Machine Learning for Efficient Assessment and Prediction of Human Performance in Collaborative Learning Environments

Pravin Chopade, Saad M Khan, David Edwards, and Alina von Davier
Artificial Intelligence and Machine Learning
ACTNext, ACT Inc.
Iowa City, IA, USA

Abstract—**The objective of this work is to propose a machine learning-based methodology system architecture and algorithms to find patterns of learning, interaction, and relationship and effective assessment for a complex system involving massive data that could be obtained from a proposed collaborative learning environment (CLE). Collaborative learning may take place between dyads or larger team members to find solutions for real-time events or problems, and to discuss concepts or interactions during situational judgment tasks (SJT). Modeling a collaborative, networked system that involves multimodal data presents many challenges. This paper focuses on proposing a Machine Learning - (ML)-based system architecture to promote understanding of the behaviors, group dynamics, and interactions in the CLE. Our framework integrates techniques from computational psychometrics (CP) and deep learning models that include the utilization of convolutional neural networks (CNNs) for feature extraction, skill identification, and pattern recognition. Our framework also identifies the behavioral components at a micro level, and can help us model behaviors of a group involved in learning.**

Keywords-Machine Learning, Collaborative Learning, Deep Learning, Computational Psychometrics, Skills, Human Behavior.

I. INTRODUCTION

Collaborative learning methods have been implemented broadly by organizations at all stages, as research recommends that active human involvement in cohesive and micro group communications is critical for effective learning [1]. In current research, an important line of inquiry focuses on finding accurate evidence and valid assessment of these micro-level interactions which supports collaborative learning. Even though there is a long practice of using mathematical models for modeling human behavior, Cipresso (2015) [2] introduced a computational psychometrics-based method for modeling characteristics of real behavior. Cipresso's [2] article provides us with a way to extract dynamic interaction features from multimodal data for modeling and analyzing actual situations.

In this paper, we propose a three-stage method to explore and study collaborative group behaviors. The first stage integrates and processes multimodal data obtained in a collaborative learning environment (CLE) that includes sensor input, audio, video, eye tracking, facial expressions, movement, posture, gestures, and behavioral interaction log data. The second stage performs feature extraction and cloud computation using computational psychometrics (CP) and convolutional neural network (CNN)-based deep learning for skill, pattern, and trend identification. Finally, the third stage uses the parameters measured in the previous two stages to understand and model group interactions, competencies, and collaborative behavior at a micro-level. The third stage uses machine learning for effective assessment and visualization of group dynamics such as correctly assessing the increase in the groups' level of shared understanding of different perspectives, and ability to clarify misconceptions.

This paper is an extension of our ongoing work [3], [4] and here we present details regarding the ML architecture for data-intensive computing and efficient assessment. Our paper is organized as follows: in Section II we briefly discuss our related work on ML for multi-modal human interaction analytics. In Section III we discuss a three-stage architecture for large-scale CLE and the layout of the functional components. Based on the architecture discussed in section III, preliminary experimentation analysis is discussed in Section IV. In Section V we discuss scalable applications of our framework for next-generation collaborative learning and assessment systems, as well as for the Department of Homeland Security (DHS) and the Department of Defense (DoD). Section VI concludes with directions for future work.

II. RELATED WORK

In the past few years, the Artificial Intelligence (AI) and Machine Learning (ML) communities have been putting their efforts into presenting and publishing advanced methods for processing and analyzing human behavior related multi-modal data. Due to page limitations, it is not possible to cover and cite all of these works, but we will provide brief highlights regarding our own work.

In our most recent work, Chopade et al [3], [4], [5], presented a framework which incorporates computational psychometrics (CP), Artificial Intelligence (AI), and a Machine Learning (ML)-based system architecture, methodology, and related algorithms to find patterns of interactions, learning, team relationships, and effective teamwork assessment of collaborative problem solving (CPS) and a collaborative learning environment (CLE).

Khan [6] presented an approach which uses multimodal telemetry data for two pilot studies from the domains of

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

collaborative learning and illustrated a framework to analyze participant behavior patterns through temporal dynamics.

Polyak et al. [7], [8] presented the application of CP for the measurement of CPS skills. They performed machine learning analysis on actual behavioral and post-game studies data. Their CPS game tasks were designed and developed based on the psychometric principles of Evidence-Centered Design (ECD) which are associated with ACT's Holistic Framework (HF) [9]. In their experiment, they performed a cluster analysis on participants' sub-skill performance scores and configurations of particular dialog responses obtained from participants' game-play data.

In this paper, we extend the foundation laid out in our recent work to implement CP, AI, and ML for effective assessment of teamwork skills. The next sections discuss in detail, the three-stage architecture for data-intensive computing and efficient assessment.

III. The Architecture

Massive data-intensive, high-performance, scalable computing is transforming our capabilities to gather and analyze data in different forms. This may lead to new inventions and discoveries in education, science, and technology. This may also impact learning and assessment (LAS) platforms. Data-intensive computing changes our thinking about education, science, and technology, by accelerating an ability to perform advanced data collection and computing [10]. Data-intensive scalable computing has a high potential for unique applications. This will be more challenging when we need to scale up the platform to handle large-scale datasets (Terabyte, Petabyte, Zettabyte scale). Recent improvements in computing have led to substantial progress towards the visualization capabilities of such data. Data analytics and visualization will serve as a vital tool for the validation of expected results by accurately identifying patterns and relationships in data. Visualization may play an essential role in understanding the big picture in group interactions within the CLE and may assist in detecting hidden factors. Convolutional Neural Network (CNN) and one of its approaches – deep learning (DL) may require the use of a highly efficient Graphical Processing Unit (GPU) implementation or for training on multiple GPUs [11], [12], [13] or for applications of this architecture to substantial learning populations.

Fig. 1 shows a possible arrangement of components for machine learning (ML) based data-intensive computing and efficient assessment. Some of the components are listed here:

A. Data Integration and Processing

B. Massive data intensive CNN (deep learning) based cloud computing and Computational Psychometrics (CP)

C. Effective Assessment (EA)

A. *Data Integration and Processing*

Establishing identities from vast volumes of CLE interaction multimodal data obtained from different sources is an essential task of data analytics and computation architectures [14], [15]. Large amounts of CLE multimodal interaction data that provide the identities of humans, machines, sensors, etc. collected from different sources will be processed through a set of solutions built upon the Hadoop data analytics platform. This arrangement considers the individual's identity, such as username, email, real name, gender, eye color, fingerprints and user's input data such as eye tracking and models of behavior [16]. A data cluster would enable massive data integration, processing, and performing of such tasks as data collection, information extraction, and storing large-size distributed datasets for long-term access. Data collection is an essential phase of acquiring data from multiple sources, categorizing it, and passing it to the next stage in the process as shown in Fig. 1. Data on humans, machines, and other entities can be categorized as structured or unstructured and incorporated into a distributed Hadoop infrastructure.

B. *Massive data intensive CNN (deep learning) based cloud computing and Computational Psychometrics (CP)*

Once data has been categorized (as described in part 'A'), we can use a computation cluster to analyze the data on a cloud platform to understand individual and group abilities. We plan to use Python/R to run deep learning in the cloud using ACT's enterprise learning analytics platform (LEAP). We plan to deploy a feature extraction algorithm including CNN based deep learning for skill, pattern, trend identification, and for achieving state-of-the-art accuracy in feature classification. We plan to update this network structure over time to make this a dynamic system.

Cloud computing moves computation closer to the data. The main advantage to this process is that this approach is scalable to hundreds of computing nodes, each providing at least a modest performance. Data-intensive cloud computing platform consists of 3 layers, i.e., map/reduce on top of Hadoop, HPC (high-performance computing) infrastructure for massive data processing and CNN deep learning-based cloud computing. For HPC, we use a method for dynamic partitioning of processes. Network updater adds new network specific data entries under the situation of any real-time events such as changes in the team and their activities.

Computational Psychometrics (CP): Collaborative problem solving (CPS) is identified as cross-cutting capabilities which is part of ACT's Holistic Framework [9]- a comprehensive description of the knowledge and 21st-century skills individuals need to know and be able to succeed at school and work [9], [17]. Advanced development in computational techniques and analytical tools has produced new pathways in CPS research [18]. Simultaneously, psychometrics researchers started developing assessments using advanced computational techniques and analytical tools which have emerged as a novel interdisciplinary field of prominent research called, "Computational Psychometrics (CP)" [19], [20], [21]. CP is a new area of learning and assessment (LAS) research, which consist of data-driven machine learning and information querying computer science methods, theory-driven psychometrics, and stochastic theory – all used in order to measure learner's latent abilities in real time [7], [19], [20], [22]. As shown in Fig. 1, computation cluster integrates CP and CLE components.

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

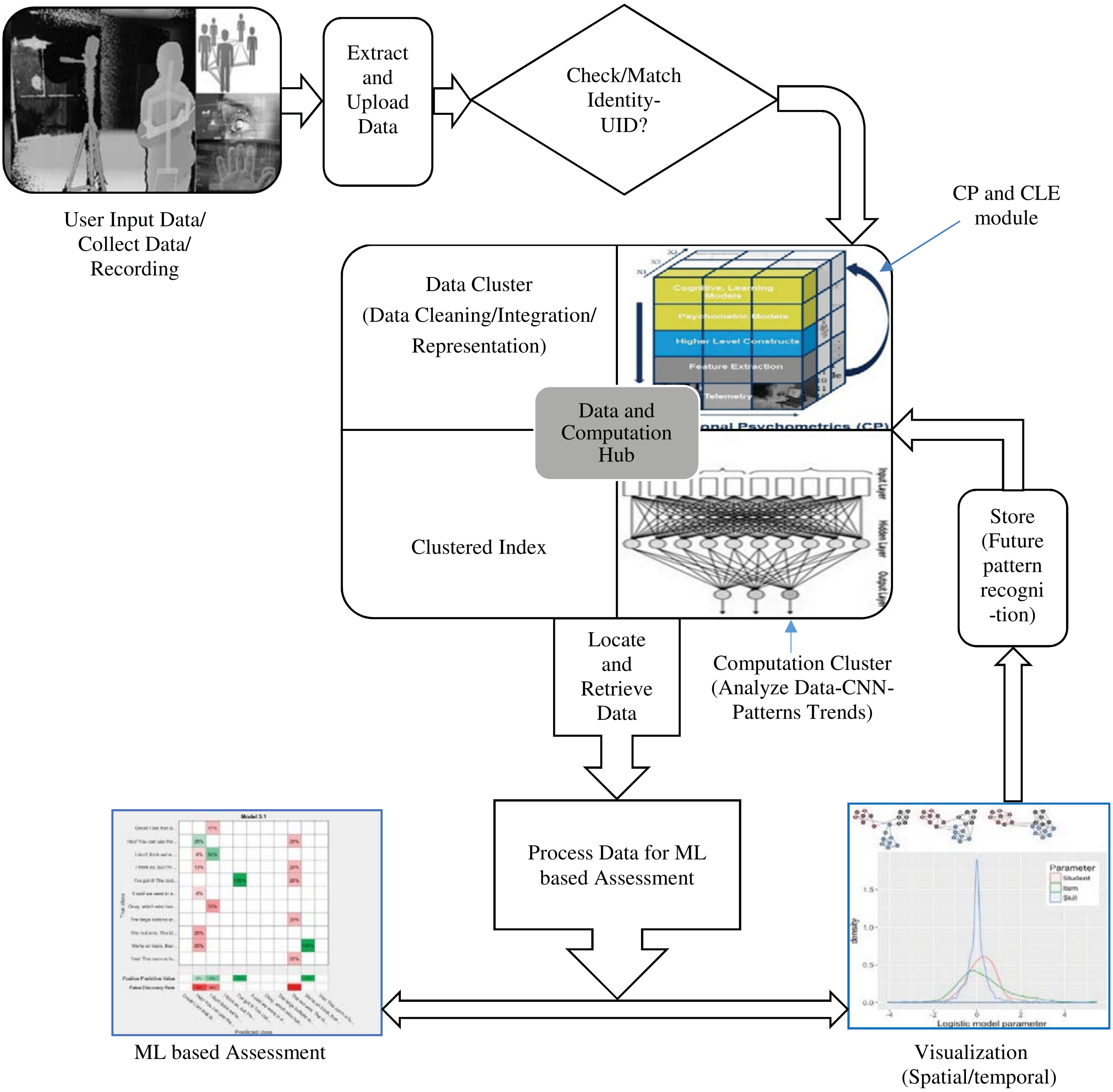

Fig. 1. Framework for ML based data intensive computing and efficient assessment. Copyright ACTNext, ACT Inc, USA

C. *Effective Assessment (EA)*

The effective assessment (EA) module locates and retrieves data after the computation is finished using the CNN and CP modules illustrated in section II-B. The CNN and CP modules carry out feature extraction, model training, and pattern identification from two player dyadic gameplay logs and behavioral data. Through effective assessment, we aim to analyze human behavior as it relates to specific situations in the game, to detect the dynamics of group behavior, such as shared understanding and engagement at the micro level.

For achieving our objectives, the EA module identifies/detects clusters of interactions, and abilities based on degree, connected components, and it performs collaborative learning skill analysis i.e., whether there is any change (increase) in group understanding for a given problem or situation. For example, it carries out real-time regular and critical event/person analyses such as situational data processing and real-time data correlation (correlation of skills, attainment level between groups). Using machine learning, we analyze collaborative learning network attributes such as users (nodes), interactions (links, weighted flow), knowledge (skills/abilities), local and global system parameters (engagement, shared understanding), behaviors, group density, and cluster formation and centrality.

The EA module also performs predictive decision-making based on group behaviors. This module stores future patterns which can be used for any exploratory analytics and effective visualization of past and present group interactions and relative performance. Biometric data is analyzed for biometric database matching and for biometric image processing. Biometric data are unique physical characteristics such as face, eye tracking/ iris, and fingerprints, which can be used for automated recognition [2], [6], [23].

Based on the three-stage architecture discussed in this section, section-IV demonstrates preliminary experimentation analysis.

IV. Preliminary Experimentation Analysis

In this section, we demonstrate some of the preliminary experimentation steps. The study participants play, 'Crisis in Space (CIS)' a collaborative problem solving web-based game published by GlassLab, Inc., a non-profit organization located in Redwood City, California [24]. CIS is a two player (dyadic) CPS game, played in a collaborative learning environment (CLE), in separate rooms communicating via Skype, and if they so choose, with the chat mechanism within the game. The setup for the rooms will be identical. They will include a laptop with Tobii software [25] installed, an external monitor with the Tobii eye tracking [25] unit mounted below the screen, with a webcam (with internal microphone) mounted above the screen. A wireless keyboard and mouse will also be connected to the laptop. As shown in Fig. 2, this study focused on collecting the following: game log data, user eye tracking, and user portrait video/audio, chat logs (conversation flow), behavioral expressions, object clicks, time in-game & between games. Our objective is to use this CPS human-human (HH) game log-data for team interaction analysis and for finding teamwork skill evidence based on the CPS construct.

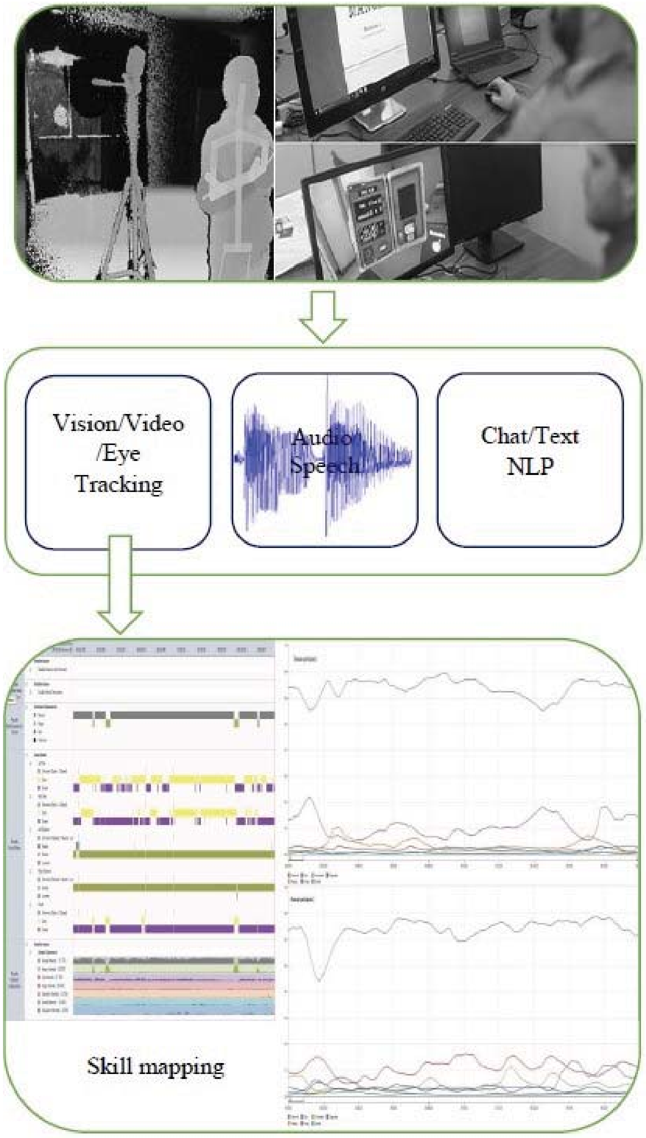

Fig. 2. CLE multi-modal data analytics and skill mapping process.

Copyright ACTNext, ACT Inc, USA

As illustrated in Fig. 2, from two-player dyadic CPS game interactions, we extract four types of data files: video, eye tracking, audio, and chat/text. We then use all these source files for identity matching or for naming convention purposes. Then we perform CNN-based machine learning analysis for feature extraction, correlation, and pattern identification. We processed video files through *Noldus* Facereader and observer analysis [26] which produces different behavioral markers and emotional states for dyadic gameplay (third block in Fig. 2). We also performed in-depth learning analysis using fine-grained data obtained from numerical and behavioral analysis with the

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

MATLAB Statistics and Machine Learning Toolbox [27], [28]. Later, we used these behavioral markers and emotional states for CPS teamwork skill mapping.

V. Scalable Applications

Our ML-based data-intensive computing framework has wide-ranging scalable applications including next-generation collaborative learning and assessment systems, DHS, Defense-Army, Airforce, Navy soldiers/Team training, and development. U.S. Army is considering learner and team-centric training, which will allow the development of mission-capable militaries, and organized teams to handle (win) in complex situations [29], [30], [31], [32].

VI. Conclusions And Future Work

In this paper, we presented a machine learning (ML)-based system architecture to identify evidence about teamwork skills from the behavior, group dynamics, and interactions in the CLE. We developed a three-stage robust architecture for data-intensive computing and efficient assessment of teamwork CPS skills.

In our future work, we will attempt to build text-based Natural Language Processing (NLP) / Machine Learning (ML) models to identify or classify various performances of CPS subskills from the chat logs, audio/video interactions data collected throughout the study. Additional feature extraction that may be used during this phase will be implemented for CNN based pattern identification. The knowledge gained in developing this baseline model will represent significant guidance for proceeding phases and potential studies to follow.

Acknowledgment

The authors would like to thank Andrew Cantine-Communications and Publications Manager, ACTNext and Jennah Davison- Intern, Communications and Publications, ACTNext for editing this work. We are also thankful to ACT, Inc. for their ongoing support as this paper took shape.

References

[1] L. Lei, J. Hao, A. von Davier, P. Kyllonen, and J-D. Zapata-Rivera. "A tough nut to crack: Measuring collaborative problem solving." Handbook of Research on Technology Tools for Real-World Skill Development. IGI Global, 2016. 344-359. Web. 28 Aug. 2018. doi:10.4018/978-1-4666-9441-5.ch013

[2] P. Cipresso, "Modeling behavior dynamics using computational psychometrics within virtual worlds," *Frontiers in Psychology*, vol. 6, no. 1725, p. 22, 6 November 2015.

[3] P. Chopade, K. Stoeffler, S. M. Khan, Y. Rosen, S. Swartz, and A. von Davier, "Human-Agent assessment: Interaction and sub-skills scoring for collaborative problem solving," in: Penstein Rosé C. et al. (eds) *Artificial Intelligence in Education. AIED 2018. Lecture Notes in Computer Science*, vol 10948, pp 52-57, June 2018 Springer, Cham, DOI https://doi.org/10.1007/978-3-319-93846-2_10

[4] P. Chopade, S. M. Khan, K. Stoeffler, D. Edward, Y. Rosen, and A. von Davier, "Framework for effective teamwork assessment in collaborative learning and problem solving," *in 19th International Conference on Artificial Intelligence in Education (AIED),* LAS, The Festival of learning, ARL workshop "Assessment and Intervention during Team Tutoring" AIED-2018, vol-2153, pp-48-59, July 2018. CEUR-WS.org/Vol-2153/paper6.pdf

[5] P. Chopade, M. Yudelson, B. Deonovic, and A. von Davier, Modeling Dynamic Team Interactions for Intelligent Tutoring, in Joan Johnston, Robert Sottilare, Anne M. Sinatra, C. Shawn Burke (ed.) Building Intelligent Tutoring Systems for Teams (Research on Managing Groups and Teams, vol. 19) Emerald Publishing Limited, pp.131 - 151, ISBN: 978-1-78754-474-1, September 2018.

https://www.emeraldinsight.com/doi/abs/10.1108/S1534-085620180000019010

https://books.emeraldinsight.com/page/detail/Building-Intelligent-Tutoring-Systems-for-Teams/?k=9781787544741

[6] S. Khan, "Multimodal behavioral analytics in intelligent learning and assessment systems," in Innovative Assessment of Collaboration. Methodology of Educational Measurement and Assessment, A. von Davier, Z. M. and K. P., Eds., Springer, Cham, 2017, pp. 173-184.

[7] S. Polyak, A. von Davier and K. Peterschmidt, "Computational Psychometrics for the measurement of collaborative problem-solving skills," in *Proceedings of ACM KDD conference*, Halifax, Nova Scotia CANADA, August 2017 (KDD2017), Halifax, Nova Scotia CANADA, 2017.

[8] S. Polyak, A. von Davier, and K. Peterschmidt, "Computational Psychometrics for the measurement of collaborative problem solving skills," *Frontiers in Psychology*, 8, pp.20-29, 2017.

http://doi.org/10.3389/fpsyg.2017.02029

[9] W. Camara, R. O'Connor, K. Mattern, and M.Hanson, "Beyond academics: A Holistic Framework for enhancing education and workplace success." *ACT Research Report Series*, 2015 (4). ACT, Inc. 2015.

[10] C. Chen and C. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," *Information Sciences*, vol. 275, pp. 314-347, 2014.

[11] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *NIPS: Advances in Neural Information Processing Systems* 25, 2012.

[12] C. Piech, J. Bassen, J. Huang, S. Ganguli, M. Sahami, L. Guibas and J. Sohl-Dickstein, "Deep knowledge tracing," *in NIPS, In Advances in Neural Information Processing Systems*, 2015.

[13] "CNN Convolutional neural network," Wikipedia,

https://en.wikipedia.org/wiki/Convolutional_neural_network

[14] DOE, "Synergistic challenges in data-intensive science and exascale computing," Department of Energy, ASCAC Data Subcommittee Report, March 30, 2013.

[15] R. Moore, B. C., R. Marciano, A. Rajasekar, and M. Wan, "Data-Intensive computing," in The Grid: Blueprint for a New Computing Infrastructure, I. Foster and C. Kesselman, Eds., New York, Morgan Kaufmann Publishers Inc, ELSEVIER, pp. 105-128, 1998

[16] S. Creese, T. Gibson-Robinson, M. Goldsmith, D. Hodges, Dee Kim, O. Love, J. Nurse, B. Pike, and J. Scholtz, "Tools for understanding identity," in *Technologies for Homeland Security (HST), 2013 IEEE International Conference*, Nov 2013.

[17] "Organization for Economic Cooperation and Development (OECD)," PISA 2015 draft collaborative problem-solving framework., 2015.

[18] G. Grover, M. Bienkowsk, A. Tamrakar, B. Siddiquie, D. Salter, and A. Divakaran, "Multimodal analytics to study collaborative problem solving in pair programming," in *LAK '16 Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 2016.

[19] A. von Davier, "Computational Psychometrics in support of collaborative educational assessments," *Journal of Educational Measurement*, vol. 54, pp. 3-11, 2017.

[20] A. von Davier, "Virtual and collaborative assessments: Examples, implications, and challenges for educational measurement.," in Invited Talk at the Workshop on Machine Learning for Education at International Conference of Machine Learning, Lille, France, 2015.

[21] A. von Davier, Ed., Measurement issues in collaborative learning and assessment, vol. 54, *Journal of Educational Measurement*, [Special Issue], pp. 3-11, 2017.

[22] A. von Davier, M. Zhu, and P. Kyllonen, Eds., Innovative assessment of collaboration, Springer International Publishing, 2017.

[23] Y. Huang and S. Khan, "DyadGAN: Generating facial expressions in dyadic interactions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW),* Honolulu, HI, 2017.

[24] GlassLab, Inc. https://www.glasslabgames.org/

[25] Tobii Eye Tracking https://www.tobii.com/

[26] Noldus https://www.noldus.com/

[27] S. Moulder, T. Sheridan, A. Dulai, and G. Rossini, "Deep Learning in the Cloud with MATLAB R2016b," Mathworks, 2016.

[28] "Mathworks Inc. MATLAB- Statistics and Machine Learning Toolbox," 2017.

[29] "The U.S. Army learning concept for training and education 2020-2040, TRADOC Pamphlet 525-8-2," U.S. Army, April 2017.

[30] I. Magnisalis, S. Demetriadis, and A. Karakostas, "Adaptive and intelligent systems for collaborative learning support: A review of the Field," *IEEE Transactions on Learning Technologies*, vol. 4, no. 1, pp. 5-20, 28 January 2011.

[31] G. Goodwin, J. Johnston, and R. Sottilare, "Individual learner and team modeling for adaptive training and education in support of the US Army learning model: Research Outline," US Army Research Laboratory, September 2015.

[32] J. Fletcher and R. Sottilare, "Shared mental models in support of adaptive, instruction for teams using the GIFT tutoring architecture," *International Journal of Artificial Intelligence in Education*, pp. 1-21, 2017.

978-1-5386-3443-1/18/$31.00 ©2018 IEEE