, we take the Eye-Chimera data set used in [21], transformed to coordinate matrices of the images. They contain a matrix of (14×2) of each look (described in each image) in which the rows are the coordinates; and the columns, the x and y axes. The five (5) first rows are coordinates of the left eye, as well as the last 2 (two); the rows from 6 to 10 are coordinates of the right eye; the coordinates of the pupil are rows 5 and 10, row 1 is the left end of the left eye, row 6 the left end of the right eye, row 11 is the upper end of the left eye, and row 13 is the upper end of the right eye (which are the variables of interest). Taking these values and performing measurements normalization to a rectangle located in the region of the eye.

The images of the taken dataset have a size of 640 px wide by 480 px high, the region of the eye is reduced to 35×18 px.

In Fig. 5, two eyes are shown taking values from a matrix of dotted positions with the pupil in the center.

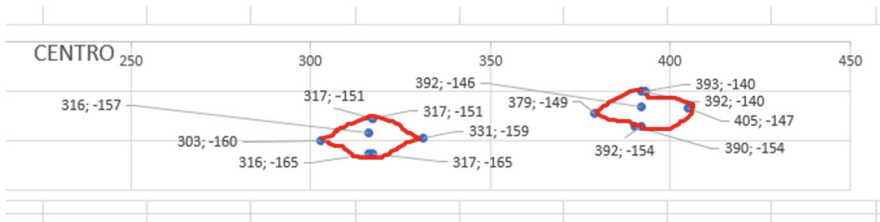


Fig. 5. Coordinates of a matrix of the eye-chimera dataset, person looking at the center [Own development].

The eye coordinates are captured in five (5) volunteers, asking a personality question for each dimension of the Big-Five (5 questions), only explaining to the person who answers the questions that will appear on the computer screen or device, the client was hidden from the interface of capturing the gaze and only the questions that had to be answered appeared and he clicked when he answered it verbally, without explaining what dimension of personality would be measured either. And at the end of the test, present the results of the quantities as shown in Fig. 5, for each question answered.

Question 1. Read and answer ×

Do you often strive to help others who need it? Do you do this in sacrifice for your own interest?

Elapsed 00:03

Start
Confirm

Fig. 6. End user interview interface. This is a question related to extroversion [Own development].

For each question, the absolute frequencies of fixation in the established coordinates are obtained, and with that the six (6) relative frequencies in each coordinate are calculated and the three (3) first major ones are taken (relative frequencies greater than 10%). The average duration of a fixation ranges between 200 and 350 ms [30].

Table 1 shows the results of the people evaluated and the sentiments corresponding obtained with the eye tracking:

Table 1. Results with absolute frequencies in the coordinates in the determined response time for the five (5) questions.

Sentiment	VD (%)	VR (%)	VC (%)	AR (%)	AC (%)	ID (%)	KI (%)	Total fixings
<i>Volunteer 1</i>								
Question 1	-	18,7	6,7	21,9	10,9	25,5	16,4	1861
Question 2	-	0,8	4,2	1,4	16,8	18,6	58,1	837
Question 3	-	2,7	2,7	2,7	2,7	2,7	86,5	37
Question 4	-	16,2	12,5	18,1	14,3	21,2	17,7	3011
Question 5	-	1,9	11,1	3,2	17,5	16,4	49,8	901
<i>Volunteer 2</i>								
Question 1	-	13,1	0,2	20,7	1,4	59,0	5,5	420
Question 2	-	9,0	19,2	11,9	22,4	12,7	24,9	1456
Question 3	-	0,0	6,8	0,2	14,5	5,0	73,5	442
Question 4	-	0,0	4,1	1,0	7,1	59,9	27,9	197
Question 5	-	11,6	11,9	13,3	15,0	16,4	31,7	293
<i>Volunteer 3</i>								
Question 1	-	1,0	5,5	8,6	16,4	39,3	29,2	384
Question 2	-	1,6	10,9	7,1	21,1	24,1	35,3	949
Question 3	-	1,5	10,9	6,8	21,2	23,4	36,1	980
Question 4	-	1,5	10,9	6,8	21,0	22,6	37,1	1042
Question 5	-	4,6	13,1	6,8	17,3	18,1	40,1	237
<i>Volunteer 4</i>								
Question 1	-	5,7	18,0	7,4	21,3	11,5	36,1	122
Question 2	-	9,0	5,4	22,3	10,8	36,7	15,7	166
Question 3	-	23,9	2,0	24,5	2,3	27,0	20,2	640
Question 4	-	2,5	25,2	2,5	28,6	9,1	32,1	2318
Question 5	-	12,9	11,8	15,1	15,6	18,9	25,7	1133
<i>Volunteer 5</i>								
Question 1	-	20,7	0,4	20,7	0,4	21,3	36,5	1630
Question 2	-	1,8	31,0	2,0	31,0	2,8	31,6	507
Question 3	-	2,0	18,1	2,2	19,9	2,4	22,2	2137
Question 4	-	42,3	12,5	45,4	0,0	58,7	41,3	489
Question 5	-	5,6	0,0	14,6	0,0	77,1	2,8	144

Each question is about identifying a dimension of the candidate for the job: Question 1, Extraversion; question 2, kindness; question 3, responsibility; question 4, emotional stability, question 5, openness to experience. The response time the person takes is not relevant in the results for the verification, although the absolute frequencies depend on it.

After training of the model with the neural network shown in Fig. 4 with the Eye-Chimera data set (885 samples in total), the results without label are processed, for this the “Tensorflowjs” tool is used.

For each point of the first 3 major ratios of the previous table, take a point closer to the curve described, of each set of points in each of those three (3) coordinates.

Table 2 shows the matrix of validation of the results, showing the percentage of success or certainty with the coordinate. The neural network throws each one of the results of the patterns to be checked with the highest corresponding values with respect to the others.

Table 2. Validation of the results in the corresponding coordinates.

Certainty Sentiment	VR	VC	AR	AC	ID	KI
<i>Volunteer 1</i>						
Question 1	0.9036851	-	0.3852791	-	0.8211777	-
Question 2	-	-	-	0.4530621	0.6093528	0.4601849
Question 3	-	-	-	-	-	0.3797783
Question 4	-	0.4514413	-	-	0.4644437	0.3347794
Question 5	-	-	-	0.3576853	0.4807395	0.3576845
<i>Volunteer 2</i>						
Question 1	0.3706563	-	0.4616654	-	0.6253502	-
Question 2	-	0.538384	-	0.5479685	-	0.5481049
Question 3	-	0.3958438	-	0.3933841	-	0.3079248
Question 4	-	-	-	0.7944366	0.8557455	0.9038013
Question 5	-	-	-	0.4073458	0.9375306	0.3971747
<i>Volunteer 3</i>						
Question 1	-	-	-	0.4683144	0.3150922	0.4673097
Question 2	-	-	-	0.4464874	0.6187378	0.4442834
Question 3	-	-	-	0.4514413	0.4640621	0.5318491
Question 4	-	-	-	0.7093528	0.6101221	0.5213944
Question 5	-	-	-	0.3812893	0.4905789	0.3982546
<i>Volunteer 4</i>						
Question 1	-	0.5383841	-	0.3167358	-	0.3474749
Question 2	-	-	0.4626554	-	0.6586387	0.5432834
Question 3	0.910508	-	0.3852791	-	0.5441523	-
Question 4	-	0.4503414	-	0.4965323	-	0.4184881
Question 5	-	-	-	0.8003421	0.3290811	0.3786671
<i>Volunteer 5</i>						
Question 1	0.912558	-	0.6461454	-	0.5992102	0.2242534
Question 2	-	0.4616415	-	0.2967152	-	0.5425104
Question 3	-	0.3994832	-	0.71549	-	0.5531846
Question 4	0.6712303	-	0.6126548	-	0.2275161	-
Question 5	0.9537871	-	0.5859972	-	0.2372161	-

5 Conclusion

In this paper, we have presented a sentiment analysis model with the use of advanced machine learning techniques, seeing that it is an emergent application area in many fields, for this case, it is applied to decision making in the process of job interviews to assess personality. There are several approaches to analyzing feelings, but in this case algorithms with neural networks were taken because of their advantages analyzed against the subject under treatment.

With the results of the personality test of each question, the person in charge of defining the future of the candidates for a job position (normally a psychologist or human resources professional), relies on the model correlating in the results of the variables to then give their final judgment in this regard.

Acknowledgments. The research work presented in this paper was partially funded by the thesis proposal of Master's Degree in Engineering, from the Politécnico Colombiano Jaime Isaza Cadavid - Colombia; entitled: "Modelo de análisis de sentimientos mediante el aprendizaje de máquina para el apoyo en la toma de decisiones en entrevistas laborales". Thanks to Yamid Asael Arenas Manso, System Engineer.

References

1. Devika, M., Sunitha, C., Ganesh, A.: Sentiment analysis: a comparative study on different approaches. *Procedia Comput. Sci.* **87**, 44–49 (2016)
2. Liu, B.: *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers, San Rafael (2012)
3. Chen, X., Wang, Y., Liu, Q.: Visual and textual sentiment analysis using deep fusion convolutional neural networks (2017)
4. Mitchell, T.: *Machine Learning*. McGraw-Hill, Boston (1997)
5. Dessler, G.: *Administración de recursos humanos, enfoque latinoamericano*, 5 edn. (2009)
6. Macan, T.: The employment interview: a review of current studies and directions for future research. *Hum. Resour. Manag. Rev.* **19**, 203–218 (2009)
7. Kron, L.: Polygraph and reliability in psychological assessment: myth or reality? (2016)
8. Pham, T.D., Tran, D.: Emotion recognition using the emotiv EPOC device. In: Huang, T., Zeng, Z., Li, C., Leung, C.S. (eds.) *ICONIP 2012, Part V. LNCS*, vol. 7667, pp. 394–399. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34500-5_47
9. Mitsuyoshi, S.: Emotion recognizing method, sensibility creating method, device, and software (2008)
10. Anta, J.: *Detección Del Engaño: Polígrafo vs Análisis Verbo-Corporal* (2012)
11. Chica, H., Escobar, F., Folino, J.: *La Entrevista Psiquiátrica Del Sujeto Simulador* (2005)
12. Petisco, J.: *Una Mirada A La Detección De Mentiras Empleando fMRI* (2015)
13. Borth D., Rongrong, J., Chen, T., Breuel, T., Chang, S.: Large-scale visual sentiment ontology and detectors using adjective noun pairs (2013)
14. Wang, J., Fu, J., Xu, Y., Mei, T.: Beyond object recognition: visual sentiment analysis with deep coupled adjective and noun neural networks (2016)
15. Baccchi, C., Uricchio, T., Bertini, M., Del Bimbo, A.: A multimodal feature learning approach for sentiment analysis of social network multimedia. *Multimed. Tools Appl.* **75**, 2507–2525 (2015)

16. Zadeh, A., Chen, M.: Tensor fusion network for multimodal sentiment analysis (2017)
17. Chinsatit, W., Saitoh, T.: CNN-based pupil center detection for wearable gaze estimation system (2017)
18. Poria, S., Peng, H., Hussain, A., Howard, N., Cambria, E.: Ensemble application of convolutional neural networks and multiple kernel learning for multimodal sentiment analysis (2017)
19. George, A., Routray, A.: Real-time eye gaze direction classification using convolutional neural network (2016)
20. Llerena, M.: Desarrollo de una metodología basada en la programación neurolingüística utilizando software educativo para mejorar el proceso enseñanza-aprendizaje. Msc tesis. Escuela Superior politécnica de Chimborazo, Ecuador (2016)
21. Florea, L., Florea, C., Vrânceanu, R., Vertan, C.: Can your eyes tell me how you think? A gaze directed estimation of the mental activity (2013)
22. Risko, E., Anderson, N., Lanthier, S., Kingstone, A.: Curious eyes: individual differences in personality predict eye movement behavior in scene-viewing (2012)
23. Rauthmann, F., Seubert, C., Sachse, P., Furtner, M.: Eyes as windows to the soul: gazing behavior is related to personality (2012)
24. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey (2014)
25. Costa Jr., P.T., McCrae, R.R.: Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Manual. Psychological Assessment Resources, Odessa (1992)
26. Hooft, V.: Intentional response distortion on personality tests: using eye-tracking to understand response processes when faking (2014)
27. Broz, F., Lehmann, H., Nehaniv, C., Dautenhahn, K.: Mutual gaze, personality, and familiarity: dual eye-tracking during conversation (2012)
28. Papoutsaki, A., Sangkloy, P., Laskey, J.: WebGazer: scalable webcam eye tracking using user interactions (2016)
29. Krishnamurthy, N.: Introduction to Computer Graphics, p. 165 (2002)
30. Tejero, P., Pastor, G., Crespo, A.: Exploración visual y movimientos oculares en conductores con distinta experiencia: una revisión (2004)