How are my Students going? A Tool to Analyse Students' Interactions on Capstone Courses

Márcia Lima[1,2] **, Awdren Fontão**[1,3] **, David Fernandes**[1] **,**
Tayana Conte[1] **, Bruno Gadelha**[1]

[1]Intituto de Computação – Universidade Federal do Amazonas (UFAM)
Av. General Rodrigo Otávio, Coroado – Manaus, AM - Brazil

[2]Universidade do Estado do Amazonas (UEA)
Av. Darcy Vargas, 1.200, Parque 10 de Novembro – Manaus, AM - Brazil

[3]Instituto de Ciência e Tecnologia (SIDIA)
Av. Darcy Vargas, 654, Parque 10 de Novembro – Manaus, AM - Brazil

msllima@uea.edu.br,{awdren,david,tayana,bruno}@icomp.ufam.edu.br

Abstract. *Computing-related undergraduate students are encouraged to participate in Project-based Learning (PBL) courses through capstone courses in order to bridge the gap between software engineering (SE) educational and industrial worlds. In these courses, students improve their skills on industrial tools and processes and engage in real-world projects. One of the challenges of this kind of courses is how to monitor students' progress. In this work, we propose a software tool based on statistical analysis and data-mining algorithms to investigate the usefulness of students' communication logs to support professors' pedagogical activities during a capstone course involving three different SE disciplines. Our results indicate the feasibility of using textual content and metadata content extracted from Slack logs to identify opportunities for the professor's intervention. A quantitative analyze reveals an average precision of* 81% *at identifying the top-5 relevant sentences registered in the log.*

1. Introduction

One of the software engineering (SE) education challenges is the gap between industry expectations and software engineering undergraduates' skills [Garousi et al. 2016]. To bridge this challenge computing-related undergraduate courses are applying project based learning (PBL) through capstone courses [Bastarrica et al. 2017]. Capstone course' main goal is to make students work in teams to develop their technical and social skills [Bastarrica et al. 2017]. In SE education, it provides means to join theoretical foundations required in software engineering with practice required for professional performance in the software industry.

It is important to insert students in scenarios that involve software design and software documentation and quality. These kinds of scenarios demand skills on industrial tools and processes that students must develop [Marques et al. 2014]. To support course activities from individual and team perspectives, professors adopted some tools commonly used in the software industry like GitHub [1] to support collaborative develop-

[1]https://github.com/

DOI: 10.5753/cbie.sbie.2019.1611

ment, Slack [2] and WhatsApp [3] in order to support team communication, Google Drive in order to support cloud storage for common development assets and Trello[4] to support the project management during that process.

Development teams in software industry adopt various communication channels [Alkadhi et al. 2018] to support the collaborative development model and coordinate tasks. Recent studies show that software development teams are increasingly using social media for communication purposes [Mushtaq et al. 2018]. When teams use informal channels such as instant messaging (IM) for communication, relevant discussions about the software development and management resides in IM log files. According to Alkadin *et al.* [Alkadhi et al. 2018] IM log files are a rich source of information and can help in identifying important issues.

During a capstone course experience, professors notice that the teams' communication log files contain design decisions related to the project, interpersonal relationship issues related to the teams, technical doubts registers, messages related to team efforts' coordinate [Neyem et al. 2017], and metadata records that can be used as a source of evidence to determine the need for a pedagogical intervention.

In this paper, we outline an analysis software tool based on data-mining (DM) algorithms to investigate whether and how we can use students' communication logs to support professors' pedagogical activities. We discuss its suitability to support professors' activities by detecting technical and no-technical students' needs, offering opportunities for very practical intervention and create more meaningful representations of communication data for professors. Specifically, this study aims at answering the following research questions:

- ***RQ 1:*** *Can we use communication tools' metadata to understand students behavior from individual and team perspectives?*
- ***RQ 2:*** *Can we use data-mining techniques to identify and extract relevant issues in students' communication log?*

2. Background

Methods for automatically support professors' activities related to students needs have been gained attention. Different tools have been developed in order to allow professors to make better-informed decisions on their instructional strategies. In order to do so, this work aims to automatically analyze and provide relevant information to professors who adopt Project Based Learning (PBL) as a pedagogical model.

PBL is a form of situated learning based on discoveries of constructivism that consider that students can gain a deeper understanding of the subjects studied when they work and reflect on such subjects. It consists of an instructional method centered on the learner [Grant 2002]. It allows an in-depth investigation of a topic. In a PBL experience, students engage in real and meaningful problems whose importance is recognized by them and are similar to the challenges encountered by professionals in the area in which students are trained. In this research, professors used PBL through capstone course to

[2] https://slack.com/
[3] https://www.whatsapp.com/
[4] https://trello.com/

enable students to exercise SE professional skills they will need in their careers after graduation course.

We intend to use a data-mining ranking algorithm to analyze the data-log from student's communication interaction. Ranking algorithms are useful for automatically sorting objects according to their relevance. Many Information Retrieval (IR) problems such as text summarization are by nature ranking problems [Gambhir and Gupta 2017]. We used a text summary technique to automatically identify and extract relevant sentences from students' communication log. Position-based measures, such as $Precision@k$ and Mean reciprocal rank (MRR), are used to evaluate the performance of ranking models [Liu 2011]. Precision is the probability that a retrieved sentence is relevant. $Precision@k$ ($P@k$) is the fraction of relevant results out of the first k returned. It captures the ranking quality for applications where only the first k results matters. MRR measure denotes the rank position of the first relevant documents. If all relevant documents are at the top of the ranked list, the MRR value is 1.

2.1. Related Works

Universities are adopting capstone courses on SE education due to their pedagogical value in minimizing the gap between academic and industrial worlds [Neyem et al. 2017]. However, these kinds of projects pose several challenges for academic teams. One of then is how to monitor teams' progress.

Ahadi *et al.* proposed a machine learning-based approach to detect students in need of assistance in an introductory programming course. Authors explored source code snapshot data recorded from students' programming process to detect high- and low-performing students [Ahadi et al. 2015]. Neyem *et al.* proposed a framework for extracting and understanding the knowledge behind the information stored on an educational software tool (EST). They mined the event-logs and derived aspects such as cooperative behaviors in a team, component and student entropy, process compliance, and verification [Neyem et al. 2017]. Özdağoğlu *et al.* also proposed a framework for understanding online learning processes applying data-mining techniques. Authors argue that the framework reveals valuable information to help professors to monitor the online learning processes [Özdağoğlu et al. 2018]. Other researchers proposed software tools to promote the educational progress accompaniment in different scenarios, like [Felix et al. 2018, Porfirio et al. 2018].

We propose a software tool that differs from the other ones by analyzing students' communication log created during a SE capstone course. Based on statistical analysis and data-mining algorithms, the software tool provides data to empowers professors with acknowledgment to make better educational decisions. In our study, we mine the conversation log and the metadata information provided by Slack.

3. Methodology

In this work, we present a data-mining based analysis tool to support professors' activities following his pedagogical planning objectives. We investigate the use of communication's logs collected from SE capstone course teams. The capstone course was designed from the demand of three courses: (1) Project Management, (2) Practice in Systems Analysis and Design, and (3) Database Practice. The courses provided a learning environment that allowed the students to experience an environment similar to the professional environment of software engineers. Students were grouped into development teams, and Scrum [Schwaber 1997] is used as a methodology for agile project management and development. In each sprint, students played different roles according to the Scrum methodology (PO, Scrum Master, Developer, and Tester). Real users with different software needs are indicated to students. To develop the projects, students need to understand the real needs of their users. During development, students have full autonomy to decide how best to meet their user requirements. At each release, new requirements may arise, and new issues to be incorporated into the projects may become necessary. At the end of the course, the students present the results of their projects to the professors and its users.

We collected and analyzed the communication log of three different teams: Team A, Team B, and Team C, all of them from the same capstone course. Students adopted Slack to supported the team's interaction. Slack is a collaboration hub where a team can work together to get things done. Students used Slack to discuss software design, development, verification, validation, testing issues, and other issues. Slack returns metadata objects for each user and channel into a workspace.

Each team is composed of 5 members. The teams' interaction occurred through in-person meetings (at class time) and the Slack tool. We collected teams' log ranging from September 13, 2018, to December 17, 2018. In this timeframe, Team A registered 75 days of conversation and exchanged 706 messages, Team B registered 81 days of communication and exchanged 1,319 messages, and Team C registered 67 days of conversation and exchanged 421 messages.

To investigate whether and how we can use students' communication logs to support professors' pedagogical activities, we propose a software tool based on statistical analysis and data-mining algorithms. We asked Professors to use the proposed tool, and we conducted a quantitative and qualitative analysis based on their quotations. By the end, we follow up by contacting the professors, to understand if the proposed software tool is useful to support their daily teaching tasks. We asked the following questions: (1) Is the proposed software tool useful to support your teaching tasks?, and (2) Would you be willing to use this tool to assist you in your next classes?

4. The Data-mining based Analysis Tool

We proposed a system that extracts data from student's communication logs to create and present statistical information in charts and tables and apply data-mining algorithms to rank the most relevant discussions recorded. The proposed approach provides a user-friendly interface for understanding students' difficulties, misunderstanding, and doubts. Using the software tool, professors can specify properties such as data sources, display formats, graphical visualizations, filters, parameters, and other options that are essential for report presentation. Figure 1 shows the overall Data-mining based analysis tool' features.

To follow the teams, the professors must identify themselves in the system, register the courses, and provide the communication log file (Figure 1(a)). Then, a chart displays the number of exchanged messages per day by the team according to the communication log evaluated. This chart provides insights to professors to select the period they are interested in to conduct the log' analysis. After that, professors have access to "Statistical

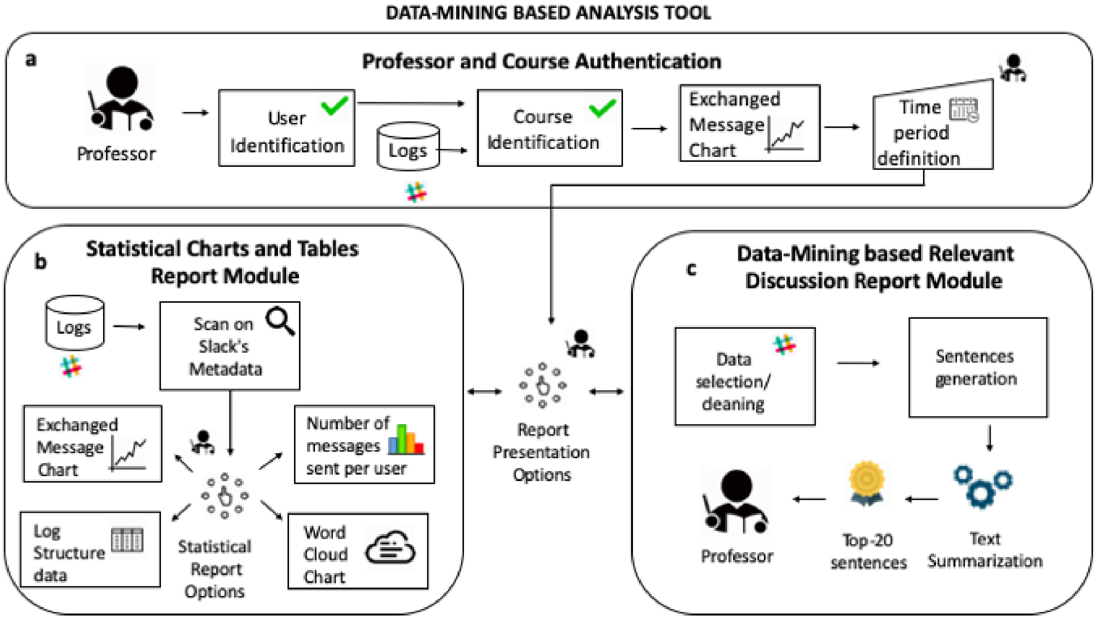


Figure 1. The Data-mining based Analysis Tool Features

Charts Reports" module and "Data-mining based Relevant Discussion Reports" module.

4.1. Statistical Charts and Tables Report Module

Measuring and reporting students behavior based on their interaction logs is one of the base components that can support professors on monitoring students' progress. The tool generates charts and tables that can support professors over all stages of the capstone course. Statistical graphs and tables are plotted based on provided Slack's metadata and text terms frequency (Figure 1(b)). The following statistical charts and tables are present in the analysis tool.

- **Exchanged Messages chart**: It depicts the distribution of posted messages over a predefined period of time on a specific channel. Professors can notice the team's interaction profile by visualizing the number of exchanged messages by the time.
- **Log Structure Data table**: It summarizes structural metadata provided by the log, like the number of Slack workspace' channels, days on which team interaction occurred, number of members by channels, and number of exchanged messages.
- **Word Cloud chart**: It provides an overview by distilling text down to those words that appear with the highest frequency in the log file. It is useful to summarize the log text content. Professors can notice the team's interaction profile by the most common terms they use.
- **Number of Messages sent per User chart**: It introduces the participant's profile via personal interaction with the team's members. Professors can notice students' engagement by analyzing the number of student posted messages.

4.2. Data-mining based Relevant Discussion Report Module

In order to extract relevant discussion from the students' communication log, we applied a text summarization algorithm (Figure 1 (c)), which is a DM algorithm used for automatically sorting objects according to their relevance. Applying text summarization algorithm in natural language requires data pre-processing in order to optimize the algorithms execution [Gambhir and Gupta 2017]. We conducted the following pre-processing steps: Splitting the entire chat log into individual sentences, removing stopwords, non-alphabetic data, and administrative messages sent by Slack. The text summarization algorithm used in this work is based on term relevance measured using Term Frequency (TF) and Inverse Document Frequency (IDF). $TF \times IDF$ measure is used to compute the relevance score of each sentence with the whole document [Haiduc et al. 2010]. We used $TF \times IDF$ to retrieve relevant sentences that could be usefully for class monitoring.

The relevant discussion determination requires prior knowledge about the specific capstone course analyzed. So the three professors, who have a comprehensive understanding of the course evaluated the results from the applied DM-based analysis tool.

5. Results

We present a data-mining based analysis tool to support professors' pedagogical activities. We developed a software tool in order to provide professors useful information by incorporating a comprehensive reporting system, designed to give an overall view of the relevant data in Slack's log from development teams composed by undergraduate students.

To answer RQ1 - *Can we use communication tools' metadata to understand students behavior from individual and team perspectives?* - We designed and developed the "Statistical Charts and Tables Reports Module". This module provides statistical data analysis based on Slack' metadata extraction and text terms frequency. To do so, the module offers different graphic features, like charts of message distribution by teams Figure 2(a); charts of message distribution by members2(c); word cloud charts based on message's texts frequency Figure 2(b); and tables explaining the data-log' structure Figure 2(d).

We also interviewed professors in order to understand what were their perceptions regarding the data mined from students' communication log:

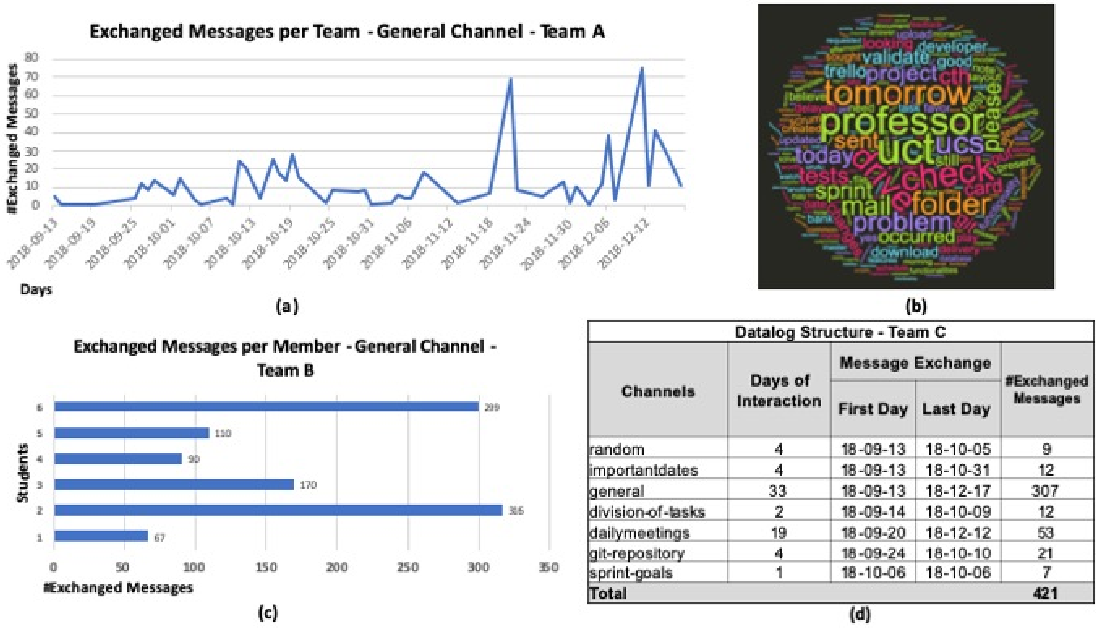
Professor-A's response: *"The data reported on tables and charts covered aspects of planning, development, project release and interaction among team members."*

Professor-B's response: *"As a Project Management professor, I could analyze teams' organization structure. I could visualize teams members' engagement by checking the number of the sent message of each student. Statistical and graphical information was useful, but the analysis of the relevant sentences should complement information regarding the capstone course."*

Professor-C's response: *"Statistical charts did not give much relevant information about the team's behavior. The graphs become useful when combined with the data retrieved by the Data-mining based Relevant Discussion Report feature."*

Answering RQ1: **Statistical Charts based on Slack metadata and the text term frequency are useful to support professors' tasks. However, they provide shallow knowledge about teams activities.**

To answer RQ2 - *Can we use data-mining techniques to identify and extract relevant issues in students' communication log?* - We designed and developed the Data-mining based Relevant Discussion Report Module as a feature of the proposed tool. This module applies a text summarization algorithm to retrieve relevant sentences from students' communication log. We investigated its suitability to support professors' activities



Channels	Days of Interaction	Message Exchange: First Day	Message Exchange: Last Day	#Exchanged Messages
random	4	18-09-13	18-10-05	9
importantdates	4	18-09-13	18-10-31	12
general	33	18-09-13	18-12-17	307
division-of-tasks	2	18-09-14	18-10-09	12
dailymeetings	19	18-09-20	18-12-12	53
git-repository	4	18-09-24	18-10-10	21
sprint-goals	1	18-10-06	18-10-06	7
Total				421

Figure 2. Different type of statistical charts and table report:(a)Message Distribution per Team,(b)Word Cloud,(c)Message Distribution per Members,(d)Datalog Structure Table.

by asking professors to classify the relevance of top-20 sentences identified by the DM algorithm, according to its usefulness on pedagogical activities.

For practical purpose, professors could not analyze all messages in the communication log. They chose different time periods according to their interest to analyze Team A, B, and C log file, as shown in Table 1. The text summarization algorithm provided a ranked list of the most important discussions made by the teams. This ranking was calculated based on the relevance score of 1,308 sentences. A total of 564 relevant sentences were identified and extracted, Table 1. The ranking algorithm was evaluated using $(P@k)$ and MRR measures. In order to measure the precision value of the top-3, top-5, and top-10 results, we calculated the $P@3$, $P@5$ and $P@10$ values.

Table 1. Evaluation metrics values

	Team	#Analyzed Periods	#Exchanged Sentences	Relevant Sentence	MRR	P@3	P@5	P@10
Professor A	Team A	4	110	68	1	91%	85%	79%
	Team B	4	249	62	1	79%	77%	80%
	Team C	4	72	56	1	100%	90%	82%
Professor B	Team A	3	186	60	1	88%	80%	73%
	Team B	4	281	65	0.62	50%	50%	47%
	Team C	5	93	77	0.80	73%	65%	53%
Professor C	Team A	4	140	73	0.87	83%	90%	87%
	Team B	4	154	80	1	91%	95%	92%
	Team C	4	23	23	1	100%	100%	80%
	AVG					83%	81%	74%
TOTAL			1,308	564				

As shown in Table 1, professor A highlighted that the top-3 sentences from each analyzed team were useful to his pedagogical tasks, achieving precision values of 91%, 79%, and 100%, respectively. Considering the top-10 sentences, professor A identified that 79%, 80%, and, 82% of the analyzed sentences were considered relevant according to each team. Professor A also classified all first ranked sentences as relevant; the MRR measure achieved its best value. From professor B perspective, 88% of the top-3 identified sentences from Team A were relevant. On the other hand, he classified just 50% of the top-3 sentences from Team B as relevant. From professors B point of view, not all first ranked sentences were relevant, as one can notice by the MRR values (0.62 to Team B and, 0.80 to Team C). Professor C highlighted that 83% of the top-3 sentences and 87% of the top-10 sentences from Team A were relevant. From professor C perspective, the approach achieved its best efficiency on Team C evaluation, where the MRR, $p@3$ and $p@5$ values reached their best values, revealing that all the top-5 ranked sentences were considered relevant according to the professors' perspective. One can notice that, in some cases, precision values from professors evaluation differs. This phenomenon is due to the fact the relevance judgment involves human inaccuracy, given that the concept of relevance depends on the evaluators' perspective.

We asked professors to describe the importance of identifying relevant sentences:

Professor-A's response: *"Relevant sentences helps the teacher to focus on elements that matter in the teaching-learning process. In this course, students must deliver a product from a project to a real customer, and relevant sentences help reduce the risk of a delivery failure."*

Professor-B's response: *"The relevant sentences extracted from students' communication log were essential to identify opportunities for intervention. As project management' professor, the relevant discussions were useful to evaluate the performance of the scrum master."*

Professor-C's response: *"Relevant sentences are useful to reveal the teams' engagement level. I could identify communication problems, technical failures, and project management activities failures. These sentences are useful to help me improve the feedback I give students at the end of each sprint. I could point out the positives and the negatives aspects related to each team."*

Answering RQ2: **We can extract relevant discussion from SE students communication log using DM techniques to support professors pedagogical activities.**

We also interviewed professors to investigate their perceptions regarding the proposed tool utility and its adoption in other courses as well. They all three showed interest.

Professor-A's response: *"It supports which topics need a "reinforcement" on our part in order to release the product. The tool helps to identify students' needs in order to improve their practical knowledge. Yes. Absolutely! I suggest adding a students' sentimental analysis report."*

Professor-B's response: *" The software tool is useful, so I could identify opportunities for intervention, giving more accurate feedback and support to students. Yes! I would certainly use it in other courses."*

Professor-C's response: *"Yes!! The proposed tool is essential to evaluate and follow-up students. It gives me information regarding students' cooperation in the project. Analyzing the reported data, I could design new classroom activities to clear the students' technical doubts."*

6. Discussion

This paper presents and reveals a promising use of ta data-mining based analysis tool in order to provide relevant information about students behavior and engagement in SE capstone course. Our findings suggest that we can use Slack's metadata information and text term frequency to create statistical charts reports to reveal the teams' organization structure, interaction patterns between team members and identify students' needs related to technical and pedagogical issues. However, professors indicated statistical charts provides shallow knowledge about teams activities.

The results also denote a positive use of text summarization algorithms to identify and extract relevant sentences from students' communication log, as [Ahadi et al. 2015] and [Neyem et al. 2017] also emphasizes. Quantitative analysis indicates that the data-mining module reached an average precision of 83%, 81% and 74% at identifying the top-3, top-5 and top-10 relevant sentences registered in the log, respectively. As highlighted by [Özdağoğlu et al. 2018], data-mining approaches are efficient in revealing valuable information to help professors to monitor and manage the learning process. However, professors' quotations revealed an interest in exploring text sentiment analysis (SA) information from the data-log. They argue that analyzing students' sentiment helps to determine their's perception regarding the course.

7. Conclusion and Future Work

In this paper, we investigate the usefulness of a software tool based on statistical analysis and data-mining algorithms to support professors' pedagogical activities. The software tool provides reports by analyzing students' communication logs. We investigated whether Slack's log data are an important source of evidence to monitor student team's progress in a PBL capstone course. Our analysis indicates that we can use message texts content and the metadata content from Slack logs to provide relevant information to capstone course professors. These data are often overlooked and are not employed during the team evaluation process. We also found that data-mining techniques can retrieve relevant information from log discussions which can be used by professors to identify opportunities for intervention. Professors also highlighted that statistical charts and relevant sentences reports are features that complement each other, providing an overview of individual and team progress. Mining these pieces of information is useful to help professors to determine opportunities for pedagogical interventions, to have insights about teams' engagement level, and to identify technical and non-technical students' needs. Professors also identified interest in adopting the proposed software tool in their future courses.

For future works, we intend to combine data extracted from other tools commonly used in the software industry and adopted by capstone course teams, like GitHub. These data could reveal a more accurate view of team engagement, e.g., by analyzing the number, and the ownership of the committed files.

Acknowledgement

We would like to thank the financial support granted by CNPq (423149/2016-4, and 311494/2017-0), CAPES (175956/2013, and PROAP 001), and FAPEAM (PPP 04/2017).

References

Ahadi, A., Lister, R., Haapala, H., and Vihavainen, A. (2015). Exploring machine learning methods to automatically identify students in need of assistance. In *Proceedings of the Eleventh Annual International Conference on International Computing Education Research*, ICER '15, pages 121–130, New York, NY, USA. ACM.

Alkadhi, R., Nonnenmacher, M., Guzman, E., and Bruegge, B. (2018). How do developers discuss rationale? In *2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pages 357–369. IEEE.

Bastarrica, M. C., Perovich, D., and Samary, M. M. (2017). What can students get from a software engineering capstone course? In *2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering Education and Training Track (ICSE-SEET)*, pages 137–145. IEEE.

Felix, I., Ambrósio, A. P., LIMA, P. D. S., and Brancher, J. D. (2018). Data mining for student outcome prediction on moodle: a systematic mapping. In *Brazilian Symposium on Computers in Education - SBIE*, volume 29, page 1393.

Gambhir, M. and Gupta, V. (2017). Recent automatic text summarization techniques: a survey. *Artificial Intelligence Review*, 47(1):1–66.

Garousi, V., Petersen, K., and Ozkan, B. (2016). Challenges and best practices in industry-academia collaborations in software engineering: A systematic literature review. *Information and Software Technology*, 79:106–127.

Grant, M. M. (2002). Getting a grip on pbl: theory, cases and recommendations. *Meridian: A Middle School Computer Technologies Journal A Service of NC State University, Raleigh, 5 (1)*.

Haiduc, S., Aponte, J., and Marcus, A. (2010). Supporting program comprehension with source code summarization. In *Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering-Volume 2*, pages 223–226. ACM.

Liu, T.-Y. (2011). *Learning to rank for information retrieval*. Springer Science & Business Media.

Marques, M. R., Quispe, A., and Ochoa, S. F. (2014). A systematic mapping study on practical approaches to teaching software engineering. In *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, pages 1–8. IEEE.

Mushtaq, H., Malik, B. H., Shah, S. A., Bin Siddique, U., Shahzad, M., and Siddique, I. (2018). Implicit and explicit knowledge mining of crowdsourced communities: Architectural and technology verdicts. *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS*, 9(1):105–111.

Neyem, A., Diaz-Mosquera, J., Munoz-Gama, J., and Navon, J. (2017). Understanding student interactions in capstone courses to improve learning experiences. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, SIGCSE '17, pages 423–428, New York, NY, USA. ACM.

Özdağoğlu, G., Öztaş, G. Z., and Çağliyangil, M. (2018). An application framework for mining online learning processes through event-logs. *Business Process Management Journal*.

Porfirio, A., Pereira, R., and Maschio, E. (2018). Inferência de conhecimento a partir da detecção automática de evidências no domínio da programação de computadores. In *Brazilian Symposium on Computers in Education - SBIE)*, volume 29, page 1553.

Schwaber, K. (1997). Scrum development process. In *Business object design and implementation*, pages 117–134. Springer.