Human Error Prediction Using Eye Tracking to Improvise Team Cohesion in Human-Machine Teams

Praveen Damacharla(✉), Ahmad Y. Javaid, and Vijay K. Devabhaktuni

EECS Department, College of Engineering, The University of Toledo, MS 308, 2801 W Bancroft St, Toledo, OH 43606, USA
Praveen.Damacharla@rockets.utoledo.edu

Abstract. The rapid increase in integration of intelligent systems in every corner of technology created an emerging field called Human-Machine Teaming (HMT). In HMT, human and machine collaborate with one another to accomplish a common goal or task. To achieve the best performance of a team, it is necessary to build trust and cohesion among all teammates (machines and humans). Furthermore, in a team, it is an established fact that a team member ability to predict fellow member's future course of action and have an accurate picture is a valuable asset and will result in better team dynamics and team performance. To realize such an ability we are proposing a human error predicting methodology that could give an intelligent system a better understanding of human actions in advance. In the proposed method, we used eye-tracking metrics such as gaze density and cognitive state to predict human errors. The results obtained with the proposed and developed methods are found to be efficient in predicting human error probability.

Keywords: Cohesion · Decision trees · Eye-Tracking
Human-Machine Teaming · Support Vector Machine

1 Introduction

The upsurge in the incorporation of intelligent frameworks in almost every aspect of human-life has laid the foundation for better possibilities pertaining to increased interactions with machines and devices. On that note, the complexity in collaboration with machines has led to the development of a new concept called human-machine teaming (HMT), an emerging field. Indeed, the evolving relationship between human and machines has become a primary intent with most of the developments aimed at addressing the complexity factors. Today, the intelligence emanating from machines has been instrumental, penetrating various elements of human life as indicated by IBM survey results [1]. The international organization claims that 75% of respondents in their survey anticipate that intelligent machines ought to have a meaningful impact on business in the next three years. Also adding to that, 70% of organizations who

R. L. Boring (Ed.): AHFE 2018, AISC 778, pp. 47–57, 2019.
https://doi.org/10.1007/978-3-319-94391-6_5

participated in the survey believed that intelligent machines could result in a higher-value work for employees if adjustments were made to achieve most effective HMTs [1]. Essentially, the machines have made it possible to push the performance horizon of humans beyond the conventional threshold (physical, emotional, and cognitive domains). Moreover, this is apparent from the results of technologies such as bioacoustics sensing, quantified self and human augmentation focusing on augmented reality, as well as voice and gesture control. Despite these, there is still a dearth of comprehensive understanding of dynamics of trust and cohesion between people and machines [2].

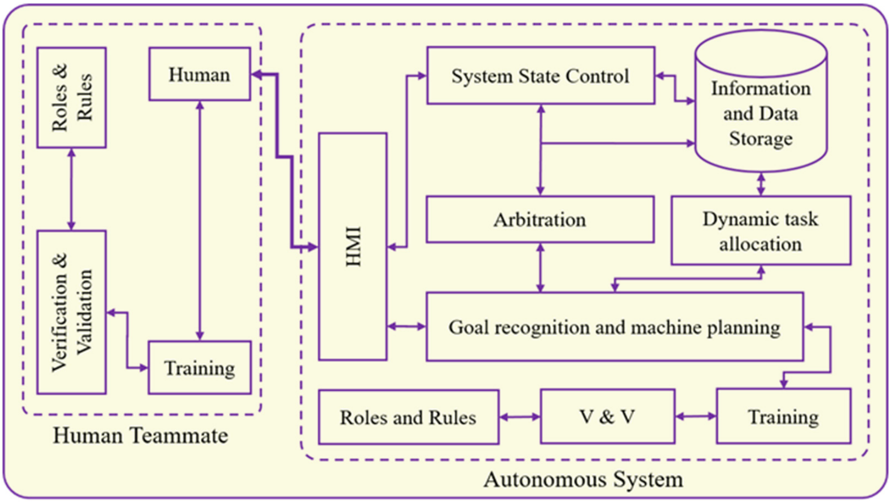

Fig. 1. Typical structure of a human-machine teaming system

1.1 Frameworks of Human-Machine Teams

Associations and processes that take place in realistic contexts characterize the customary HMT framework (Fig. 1). Additionally, the machines interact with the humans through various interfaces and controls depending on the type of system application. These interfaces mostly entail panels and indicators for display in addition to decision support tools. Conversely, the controls help humans in automation and system operation and provide a room to standardize ways for implementing human intentions and strategies [3].

Despite the practicality of this explanation, it only refers to the most typical HMTs such as autonomous assembly chains used in factories. The HMT does not allude to the contemporary human-machine interaction which is a multi-disciplinary field focusing on facets such as artificial intelligence, robotics, voice control interfaces, and social sciences. Today, HMTs have a common share of collaborative efforts from both entities with a mutual goal of accomplishing a specific goal and task. Nonetheless, despite having an intelligent framework, the machines do not have comprehensive control over a task [4]. Therefore, the task accomplishment requires leveraging best skills of human and machine. For instance, humans have optimal skills such as Correlation identification, Perspective & Emotional intelligence while the machines are more dominant in Scouring of Data, All-pairs testing, and Classification. Some of the common machines used in HMT include hardware robots with artificial intelligence based operating systems or complete software agents [4]. Moreover, the completion of the multi-agent setup includes linked data resources that positively contribute to the interactions.

Machines are the simulated agents with capabilities of perception and action in the natural world. In the modern setting, their utilization has become a paramount factor in technologically advanced societies that play equally significant roles in crucial domains [2]. In the context of HMTs, a range of elements comes into play in determining the degree at which human trusts the machine and their level of productivity in dealing with specific tasks. Some of these elements include the reliability of the machine, the self-confidence of the humans, and the synchronized workload [3].

Certain tasks require the machines to classify objects, identify, and detect humans and their emotions. Typically, the urge for energetic dimensions prompts the expansion of all the sub-fields of human-machine operations. Moreover, there are increased possibilities and potential for HMT [5]. Consequently, with the expansion of human-machine cohesions, the current and advanced technologies could facilitate scenarios where the team members collaboratively assist each other in control, coordination, and achievement of tasks without any errors [5].

Both humans and machines can be erroneous, particularly in circumstances with high levels of fuzziness. Therefore, a better standardization of cohesion between machines and human is a significant necessity for a good team performance. Commendably, industrial machines have been a major integration into industrial assembly lines and have indicated collaborative effectiveness with humans. However, as part of the team's functionality, it is significant that team members have the forecast for courses of action of fellow members for improved accuracy through minimal errors [5, 6].

1.2 Overview of the Paper

Despite the prevalence of multiple protuberant approaches that utilize facial lexes to identify prompt errors or task-related performance errors, HMTs lack the capabilities of forecasting errors during task execution. Resolutely, the paper entails the proposition of an approach to exploit eye tracking with eye point data as a means for measuring the concentrated area followed by utilization of data to forecast both the type and percentage of errors. Notably, the proposed plan utilizes a chronological method to measure cognitive state to predict human errors. In Sect. 2, we briefly discuss different human error identification technologies studied over the course of time, and in Sect. 3 we propose an error prediction method and discuss its implementation. Section 4 discusses results and effects of results, and Sect. 5 concludes the research work.

2 Background and Related Work

Over the years, there have been multiple publications of work focusing on the prediction and identification of human errors in completing a task or interacting with a complex machine. These literature could be efficient in informing humans and influencing efficient interactions with the machines. In 2015, Mu et al. highlighted that the increasing dependency on machines was marked with augmented accidents emanating from human error [7]. This applied as the basis for indicating the need for a human reliability analysis. In their paper, they constructed a prediction method of human error probability through the integration of performance shaping factors and Bayesian networks theory. The results of the study showed that the method could effectively address challenges on correlation and the dissimilar weight.

The robotics department of MIT has also been one among a few research institutions at the forefront making developments on the subject matter. Their study aimed at discovering how interface design could influence the calibration of trust and reliability between humans and machines. The primary aspects of interest were the probable advantages of predictive alarm systems that could identify human errors before they occurred [8]. Contrary to the traditional binary alarms that indicate solely ‘threats’ or ‘no threats’, the predictive alarms would give additional details on the confidence levels and the insistence of the forewarned error. Paper concluded that the predictive alarm would be influential in helping alleviate the ‘cry wolf’ effect as a sensation popular in high-risk industries operating with both humans and machines [8].

In a different study, Yang et al. formulated a method entailing human respondents performing a memory and recognition task with the help of an automated decision aid. In the study, the respondents were asked to memorize images and were later asked to recognize them in a gallery of relatively similar images [9, 10]. The automated decision aid helped the respondents by recommending image to select, which was very reliable for some of the participants but deemed undependable for others. The results of the study showed that the respondents under-trusted highly reliable aids and over-trusted the unreliable ones. In conclusion, the study showed that cohesion between humans and machines evolves and stabilizes over time, which positively influences the prevalence of errors [9, 10].

Resolutely, based on recent studies, the identified works which made positive contributions to the subject are presented in Table 1. However, the study focuses more on the prediction and identification of human errors in a task. The studies are aimed at developing cohesion process between humans and machines but fail to indicate how it can be improved through a form of contemporary interaction where machines also get to observe humans back. This leaves an unexplored area in HMTs warrants a two-way communication between human and machine through mutual feedback.

Table 1. Comparison of human error prediction studies to improve HMTs

Research study	Area	Results and conclusions
An increasing dependency on machines Vs human error [10]	Construction of a prediction method of human error probability through the integration of performance shaping factors and Bayesian networks theory	The method could effectively address challenges on correlation and the dissimilar weight using Bayesian networks. Can be concluded, the method predicts human errors though actions sequence data
The probable advantages of predictive human error alarm systems [11]	Discovering how interface design could influence the calibration of trust and reliability between humans and machines	The predictive alarm would be influential in helping alleviate the 'cry wolf' effect as a sensation popular in high-risk industries operating with both humans and machines
Humans calibrate their trust in machines based on past experiences [12]	Memory and recognition task with the help of an automated decision aid	The cohesion between humans and machines evolves and stabilizes over time, which positively influences the prevalence of errors

3 Proposed Methods

The primary reason for selecting Eye-tracking mechanisms for human errors detection is its possibility to develop methods task independently and can be adapted to most of the environments [13]. Studies show that eye tracking methods can track eye motions to the highest degree of accuracy possible. The methodology section looks at gazing eye mechanisms and measurement of cognitive eye movement to estimate error-making possibilities. The method section will base the findings of the error prediction mechanism on the outcomes of experimental processes derived by using gazing mechanisms and webcam applications to identify incidences of errors.

3.1 Video Monitoring Based Gaze Tracking

Video-based method of eye tracking provides a suitable method of gaze tracking. The experiment relied on non-invasive and video-based tracking method to attain accurate results of error prediction prior to errors made. The non-invasive method relies on a multi-camera or mono-camera eye tracker that most common systems are equipped with (Fig. 2). A major benefit of this methodology is that it entails minimal mathematical details, which makes the calculation processes and methods of generating results less complex.

The first step was to illuminate the eye with an infrared light source as it was essential in the production of glints on the cornea. Here, the reflection generated (corneal reflection), was a useful output source to help identify the reference point for the estimation of the gaze. The glint vector produced by the pupil remains constant

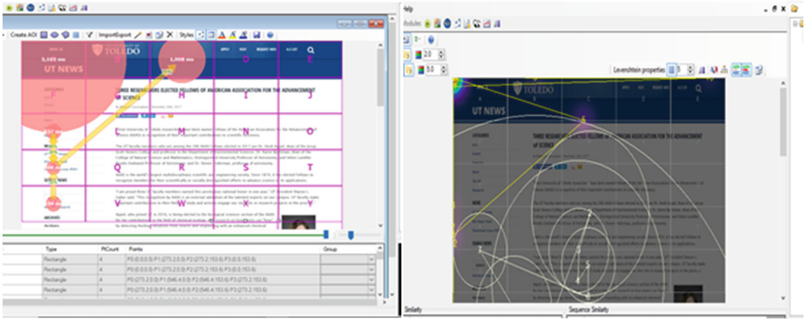

Fig. 2. OGMA eye-tracking working and setup

irrespective of the eye and head movement. But, employing a single camera may not produce the desired illumination effect. The above gazing method bases its working mechanism on the ability to trace the directions of eye movements [14]. Accordingly, the process provides evidence on the ability to generate different sets of images for different eye movements. When the eye is captured on various actions, it is simpler to develop feasible mechanisms upon which the simulation of errors can also be identified. The method is implemented using stranded software library called Open Gaze and Mouse Analyzer (OGMA), with attention maps calculated as the Gaussian distribution of each fixation in a stimulus slide using Eq. 1 [14].

$$f(x, y) = \frac{e^{-\frac{x^2+y^2}{2\sigma^2}}}{2\pi\sigma^2} x, y \in [-s, s] \tag{1}$$

Suitable outcomes on the assessments were dependent on the results garnered in the process. For instance, irrespective of whether a post-completion error occurred or not, the data collected in the process would be useful in generating a fixation measure to conduct better prediction capacities of alerting human beings on the nature of errors that they commit.

3.2 Predicting Human Error Through Cognitive State Measurement

Eye movement measures are dependent on the cognitive state of the individual. The fact that the eye movements are an essential aspect in evaluating human abilities to avoid making errors implies that a better understanding of the cognitive state of a person would yield superior outcomes. We analyze the commonly used mean pupil diameter change (MPDC) [15], to evaluate cognitive state, which is calculated as the average for the six curvy road segments of the difference between the mean pupil diameter in a given curvy segment and the overall mean pupil diameter for a given subject. Polar research results proved that MPDC is a function of cognitive state [13, 15].

By using MPDC values and fixation values obtained through Eq. 1 we would be able to predict potential human error, because, a human error is going to be a function of human cognitive state and human gaze point on a screen. Other parameters that may play a role for a human to make an error are natural skills, knowledge, experience, environment, and auxiliary systems. For the experiment, simulation condition the paper considers as other parameters effect are standardized and represented with single variable value.

3.3 Data Generated and Processing

Using two major factors that from eye tracking data gaze and MPDC we will be able to predict the probability of distraction or probability of human going to commit error assuming rest of the conditions are ideal. Eye tracking data generated from Eq. 1 is a density value that ranges between to 0 to positive making higher the value better the human viewing targeted area. Whereas, MPDC ranges between −0.8 mm to 1.2 mm [13, 15] significantly in general population. According to studies, the larger the pupil diameter the more the cognitive load on human, and the more it directly relates to a higher probability for a human to make a mistake.

We normalized gaze values from 0 to 1 and MPDC values inverted to get direct relations from chances for humans to make errors that can be represented by Eq. 2. Using Eq. 2 we have generated different sets of synthetic data that is used to train machine learning prediction models to predict human error chances.

$$\begin{gathered} \textit{human error chances} = (\mathrm{gaze} * \mathrm{w1}) + (-\mathrm{MPDC} * \mathrm{w2}) \\ \mathrm{weights}\ \mathrm{w1} + \mathrm{w2} = 1 \end{gathered} \tag{2}$$

3.4 Prediction Methods

Support Vector Machines is a classification method that maps data points onto a different dimensional space that includes high dimensional space. In this technique, a support vector network maps a known two input factors into an unknown data space through nonlinear mapping techniques [16]. In general, SVMs offer fast training in a distributed manner with the efficient use of processing resources. Due to the relatedly binary nature of SVM outputs, this technique will be one of the optimal fit for our data. SVM is implemented with regression learning in this paper for our data. In addition to the training and validation accuracy metrics an additional metric, known as the fit score, can be applied to such binary classifiers as SVMs.

Decision Tree Learning. A method used classify data by sorting down the tree from the root to child nodes. Each node in tree specifies a test case and each child node represents one of the possible value for that test case. An instance is classified from root to tree, then move down to tree branches [17]. Decision tree learning is implemented with regression learning in this paper for our data. Regression tree best suits for a continuous data type which makes this model a good fit for the output data type of this research.

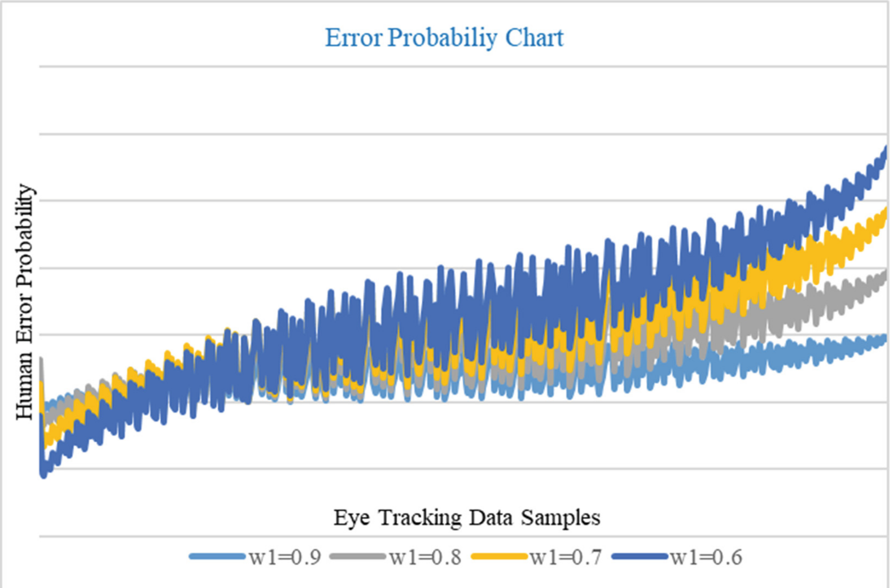

Fig. 3. The graph presents eye tracking with corresponds to human error possibility with various weights for gaze and MPDC

4 Results and Discussion

4.1 Results

This section provides the results for the MATLAB simulation data and python implementation with respect to both accuracy and computational runtime. In addition, discussion prediction methods are given. Also discoursed are the key distinctions and observations between human error detection capabilities through machine learning. Figure 3 presents data generated by MATLAB statistical methods with weights w1 from 0.9 to 0.6, it was assumed that focusing on the different surface than desired is a greater reason for a human to make more mistakes in a critical task execution than the cognitive load at that time. However, for the sake of experiments we have provided various weights, however, the combination of two weights is always equal to one. One of the critical points that needs mention regarding Eq. 2 is it derived in a way such that greater the result of the equation is, greater the chances for human to make more errors.

Figure 4(a) and (b) provide a comprehensive analysis of the SVM model. 4(b) provides the respective training and validation accuracies by γ value and by penalty term. Figure 4(a) provides the computational runtime with respect to γ and penalty. The results show in SVM the model trained to predict human error results in near 98% accuracy.

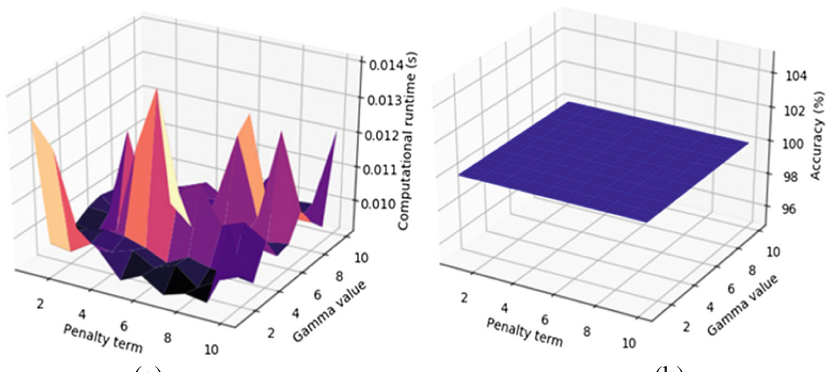

Fig. 4. (a). SVM model computational characteristic and (b) accuracy characteristics.

Figure 5(a) and (b) provide a comprehensive analysis of the decision tree model. 5(b) provides the respective training and validation percentile accuracies by minimum sample by leaf and sample for the split. Figure 4(a) provides the computational runtime with respect to minimum sample for leaf and split.

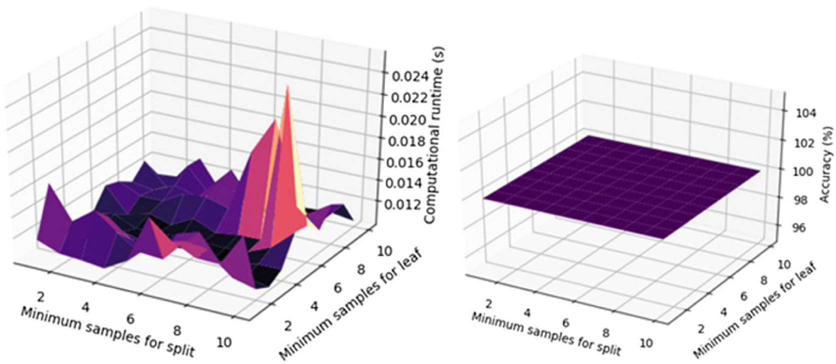

Fig. 5. (a). Decision trees model computational characteristics and (b) decision tree accuracy characteristics.

The first method of understanding the cognitive mechanisms of human intuition indicated that better results could be yielded when implemented to facilitate real-time tracking. The efficiencies of machine learning models to predict human error probability through eye tracking prove to be effective from Figs. 4(b) and 5(b). However, there is practical validation still remaining to make any conclusion on efficiency in predicting human errors in real tasks. With that, there is still need for a study in human-machine teaming affected by this through eye tracking before we conclude this method to be effective. Those studies still remaining to be done. For now, according to our results, it can be concluded that machine learning based methods such as SVM and Decision tree methods are effective in predicting human errors only through eye tracking data.

4.2 Discussion

To recap the proposed method in simple terms, the paper presents a study about predicting human errors in HMT environments through eye tracking data. Gaze density and MPDC are two important features extracted from eye tracking data and utilized in error predicting model development. Gaze density represents human eye physical location on a two-dimensional space which is directly proportional to human error probability whereas, MPDC gives an inverse relation of congestive load and another factor in human error probability. Both are statistically used to predict chance where a human is going to make an error to alert the person. However, failure to look at the fixated points implies that the user is highly likely to make an initial error at the first instance. Nonetheless, the results and studies showed that there are incidences when post-completion errors diminish accuracies, thereby leading to inevitable events. The efficiency of the process is also a by-product of different eye measures, which are developed to feasibly provide relevant results of the movements parallel with the performance of a task. The predictors from the regression model may vary depending on the degree of the error change and the actual probabilities notably predicted.

5 Conclusions and Future Work

Real-time error prevention mechanisms can be attained after effective learning of the way they occur. The human eye is often exposed to various intuitions all of which have a significant effect on the possibility of making an error. First, the paper illustrated that the ability to predict real-time eye tracking and data collection. This paper presents SVM and decision tree based prediction models that can use eye-tracking data to predict human error probability. Besides, the ability to gauge human intuitions provides a feasible support user interaction in various ways. Prediction models can be used in alerting human before committing a mistake. Human error models will be helpful to machines to understand the human state that bonds the cohesion between human and machine in HMTs. Future work that needs to be conducted in implementing the model in real time to test human subject data for validation. Human studies need to be conducted to understand the practical feasibility of human alert system in HMTs, and performance studies have to be conducted in HMTs with this feature included.

Acknowledgments. The University of Toledo and Round 1 Award from the Ohio Federal Research Jobs Commission (OFMJC) through Ohio Federal Research Network (OFRN) fund this research project; authors also appreciate support of the Paul A. Hotmer Family CSTAR (Cybersecurity and Teaming Research) Lab and EECS (Electrical Engineering and Computer Science) Department at the University of Toledo.

References

1. IBM: The human-machine interchange: how intelligent automation is reconstructing business operations. Business consulting (2017) public.dhe.ibm.com/common/ssi/ecm/gb/en/gbe03879usen/human-machine-interchange.pdf
2. Guszcza, J., Lewis, H., Evans-Greenwood, P.: Why Humans and computers think better together. Deloitte Insights (2017). https://www2.deloitte.com/insights/us/en/deloitte-review/issue-20/augmented-intelligence-human-computer-collaboration.html
3. Ziemke, T., Schaefer, K.E., Endsley, M.: Situation awareness in human-machine interactive systems, pp. 1–2 (2017)
4. Eddy, N.: Relationship between humans, machines must evolve: gatner. E-Week. http://www.eweek.com/small-business/relationship-between-humans-machines-must-evolve-gartner
5. Aponso, B.: HAT tricks: understanding human autonomy teaming through applications (2017)
6. McCourt, M.J., Mehta, S.S., Doucette, E.A., Curtis, J.W.: Human-machine teaming for effective estimation and path planning. In: Micro-and Nanotechnology Sensors, Systems, and Applications VIII, vol. 9836, p. 98361 W. International Society for Optics and Photonics (2016)
7. Mu, L., Xiao, B.P., Xue, W.K., Yuan, Z.: The prediction of human error probability based on Bayesian networks in the process of task. In: 2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 145–149. IEEE (2015)
8. Ramakrishnan, R., Zhang, C., Shah, J.: Perturbation training for human-robot teams. J. Artif. Intell. Res. **59**, 495–541 (2017)
9. Yang, X.J., Unhelkar, V.V., Li, K., Shah, J.A.: Evaluating effects of user experience and system transparency on trust in automation. In: HRI, pp. 408–416 (2017)
10. Baluja, S., Dean, P.: Non-intrusive gaze tracking using artificial neural networks. In: Advances in Neural Information Processing Systems, pp. 753–760 (2014)
11. Dzeng, R.J., Lin, C.T., Fang, Y.C.: Using eye-tracker to compare search patterns between experienced and novice workers for site hazard identification. Saf. Sci. **82**, 56–67 (2016)
12. Goswami, A., Gursimran, W., Mark, M., Ganesh, P.: Using eye tracking to investigate reading patterns and learning styles of software requirement inspectors to enhance inspection team outcome. In: Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, p. 34. ACM (2016)
13. Maguire, R.: Validating a process for understanding human error probabilities in complex human computer interfaces. Complexity in Design and Engineering, pp. 313–326 (2005)
14. Voßkühler, A., Nordmeier, V., Kuchinke, L., Jacobs, A.M.: OGAMA (open gaze and mouse analyzer): open-source software designed to analyze eye and mouse movements in slideshow study designs. Behav. Res. Methods **40**(4), 1150–1162 (2008)
15. Piquado, T., Isaacowitz, D., Wingfield, A.: Pupillometry as a measure of cognitive effort in younger and older adults. Psychophysiology **47**(3), 560–569 (2010)
16. Guenther, N., Schonlau, M.: Support vector machines. Stata. J. **16**(4), 917–937 (2016)
17. Su, J., Zhang, H.: A fast decision tree learning algorithm. In: AAAI, vol. 6, pp. 500–505 (2006)