A Review of Multimodal approach for Rating prediction

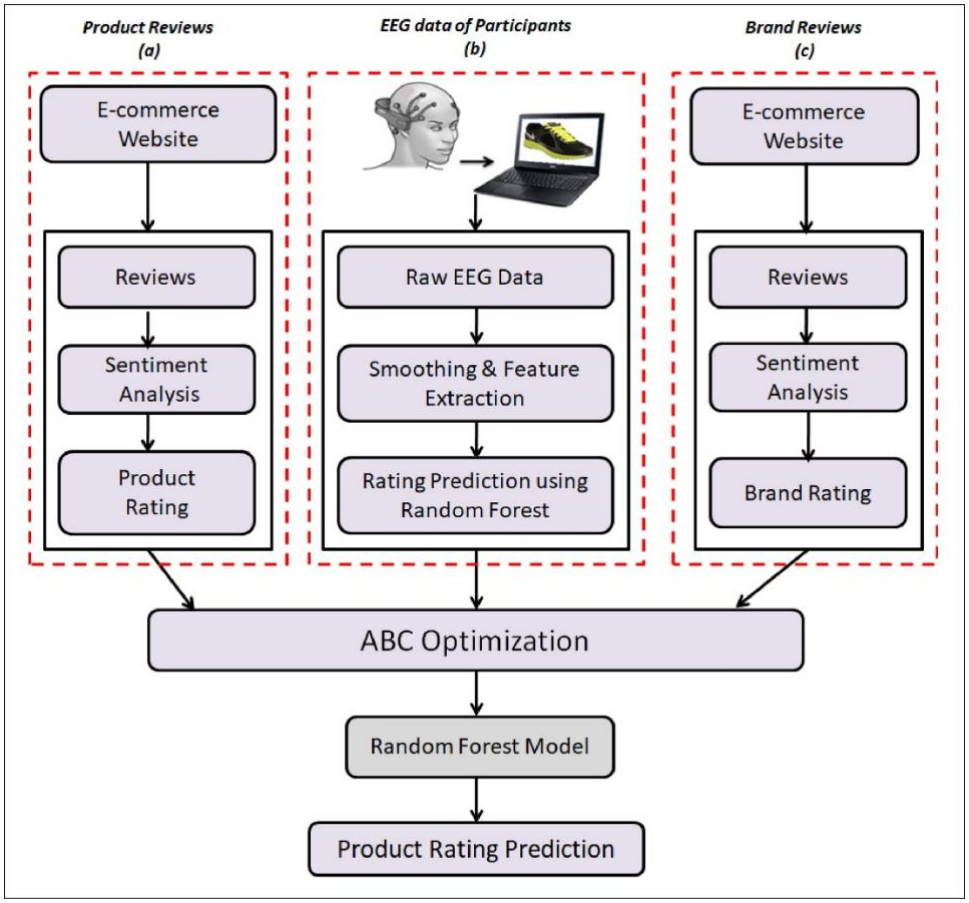
Introduction:

Advertisers look for new ways other than conventional advertisement mediums like television, questionnaire, attitude or verbal communication, to market their products in the changing era of digitization. But, such techniques may fail in predicting the consumer mindset towards the product since customers may provide wrong feedback during such communications. Some different methods of advertisement are social networking sites like Facebook, Amazon, Flipkart, etc., where purchasers can express their views as a text for the product. The technique of getting feedback using brain waves is called 'Neuromarketing', where the psychology of the customer is understood. If right and proper feedback is obtained, it helps in reducing the expenditure on product advertisement. Instruments like functional Magnetic Resonance Imaging (fMRI) and Magneto Encephalography (MEG) are often used in Neuro-marketing to capture the brain activity. MEG instrument provides a map of neuron activity in spatio-temporal information. However, there are technical complexities to configure such devices also since it is expensive, it is not generally preferred by researchers in Neuro-marketing. Due to these reasons, the work considered EEG device as it offers high-resolution data with lower cost in comparison to fMRI and MEG instruments in addition to the maintenance and support costs also being reasonable. Similarly, Opinion mining is a trending research area through which marketers can guess the likeability of their product amongst the customers by targeting marketing and customer services. It uses NLP tools and techniques to extract and study both , the affective as well as subjective information. Sentiment analysis involves calculation of a compound score and polarity of a piece of text, categorizing it into three categories: positive, neutral and negative. Multi-modal frameworks have started gaining popularity in Human Computer-Interaction (HCI) domain to fuse two or more modalities in order to achieve better results. The advantage of multiple modalities helps in increasing usability where the weaknesses of one modality are offset by the strengths of another. Such systems offer a conducive environment to their users and recognize the inputs from different modalities. It then combines them according to temporal and contextual constraints to increase their interpretability. Multi-modal systems are composed of multiple uni-modal solutions. therefore, optimization algorithms can be used to fuse multiple solutions into a single best solution. It involves the design of an objective function which minimizes the cost of the production or maximizes the efficiency of production. Optimization algorithms are executed iteratively till an optimum solution is found. ABC is preferred since it has less number of parameters and its capability to produce global optimal solutions.

The paper reviewed combines EEG signals and sentiment analysis for rating prediction of e-commerce products.

Rating Prediction of E-Commerce Products:

The approach proposed in the paper was as follows:



A. Preprocessing of EEG Signal :

Smoothing is performed by using Savitzky-Golay (S-Golay) filter. The general equation of S-golay filter:

$$R_k = \sum_{i=-\frac{q-1}{2}}^{\frac{q-1}{2}} c_i \; S_{k+i}, \qquad \frac{q-1}{2} \leq k \leq p - \frac{q-1}{2}$$

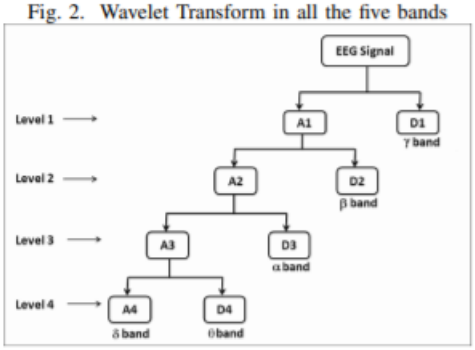
where S_k denotes a signal at instant k, q denotes the frame span, c_i represents the set of q convolution coefficients and R is smoothed signal. The value of c_i is calculated by using frame span q. The equation used is Quadratic polynomial with frame span of size 5 , given by:

$$Y_j = \frac{1}{35}(-3y_{j-2} + 12y_{j-1} + 17y_j + 12y_{j+1} - 3y_{j+2})$$

B. Wavelet Transform:

The signal is decomposed by passing a Low Pass (L) and High Pass (H) filters into different frequency bands. Detail (D) and Approximation (A) coefficients are the output of H and L, respectively. The A1 coefficients are further divided into second stage and the procedure is repeated until the signal is decomposed upto an expected level. Signal is decomposed into 5 bands, γ (32–100 Hz), β (13–22 Hz), α (8–13 Hz), θ (4–8 Hz), and δ (1–4 Hz). Three statistical features ,i.e., Mean, Standard Deviation, and Energy are extracted for each band waves that are used for rating prediction.

Fig. 2. Wavelet Transform in all the five bands



C. Random Forest Regression:

The output of regression tree is an average of all individual tree outcomes. Root-mean-square error (RMSE) is calculated while RF tree is grown by using out of bag method. RMSE and R-square (R^2) can be computed by using the equations shown.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2}$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2}$$

Where,

$\hat{y}$ – predicted value of y

$\bar{y}$ – mean value of y

D. Opinion Mining:

VADER approach views the sentiment score of each word in the review and finally produce the score of the whole review. From each review, VADER produces four sentiment components. Out of these the first three are: positive, neutral, and negative. The fourth component is a compound score which is a normalized score of first three components. The range of compound score varies from -1 to 1, where -1 represents least preferred and 1 represents most preferred choice of products. Later, the score is calculated into 1 to 5 rating scale using :

Rating = $1 + 2 * (1 + x)$, where x is a compound score.

E. ABC Optimization:

The initialization of food source is calculated using ,

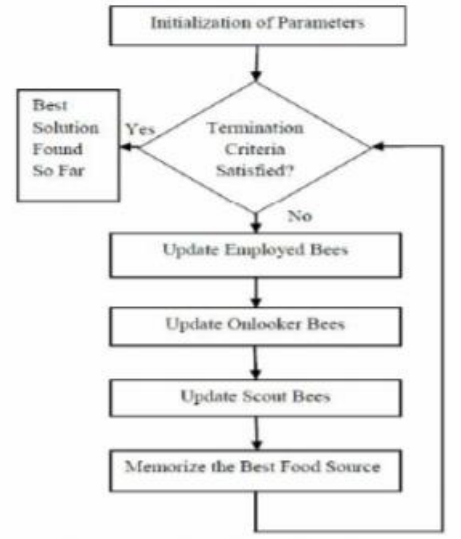
$x_{mi} = l_i + rand(0, 1) * (u_i - l_i)$, where x_m represents the m^{th} food source in the swarm and i taken from 1,2, ...D, here D is the number of variable in the optimization problem. l_i and u_i are bounds of x_m in i^{th} direction. rand (0,1) is the random number range between 0 and 1.
The neighborhood food (v_m) source can be selected by using : $v_{mi} = x_{mi} + \phi_{mi} * (x_{mi} - x_{ki})$, where k represents the number of parameter to optimize and m_i is a random number between -1 to 1.
After selecting neighborhood food source, their fitness value is calculated using fitness function. The fitness value of a solution can be calculated using :

$$fit_i = \begin{cases} \frac{1}{1+f_i} & if\ f_i \geq 0 \\ 1 + abs\,(f_i) & if\ f_i < 0 \end{cases}$$

The onlooker bees wait in dancing area to take decision for selecting the food source. They select a food source based on the probability (Pi) of fitness values provided by employed bees by:

$$P_i = \frac{fit_i(\eta_i)}{\sum_{i=1}^{D} fit_i(\eta_i)}$$

Fig. 3. Phases of Artificial Bee Colony Algorithm



CONCLUSION :

Multimodal fusion helps in better valence score determination compared to unimodal schemes prevalent. ABC Optimization process further improves its accuracy. EEG signals are smoothed and filtered. Their features are extracted using DWT algorithm. Sentiment analysis is performed by using VADER. RMSE= 0.29 and R^2= 0.72 values are obtained. θ band waves are most useful for prediction of local ratings.

REFERENCES :

[1] S Kumar, M Yadava, P Roy, Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction, Information fusion(52). (2019) 41–52.

[2] S. Kumar, M. Yadava and P.P. Roy, R.Saini, D P Dogra, Analysis of EEG signals and its application to neuromarketing , Multimed. Tools Appln. 76 (18)(2017) 19087-19111.

[3] H. Gauba, P. Kumar, P.P. Roy, P. Singh, D.P. Dogra, B. Raman, Prediction of advertisement preference by fusing EEG response and sentiment analysis, Neur. Netw. 92 (2017) 77–88.

[4] A. Gupta, H. Sahu, N. Nanecha, P. Kumar, P.P. Roy, V. Chang, Enhancing text using emotion detected from EEG signals, J. Grid Comput. (2018) 1–16.

[5] B. Kaur, D. Singh, P.P. Roy, Age and gender classification using brain–computer interface, Neur. Comput. Appl. (2018) 1–14.

[6] S. Saini, L. Dewan, Application of discrete wavelet transform for analysis of genomic sequences of mycobacterium tuberculosis, Springerplus 5:64 (1) (2016) 1–15.

[7] R. Polikar, C. Tilley, B. Hillis, C.M. Clark, Multimodal EEG, MRI and PET data fusion for Alzheimer's disease diagnosis, in: Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), 2010, pp. 6058–6061.