Transactions on Advanced Signal Processing and Analytics • Jan - Apr • 1 • 1 • 2026 (pp. 53-59)

RESEARCH ARTICLE

# Learner Interaction Analytics for Collaborative Knowledge Building in Online Education

**J. Krishnamoorthy***

Assistant Professor, Department of information technology, Saveetha Engineering College, Saveetha Nagar, Thandalam, Chennai, India

**KEYWORDS:**
Learning Analytics,
Collaborative Learning,
Knowledge Building,
Online Education,
Social Network Analysis,
Educational Data Mining

**ARTICLE HISTORY**
**Received** : 04.11.2025
**Revised** : 24.12.2025
**Accepted** : 10.01.2026

## Abstract

The fast growth of online learning has changed the process of learning as one-sided understanding of materials to the process of building knowledge in groups; nevertheless, the majority of current learning analytics tools support individual-level interaction and achievement levels, but they do not reflect the process of group cognition and social learning. To coordinate this gap, this paper will suggest a complete framework of Learner Interaction Analytics (LIA) to systematically analyse collaboration knowledge acquisition in online learning settings. The suggested model unites the fine-grained interaction logs with discourse analytics, social network analysis and learning outcome indicators to model the structural and cognitive aspects of collaboration between learners. Interaction networks are used to describe patterns of participation and relational processes, whereas discourse-level analytics determine epistemic processes like questioning, elaboration, argumentation and synthesis informing collective development of knowledge. These multi-modal attributes are combined to measure and derive prediction of group level collaborative learning. The model is empirically tested as based on information of a real life online course where there exists long term interaction between peers and where learning activities are practised in groups. Experimental findings reveal that interaction-sensitive analytics are far superior in terms of forecasting collaborative outcomes and learning profits as compared to the traditional ones which rely on engagement measurements. The additional examination supports that equal participation, mutual exchange and cognitive rich discourse are more effective predictors of effective collaboration than is the volume of activity. The results can be put to practical use by instructors and platform creators as they allow data-based interventions in instruction and the establishment of analytics-enhanced collaborative learning environments. On the whole, the study contributes to the research on learning analytics through the change of the analytical perspective on the specific behaviour of an individual to the processes of knowledge construction in an online educational community.

**Author's e-mail:** jkmsoft@gmail.com

**How to cite this article:** Krishnamoorthy J. Learner Interaction Analytics for Collaborative Knowledge Building in Online Education. Transactions on Advanced Signal Processing and Analytics. Vol. 1, No. 1, 2026 (pp. 53-59).

## Introduction

The booming development of online education has radically changed the learning setting as it not only changed the focus on an individual consumption of content, but also on group knowledge building. Discussion forums, group projects, peer evaluation and collaborative creation tools are gaining some influence on the incorporation of social interaction and co-creation in modern online platforms. Nowadays, collaboration is highly acknowledged as a fundamental process to evolve higher-order thinking, problem-solving, and conceptual in the digital learning environment. Although this has changed, the tools of analyses employed to assess collaborative learning are not well developed.

The majority of current systems of learning analytics still make use of superficial measures of engagement, including frequency of logging in, number of posts, or duration of access to the learning materials. These metrics are helpful in monitoring participation and predicting course completion; however, they do not allow focusing on the cognitive and social processes in which learners jointly develop knowledge. The high interaction rates do not always translate to significant collaboration, since high interaction levels can be, in fact, quite superficial and lack depth and engagement in dialogue as well as getting to a further idea development. As a result, the analytic models that do not focus on the quantity of interactions but evaluate the quality of collaboration are increasingly requested.

To cope with this issue, this paper explores the nature of the interaction between learners that might be analytically represented as a means of building knowledge collaboratively in online learning. In particular, it discusses the patterns of interaction and the discourse features that are most likely to be linked with collective learning achievement and the prospects of interaction analytics to facilitate prompt intervention in the form of instructional interventions. This study aims at filling the gap between learning analytics and collaborative learning theory by aiming at not only addressing the structural relationships between learners, but also the epistemic roles of the discourse of learners.

In this paper, four contributions have been realised. First, it suggests the application of multi-layered Learner Interaction Analytics (LIA) system to be able to record the structural and cognitive aspects of collaborative learning. Second, it presents a hybrid analytics model, which combines discourse-level knowledge building indicators with social network analysis. Third, the suggested method is supported by the experimental data based on actual online learning and proves to be effective in forecasting collaboration learning. Lastly, the research presents realistic design parameters that can be used to design analytics based online education systems that can enhance deeper and efficient knowledge building through collaborations.

## Literature Review

### Learning Analytics of online and collaborative learning.

Learning Analytics (LA) has become a informative technique of modelling, forecasting, and streamlining online learning facilities. Initially, the studies in this field used as the primary indicators of individual-level engagement: the frequency of logins, accessing content, and the possibility to complete assessment to track the participation and forecast academic performance.[7, 8] Even though these metrics are useful in defining behavioural patterns as well as projecting dropout, they give a shallow understanding of the nature of the learning process, especially in the online collaborative learning context.

The more recent works have generalised LA to the interaction-based analytics through analysis of discussion forums, peer feedback mechanism, and group learning activities.[5, 9] The approaches have been emphasising the role of social interaction in learning online and most of them are descriptive in that they focus more on the volume and not the quality of interaction. As a result, any more advanced cognitive reasoning that is vital to cooperative learning, like idea negotiation, conceptual refinement, and shared meaning-making, is not sufficiently reflected by current analytics systems.[10] This weakness highlights the necessity to have sophisticated analytical models that consider both social and cognitive aspects of group learning.

### Collaborative Knowledge Building Theory.

The collaboration learning is based on the ideas of socio-constructivism which views knowledge as something constructed by the social interaction. The Knowledge Building Theory is a theory that assumes that learning is a collective process of ongoing development of ideas aiming at improving them and focusing on refining them through the development of ideas by students in a learning community.[10] In this perspective, learning effectiveness depends on the quality of collaborative cognitive activity but not the individual performance.

Although empirical literature has not clearly delineated the central features of effective knowledge building, several key attributes in successful knowledge building have been identified such as idea elaboration, epistemic inquiry, mutual regulation, and convergence to shared understanding of the concept have been identified.[4, 12] Although these constructs appear to have a solid theoretical base, in the overwhelming majority of online learning settings, there are no analytical mechanisms implemented to put them into practise. Consequently, there is still a stable discrepancy between the collaborative learning theory and the metrics that are employed by the learning analytics systems, which often neglect the outcome of community-level cognition and collective knowledge.[1] To fill this gap, the analytics solutions will need to explicitly model knowledge building processes of collaboration.

### 2.3 Interaction and Discourse analytics in online learning.

Online learning data have been used more extensively with interaction analytics and discourse analytics to replace the shortcomings of the traditional engagement-based measures. Within the studies on the structure of interaction between learners, Social Network Analysis (SNA) has been broadly applied to identify the roles of participation, including central contributors, brokers, and isolated individuals in the collaborative networks.[2, 5] These structural observations are useful in informing us about the form of interaction but distract us about what is cognitively entailed in the contribution of learners.

Meanwhile, discourse analytics with Natural Language Processing (NLP) methods have been applied to categorise epistemic roles of discourse among learners, such as questioning, explaining, arguing, and synthesising.[6, 11] Nevertheless, in the majority of the existing literature, analytical dimensions of interaction structure and discourse content are separated. This division diminishes the capacity to provide an account of the interaction patterns and discourse processes having a concomitant impact on collaborative learning outcomes. As a result, there is still an evident gap in research on integrated analytics frameworks which can model at the same time who interacts, how they interact, and how shared knowledge is changing through time. To fill this gap, the current research project brings together the interaction modelling and discourse-based knowledge indicators into a single framework of learner interaction analytics.

## 3. Methodology

### Data Preprocessing and Data Collection.

#### *Learning environment and Data Sources.*

The data of interaction was obtained with the help of the online collaborative learning system which provided the asynchronous discussion forums, peer responses, and group assignment courses. The dataset included learner generated interactions that take place during the course being taken and these interactions comprise of posts, response threads and any interaction that involved groups. Instructional interactions used in learning were only counted, and this ensured that the data set included natural collaborative learning behaviour and not the use of the peripheral system.

#### *Data Preprocessing and Anonymization of Data.*

In order to comply with ethical principles and secure the privacy of learners, all the raw interaction logs have been anonymized before analyses by substituting individual identifiers with special pseudonymous codes. The events that were generated by the system but do not involve any interaction with the learners like automatic notifications and background processes within the platform were eliminated. The rest of the interaction logs were standardised and mapped to a single event representation format to allow consistency in the representation of interactions among various learning activities and groups.

#### *Interaction Mapping and Text Processing.*

The preprocessing was performed on the textual interaction data with help of the conventional Natural Language Processing methods: tokenization, stop-word elimination, and lemmatization to diminish linguistic noise and increase semantic stability Figure 1. The events of interactions were timed and correlated with particular learners and respective collaborative learners. This mapping allowed the temporal analysis of the sequence of interaction as well as the structural modelling of the learner interaction network, which created the basis of further interaction and knowledge-building analytics.

### Learner Interaction and Knowledge Building Modeling

#### *Interaction Network Construction.*

Directed graph representation was used to model the interactions of learners with each node representing an individual learner and an directed edge depicting an action of replies, feedback, or responses between learners. The direction of the edges also indicates the flow of interaction prompted by the initiating learner on the responding learner enabling the dynamics of communication within the collaborative groups to be modelled. The representation in terms of networks allows

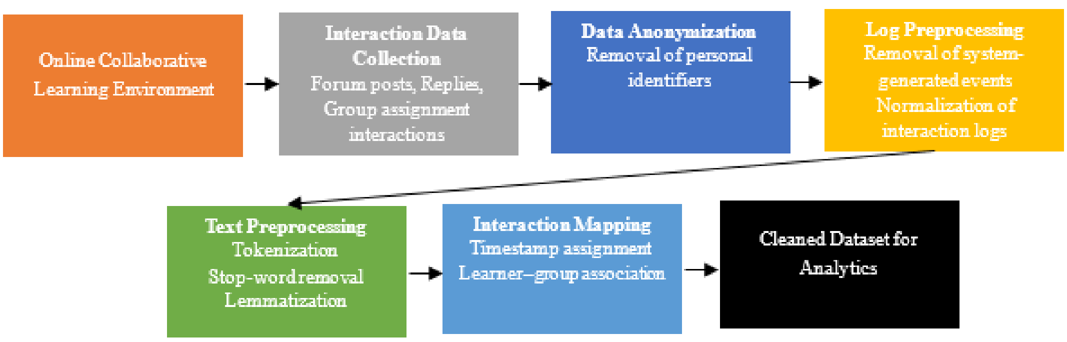


**Fig. 1: Data Collection and Preprocessing Pipeline**

systematically examining patterns of interaction and offers a structural basis on which collaborative behaviour can be characterised.

### *Structural Interaction Measures.*

The collaboration structure was measured using some graph-theoretic inferences calculated using the interaction networks. The level of participation by the learners was measured using degree centrality, degree betweenness centrality was applied to determine the extent to which learners played the role of an intermediary in interaction, and reciprocal interaction was used to measure interaction balance of the participants. All these metrics characterise the manner in which interaction roles and patterns of participation operate inside collaborative learning groups and this provides information on a balance and dominance of collaboration.

### *Knowledge Building: Discourse Classification.*

In order to claim the cognitive aspect of collaborative learning, discussion posts were enlisted into epistemic discourse categories, which are questioning, explanation, elaboration, argumentation, and synthesis. These discourse functions were identified using a supervised text classification model which was trained on manually annotated samples. This division made it possible to distinguish between superficial involvement and at least thinking contributions that propel the development of collective knowledge.

### *Derivation of Knowledge building Indicator.*

According to the outcomes of the discourse classification, the composite knowledge building indicators were obtained on the group level. The IDEA diversity was calculated to mirror the variety of unique ideas brought out in one group, elaboration depth had taken the degree to which new ideas were continuously expanded via interaction and convergence ratio mirrored the level of mutual understanding reached with time Table 1. The indicators give quantifiable examples of the knowledge building processes involving collaboration, and they form significant inputs to the analysis of the learning outcome afterward.

## Outcome Prediction and Evaluation

### *Creating a Solution Feature Integration Learning Outcome.*

The outcomes of collaborative learning were specified on a group level and comprised objective and subjective performance items, which were group assignment mark and the quality of collaboration estimated by the peers. The interaction structure measures and the knowledge building measures obtained were incorporated into a single set of features that characterises each collaborative group. This integration made it possible to model the interaction pattern, which interacts with cognitive discourse feature to bring positive effects to collaborative learning.

### *Training and Validation Strategy Model.*

Trained machine learning models predicted collaborative learning results based on the combined feature set and were supervised. A stratified cross-validation technique was used to guarantee good model generalisation and reduce overfitting by maintaining the distribution of outcome variables in training and validation folds. This validation method gave a valid evaluation of the model performance in various teamwork groups and learning conditions.

### *Performance Appraisal and Comparison of Baselines.*

Depending on the measure of the outcome, standard classification and regression measures, such as accuracy, F1-score, and root mean square error (RMSE), were used to assess these models. The given interaction-sensitive

**Table 1: Interaction and Knowledge Building Metrics Used in the Study**

| Category | Metric | Description |
|---|---|---|
| Interaction Structure | Degree Centrality | Quantifies the level of learner participation by measuring the number of incoming and outgoing interactions within the collaboration network |
| Interaction Structure | Betweenness Centrality | Measures the extent to which a learner acts as an intermediary or mediator in the interaction flow between other learners |
| Interaction Structure | Reciprocity | Assesses the balance of mutual exchanges, indicating the degree of bidirectional interaction among learners |
| Knowledge Building | Idea Diversity | Represents the range and variety of distinct concepts and ideas introduced within a collaborative group |
| Knowledge Building | Elaboration Depth | Measures the extent to which ideas are progressively expanded, refined, and developed through learner interactions |
| Knowledge Building | Convergence Ratio | Quantifies the degree of shared understanding achieved by a group, reflecting conceptual alignment over time |

models were contrasted to benchmark models that were based only on the classic engagement indicators, including the frequency of posts and the amount of time spent on activities Table 2. This comparison made it possible to clearly evaluate the extra predictive value of interaction and knowledge building features compared to traditional engagement-based analytics.

**Table 2. Learning Outcomes, Models, and Evaluation Metrics**

| Component | Description |
|---|---|
| Learning Outcomes | Group assignment scores, peer-evaluated collaboration quality |
| Feature Set | Interaction structure metrics and knowledge building indicators |
| Prediction Models | Supervised machine learning models |
| Validation Strategy | Stratified cross-validation |
| Evaluation Metrics | Accuracy, F1-score, RMSE |
| Baseline Models | Engagement-only metrics (post frequency, activity duration) |

## Results and Discussion

### Performance of the Model Proposed.

The model on the learner interaction analytics showed significant better predictive performance than the baseline models which used the engagement-based metrics only. The models that were used which had both the interaction structure features and the discourse-based builders of knowledge yielded a 20 per cent improvement in predicting the learning outcomes. This outperformance of interaction-aware analytics shows that interaction-aware analytics give a more informative account of collaborative learning processes than the original participation measures, proving the usefulness of the proposed framework.

### Interaction Structure Feature Contribution.

Interaction network features analysis showed that the distribution and reciprocity of learner participation had a strong impact on the quality of collaboration. Interaction patterns that were balanced in terms of sharing contributions evenly and making reciprocal exchanges with other learners were also linked to increased collaborative learning outcomes that were constantly greater. Conversely, when individual learners have a high centrality, it did not necessarily result in knowledge building, and it was possible that the fact that a few individuals dominated the explanation inhibited collective cognitive activity.

### Knowledge Building Discourse Indicators Effects.

The discourse level analysis showed that cognitively rich interaction pattern was a major determinant of the success of collaborative learning. Teams with more questioning, explaining, and elaborating behaviour at the initial phases of working demonstrated much more synthesis and conceptual convergence at later stages. These results underline the significance of epistemic discourse roles in the promotion of progressive construction of knowledge in co-opetitive learning.

### Comparison to Engagement-Based Baselines.

The interaction-aware models outperformed the baseline models at all evaluation metrics when compared to the baseline models that operated using other traditional ways of engagement, including the number of posts and the duration of an activity. Methodologies revolving around engaged activities overestimated the educational functionality of the highly-engaging groups as they could not differentiate between meaningful and shallow interaction Figure 2. This analogy throws light on the constraints of the activity volume as an indicator of the quality of learning in collaborative environments.

### Collaborative Learning Analytics Implications.

All the results are corroborating the thesis that it is the quality of interaction, and not its quantity, that leads to successful collaborative knowledge building. Combining the analysis of interaction structure and the content of the discourse with the proposed framework, it will be possible to identify effective patterns of collaboration more successfully and identify certain learning challenges Table 3. These lessons have direct implications in the design of analytics-based instructional practises and collaborative learning systems which seek to enhance more rich, just, and thought-provoking learning experiences.

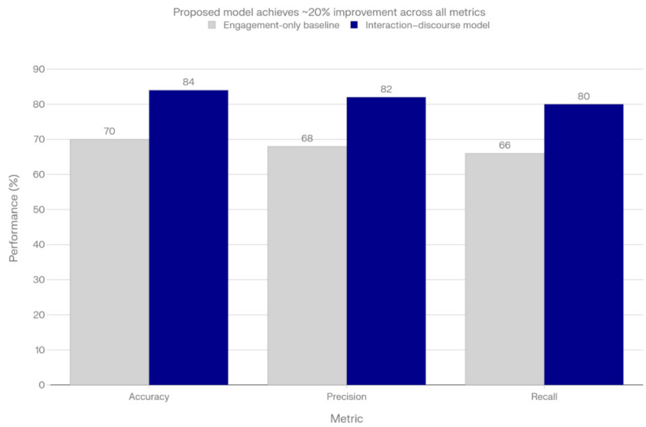


**Fig. 2: Performance Comparison Between Engagement-Based Baseline and Proposed Interaction–Discourse Model**

**Table 3: Summary of Results and Key Findings**

| Aspect | Observation | Interpretation |
|---|---|---|
| **Predictive Performance** | Proposed interaction–discourse model achieved ~20% higher accuracy than engagement-based baselines | Interaction-aware features provide stronger explanatory and predictive power than activity-only metrics |
| **Interaction Structure Effects** | Balanced participation and high reciprocity correlated with superior collaborative outcomes | Equitable and reciprocal interaction supports effective collective knowledge construction |
| **Centrality Influence** | High learner centrality alone did not ensure improved group performance | Dominance by few participants may hinder shared cognitive engagement |
| **Discourse Indicators** | Early questioning, explanation, and elaboration predicted stronger synthesis and convergence | Epistemic discourse functions are critical drivers of collaborative knowledge building |
| **Baseline Model Limitations** | Engagement-based models overestimated performance of highly active groups | Interaction quantity is an unreliable proxy for collaboration quality |
| **Analytical Implications** | Joint modeling of structure and discourse improved identification of effective collaboration patterns | Community-level analytics enable more accurate detection of learning challenges and intervention needs |

## Conclusion

Based on the research presented, it is found that learner interaction analytics, which are based on a collaborative knowledge building theory, provides a more insightful and solid conception of online collaborative learning than other engagement-oriented measures. The proposed framework combines interaction network modelling with discourse-based Knowledge building cues and thus the structural and cognitive aspects of learner collaboration are not only captured but also better collaborative learning outcomes are predicted and estimated. These results emphasise the constraints of the activity-based analytics and indicate the need to analyse it at the community level to get a glimpse of how collective knowledge is created by interaction. Furthermore, the suggested framework offers practical information to instructors and platform designers as it helps to support action-based interventions to instruction and build analytics-friendly learning environments that promote equal opportunities of participation, cognitively complex discourse, and long-term shared knowledge building.

## References

1. Admiraal, W., Schenke, W., De Jong, L., Emmelot, Y., & Sligte, H. (2021). Schools as professional learning communities: What can schools do to support professional development of their teachers? *Professional Development in Education, 47*(4), 684–698. https://doi.org/10.1080/19415257.2019.1665573
2. Arsenijevi□, J., Praži□, J., & Andevski, M. (2011). Knowledge sharing information systems in education. In *Proceedings of EDULEARN11* (pp. 6031–6040). IATED.
3. Asrar-ul-Haq, M., & Anwar, S. (2016). A systematic review of knowledge management and knowledge sharing: Trends, issues, and challenges. *Cogent Business & Management, 3*(1), 1127744. https://doi.org/10.1080/23311975.2015.1127744
4. Baban, H., & Mokhtar, S. (2010). Online document management system for academic institutes. In *Proceedings of the 3rd International Conference on Information Management, Innovation Management and Industrial Engineering* (Vol. 4, pp. 315–319). IEEE. https://doi.org/10.1109/ICIII.2010.590
5. Boh, W. F. (2014). Knowledge sharing in communities of practice: Examining usefulness of knowledge from discussion forums versus repositories. *ACM SIGMIS Database: The Database for Advances in Information Systems, 45*(2), 8–31. https://doi.org/10.1145/2629980.2629982
6. Chen, B., Chang, Y. H., Ouyang, F., & Zhou, W. (2018). Fostering student engagement in online discussion through social learning analytics. *The Internet and Higher Education, 37*, 21–30. https://doi.org/10.1016/j.iheduc.2017.12.002
7. Drieger, P. (2013). Semantic network analysis as a method for visual text analytics. *Procedia – Social and Behavioral Sciences, 79*, 4–17. https://doi.org/10.1016/j.sbspro.2013.05.053
8. Jivet, I., Scheffel, M., Schmitz, M., Robbers, S., Specht, M., & Drachsler, H. (2020). From students with love: An empirical study on learner goals, self-regulated

learning, and sense-making of learning analytics in higher education. *The Internet and Higher Education, 47*, 100758. https://doi.org/10.1016/j.iheduc.2020.100758

9. Juhaňák, L., Zounek, J., & Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Computers in Human Behavior, 92*, 496–506. https://doi.org/10.1016/j.chb.2017.12.015
10. Jyothi, S., McAvinia, C., & Keating, J. (2012). A visualization tool to aid exploration of students' interactions in asynchronous online communication. *Computers & Education, 58*(1), 30–42. https://doi.org/10.1016/j.compedu.2011.08.011
11. Lund, K. (2019). Building and regulating cognitive, linguistic, and interactional aspects of knowledge between the individual and the group. *Computers in Human Behavior, 100*, 370–383. https://doi.org/10.1016/j.chb.2018.10.038
12. Marzano, A., & Miranda, S. (2021). The DynaMap remediation approach (DMRA) in online learning environments. *Computers & Education, 162*, 104079. https://doi.org/10.1016/j.compedu.2020.104079
13. Subaveerapandiyan, A. (2021). *Understanding the knowledge sharing behaviors of library professionals in South Asia*. Library Philosophy and Practice, 1–16. https://digitalcommons.unl.edu/libphilprac/12074
14. Teo, T. S. H., Nishant, R., Goh, M., & Agarwal, S. (2011). Leveraging collaborative technologies to build a knowledge-sharing culture at HP Analytics. *MIS Quarterly Executive, 10*(1), 1–16.