Hadyaoui and Cheniti-Belcadhi
Smart Learning Environments (2023) 10:43
https://doi.org/10.1186/s40561-023-00262-w

Smart Learning Environments

RESEARCH Open Access

Ontology-based group assessment analytics framework for performances prediction in project-based collaborative learning

Asma Hadyaoui[1*] and Lilia Cheniti-Belcadhi[1]

*Correspondence: asmahadyaoui@gmail.com

[1] ISITC, PRINCE Research Laboratory, Sousse University, No 1 Route Principale, H. Sousse Sousse 401, Hammam Sousse, Tunisia

Abstract

This article introduces an ontology-based framework for group assessment analytics that investigates the impact of intra-group interactions on group performance within the context of project-based collaborative learning (PBCL). Additionally, it aims to predict learners' performance based on these interactions. The study involved 312 first-degree students specializing in transportation and technology engineering. The framework collects interaction data from discussion forums and chat rooms, conducts comprehensive data analysis, and constructs prediction models using supervised learning methods. The results unequivocally demonstrate that intra-group interactions significantly affect group performance in PBCL. The prediction model, with an accuracy metric of 0.92 and a final test score of 0.77, supports the credibility of the findings. Notably, the framework utilizes an ePortfolio specifically designed for group assessments, effectively managing both assessment and group data. This framework provides educators with a robust tool to assess group performance, identify areas requiring improvement, and contribute to shaping informed student learning outcomes. Furthermore, it empowers students by enabling them to receive feedback on their collaborative efforts, fostering enhanced interaction skills. These findings carry significant implications for the development and implementation of PBCL environments, offering educators valuable insights for evaluating student progress and making strategic decisions.

Keywords: Collaborative learning, Project-based collaborative learning (PBCL), Group assessment, Analytics framework, Intra-group interactions, Ontology-based, Supervised learning, EPortfolio

Introduction

In response to the global COVID-19 pandemic, a substantial shift toward digitalized educational approaches has unfolded. This transformation is characterized by a growing reliance on technological tools to facilitate and mediate student interactions, as encapsulated by the term Computer-Supported Collaborative Learning (Computer-Supported Collaborative Learning (CSCL)). At the core of CSCL lies an exploration of the intricate interplay between social interaction and computational aids. In this context, computational tools act as mediators, facilitating cooperative sensemaking

among participants. The effectiveness of this sensemaking process hinges on the dynamic interactions between participants and the computational tools. Through this detailed analysis of mediation, the foundational CSCL mechanisms come to light, drawing upon the realms of social science, learning theory, and computer science, as eloquently elucidated by Kaliisa et al. (2022).

Expanding on this evolving educational environment, our method for forecasting learner performance in collaborative project assessments aligns with the tenets of CSCL. This ability to predict equips educators with the means to identify students encountering difficulties in the initial phases of collaborative learning, enabling timely interventions and tailored support to improve the collective performance. Furthermore, delving deeper into this educational paradigm, Laffey et al. (2023) emphasizes the increasing adoption of collaborative learning systems, particularly CSCL, to enhance performance and learning outcomes in both educational and business settings. Central to this effort is the cultivation of key learning objectives, such as critical thinking, problem-solving skills, and effective communication abilities. These competencies are nurtured within authentic community activities, spanning professional domains and academic disciplines, underscoring CSCL's versatile potential to enrich educational and professional journeys. Hidalgo-SuÃrez et al. (2023) indicated a significant increase in student success rates due to collaborative learning, with a 17 percent higher achievement rate compared to individual work. Moreover, the study underscored collaboration's positive impact on social skills, encompassing friendship, motivation, and group cohesion. As Allaymoun (2021) highlights, CSCL facilitates learning through collaborative engagement and the social construction of knowledge, bolstered by seamless integration with information technology. It offers an array of technological tools that empower students to actively participate and engage within a virtual learning environment. However, prior research has predominantly concentrated on learning outcomes, often neglecting the intricate processes of collaborative learning.

The latter, in turn, can present an array of challenges. These challenges encompass attitudinal disparities and potential conflicts among coordinators arising from differing expectations and priorities. They also extend to variations in methodologies for data collection and issues surrounding trust, as underscored by the research conducted by Rodriguez-Ferradas et al. (2023). In a related vein, Ma et al. (2023) has delved into the intricacies of online collaborative learning, revealing challenges related to interaction quality and motivation. These challenges were rooted in unmet expectations, leading to feelings of isolation and burnout fueled by a lack of comprehensive understanding of fellow learners' progress. These hurdles often contribute to the low performance observed in collaborative projects. Therefore, comprehending the root causes of diminished performance empowers stakeholders to formulate targeted interventions aligned with project-specific requisites.

A comparative analysis conducted by Chen et al. (2022) explored two student cohorts, identified as high-performing and low-performing, within a secondary school setting in Singapore. The study unveiled three pivotal distinctions in collaborative argumentation between these groups: variance in social interaction volume, diversity of interactive patterns, and sequencing of contributions to group endeavors and information acquisition.

As such, acknowledging intra-group interactions assumes equal significance in cultivating a more efficient and cohesive collaborative eLearning environment.

A recent study conducted by Vlachopoulos et al. (2021) at a prominent Australian university delved into the application of a team-based learning strategy, with implications for educators aiming to amplify student engagement, enrich experiences, and optimize learning outcomes via collaborative learning. The study unveiled the pivotal roles of cognitive engagement, information dissemination, and reflective contemplation in fostering students' knowledge construction. In the CSCL context, the social dimension is integral to the learning process. It is therefore crucial to consider it when evaluating student groups, as it represents the group's cohesion and the relationships among its members.

One of the most crucial tools for assessing student engagement and course success in such an environment is the student's ability to connect and be socially present (Dascalu et al. 2014). Social presence refers to the exceptional degree to which individuals are engaged in contact and communication. In this study, "social presence" pertains to intra-group interactions among learners in the online learning environment through discussion forums. These platforms allow them to connect with group members and actively contribute their views. Communication and collaboration among group members constitute intra-group interaction, encompassing activities like brainstorming, problem-solving, and decision-making. Herrera-pavo (2021) demonstrated the advantages of utilizing virtual forums, online training, and communication tools to enhance collaborative efforts. Synchronous communication technologies, such as forum discussions, provide students with the immediate ability to evaluate the value of their ideas and inquiries. This real-time assessment fosters increased interaction among individuals with shared interests and promotes enhanced access to information, thereby facilitating learning. Through this mode of interaction, learners engage in the reciprocal exchange of knowledge, benefiting from one another's insights.

In tandem with this, the ability to forecast student performance equips educators and project leaders with a proactive means of identifying potential areas of concern. This enables the targeted implementation of strategies aimed at addressing these concerns. Anticipating internal production challenges, such as disruptions in data flow and communication, empowers project teams to formulate contingency plans and establish efficient communication protocols. These measures effectively mitigate negative impacts, ensuring the progression of tasks within established timelines. This state of preparedness not only enhances project efficiency but also contributes to the seamless execution of tasks, thereby fostering improved project outcomes.

Furthermore, the incorporation of recommended approaches, encompassing both the internal and academic dissemination of project outcomes, plays a central role in cultivating transparency and expediting the exchange of knowledge within the team. Sharing discoveries and advancements among team members and the wider academic sphere cultivates a strong ethos of ownership and responsibility, invigorating participants to be actively engaged in steering the project towards success. This dynamic interplay between the capacity for foresight and the judicious execution of strategies serves as the foundational framework underpinning the landscape of collaborative eLearning environments. Such an approach not only empowers learners and facilitators to attain their objectives with heightened effectiveness but also triggers a surge of

motivation and commitment to the project's fruition. By establishing a symbiotic relationship between predictive potential and the dynamics of intra-group interactions, collaborative eLearning environments can flourish, granting learners and facilitators alike the means to realize their aspirations with unparalleled efficiency and impact. This fusion of predictive prowess with an appreciation of intra-group dynamics forms the bedrock for the flourishing of collaborative eLearning realms, endowing learners and facilitators with the tools to realize their ambitions with exceptional efficiency and efficacy. To achieve this comprehensive approach, it is paramount to first establish a solid foundation by elucidating the core tenets of our methodology. This involves delving into the essential principles underpinning Student Learning Analytics (SLA), a discipline that harnesses data analysis techniques to gain deeper insights into students' interactions and learning dynamics within online collaborative frameworks (Kaliisa et al. 2022). By employing this methodology, educators are empowered to track students' progress, identify potential areas of concern, and develop targeted interventions to enhance learning outcomes.

In parallel, the implementation of Pedagogical-Based Collaborative Learning (PBCL) brings into focus the pivotal role of communication, problem-solving, and critical thinking skills. In a PBCL approach, students work collaboratively under the guidance of an instructor or mentor to accomplish assignments. This collaborative pedagogy not only enhances motivation, engagement, and social learning opportunities but also nurtures a sense of collective ownership over the learning process. When integrated with SLA, PBCL provides educators with a potent toolkit for elevating learning outcomes. By scrutinizing data on students' interactions, contributions, and feedback within collaborative environments, teachers can refine their instructional strategies, identify areas for improvement, and tailor their teaching methods to align with the unique needs of their students.

Building upon these foundational principles, our study takes a step further by introducing a novel methodology for collecting and analyzing assessment and group interaction data within the framework of Pedagogical-Based Collaborative Learning (PBCL). While our previous research (Nouira et al. 2018) predominantly focused on individual learner assessment analytics, our current investigation shifts its lens towards the realm of group-based learner assessment analytics. Incorporating both Student Learning Analytics (SLA) and assessment ePortfolios, our approach offers a more holistic perspective on groups' project trajectories. SLA aids in pinpointing areas where groups may encounter challenges or excel, while group assessment ePortfolios furnish tangible evidence of their advancement and accomplishments. This potent amalgamation empowers us to tailor project instructions, offer precise feedback, and guide groups towards the attainment of their learning objectives.

Given these factors, our investigation delves into the following probing questions: To what extent do the quantity and frequency of online engagement, along with the quality of online interactions, affect group achievement? How do intra-group interactions impact assessment outcomes? Can we predict group performance based on these interactions? To comprehensively explore these aspects, we scrutinize the nature of group interactions, delve into their efficacy, and investigate potential avenues for predicting and optimizing group performance. This entails not only

understanding the variables at play but also devising an optimal format for storing and utilizing the wealth of data derived from group interactions.

In pursuit of addressing these inquiries, our study introduces a comprehensive approach for the collection and analysis of assessment and intra-group interaction data within the framework of Pedagogical-Based Collaborative Learning (PBCL). As a natural extension of this method, we advocate for the organization and management of this acquired data through the implementation of an ePortfolio dedicated to assessment purposes.

Recent studies, such as the one by Mudau (2022), underscore the importance of e-assessment facilitated by ePortfolios as a means to promote constructivist learning paradigms. This pedagogical approach empowers students to actively construct their knowledge through innovative learning and evaluation activities. Moreover, the research highlighted by Reforming higher education (2023) emphasizes that ePortfolios are a high-impact strategy fostering student autonomy and active engagement in the learning process. These interactive tools are not only harnessed for the assessment of educational achievements across various learning settings but also for accrediting graduation programs and seamlessly integrating into a university's information-educational ecosystem.

In the context of unforeseen events such as the COVID-19 pandemic, as highlighted by Mat Razali et al. (2023), ePortfolios emerge as versatile and authentic assessment tools. Their adaptability supports continuous learning and equips educators to effectively navigate unexpected disruptions to traditional education. This underscores the potential of ePortfolios to serve as dynamic instruments for enhancing both student learning experiences and teacher preparedness in times of uncertainty.

The remaining sections of this article are organized as follows: In the second section, we will examine the concepts and literature about collaborative learning, especially within the PBCL setting, as well as the role of SLA. The section that follows describes our Group Assessment Analytics Framework (GAFF) and its architecture. In Sect.Social learning analytics (SLA), we delve further into the Ontological Group Assessment ePortfolio (OnGAsseP). Then, in Sect. Description of the group assessment analytics framework (GAAF), we discuss our analytics engine based on prediction models, which we developed and utilized to study the impact of intra-group interactions on project assessment outcomes. Included in the discussion of our study's methodology are the measures we took to reach our conclusions. In Sect. The interface (online collaborative platform), we present the results, draw conclusions, and wrap up the findings.

Literature review

This section examines the collaborative eLearning environment, with a focus on the PBCL context. Subsequently, we will provide a concise literature review on the concepts of learning analytics and social learning analytics.

Collaborative learning and PBCL context

As defined by Dillenbourg (1999), collaborative learning entails "a situation in which two or more people learn or attempt to learn something together, and collaboration involves the mutual engagement of participants in a coordinated effort to solve problems together." The rapid advancements in computer-supported communication during the

late 1980 s gave rise to a new discipline in the 1990 s, now recognized as Computer-Supported Collaborative Learning (CSCL). This field stands as a pivotal facet of computer-supported learning, striving to enhance learning outcomes and facilitate collaborative work that enables learners to exchange ideas and present their viewpoints (Lipponen 2002).

Numerous studies have been conducted in the realm of CSCL to explore the efficacy of collaborative learning strategies and technology-supported learning environments (Hadyaoui and Cheniti-Belcadhi 2022). In their meta-analysis, Chen (2018) synthesized existing evidence comparing the academic achievement effects of project-based learning and standard education. Their findings indicated that project-based learning has a moderately to strongly positive impact on students' academic achievement compared to traditional education. Similarly, Mhlongo et al. (2020) evaluated the effectiveness of collaboration within a hackathon setting for teaching computer programming to IT students, demonstrating that collaborative learning experiences contribute to enhancing both technical and soft skills.

In the context of PBCL, Apeanti (2021) conducted an extensive analysis of students' experiences, collaboration levels, and challenges when learning computer programming. The study illuminated that students not only acquired new ways of working collaboratively but also gained a better grasp of the technical principles of the course. Interestingly, PBCL was found to be more effective for teaching lower-level undergraduate programming courses than higher-level ones. This trend was consistent with the findings of Yeom et al. (2022), who investigated whether the PBCL strategy could enhance students' performance in an introductory programming course. The research demonstrated that PBCL has the potential to sustain students' interest, particularly among women, and enhance their learning experience in an introductory programming course. The exceptional circumstances presented by the COVID-19 pandemic prompted an investigation by Zarzycka et al. (2021), which delved into the factors influencing communication and collaboration within a remote learning setting, both within the virtual classroom and beyond. This study also scrutinized the role of social media in this context, including considerations regarding group size. The findings of the research indicated that utilizing smaller groups yields superior results, a conclusion supported by several studies, among them Panadero and Järvelä (2015). Additionally, there is mounting evidence to suggest that online learning environments that incorporate small-group sessions can generate positive outcomes. Consequently, the size of a group significantly affects students' academic performance, self-esteem, and access to social support, all of which contribute to improved learning outcomes (Region 2015). This insight motivated us to choose the small-group approach for our experiment to align with these favorable outcomes. Despite the extensive research conducted on team dynamics and their effects on performance, there is a noticeable lack of literature that delves into the influence of group dynamics on performance. Particularly, the crucial aspect of how interactions impact group performance and learning hasn't received the attention it deserves. While numerous studies have explored the impact of mixed-gender teams on performance, the role of group dynamics within these scenarios remains largely uncharted territory. This gap in the existing literature underscores the necessity for further research to gain a deeper comprehension of the intricate relationship between group dynamics, interactions, and

performance within a team context. Through investigating these factors, researchers can develop more targeted strategies for constructing successful teams and cultivating optimal performance outcomes.

Social learning analytics (SLA)

According to Shum and Ferguson Shum et al. (2014), Social Learning Analytics (SLA) encompasses the collection and evaluation of digital artifacts and online interactions produced by students in both formal and informal contexts, aiming to analyze their activities, social behaviors, and knowledge generation within a social learning environment. SLA has emerged as a potential approach to offer insights and guide instructional decisions by extracting concealed information from vast amounts of educational data retrieved from Computer-Supported Collaborative Learning (CSCL) settings, such as Learning Management Systems (LMS) and wikis. As students engage with online activities, they generate log files within the LMS, and through the utilization of LA, additional latent information about students can be unearthed within the online learning environment. This has a significant role in identifying challenges and enhancing the learning environment (Na and Zaidatun 2022). This role gains even more significance in light of current challenges in higher education, which require active student participation to foster critical thinking, cooperation, and self-regulation. Hence, SLA facilitates both summative and formative assessments, enabling continuous, near-real-time feedback to support students' self-regulated learning and educators' interventions (Isohätälä et al. 2017).

Addressing some of the debates in this domain, Yadav et al. (2021) provided an analytical framework of measures for evaluating the quality of cooperation in a wiki-based collaborative learning environment. This framework included factors like student input, participation, transactivity, and social dynamics. Moreover, Afacan Adanir (2019) undertook a study to identify the themes of chat discussions held by groups of students engaged in collaborative study within the online CSCL environment of Virtual Math Teams (VMT).

A systematic review conducted by Kaliisa et al. (2022) illustrated that the social constructivist perspective was frequently employed to elucidate students' learning behaviors in the majority of SLA studies aimed at comprehending students' learning processes. Despite the substantial data available in the literature regarding the effects of SLA, all researchers concur that further investigation is warranted.

However, while a wealth of literature exists on the use of SLA, there remains a gap in research concerning the exploration of intra-group interactions and their impact on group performance. While studies have shown the influence of SLA on individual learning outcomes, less attention has been paid to how SLA affects group dynamics and performance. By scrutinizing the interactions among group members within SLA environments, researchers can gain a deeper insight into the social processes at play and their influence on group performance. Furthermore, this gap in the literature underscores the imperative need for more comprehensive research that delves into the effects of SLA on group assessment outcomes, particularly with a focus on the social dimensions of learning within online environments.

Description of the group assessment analytics framework (GAAF)

The Group Assessment Analytics Framework (GAAF) plays a pivotal role in the aggregation, processing, and analysis of data from diverse assessment sources. This data subsequently informs insights into students' performance and progress. The GAAF encompasses a blend of software tools and systems that facilitate the collection, storage, and visualization of assessment data. This capability empowers the GAAF to harness data from a range of sources, including formative and summative assessments, as well as group interactions, offering a holistic perspective on students' learning journeys.

The proposed architecture introduces four key components that synergistically enhance student collaboration and assessment. These components are visually represented in Fig. 1.

Interface (Online Collaborative Platform): This platform prioritizes communication, cooperation, and coordination, thereby facilitating team members in sharing of ideas, collaborating on projects, and fulfilling tasks.

Learning Record Store (LRS): As the second component of the system, the LRS retrieves information from assessment tools and learning management systems. Prior to analysis, data is aggregated and preprocessed, thereby fostering an enhanced comprehension of team performance.

Ontological Group Assessment ePortfolio (OnGAsseP): The third component enables the collection and organization of evidence that underscores the proficiency of a group of learners across diverse learning domains. This function is crucial in enabling a comprehensive analysis of group performance, spotlighting strengths, and pinpointing areas for enhancement. The integration of an ontology-based framework ensures consistent and intelligible data structuring, thereby enabling more dependable insights and predictions. Ultimately, the OnGAsseP stands as a critical tool for providing intelligent assessment, equipping educators with the requisite information to make judicious decisions regarding the enhancement of student learning outcomes.

Analytical Engine: The fourth component engages in the analysis of preprocessed data to unveil the system's success in terms of group performance, learning outcomes, and assessment. Leveraging machine learning, statistical analysis, and data visualization, this component scrutinizes the data. Notably, machine learning algorithms assume significance as they empower the system to learn from the data and forecast future outcomes. Consequently, the OnGAsseP and the analytical engine collaborate to deliver a comprehensive evaluation of group performance, enabling instructors to better support their students and elevate academic accomplishments.

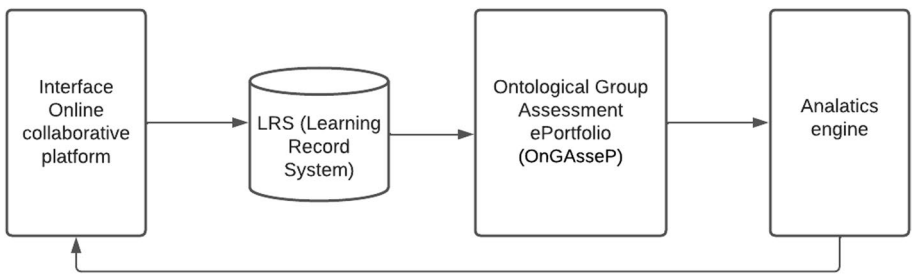

Fig. 1 GAAF architecture

Consequently, educators can identify concealed trends, patterns, and insights, harnessing this information to make informed decisions about refining student learning outcomes.

The interface (online collaborative platform)

The interface assumes a pivotal role within the ontology-driven intelligent collaborative framework for assessment in a collaborative eLearning environment. The relationship between the platform's users, which encompass instructors and students, and the array of functionalities it offers, can be likened to a critical link. Multiple factors contribute to the interface's significance.

First and foremost, the domain under consideration is centered on User Experience (UX) considerations. A well-crafted interface stands as a testament to a positive user experience, as it amplifies user satisfaction and engagement with the platform. The interface embodies attributes of simplicity, accessibility, and visual appeal, thereby facilitating seamless navigation and utilization of diverse functionalities by both learners and instructors.

Furthermore, the interface boasts a user-centric design that streamlines efficient navigation and information retrieval, thereby mitigating the chances of users becoming disoriented or struggling to locate specific content. A streamlined arrangement of tasks and resources not only economizes time but also expedites the attainment of goals for all stakeholders.

Additionally, the interface has a notable influence on communication efficacy. It's crucial to emphasize the notion of real-time chat and messaging. For these functionalities to perform optimally, it's imperative that the interface demonstrate a high degree of user-friendliness. Effective communication among team members facilitates a seamless collaborative environment, fostering cooperation without impediments.

Equally significant is the role of collaborative work. This is where the Collaborative Program Editor comes into play. An interface that facilitates real-time collaborative editing of shared documents by multiple users becomes indispensable. This approach empowers teams to actively participate in collaborative endeavors, collectively progressing towards their shared objectives.

Promoting student engagement and motivation is a pivotal factor in facilitating effective learning. A well-crafted interface holds the potential to achieve this outcome. When educational materials are presented in a visually appealing and interactive manner, the probability increases that students will sustain their interest and actively participate in the learning process.

Lastly, the concept of user adoption comes into play. An interface that is intuitively comprehensible and designed with user-friendliness as a priority fosters enhanced user engagement and adoption. The reduction of the learning curve facilitates the seamless integration of new users, enabling them to swiftly harness the range of innovative functionalities available.

The platform seamlessly integrates several essential features with the aim of enhancing team collaboration. These features encompass:

Real-time chat and messaging: an overview

The real-time chat and messaging features serve as the primary mode of communication among team members. Students have the capability to interact with each other, exchange digital messages, share electronic files, and engage in discussions within the scope of their designated project groups. Figure 3 offers a visual depiction of the interface that facilitates smooth message exchanges among group members. The swift sharing of information enabled by instant communication platforms enhances the efficient dissemination of ideas and the proficient resolution of issues.

Collaborative program editor

CollabLearn offers a collaborative programming editor that enables simultaneous collaboration among team members on project-related documents, including reports, presentations, and code files. The incorporation of real-time synchronization ensures that changes made by one team member are instantly visible to others, thereby enhancing the editing process and promoting seamless collaboration.

Task allocation and monitoring of progress

The interface not only facilitates efficient task allocation but also enables comprehensive monitoring of interactions among group members. Through this platform, the team leader can readily assign specific tasks to individual members, fostering a transparent distribution of responsibilities. Additionally, students can actively track task progress, gaining valuable insights into project development and enhancing overall team coordination. This dual functionality of task allocation and interaction monitoring is exemplified in Fig. 4, which presents a dashboard showcasing the interactions carried out by group members.

Ontological group assessment ePortfolio model (OnGAsseP)

According to Sarwandi and Wibawa (2022), ePortfolios are digital data collections of trainees that can support learning by providing a method to organize, archive, and present individual or group work. In the literature, there are different studies that deal with ePortfolio models. Some of these studies are based on Semantic Web technologies, presenting ePortfolio components in an ontological form Ghedir et al. (2018). Therefore, every piece of data in the ePortfolio is described using semantic web formalisms to maximize the benefits of sharing and automatically processing assessment data. Only by explicitly defining the meaning of assessment data can people and computers work effectively. Our recommended ePortfolio for group evaluation strongly depends on the modeling of assessment data produced by our group assessment analytics. We therefore concentrated all the data by creating an ontological model for assessment analytics to ensure a standardized representation of assessment data. We relied on the Performance Application Programming Interface (PAPI) specs standard during the construction of our ontological group assessment ePortfolio. This specification intends to facilitate the representation, retrieval, and exchange of learner models between diverse educational systems by providing minimal learner information. In addition, it provides researchers or developers wishing to design a learner model with

a basis for the development of learner models as well as a standardized and expanding data source (Zine et al. 2019). To further complement our proposed ePortfolio, we have modified the PAPI requirements to cover the social dimension as well as the member interactions and the assessment results. As shown in Fig. 2, to develop the OnGAsseP, we followed the following structured process:

- Define the assessment objectives: The first step was to define the goals and objectives of the group assessment. This information guided the selection of the PAPI tests and the interpretation of the results. The assessment objectives could be to identify the strengths and weaknesses of the group, identify potential areas for improvement, or evaluate the effectiveness of a particular intervention.
- Select the PAPI tests: Based on the assessment objectives, we selected the PAPI tests to provide the relevant data. The PAPI tests measure a range of personality traits and behavioral preferences.

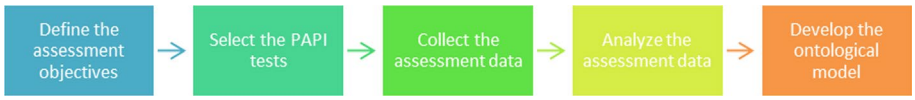

Fig. 2 Description of the development process of the OnGAsseP

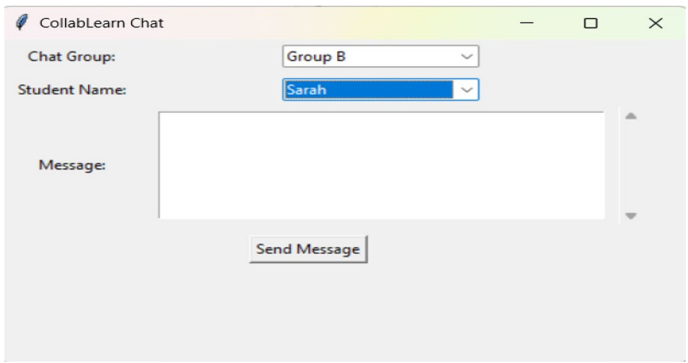

Fig. 3 Interface for seamless message exchanges among group members

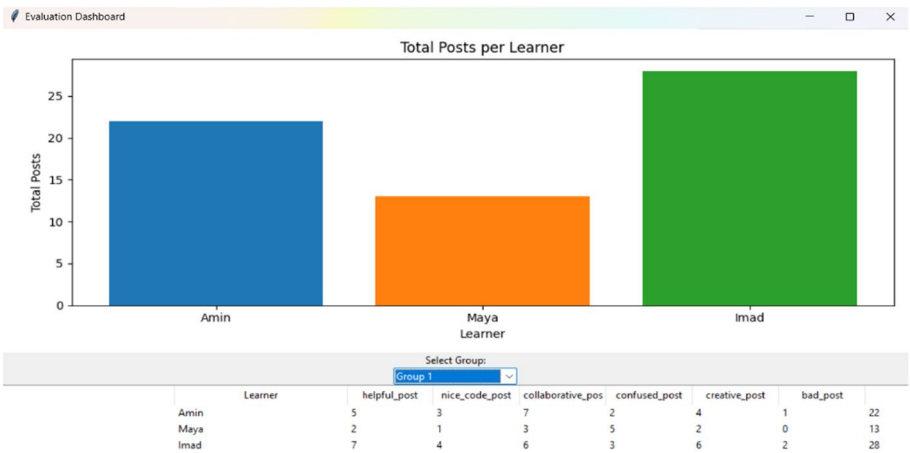

Fig. 4 Interactive dashboard visualizing data interactions

- Collecting Assessment Data: Our initial step involved the selection of PAPI tests, which were then administered to the group members to gather essential assessment data.
- Analyzing Assessment Data: With the assessment data in hand, we embarked on a thorough analysis, meticulously scrutinizing the data to uncover noteworthy patterns and trends within the group's personality traits and behavioral preferences. This analytical process formed the bedrock for our subsequent assessment outcomes.
- Developing the Ontological Model: Building upon the insights derived from the assessment data analysis, we proceeded to construct an intricate ontological model. This model serves as a visual representation of the intricate interplay between various key concepts and relationships within the realm of group dynamics and personality.

Consequently, our efforts culminated in the creation of the Ontological Group Assessment ePortfolio (OnGAsseP), a robust tool designed to comprehensively evaluate the dynamics and performance of groups, pinpointing potential avenues for enhancement. The hierarchical structure of the OnGAsseP model is elucidated in Fig. 5.

In the ensuing sections, we will delve deeper into the specifics of each class encompassed within the hierarchical structure of the ontological group assessment ePortfolio:

- Group: A foundational class that encompasses vital details about the assessed group. It delves into sub-classes, providing insights into the group's size, composition, objectives, strengths, weaknesses, areas of improvement, and possible interventions. These sub-classes cover diverse aspects such as the group's member count, gender distribution, goals, and overall strengths and weaknesses. Additionally, they highlight specific domains for improvement and offer potential strategies for addressing these areas.
- Personality Traits: This class captures the quintessential five personality traits gauged by the PAPI tests: openness, conscientiousness, extraversion, agreeableness, and neuroticism. These traits bear a significant influence on the dynamics and performance of the group.
- Another pivotal class that encapsulates the behavioral preferences gauged by the PAPI tests. These encompass decision-making style, leadership approach, communication patterns, and teamwork tendencies. Understanding these preferences can reveal potential sources of conflict or areas that warrant improvement within the group's dynamics.
- PAPI Tests: This class serves as a conduit for the specific tests employed to amass the assessment data. Each distinct test is meticulously tailored to gauge a particular personality trait or behavioral preference.
- Projects: Within the Projects class, we explore the array of project options presented to the group within the context of PBCL.
- Assessment Data: Illustrated in Fig. 6, this class acts as a repository for both the raw and meticulously analyzed data derived from the PAPI tests. The raw data captures the individual responses of group members to the PAPI tests, while the analyzed data encompasses the outcomes of statistical analyses that unveil the

Fig. 5 The hierarchical classes of the OnGAsseP model

group's response patterns. Moreover, this class encompasses components pertaining to group achievements in various projects. Notably, CriticalThinking and Creativity components gauge critical thinking levels and creative merit in group projects, respectively. These components culminate in the computation of the FinalScore, providing an overarching assessment of group achievements. Finally,

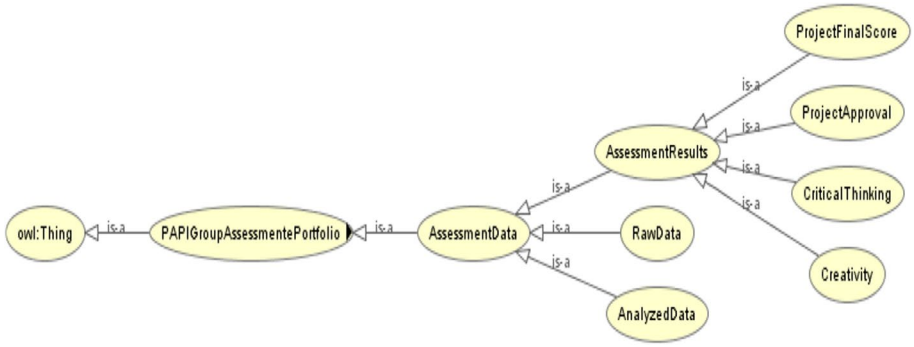

Fig. 6 The sub-classes of the AssessmentData class

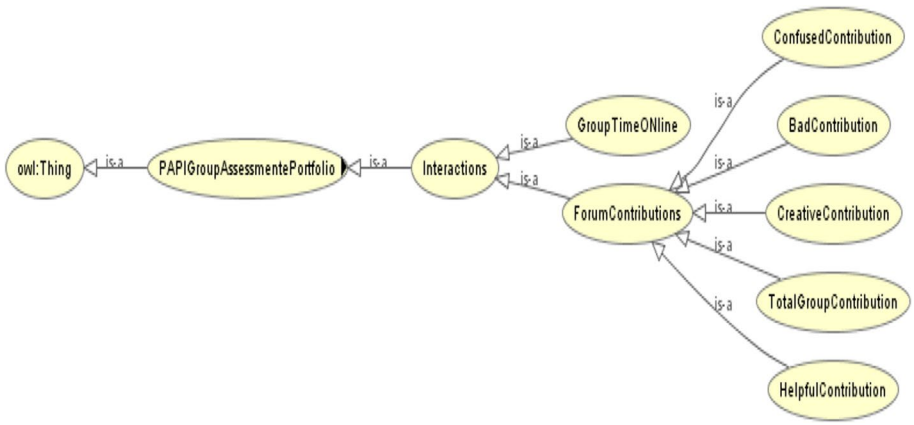

Fig. 7 The sub-classes of the interactions class

the Approved component guides the determination of conclusions based on the final score.

- Interactions: As shown in Fig. 7, this class includes two main components: Group Time Online and Group Contributions. Group Time Online refers to the total amount of time spent online by all members of the same learning group. Group Contributions describe the various messages and interactions made by the group members in the discussion forum. Before posting a message, students can indicate the type of message they intend to share. The different types of contributions are categorized into four groups: Helpful, Confused, Creative, and Negative. Helpful contributions involve providing solutions to questions or problems raised by other group members. Participants who experience difficulties completing their tasks as part of a larger collaborative project make confused contributions. Creative contributions may include sharing resources, providing tips, or planning synchronous meetings with other group members. Negative contributions are messages that have no attachments or valuable content.
- Assessment Objectives: This class represents the goals and objectives of the assessment. It includes sub-classes that provide information about the selection of PAPI tests and the interpretation of the assessment results.

- Interventions: This class represents potential interventions that can be used to address any areas for improvement identified by the assessment. These interventions can include instructor feedback, team-building exercises, leadership training, communication workshops, and other strategies aimed at improving group dynamics and performance.

Analytic engine

The Analytical Engine stands as a pivotal cornerstone within the framework of our intelligent collaborative assessment, a cornerstone meticulously outlined within this study. Its role is nothing short of the central processing unit, charged with the intricate task of dissecting and comprehending the intra-group interactions unfolding among online participants, all with the overarching aim of prognosticating group success in the realm of collaborative projects. As the collaborative phase of a project unfolds, this engine springs into action, employing an array of intelligent models and algorithms. Its mission: to extract pearls of wisdom from the interactions that weave the tapestry of the group. Armed with participant-generated data, its purpose spans the breadth of crucial research queries. It embarks on a quest to unveil the reverberations of these interactions on project evaluation while contemplating the very feasibility of their deployment in the prophesying of project outcomes. Elegantly tailored to manage vast datasets, the Analytical Engine takes the reins with finesse. It shoulders the responsibility of processing information hailing from diverse groups of undergraduates, all enrolled in the inaugural year of the Transportation Technology and Engineering degree program. To this end, the engine employs a symphony of cutting-edge machine learning techniques, including supervised learning methods such as Decision Trees, K-Nearest Neighbor algorithms, and multiple regression algorithms (such as simple linear regression, Ridge regression, and Lasso regression).

Methodology

This study aims to develop models that can predict the success or failure of a project in a PBCL setting based on group dynamics. To achieve this goal, the following research questions were formulated:

1 How do intra-group interactions affect the learning skills and achievements of groups in a PBCL context?
2 What is the impact of various types of group dynamics on the reliability of project outcomes?
3 Can online intra-group interactions be used to predict the assessment result of a collaborative project?

To investigate the impact of group dynamics on the project's overall grade, the following hypotheses were proposed:

- Hypothesis 1: Intra-group interactions in a PBCL context have a significant relationship with the groups' assessment results.

- Hypothesis 2: All types of interactions equally impact the project's outcome.
- Hypothesis 3: The project's approval and final score depend on intra-group interactions.

Study approach

The present study employed a skill-based experimental approach to determine the impact of intra-group interaction in a PBCL context on the project assessment result in terms of groups' learning skill achievement, the project's approval or not, and its final score. The skill-based approach was used because it is consistent with the goals of PBCL, which are to help students improve their problem-solving, critical thinking, collaborative, and communication skills. The approach can assess whether PBCL is successful in reaching its goals by focusing on skill attainment. The skill-based method additionally enables the quantification of learning outcomes. It offers a more objective and concrete evaluation of the effects of PBCL by tracking students' skill accomplishments. It can also find areas for improvement and guide changes to PBCL programs by evaluating skill achievement. This makes it easier to keep PBCL's skill-development efficacy from declining.

We've decided to use the P21 Framework for 21st-Century Learning to evaluate the results of our proposed group project, especially the first two skills. 21st Century Skills shows the P21 Framework for 21st Century Learning, which was made with help from teachers, education experts, and business leaders (Aghazadeh et al. 2019). The framework also focuses on the "soft skills" and "support systems" that students need to be productive members of society. We used a scale from 0 to 10 to assess each skill gained during the project:

- The group's ability to think critically and solve problems,
- The group's qualities of creativity and originality.

And the final score for the project is based on a scale from 0 to 20 since it is the sum of the two previous scores.

Participant population

In the first term of the school year 2022–2023, a total of 312 students were enrolled in the "Python programming" course at the Higher Institute of Transport and Logistics of Sousse, which is a part of the University of Sousse in Tunisia. To facilitate the learning process, the students were organized into 60 groups, with 4 students in each group, and 24 groups with 3 students each. The process of grouping participants was meticulously executed based on their scores, with a deliberate intent to cultivate a diverse blend of abilities and backgrounds within each group. To initiate the experiment, a pre-test was thoughtfully administered to each group, serving a dual purpose: firstly, to gauge the students' programming proficiency, and secondly, to ensure that each group consisted of individuals spanning a spectrum of programming skills. This methodological approach was adopted to nurture heterogeneity within the groups, thus enriching the potential for collaborative learning encounters. During the experiment, the collaborative assessment

framework seamlessly integrated the utilization of Kahoot!, a digital learning platform. This inclusion of Kahoot! yielded favorable outcomes across multiple dimensions of the classroom milieu. It catalyzed an improved classroom dynamic by fostering interactive and captivating activities. Students enthusiastically participated and exhibited heightened motivation during the learning sessions, which subsequently translated to elevated learning achievements. Student feedback corroborated their satisfaction with the incorporation of Kahoot! as an integral facet of their learning journey. The interactive nature of the platform infused an element of enjoyment and enthusiasm into the learning process. Moreover, Kahoot! presented an added advantage in real-time performance monitoring. The teacher gained the capability to simultaneously track correct answers provided by students, furnishing invaluable insights into the collective comprehension and progress of the groups, as underscored by Ben (2022).

Research design

We adopted the Analysis, Design, Development, Implementation, and Evaluation (ADDIE) model proposed by Branch (2009) to meticulously plan and execute our study. This model served as a structured and methodical guide, ensuring that each phase of our research was well-considered and carried out effectively. The ADDIE paradigm is well-known for its efficacy in instructional design and research, providing a solid framework for the development of effective learning experiences. Its adaptability makes it an invaluable asset in a variety of educational settings. In our particular context, the application of the ADDIE model enabled us to precisely define the objectives and aims of our research. We were able to develop a detailed strategy for the Project-Based Collaborative Learning (PBCL) activity, ensuring a structured and organized approach. In addition, the model facilitated the evaluation of intra-group interactions and their influence on the overall project outcomes. By adhering to the ADDIE model, we gave our research a sense of methodical and deliberate direction, which ultimately contributed to the success of our study.

Analysis stage

We conducted a needs assessment to determine the learning needs and objectives of the study participants, including identifying the specific skills that we wanted to assess, as well as the criteria for determining the success of the project. The primary purpose of this study was to evaluate the impact of group interaction on PBCL environments that will be utilized to build a Python programming project. The participants were first-year Transportation Technology and Engineering undergraduates enrolled in the programming curriculum. The participants were familiarized with the educational content, the fundamental concepts of Python programming (the project is proposed after eight weeks of classes), as well as the online platform, the Modular Object-Oriented Dynamic Learning Environment (MOODLE), that will be used to collaborate and communicate with the members of their groups, consult and share resources, and submit their work. We used Moodle to put our PBCL strategy into action because it is the most widely used Learning Management System (LMS) in Tunisia. Moodle was created to incorporate various facets of modern education. It was created using open-source software and is currently under global development (Jan et al. 2018).

Design stage

PBCL was used in the research. The project's main goal was to give students the knowledge and abilities needed to solve an issue before suggesting a Python implementation of the solution in the form of functions. The project's goals and the crucial information and skill sets that must be acquired through project activities were described. It was confirmed that students have learned the fundamentals of Python programming.

Development stage

In this stage, the project's main task—to implement a Python program—was created as a Moodle activity. A discussion forum was created for each group of the sample on Moodle to facilitate intra-group interaction and resource exchange as a consequence of promoting more online communication and collaboration. The learning content that was programmed was made available on the learning management system Moodle.

Implementation stage

A pre-test was conducted for all the students in the sample using online quizzes published using a gamified approach using Kahoot! The sample was divided into groups of three to four. The purpose of the assessment is to determine whether or not the student has acquired the necessary knowledge and abilities to complete the project. The pre-test was important to identify any barriers that might prohibit the conduct of the programming project.

Evaluation stage

This stage involved the evaluation of all of the previous stages, including the design and implementation of the treatment. The feedback from the pre-test was considered an essential element in preparing the programming project.

Data collection instruments

The data collected includes logs of the writing process of logs from Moodle about forum groups' time online and the forum's contributions.

Learning skills and achievement project

A closed-ended assessment grid, as shown in Table 1, was prepared to measure groups' achievement of the learning skills related to the project based on 21st-century competencies. The four Cs are by far the most sought-after: Critical thinking, Creativity, Collaboration, and Communication (Aghazadeh et al. 2019). Learnability refers to the capacity to acquire new information. These skills are essential in any field, so it makes sense that more teachers are aware of them. Their relative importance also differs from one person to another and from one career path to another. In our research, we are interested in the two first pillars of 21st-century competence: Critical thinking. Critical thinking includes posing pertinent questions, obtaining and creatively sorting important information, linking new information to existing knowledge, reexamining

beliefs, and reasoning, and drawing trustworthy and dependable conclusions (Ramadhan et al. 2021). Creativity. Creativity is the capacity to generate new ideas and integrate old ideas in novel ways to produce novel solutions to problems (Turnbull et al. 2010).

Procedure

As aforementioned, participants were divided into 84 small groups of 3–4 students each (60 groups of 4 and 24 groups of 3). A discussion forum for each group was set up on Moodle. The participants were also introduced to the available learning resources on Moodle. All groups started to communicate via the discussion forums to define their sub-tasks as part of the total project and organize how they would accomplish them. We were available to assist and scaffold the students as needed, as well as facilitate any difficulties they may have encountered. After the experiment was over, a review of the project's success and a look at the contributions on the discussion forum were done. Figure 8 depicts a sample of submissions by students to the planned forum. Spanning 6 weeks within the initial term, the core experiment unfolded. Data harvested from the full array of achievement projects and threaded forum discussions underwent meticulous pre-processing, poised for the litmus test of hypotheses via the harnessing of machine learning algorithms.

This study relies heavily on Machine learning (ML) algorithms. Specifically, we have centered our experiments on machine learning methods for supervised learning. ML is used in a variety of fields to address complex problems that cannot be easily solved using computer-based methods (Ben-David 2014). First, to identify the impact of the intra-group interactions on groups' achievement and collaborative project outcomes, Exploratory Data Analysis (EDA) was used. EDA is the important process of doing preliminary research on data to find patterns, find outliers, test hypotheses, and confirm hypotheses using summary statistics and graphical representations. EDA gathers information, gives the data more meaning, and gets rid of strange or unnecessary values. This enables a machine learning model to predict our data set more accurately. This results in more precise predictions or classifications. In addition, it assists us in selecting a superior Machine Learning model. We relied on supervised machine learning to identify patterns and generate predictions.

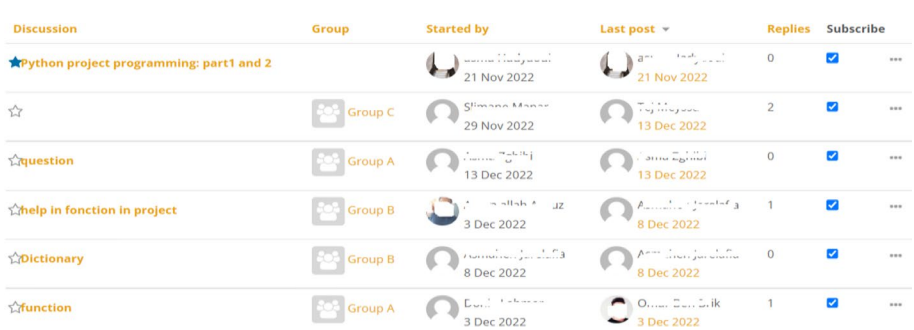

Fig. 8 Groups' discussion forum

Results

Features related to intra-group interactions to predict group project approval

To assess the influence of intra-group interactions on project approval, our study utilized Attribute Selection Measures (ASM) to partition the data, followed by the implementation of the Decision Tree algorithm. The DT algorithm identified the optimal attribute for predicting project approval from the established group interaction criteria, which included total group contribution, useful contributions, creative contributions, poor contributions, and group connection time. Among these criteria, the number of confused contributions emerged as a major predictor of project evaluation outcomes. Figure 9 visually represents this significance as a decision node, dividing the dataset into two distinct portions. To measure the risk of mistakenly identifying a feature when chosen randomly, the algorithm calculated the Gini index (gini) multiple times during each iteration. Mathematically, the Gini index is expressed as:

$$\text{Gini Index} = 1 - \sum_{i=1}^{n} P_i^2$$

where P_i is the probability of an element belonging to a specific class. Additionally, in order to highlight the importance of the features used in our study, we presented the weights of the criteria that contributed to the assessment of the collaborative project using the Decision Tree method. Essentially, this means that we calculated the relative importance of each feature (total group contribution, useful contributions, creative contributions, poor contributions, group connection time, and number of confused contributions) in predicting the outcome of the project assessment, and then we displayed these weights in Fig. 10. By doing so, we were able to demonstrate which features had the most impact on the project assessment outcome and which ones were less significant.

Our data analysis revealed that certain interactive features had a more significant impact on project assessment outcomes than others. Specifically, the number

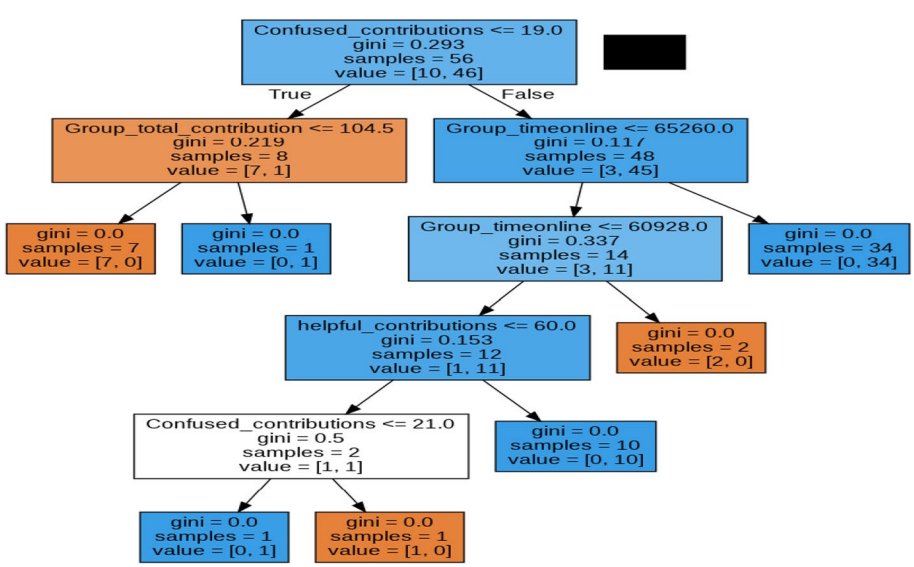

Fig. 9 Decision tree visualization

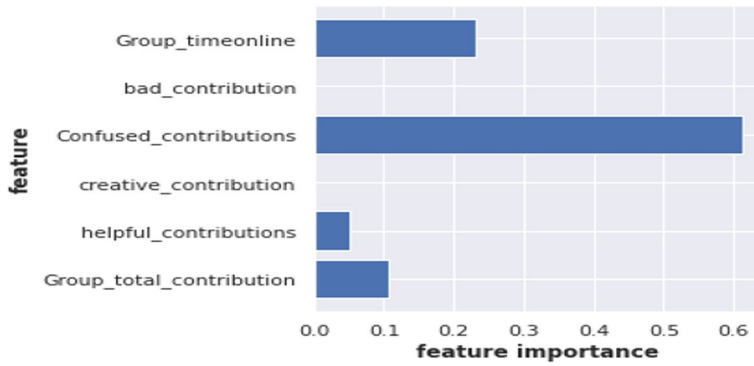

Fig. 10 Feature importance

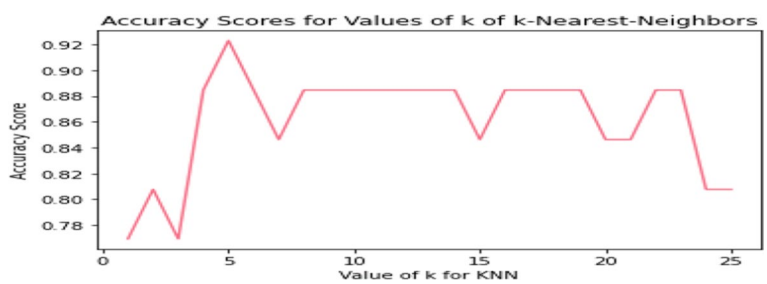

Fig. 11 The accuracy scores for values of K of KNN predictions

of confused contributions was found to be the most important factor in determining whether a project would be approved or not, while the rates of creative and poor contributions did not appear to be relevant predictors.

Predictions of the outcome of the project assessment based on intra-group interactions

Our research aims to predict the performance of a group of learners on a collaborative project. This predictive capability will enable early identification of potential issues and allow for proactive measures to be taken to prevent them. To achieve this goal, we require an effective prediction algorithm that can produce improved prediction outcomes. In our study, we observed that the data points were divided into two groups based on the validation of the project result. This led us to the use of the K-Nearest-Neighbor algorithm, also known as K-Nearest Neighbor (KNN), for prediction. The KNN algorithm classifies new data based on its similarity to the K-nearest neighbors of each group member. Since the algorithm bases its predictions on the nearest neighbors, we began by determining the precise number of neighbors to take into account. To confirm that the optimal value was chosen, we plotted, the accuracy scores of the prediction algorithm for several values of K to determine the best option.

We utilized $K = 5$ for our KNN algorithm, as it demonstrated the highest accuracy, as shown in Fig. 11. Next, we trained the model with our data and assessed its overall accuracy. To evaluate the effectiveness of our prediction method, we compiled a metric report, which is depicted in Fig. 12. In Machine Learning, it is crucial to consider performance measurements, which are often selected based on the research objective. These metrics include:

	precision	recall	f1-score	support
0	1.00	0.75	0.86	8
1	0.90	1.00	0.95	18
accuracy			0.92	26
macro avg	0.95	0.88	0.90	26
weighted avg	0.93	0.92	0.92	26

Fig. 12 Metric report of the KNN algorithm

Table 1 Table of skills and approaches required for the project

Skills	Approaches	Project's skills
Critical thinking	Uses logic and reasoning effectively as appropriate to the situation	Implementing a Python program to solve the proposed problem: Creating a list Adding a new item to the list Modifying an item Sorting the list depends on: Age Baccalaureate average Name
Creativity	Creates novel and useful concepts by integrating current information and resources	Using menus to present the multiple functions Using QtDesigner to make attractive interfaces

- Precision: The proportion of correct positive predictions out of the total number of positive predictions.
- Recall: The proportion of correct positive predictions out of the total number of actual positives.
- Support: These values indicate the number of groups in the test dataset that belong to each class.
- The F1 Score is a weighted harmonic mean of precision and recall, with a higher score indicating a better model performance towards:

$$\text{F1 Score} = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

The effectiveness of our model is evaluated based on its ability to predict, using unseen data, whether the collaborative project will be approved or not. According to the results presented in Fig. 10 of our report, the prediction model performs reasonably well, achieving an accuracy of 0.92. We want to find out if we can predict the group project's final score by looking at the interaction patterns between members in order to further improve the accuracy of our predictions. In pursuit of this question, we have opted to perform a regression analysis.

Prediction of the project's final score using regression

In pursuit of predicting the final score of group projects, we employed regression analysis as a powerful statistical tool to uncover relationships between variables. Our

study aimed to achieve two primary objectives: (1) forecast and predict the final score based on independent variables; and (2) unveil causal relationships between these independent variables and the dependent variable (final score). Notably, regression analysis primarily reveals relationships between dependent and fixed factors within a dataset. Regression analysis is a statistical tool used to identify relationships between variables (Maulud and Abdulazeez 2020). However, it is essential to note that regression analysis alone can only reveal relationships between dependent variables and a fixed selection of factors within a dataset. In our case, the independent variable "X" can have a single value (confused contributions) or multiple values (various interactions among group members), and it predicts the dependent variable "Y," which represents the final score. To determine the best regression algorithm that matches our predictive model, we explored several regression algorithms. Table 2 summarizes our prediction work and the performance of each regression algorithm.

In linear regression, we assume that the relationship between the input and output variables is linear, which means that the output variable can be represented as a linear combination of the input variable(s) plus an intercept term. Specifically, we assume that $Y = aX + b$, where Y is the output variable (i.e., final project grade), X is the input variable(s) (i.e., confused contributions and interaction among group members), a is the slope or coefficient of the input variable(s), and b is the intercept term. To find the best regression algorithm for predicting the final project grade, we have tested five different regressions: simple regression, simple ridge regression, multilinear ridge regression, simple lasso regression, and multilinear lasso regression. To evaluate the performance of these regressions, we compared their prediction scores for both the training and test data. However, we are more interested in the test results because they demonstrate the accuracy of the prediction based on new and unseen data. Based on our experiments, we have found that linear lasso regression is the best regression algorithm for predicting the final project grade. This regression model allows for 77% correct predictions when considering the confused contribution criterion.

The successful identification of the best regression algorithm contributes significantly to our ability to predict the final scores of group projects based on various intra-group interactions. This finding has implications for enhancing project evaluations and improving the overall collaborative learning experience.

Table 2 Table of results for linear and multilinear regression algorithms

Algorithm	Linear regression	Multilinear regression	Y	X	
Simple linear regression	x	Final score	Confused contribution	0.95	0.61
Ridge	x	Final score	Confused contribution	0.86	0.75
Ridge	x	Final score	Group_total_contribution, helpful_contributions, creative_contribution, bad_contribution, Group_timeonline	0.35	0.18
Lasso	x	Final score	Confused contribution	0.90	0.77
Lasso	x	Final score	Group_total_contribution, helpful_contributions, creative_contribution, bad_contribution, Group_timeonline	0.35	0.19

Discussion

By using ontologies as the foundation, we have developed a formal approach to describing assessment and group data in a structured and standardized manner. Our semantic web strategy, which integrates ontologies and eLearning standards, supports data reuse and interoperability. Using an ontology has several advantages, including providing a standardized and semantically rich representation of data, which can be used to integrate data from different sources for decision-making and assessment. Recent studies have shown that semantic web technologies are useful for building e-learning platforms, such as the work by Halimi (2021), which presented a method of assessment using semantic analytics for assessing learners' competencies. In this study, semantic representations were used to model all knowledge about students and their competencies. Another study by Zine (2019) included IMS-LIP, IMS-ACCLIP, and IMS-RDCEO criteria in the suggested learner model to enhance representation and meet adaptation criteria and needs.

By employing OnGAsseP and supervised machine learning methods, we have created predictive models that can anticipate the outcome of group assessments. This demonstrates the potential of collaborative learning in higher education, as it fosters knowledge co-construction and skill development through interaction, leading to more dynamic learning processes. These findings are consistent with previous studies, such as those conducted by Moreno-Guerrero (2020), which further support the benefits of collaborative learning in higher education.

The study's results revealed that when the number of contributions related to each group in the forum discussion is higher, the project assessment results are more indicative of the learners' success in developing critical thinking and creative abilities. Consequently, the first hypothesis was supported, which is consistent with previous research. For instance, Hernandez-Selles (2019) emphasized the importance of interaction, including both teacher-student and peer-to-peer communication, in the classroom. Similarly, Qureshi (2021) confirmed these findings by demonstrating that collaborative learning and engagement, along with social factors, enhance students' learning activities and should be encouraged in higher education institutions. Furthermore, we found that confused contributions significantly impact project success and the acquisition of new skills. These findings were supported by the classification report of the Decision Tree. However, the data showed significant differences in the impact of interactive features on project outcomes, with confused contributions being the most critical variable. As a result, the second hypothesis was rejected, rendering bad contributions and creative contributions insignificant.

These findings allowed us to achieve high accuracy in predicting the validation of students' projects, with a success rate of over 92%. Furthermore, we utilized a regression model that relied on the number of confused contributions to predict the final project grade. These results suggest that a collaborative project in a PBCL context will be more successful if the groups of learners generate a higher number of questions. The act of questioning demonstrates intellectual curiosity and a willingness to work towards finding solutions, contributing to better project performance and group communication. In collaborative settings, questioning skills are essential for fostering creativity and critical thinking.

Conclusion

The study aimed to develop predictive models for group assessment results based on intra-group interactions in a PBCL context. The research utilized Machine Learning techniques, including supervised learning methods such as Decision Trees, K-Nearest Neighbors (KNN), and multiple regression algorithms, to determine the impact of within-group interactions on project assessment outcomes. The results revealed a significant impact of intra-group interactions, particularly the contributions of group members to the discussion forum that appeared confused, on the final project assessment outcomes. The research can assist in monitoring learners' assessments in a collaborative learning setting to enhance group outcomes. The findings can also predict in advance that a collaborative project would not be validated by evaluating group contributions in the discussion forum, which is a remarkable achievement that prompts further investigation on how to prevent such failures.

To set the stage for future research, the study's limitations need to be identified. One limitation is the small sample size, which could be expanded to enhance the generalizability of the results beyond first-year programming students. Another limitation is the need for a larger dataset in terms of the number of features, which could enable comparative research on the actions of male- and female-dominated groups and offer insights into group behaviors based on group composition. These limitations present opportunities for future research to improve the study's approach and expand its generalizability.

Acknowledgements
Not applicable.

Author contributions
AH: Description of the Group Assessment Analytics Framework, Ontological modeling, Research methodology design, data collection, and analysis, writing of the introduction, literature review, the description of the Group Assessment Analytics Framework, Ontological modeling, methods, results, and discussion sections of the manuscript. LC-B: Project supervision, methodological guidance, critical revision of the manuscript, and contribution to the writing of the introduction, the description of the Group Assessment Analytics Framework, Ontological modeling, methods, results, and discussion sections.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The author have read and approved the final version of the manuscript.

Competing interests
The author(s) declared no potential competing interests with respect to the research, authorship, and/or publication of this article.

Received: 23 May 2023 Accepted: 13 September 2023
Published online: 26 September 2023

References

Afacan Adanir, G. (2019). Detecting topics of chat discussions in a computer supported collaborative learning (CSCL) environment. *Turkish Online Journal of Distance Education, 20*(1), 96–114.
Aghazadeh, S., Ho, J., Ng, B., Khng, K. H., Roock, R. D., See, T. S., et al. (2019). *Assessment of 21st century skills*. Nanyang Technological University, Office of Education Research, National.
Allaymoun, M. H. (2021). Analysis of CSCL chats for cognitive assessment and individual participations. *International Journal of Computing and Digital Systems, 10*, 181–190.

Apeanti, W. O. (2021). Learning computer programming using project-based collaborative learning: Students' experiences, challenges and outcomes. *International Journal for Innovation Education and Research, 8*, 191–207.
Ben, S. (2022). ScienceDirect using the E-learning gamification tool Kahoot ! to learn chemistry principles in the classroom. *Procedia Computer Science, 207*, 2667–2676.
Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press.
Branch, R. M. (2009). *Instructional design: The addie approach*. New York: Springer.
Chen, C.-H. (2018). Educational research review. EDUREV 263: Elsevier Ltd.
Chen, W., Su, J., Lyu, Q., Chai, A. S. C., & Toh, W. L. (2022). Interaction and monitoring matter: Comparison of high and low-performing groups in cscl (Vol. 1).
Dascalu, M.-I., Bodea, C.-N., Lytras, M., Ordoñez, P., & Pablos, D. (2014). Computers in human behavior improving e-learning communities through optimal composition of multidisciplinary learning groups. *Computers in Human Behavior, 30*, 362–371.
Dillenbourg, P. (1999). Introduction: What do you mean by "collaborative learning"? *Collaborative Learning: Cognitive and Computational Approaches, 1*, 1–19.
Ghedir, A., Cheniti-Belcadhi, L., El Khayat, G., & Said, B. (2018). Ontological models for an eportfolio assessment skills through serious game. In (p. 1–8).
Hadyaoui, A., & Cheniti-Belcadhi, L. (2022). Towards a context-aware personalized formative assessment in a collaborative online environment. In *19th IEEE/ACS International Conference on Computer Systems and Applications, AICCSA 2022*, Abu Dhabi, United Arab Emirates, 5–8, 1–6.
Halimi, K. (2021). Students' competencies discovery and assessment using learning analytics and semantic web. *The Australasian Journal of Educational Technology (AJET), 37*(5), 77–97.
Hernández-Sellés, N., Muñoz-Carril, Pablo-César., & González-Sanmamed, M. (2019). Computer-supported collaborative learning: An analysis of the relationship between interaction, emotional support and online collaborative tools. *Computers and Education, 138*(April), 1–12.
Herrera-pavo, M. Á. (2021). Collaborative learning for virtual higher education. *Learning, Culture and Social Interaction, 28*, 100437.
Hidalgo-SuÃrez, C.-G., Bucheli-Guerrero, V. A.-c.-A. A., & OrdoÃ-Erazo, H.-A. (2023). RÃ basada en competencias de aprendizaje en un curso CS1 para evaluar actividades de programaciÃ CSCL. Revista cientÃfica, 134–146.
Isohätälä, J., Järvenoja, H., & Järvelä, S. (2017). Socially shared regulation of learning and participation in social interaction in collaborative learning. International Journal of Educational Research, *81*, 11–24.
Jan, H., Amin, S. Noor-ul., & Matto, M. I. (2018). Modular object oriented dynamic learning environment. *Journal of Applied Research in Education, 33*(1), 1–11.
Kaliisa, R., Rienties, B., Mørch, A. I., & Kluge, A. (2022). Social learning analytics in computer-supported collaborative learning environments: A systematic review of empirical studies. *Computers and Education Open, 3*, 100073.
Laffey, J. M., Musser, D. R., Espinosa, L., Remidez, H., Gottdenker, J. S., & Hong, R.-Y., et al. (2023). Cscl for schools that learn. In *Computer Support for Collaborative Learning* (pp. 111-118).
Lipponen, L. (2002). Exploring foundations for computer-supported collaborative learning. *Proc CSCL, 2002*, 72–81.
Ma, X., Liu, J., Liang, J., & Fan, C. (2023). An empirical study on the effect of group awareness in cscl environments. *Interactive Learning Environments, 31*(1), 38–53.
Mat Razali, N. F. Y., Hamid, J., & Hashim, H. (2023). E-portfolio as an assessment tool in teaching and learning: A survey of teacher's perceptions. *Journal of Contemporary Social Science and Education Studies (JOCSSES), 2*(2), 64–72.
Maulud, D. H., & Abdulazeez, A. M. (2020). A review on linear regression comprehensive in machine learning. *Journal of Applied Science and Technology Trends, 01*(04), 140–147.
Mhlongo, S., Oyetade, K. E., & Zuva, T. (2020). The effectiveness of collaboration using the hackathon to promote computer programming skills.
Moreno Guerrero, A., Rodríguez, C., García, G., & Navas-Parejo, M. (2020). Educational innovation in higher education: use of role playing and educational video in future teachers' training. *Sustainability, 12*, 2558. https://doi.org/10.3390/su12062558
Mudau, P. K. (2022). Lecturers' views on the functionality of e-Portfolio as alternative assessment in an open distance e-learning. *International Journal of Educational Methodology, 8*(1), 81–90.
Na, S., & Zaidatun, K. (2022). Learning analytics in online learning environment: A systematic review on the focuses and the types of student - related analytics data. *Technology, Knowledge and Learning, 27*(2), 405–427.
Nouira, A., Cheniti-Belcadhi, L., & Braham, R. (2018). A semantic web based architecture for assessment analytics. In *Proceedings—International Conference on Tools with Artificial Intelligence, ICTAI*, 2017-November, 1190–1197.
Panadero, E., & Järvelä, S. (2015). Socially shared regulation of learning: A review. *European Psychologist, 20*(3), 190–203.
Qureshi, M. A., Khaskheli, A., Qureshi, J. A., Raza, A., & Yousufi, S. Q. (2021). Factors affecting students' learning performance through collaborative learning and engagement. *Interactive Learning Environments, 0*(0), 1–21.
Ramadhan, M. A., Maulana, A., Saka, N., & Setiani, O. (2021). Online learning collaboration with assisted learning models in technical drawing subjects for vocational students. 640(Iccie), 63–68.
Reforming higher education with ePortfolio implementation, enhanced by learning analytics. (2023). Computers in Human Behavior, *138*, 107449.
Region, E.-C. (2015). Effects of group size on students mathematics achievement in small group settings. *British Journal of Education, 6*(1), 119–123.
Rodriguez-Ferradas, M. I., Sanjurjo-San Martin, E. L., & Alfaro-Tanco, J. A. (2023). Relevant factors influencing cognitive distance in the performance of collaborative research projects. *International Journal of Qualitative Methods, 22*, 1–15.
Sarwandi, Wibawa, B., & Wibawa, R. (2022). Usage of e-portfolio as an assessment tool in physics learning. *Journal of Physics: Conference Series, 2165*(1), 3–8.
Shum, S. B., Ferguson, R., & Shum. (2014). Technical Report KMI-11-01. Journal of Learning Analytics (June).
Turnbull, M., Littlejohn, A., & Allan, M. (2010). Creativity and collaborative learning and teaching strategies in the design disciplines. *Industry and Higher Education, 24*(2), 127–133.

Vlachopoulos, P., Jan, S. K., & Buckton, R. (2021). A case for team-based learning as an effective collaborative learning methodology in higher education. *College Teaching, 69*(2), 69–77.
Yadav, A., Mayfield, C., & Hu, H. H. (2021). Collaborative learning, self-efficacy, and student performance in CS1 POGIL. 775–781.
Yeom, S., Herbert, N., & Ryu, R. (2022). Project-based collaborative learning enhances students; Programming performance. In *Proceedings of the 27th acm Conference on on Innovation and Technology in Computer Science Education* vol. 1 (pp. 248–254). New York, NY, USA.
Zarzycka, E., Krasodomska, J., Mazurczak-Mąka, A., & Turek-Radwan, M. (2021). Distance learning during the covid-19 pandemic: Students' communication and collaboration and the role of social media. *Cogent Arts & Humanities, 8*(1), 195–228.
Zine, O., Derouich, A., & Talbi, A. (2019). IMS compliant ontological learner model for adaptive e-learning environments. *International Journal of Emerging Technologies in Learning (iJET), 14*(16), 97. https://doi.org/10.3991/ijet.v14i16.10682

Publisher's Note