**Detecting Free-riding effect variations through peer review assessments**

**2021-234**

NAVANJANA E.H.D.T.D.

IT18003642

BSc(Hons) in Information Technology

Specialized in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

October 2021

**Detecting Free-riding effect variations through peer review assessments**

**2021-234**

NAVANJANA E.H.D.T.D.

IT18003642

Individual thesis report

Department of Information Technology

Sri Lanka Institute of Information Technology

October 2021

# Declaration

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

|  |  |  |
| --- | --- | --- |
| Name | Student ID | Signature |
| Navanjana E.H.D.T.D. | IT18003642 | **A picture containing insect  Description automatically generated** |
|  |  |  |
|  |  |  |
|  |  |  |

The supervisor/s should certify the proposal report with the following declaration.

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor: Date : 2021.10.8

# **Acknowledgment**

There are many people I wanted to mention here particularly for the acknowledgment for my research theirs with their direct and indirect contributions. first I would like to thank our project supervisor Ms. Uthpala Samarakoon for accepting our group to supervision and the guidance through this critical period with the pandemic. the motivation and the encouragement through progress meetings and knowledge sharing to improve our research work helps us to achieved quality research studies and good performance in proposal reports, progress presentations, and demonstrations.

Next, I would like to thank you for all the article writers is a different platform like medium, towards machine learning, and many more. I learned soo much from those articles for my research studies. some of the articles help me in the difficult situations when I seek the methodologies to my algorithms. After articles, related videos are published related to those articles by the same or the different people on the web. Your contribution to those research areas improves my research studies.

Next all the reference research papers, thesis, presentations related to those from the different platforms and the different universities. thank you for making them publicly accessible to all the students in the world. when it comes to formulas, and in-depth methodological explanations, and process guidance those reference research papers and research studies contribute soo much meaning to my research studies with this research thesis. finally, my research team members, thank you all for the positive environment created for the research project and the help and guidance with this pandemic time. I would like to thank our team leader for the leadership and all the development contribution made for the project with machine learning environments, presentations, and templates for the research project.

# **Abstract**

Outcome-based education has made so much contribution to the active learning environment in the undergraduate projects all these years. Every infrastructure needs improvements to adapt and survive for a better future. It also applied to outcome-based education. with the project and problem-based learning approach, outcome-based education adapts the undergraduate project modules significantly. but in the evolving education infrastructure can be biased to the polarities define within its framework. some of the biases can be beneficial like student-centered learning. But in this research study, we are focusing on the disadvantage of getting biased to the side of result -orientation on key performance indicators which leads to discouraging the active learning environment. In this research study, we developed the peer-review assessments to represent the process-oriented approach of the key performance indicators in the learning outcomes. student have the opportunity to maintain their review templates about his team mates performance and the algorithm will provide the free-riding effect status with dominant topic indicates from the review with its sentiment score comparing student review scores which will lead to supervision decision making to improve and maintain the active learning environment with positive feedback culture. In the methodology component following main and helper functions to calculate the coherence drop to detect a free-riding effect, compare sentiment scores with review numeric scores collect from student reviews and provide an informative visualization about the project timeline based on monthly reviews. there few review types introduce to discourage hate speech and review bombing to the component. hyper parameters tuned specialized to the student to get the optimal document coverage with their review approaches.

# **Table of content**

[Declaration i](#_Toc84599385)

[**Acknowledgment** ii](#_Toc84599386)

[**Abstract** iii](#_Toc84599387)

[**Table of content** iv](#_Toc84599388)

[**List of Figures** vi](#_Toc84599389)

[**List of tables** viii](#_Toc84599390)

[**List of abbreviations** ix](#_Toc84599391)

[**1.** **Introduction** 1](#_Toc84599392)

[**A.** **Background and literature survey** 3](#_Toc84599393)

[**B.** **Research Gap** 11](#_Toc84599394)

[**2.** **Research Problem** 16](#_Toc84599395)

[**3.** **Research Objectives** 19](#_Toc84599396)

[**4.** **Methodology** 20](#_Toc84599397)

[**Data collection** 24](#_Toc84599398)

[**Data preparation** 27](#_Toc84599399)

[**Data pre-processing** 28](#_Toc84599400)

[**data preprocessing framework** 28](#_Toc84599401)

[**Noise removal** 29](#_Toc84599402)

[**Normalization** 30](#_Toc84599403)

[**Topic** **Model** 36](#_Toc84599404)

[**Pre-requisites** 36](#_Toc84599405)

[**Model build** 37](#_Toc84599406)

[**Model training** 38](#_Toc84599407)

[**Base model training** 38](#_Toc84599408)

[**Model evaluation** 39](#_Toc84599409)

[**Perplexity Score** 39](#_Toc84599410)

[**Coherence Score** 39](#_Toc84599411)

[**Parameter tuning** 41](#_Toc84599412)

[**Model hyper parameters** 41](#_Toc84599413)

[**Model parameters** 41](#_Toc84599414)

[**Sensitivity tests** 41](#_Toc84599415)

[**5.** **Commercialization** 42](#_Toc84599416)

[**Multi-language support** 42](#_Toc84599417)

[**Multicore environment support** 42](#_Toc84599418)

[**Scalability and efficiency** 42](#_Toc84599419)

[**6.** **Testing & Implementation** 43](#_Toc84599420)

[**Data collection** 43](#_Toc84599421)

[**Reading data** 44](#_Toc84599422)

[**Data processing** 45](#_Toc84599423)

[**Pre-processing stage** 45](#_Toc84599424)

[**Sensitivity Tests and hyper parameter tuning** 46](#_Toc84599425)

[**Results** 48](#_Toc84599426)

[**Research Findings** 56](#_Toc84599427)

[**Discussion** 60](#_Toc84599428)

[**7.** **Conclusion** 61](#_Toc84599429)

[**8.** **References** 63](#_Toc84599430)

[9.**Appendices** 68](#_Toc84599431)

# **List of Figures**

[Figure 1 multi-process project timeline[1] 1](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596759)

[Figure 2 importance of Outcome-based education[28] 3](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596760)

[Figure 3 Outcome-based education model 5](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596761)

[Figure 4 Constructive alignment 6](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596762)

[Figure 5 relations between feedback culture in organizations infrastructure[17] 12](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596763)

[Figure 6 Sentiment analysis categories[27] 13](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596764)

[Figure 7 Outcomes-based learning framework [26] 16](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596765)

[Figure 8 Active learning[29] 17](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596766)

[Figure 9 Student review path structure 20](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596767)

[Figure 10 Review types and usage 21](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596768)

[Figure 11 Numeric review scale description 22](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596769)

[Figure 12 Seven steps of machine learning approach 22](#_Toc84596770)

[Figure 13 Two major paths of the main algorithm 23](#_Toc84596771)

[Figure 14 Access range comparison in the cloud space 25](#_Toc84596772)

[Figure 15 Access range comparison in the cloud space 26](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596773)

[Figure 16 Sentiment analysis data preparation 27](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596774)

[Figure 17 The text data preprocessing framework 28](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596775)

[Figure 18 Importance of noise removal process 29](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596776)

[Figure 19 Standard Normalization 30](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596777)

[Figure 20 Duplicate processes to validations 31](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596778)

[Figure 21 Stemming -suffix removal 32](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596779)

[Figure 22 Stemming -prefix removal 33](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596780)

[Figure 23 Stemming -infix removal 33](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596781)

[Figure 24 Stemming -circumfix removal 33](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596782)

[Figure 25 Lemmatization process 35](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596783)

[Figure 26 LDA topic model process 36](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596784)

[Figure 27 LDA Multicore process 37](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596785)

[Figure 28 Perplexity score 39](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596786)

[Figure 29 Umass calculation 40](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596787)

[Figure 30 pre-processed sample with word cloud comparison 45](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596788)

[Figure 31 Cv score calculation of the LDA base model 50](#_Toc84596789)

[Figure 32 Umass calculation of the base model 50](#_Toc84596790)

[Figure 33 Initial intertopic distance map 51](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596791)

[Figure 34 Term-topic distribution 52](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596792)

[Figure 35 Intertopic distance map first iteration 53](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596793)

[*Figure 36 Term-topic distribution first iteration* 54](#_Toc84596794)

[Figure 37 Altered coherence progression 56](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596795)

[Figure 38 Expected coherence progression 56](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596796)

# **List of tables**

[Table 1 Comparison between two different education frameworks 4](#_Toc84596740)

[Table 2 CSV template of the reviews 24](#_Toc84596741)

[Table 3 Bigram generation 34](#_Toc84596742)

[Table 4 Trigram generation 34](#_Toc84596743)

[Table 5 Stable timeline month naming convention 43](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596744)

[Table 6 Unstable timeline month convention 44](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596745)

[Table 7 Sensitivity test results part 1 46](#_Toc84596746)

[Table 8 Sensitivity test results part 2 46](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596747)

[Table 9 Sensitivity test results part 3 46](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596748)

[Table 10 Sensitivity test results part 4 47](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596749)

[Table 11 Sensitivity test results part 5 47](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596750)

[Table 12 Sensitivity test results part 6 47](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596751)

[Table 13 LDA base model parameters 48](#_Toc84596752)

[Table 14 Dominant topic coverage 58](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596753)

[Table 15 Topic number coverage 58](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596754)

[Table 16 Sentiment score coverage 59](https://d.docs.live.net/da69c3750411182d/Documents/individual%20thesis.docx#_Toc84596755)

[Table 17 Basic review flow 59](#_Toc84596756)

[Table 18 Intermediate review flow 60](#_Toc84596757)

[Table 19 Complex review flow 60](#_Toc84596758)

# **List of abbreviations**

|  |  |
| --- | --- |
| Abbreviation | Description |
| PMI | Project Management Institute |
| PMBOK | Project Management Body of Knowledge |
| NLTK | Natural Language Tool Kit |
| OBE | Outcome-Based Education |
| PBL | Problem-Based Learning |
| HMM | Hidden Markov Model |
| LDA | Latent Dirichlet Allocation |
| AWS | Amazon Web Services |
| ITP | Information Technology Project |
| SLIIT | Sri Lanka Institute of Information Technology |
| XML | Extensible Markup Language |
| JSON | JavaScript Object Notation |
| HTML | Hyper Text Markup Language |
| KPI | Key Performance Indicators |
| API | Application Programming Interface |
| OBL | Outcome-Based Learning |
| OBT | Outcome-Based Teaching |
| RE | Regular Expressions |
| NPMI | Normalized pointwise mutual Information |
| PP | Perplexity probability |

# **Introduction**

Project management is a core application of addressing multiple skill sets to different ranges to reaches its optimal requirements. Project management has a multi-process timeline as described in the project management body of knowledge[1].project management timeline starts with project initiation. It will go through the planning stage, execution stage, and control stage before the end of its timeline. In this research component, we are adapting and redesigning the stages from industrial project management timeline to undergraduate project modules.

Timeline

Description automatically generated

Figure 1 multi-process project timeline[1]

The main industrial component we are adapting to undergraduate project modules is project health monitoring and its measurements to maintain optimal project health continuously. In industrial project health monitoring, they are using project health indicators to detect the health and status of key performance indicators. The value of project key performance indicators can be vary depending on the project module and scope of individual functions.

Some undergraduate project modules may have budget requirements but most project modules do not priorities it since it’s not the main objective of student project modules. Some project modules have a limited technology stack defined with its module outline, some project modules have a limited scope. Regardless of the project module and priorities, some key performance indicators are common across all these project modules. As a common project KPI, we can get Project Schedule, Project Technology stack, Project Learning Outcomes, Project Risks and Issues, Project Change Requests, Project Action Items, and Project Quality.

In both industrial and academic projects, KPI will function as a performance management tool and as well as a motivational tool. We are using gamification to maintain project KPI and it will keep the project health at an optimal level. Gamification is the process of motivating and encourage students' engagement with project activities to achieve the optimal KPI with game-like mechanics. In the industry, gamification is used to motivate employees to complete activities. [1]

Monitoring the health of the project continually through a review process is the best practice of a project management timeline. A proper health monitoring process can improve project key performance indicators by reducing symptoms of the sick project.it will eventually increase the success rate of the project module. There are common symptoms of a sick project which can identify through health check processes such as confusion, failing vision, scope sickness. Original vision failing due to conflicting priorities is a common symptom of project development with weak health monitoring. Scope sickness can occur due to an incomplete or ill-defined change management process. Confusion can occur due to lots of reasons such as absent leadership, blind task followings, roles, and responsivity conflictions.

in industrial projects, this health check process is performed by various parties depending on the project's nature.it can be a peer manager, external entity, or project sponsor. Since undergraduate project modules are varied and cant specifically attach a human resource like the industry we divide the health check process to peers of the project group through brief continuous reviews functionality.

## **Background and literature survey**

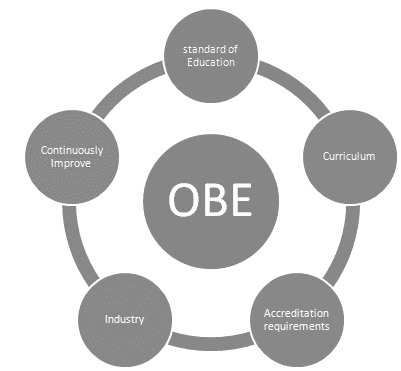
In Project management timelines review assessments and reports, generation is used for multiple purposes by multiple entities from the beginning to end of the projects. In universities, undergraduate project modules design to practice and educate students for the industrial project environments. Because of that undergraduate projects and industrial projects have lots of common methodologies. But most of the undergraduate project management systems barely implement practices related to project health even after the projects because of the focus on different learning outcomes, not project KPIs.

Figure 2 importance of Outcome-based education[28]

Since most undergraduate modules following outcome-based education. Outcome-Based Education is focusing and organizing the module objectives as learning outcomes to evaluate students learning experiences with the different evaluation methods.[2] undergraduate project modules partially shift from Outcome-based learning to Problem and Project-based learning to satisfy the learning outcomes.

|  |  |
| --- | --- |
| **Content-Based Education** | **Outcomes-Based Education** |
| Passive students | Active learners |
| Assessment procedure   * examination driven * grade driven * exam & grade driven | * Continuous assessments |
| Rote learning | Critical thinking, reasoning |
| Content-based module knowledge | Integration knowledge,  learning relevant real-life situations |
| Teacher centered | Learner-centered |
| Pre-defined syllabus | Tangible learning paths |
| Supervisors responsibility | Learners responsibility |
| Emphasis supervisors vision | Emphasis Learners vision |
| Fixed timelines | Flexible timelines |

Table 1 Comparison between two different education frameworks

As above table outcome-based education evaluation goes through continuous assessments. But in undergraduate project modules, these continuous assessments altered into the problem and project-based learning events such as proposal reports, proposal presentations, progress reports, progress presentations, final reports, final presentations. There are no pre-defined or existing methodology to monitor the learning path or development path of the undergraduates through these events. Because of that above-mentioned learning experience can be wasted with a lack of monitoring. Project health is an essential factor to empower these outcome-based education learning experience benefits such as critical thinking, integration knowledge, responsibilities of the learners, and emphasis on their proposed vision through flexible timelines. If there is no proper way to empower outcome-based education it will act as a traditional content-based education system which will be a huge disadvantage to students who are facing an industrial environment.

Because complete shift will lead to having tangible outcomes and expected outcomes to have many variations which are difficult to evaluate.[3] but lack of project monitoring in outcome-based education through project life cycle will affect its aspects which are objectives and outcomes.

In outcome-based education, there are three types of outcomes sub-divided into different categories. The first type is Knowledge which is the category students developed before they getting in-depth experience or have an interest in a particular stream or development area. This is where both academic and industrial guidance comes in. It is both industrial and institutional responsibility to provide proper information for the students. The institutional or industrial interests of decisions making on the teaching approach independent situation to them.in current outcome-based education, there are no methodologies to evaluate its aspects. Weak knowledge transferring methods will discourage the student's outcomes.[4]

Diagram

Description automatically generatedThe second outcome category is a skill that will heavily depend on first category knowledge. Weak knowledge transferring methods will affect students' skill sets which will fail the outcome-based education process. Third and the last outcome is attitude. the positive attitude of the student is an outcome that is expected on outcome-based education. Attitude can be independent and depend on their knowledge and skills. [4]

Figure 3 Outcome-based education model

When building an Outcome-based Education syllabus industry and institutes follow a process called Constructive alignment which is creating an active learning environment that supports student learning activities defined for achieving the expected learning outcomes by module outline.

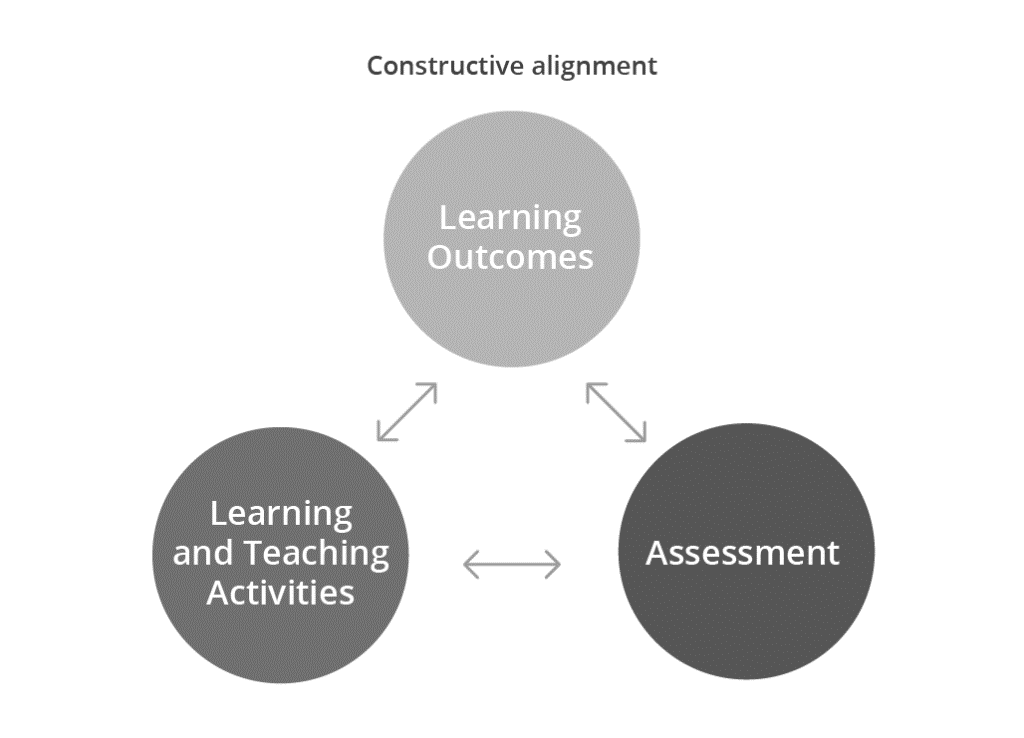
 constructive is about what student learning path to achieve the learning activity. alignment is about the method of teaching component which is an aspect of a teaching system. to monitor the learning path there no methodologies implement in student project modules as mentioned above which will affect benefits students can get from Outcome-based education through an active learning environment. Clarity is one of the main benefits of outcome-based education. The emphasis on outcomes establishes clear expectation points of what students must be completed at the end of the module competence. Students will understand what is expected of them, and supervisors will understand what they need to teach and guide their students during the course. Aside from years of study, clarity is difficult when teaching and guidance are done as a project team.

Figure 4 Constructive alignment

Clarity allows each project member to understand what has to be completed in each phase of the project timeline or at each level for students to grow. Once the outcome has been determined, the persons who are outlining and planning the curriculum are required to work. They must determine what knowledge and skills will be necessary to achieve the desired result.

Flexibility is another benefit of Outcome-Based Education which empowers student-centered learning methodologies. To teach and guide a student, any method can be used. Outcome-Based Education does not require the supervisor to use a specific technique of instruction. Outcome-Based Education is a model of student-centered learning. Using either strategy, supervisors are supposed to guide and assist students in mastering the content.

Comparison can be taken as another benefit of outcome-based education. Based on Outcome-Based Education, several institutions are compared. Individual student achievement can be utilized to determine what outcome they have achieved. At the institutional level, institutions can be compared their module outlines by examining their shared outcomes and identifying areas for development based on the outcomes attained at competitive modules from different learning paths. But none of the institutes or different modules in one institute not following outcome-based education monitoring there is no proper way of comparison to see the benefits of the existing outcome-based systems. It’s difficult to compare different modules but common KPIs and learning paths can be compared with proper monitoring methods.

The last core benefit of Outcome-based education is participation. Student participation in the project events is an important aspect of Outcome-Based Education. Students are expected to perform self-learning to fully comprehend the topic. Students who are more involved in their study feel more responsibility for their learning which will empower an active learning environment among their project members. According to Spady’s Outcome-Based Education paradigm, there are four principles. The four core principles are Clarity of focus, Expanded opportunities, High expectations, and Design down.

The clarity of focus means that everything supervisors do must be clear in terms of what they want their project group students to know, understand, and be able to do. In other words, the supervisor should focus on assisting his students in developing their knowledge, skills sets, and personalities that will enable them to attain the specified planned learning outcomes. To help and assist with students' skill sets and their personality supervisor has to have a basic idea about his project group students. But with existing outcome-based education system only focus on results and module content development.

Expanded opportunities describe how supervisors must work hard to create more possibilities for all their students. This theory elaborates that not all students can learn the same technologies in the same learning path or at the same learning rate. Most students, however, will achieve high levels if given sufficient opportunities through the project’s flexible timeline. but with an outcome-based education system that replaces continuous assessments with project-based evaluation events can’t give the students mentioned expanded opportunities in flexible project timelines. This is linked to the third core principle of outcome-based education, which is high expectations, according to this theory, prior learning events can reinforce the success rate of the current learning events of the students, provides motivation, and increases self-confidence. This theory also implies that supervisors should set high, challenging performance goals to motivate students to engage in-depth in what they are studying. Helping students achieved high standards is tied to the practice that successful learning events foster further upcoming learning events as mentioned above. The last principle of outcome-based education is designing down. The designing down describes the outline of the module design must initiate a clear learning outcome of the targeted KPIs that students are to acquire by the end of the project timeline. After that, all instructional decisions are taken to guarantee that the desired result is achieved.

In current project modules and other undergraduate modules, feedbacks or reviews function as a summative assessment method. which means it gathers data or information at the conclusion or the end of the course, project, or module. which is not helping the current project performance management. There are several types of assessment methods in outcome-based education frameworks. The first one is formative assessments. It is a collection of information about students learning paths and their progression to improve their learning rate. The second one is which we described above with the existing review methodology. The third assessment method in outcome-based education is criterion-referenced assessment. This assessment compares student scores with specifically defined standards. in another word, it is a score for an identified learning outcome.

There are other assessment methods like alternative assessments which are direct observation of student performance. Authentic assessments are about real-life experiences like recorded learning processes, performance tracking, new knowledge integrations. The last assessment type is performance assessment which is a demonstration of the student's learning process which reflects his skillset or knowledge.

All these types of assessments are used to monitor and evaluate the performance management of a project which helps maintain the project health through the project timeline. Performance management is a methodology of managing a project timeline performance. It can have a significant impact on a project's success or failure. It can be used to assess the performance of an entire module learning outcomes, a project supervisor, an individual contributor, or even the process of developing a product or service through that project. It involves KPIs that will assist in ensuring that learning outcomes are continuously met in an effective manner. These action plans involve planning and setting expectations as we mentioned on one of the outcome-based education principles, developing performance capacity, continuously monitoring performance, rating performance in a summary format on a regular basis, and rewarding good performance.

This performance management and project health monitoring can fill the incomplete principles of the existing project module’s outcome-based education framework that are lacking. The performance management process through various performance indicators.

Performance indicators are pre-defined assessment points that evaluate a project's member's performance and success rate. They are the goals that should be achieved in order to offer the most value to a project and the project module. They are methodologies for evaluating the performance of a project and its team members. Performance indicators are created to have an impact on the entire project timeline. As a result, selecting the proper performance indicators is dependent on a thorough understanding of what is essential to the project module.

The project module may use performance indicators to evaluate its own success rate. There are two types of performance indicators. Result-oriented indicators are the common indicators we can see in existing project modules. Its focus is on key outputs of the process and related critical success points. Process-oriented indicators are passively functioning or not found at all in existing project modules. Some of the project modules collect and maintain Gantt charts, log books, Kanban boards, status documents like process-oriented resources but not consider as a process-oriented indicator to continue to add up the performance management process.

Most of the project modules only consider critical success factors as result-oriented indicators. Most result-oriented indicators are bind with delivery management and deadlines because of this process continuous improvement only encourage by deadlines and supervision. In this research study, we are addressing these continuous improvements with process-oriented indicators to manage the performance and health of the project timeline.

## **Research Gap**

Outcome-based education theory has long history and broad research studies related to it. Outcomes-based education started with its founder William G. Spady in 1994. from there, various types of research studies were carried out by researchers related to outcome-based education and related theories.

Most of the outcome-based education research studies trying to go with the curriculum level[5].other research studies related to outcome-based education focus on the program level and module levels on higher education and undergraduate environments[10]. Some of the research studies go to the national level overview based on existing policies of the countries related to outcome-based education.[7]

Existing outcome-based education research studies can be divide into four major studies based on the context of the research studies. There is outcome-based learning which addresses principles of the theory and impacts generate from those principles to the systems. There is outcome-based teaching which focuses on teaching psychology with constructive alignment theories.[11] There is outcome-based education that explores the overview and existing systems.[8]there is outcome-based education focus on students.[9]

But most of the research studies are based on and favor to teaching rather than learning and the general side of the outcome-based education systems.[6] so there is significant research on outcome-based teaching than outcome-based learning or outcome-based education sections.

An outcome, based-based education research studies are also more biased into improving the module framework rather than learning process even in the researches based on undergraduate sided studies. There are more favored to focus on outcome-based education with constructive alignment, curriculum, continuous improvement, quality assurance, and educational acceleration of the module infrastructure focus on improving upcoming undergraduates learning experience and improving teaching methodologies that exist in the current outcome-based education infrastructure.[10]

In performance management of undergraduate project modules, some of the research studies implement novel systematic approaches without the outcome-based education frameworks.[12] Also in performance management academic perspective is more favored. performance factors compare with multiple types of scores according to academic perspective carried on[13]

In performance management comes to KPIs and other performance indicators most common approach is employee and management relationships. KPIs and other performance indicators are evaluated in business and industrial levels more than the academic level which including higher education and undergraduate studies.[14]

Even in the carried research studies about undergraduate project modules with consideration of key performance indicators more favored to result-oriented indicators to support the existing outcome-based education.[15]other research studies focus on the psychology side of the existing key performance indicators in the education system with cognitive thinking.[16]

In some research studies performance management process extend to longitudinal and continuous performance management process events with individual’s review orientations and try to demonstrate a feedback culture about review driven solution.

Diagram

Description automatically generated

Figure 5 relations between feedback culture in organizations infrastructure[17]

Above mentioned research study presented a theoretical framework more similar to this research study. but it goes through continuous professional development which is considering feedback results to generate a feedback culture among organizations and the relationship between feedback culture and feedback usage. It is trying to diagnose the feedback process and improve the feedback process with feedback culture.[17]

There are research studies connecting performance and well-being with feedback environments.[18]they proved the positive feedback environments have higher perceptions and satisfaction rates.it also recommended developing supportive feedback environments.it explores about feedback environment scales with seven dimensions. Most of these types of research studies were conducted their research studies to improve the review process improve the quality of feedback given and collected through the procedures.[19]

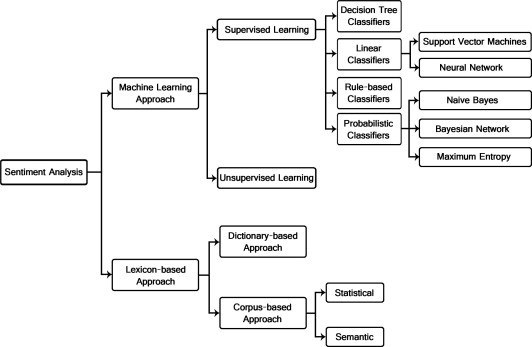
When it comes to in-depth analysis of feedback procedures. Sentiment analysis is the most common research topic in the analysis of student feedback.[20]Sentiment analysis is broadly divided into three major sections in the existing research literature. There are machine learning-based sentiment analysis, lexicon-based sentiment analysis, and hybrid sentiment analysis.[20]

Figure 6 Sentiment analysis categories[27]

In machine learning-based sentiment analysis can be divided into the supervised learning and unsupervised learning methods.in unsupervised learning, the methodology does not require a dataset to annotate sentiment labels in sentiment analysis. Text content polarity is determined by point-wise mutual information procedures which is aggregating the adjectives and adverbs of the selected phrases.[21] there are many variations and approaches to unsupervised learning using for the sentiment analysis to review textual context.[22]

in supervised machine learning approach of the sentiment analysis in textual context have training classifiers with linguistic features representing them in extraction methods. The labeled dataset is a major requirement to train the classifiers. There are lots of linguistic features using in the supervised learning method with sentiment analysis such as n-grams, word representations, part of speech tags, punctuations and emoticons. there are several ways to train the model in supervised learning with sentiment analysis such as naïve Bayes, maximum entropy, support vector machine[23]

in the lexicon-based approach of sentiment analysis are using the lexicon of the related sentiment to determine the polarity of textual content of the reviews or feedbacks. Word lists collect by lexicon or dictionary depend on the process which is associated with sentiment polarity. Lexicon generation can be a manual or automatic process depend on sentiment analysis process requirements. [23] one disadvantage of the lexicon-based sentiment analysis is the given word’s contextual value and domain-associated semantic representation is getting ignored.in the following research study with lexicon-based sentiment analysis suggest use of domain-specific sentiment lexicon methodology represent better outcomes than any other general-purpose sentiment lexicon approaches. [24]

Except for these two approaches, there is a hybrid approach that uses both the above approaches to sentiment analysis with sentiment lexicon and machine learning methodology.[25] This research study also represents the hybrid approach to sentiment analysis with a combination of machine learning approach with sentiment dictionaries to determine the sentiment scores and orientations of the given feedbacks.

Most sentiment analysis studies mainly focus on sentiment classification, which is trying to define the sentimental polarity of a text which is negative or positive. In the past years of researches, sentiment classification methodologies have been applied in different datasets, including product reviews [6, 8, 9, 11, 12, 13, 14], tweets [10, 14, 15, 18], news articles [16], movie reviews [9, 13, 17, 18]. The approaches used for sentiment classification can be divided into lexicon-based approaches [13, 19] or machine learning-based approaches [6, 14, 17, 18], the neural network models [9, 10, 12, 14, 15, 18], have proven significant performance. Based on our knowledge, sentiment classification techniques have not been applied in the aspects of peer reviews for undergraduate project management processes.

Few research studies have tried to automate the prediction of peer reviews’ effectiveness and enhance the quality of peer review assessments. In these research studies, the peer review assessments are not limited to peer reviews for research papers but include peer reviews for students’ work. For example, [7] presented machine learning methods applied for classifying peer review comments in writing, and the support vector machine has proven a significant performance. [20] proposed a system for generating automatic assessments of reviewing performance depends on problem localization at the reviewer level, and shown the possibility of detecting reviewers who have minor problem localization in their reviewing. [21] further showed that the utility of generic features in predicting review effectiveness depends on different review types. [22] used effective data preprocessing techniques with latent semantic analysis and cosine similarity to determine the tone, quality, and review comments count. [23] proposed to use a decision-tree-based classifier to detect the review’s quality. These tasks are related to this research study but different from the sentiment analysis tasks focus on this research study which is aspect-based sentiment analysis based on undergraduates peer-review assessments.

# **Research Problem**

In existing Outcome-based education-centered project modules over usage of result-oriented indicators as learning outcomes leads to different alterations and downsides of its own infrastructure. Because of that the very own module acting against its core objectives such as empowering traditional education framework instead of moving from it to project and problem-based learning.

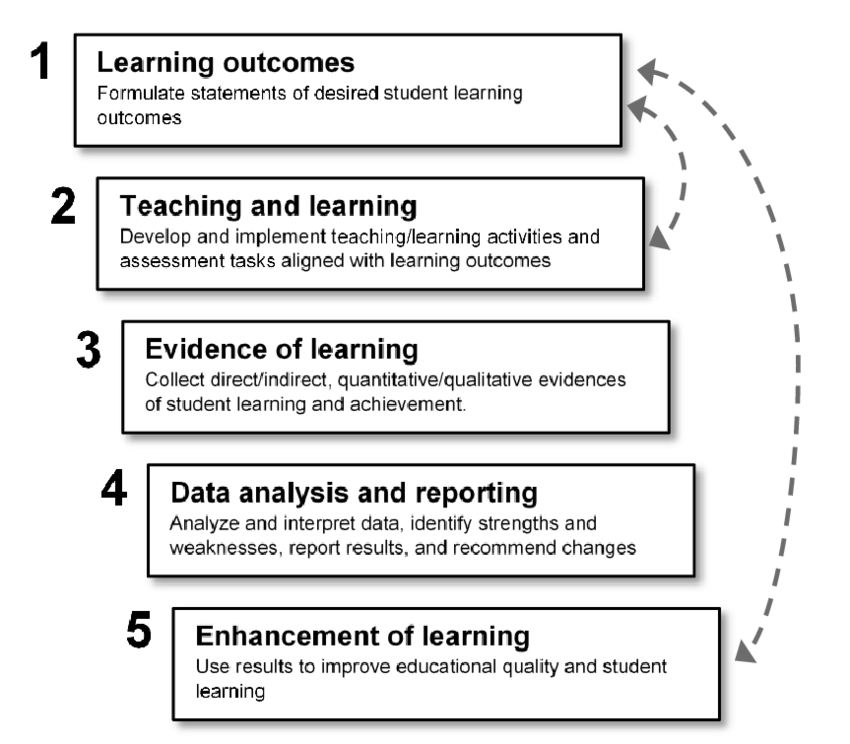


Figure 7 Outcomes-based learning framework [26]

This biased indicator usage is affecting to the outcome-based framework and its functioning as an incomplete and unstable framework with existing project modules. Because of this incomplete framework functionality, other core methodologies which are supported by the outcome-based education and supported to outcome-based educations is passively functioning or not functioning at all. Other core methodologies such as active learning and active learning environment. If the active learning environment which is bind to the project module becomes a passive component all the sub methodologies based on it will have also become passive functioning components.

collaborative learning, cooperative learning, problem-based learning, project-based learning will be discouraged by the result-oriented indicators dominance effect because these core methodologies rely on process-oriented indicators through project timeline.

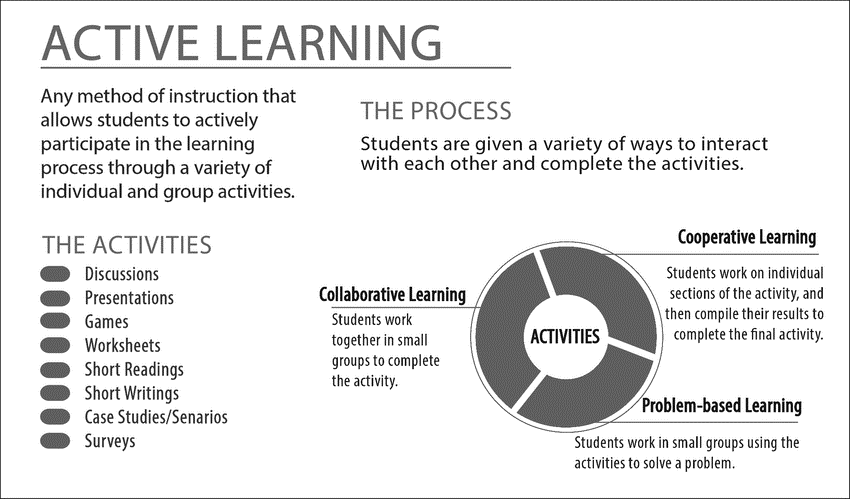


Figure 8 Active learning[29]

Because of this passive functioning effect of the learning environment will transform to an altered traditional education framework. in this altered education framework, we have new types of symptoms generated with the nature of the project modules.

In this research study, we are going through major symptoms of this alerted education framework which are the free-riding problem or free-riding effect on the project timeline and the project group. free-riding effect or problem generate when one or few members of the project group lack of contribution which lead to unfair and misleading grading of the whole project group.

Research conducted with undergraduates indicates that participation in group projects promotes their academic achievement, persistence in university, and positive mindset about learning also existing research studies recommend a peer review process to discourage free-riding in the projects.[30]There is group psychological theory bind with role models in the project group which describes the free-riding can passively and actively occur based on the orientation of the group members behaviors and groups behavior with their project supervisor.[31]

Undergraduates are more likely to expect positive outcomes from group experiences when supervisors provide them with information and guidance about teamwork.[35]

The undergraduates were particularly frustrated when they believed that the supervisor had poor teamwork skills or a lack of responsibility for helping their groups.[34] this can occur with the project timeline due to a lack of process-oriented outcomes.

The supervisor needs to stay tuned with his/her assigned project groups. Supervisors' help will be less effective if she doesn’t have sufficient information about the group and the group members to improve their project management skills. Existing project modules don’t offer process-oriented monitoring assessments to fulfill these requirements.

Group members who shirk their obligations in the hopes of benefiting from the work of others are often referred to as social loafers or free riders.[32] A free-riding problem in group work happens when one or several members of a group low contribution to a group project that if the same grade is given to all members in the group, the grade would be misleading and unfair. Peer assessment can be used to empower undergraduates by giving them to have control over their marks by allowing them to justify and explain their work. Peer assessment can reduce the possibility of the free-riding problem as the anticipation of receiving negative feedback can be enough for a student to modify his or her behavior.[33]The free-riding problem will remain with undergraduate project modules every year if there are no solutions to monitor it or discourage it.

# **Research Objectives**

In this research study, the main objective is to build and maintain the proposed peer-review assessments component. This component is a helper function to maintain process-oriented indicators defined within key performance indicators on the project module outline and the expected learning outcomes.it fulfills the incomplete framework functionality of the outcome-based education and empowers the active learning environment of the project timeline.

To maintain positive feedback culture we are exposing review reports to supervisors of the project modules. At the input end, there will be project members of the project group and at the output end, there will be the supervisors of the project groups. The peer-review component will assess the review in multiple ways to generate and visualize the optimal profile of the given review to its algorithm.

# **Methodology**

The peer-review assessment component collects monthly student reviews to initiate the peer-review assessments with its main algorithm and other supportive algorithms. In the reviewing process project group member has to review all his project teammates monthly. review process time consumption will depend on the project team member count of the project team and project modules month allocations.

In the review process, all the team members have the fair ground with their team roles which means the project team leader will assess like all the other project team members of his project team. in the beginning instructions and guidelines to the proper review input will be given to the project team members of project modules through the web application and cloud space.

Month 1

Student 1

Student 2

Student 3

Month 2

Month 3

Figure 9 Student review path structure

Review

Textual

Numeric

Supportive Algorithms

Main Algorithms

Figure 10 Review types and usage

The peer-review process is collecting few types of reviews from each project member for its main algorithm and supportive algorithms. Overall the textual review is the prioritized one. Besides, those numerical reviews are also requested with the textual review to use in the supportive algorithms.

The review process is time-consuming and it should be. It is designed to be time-consuming on purpose to encourage an active learning environment. it doesn’t have a strict and fixed deadline. But the end of the month is required to be submitted as a responsibility of the project member.

Numeric review scores represent the positivity and the negativity of the project teammate's experience with a reviewed team member. Also, it represents the current situation with him as a project team member with the decision making in the project timeline which describes these scenarios in figure 11.

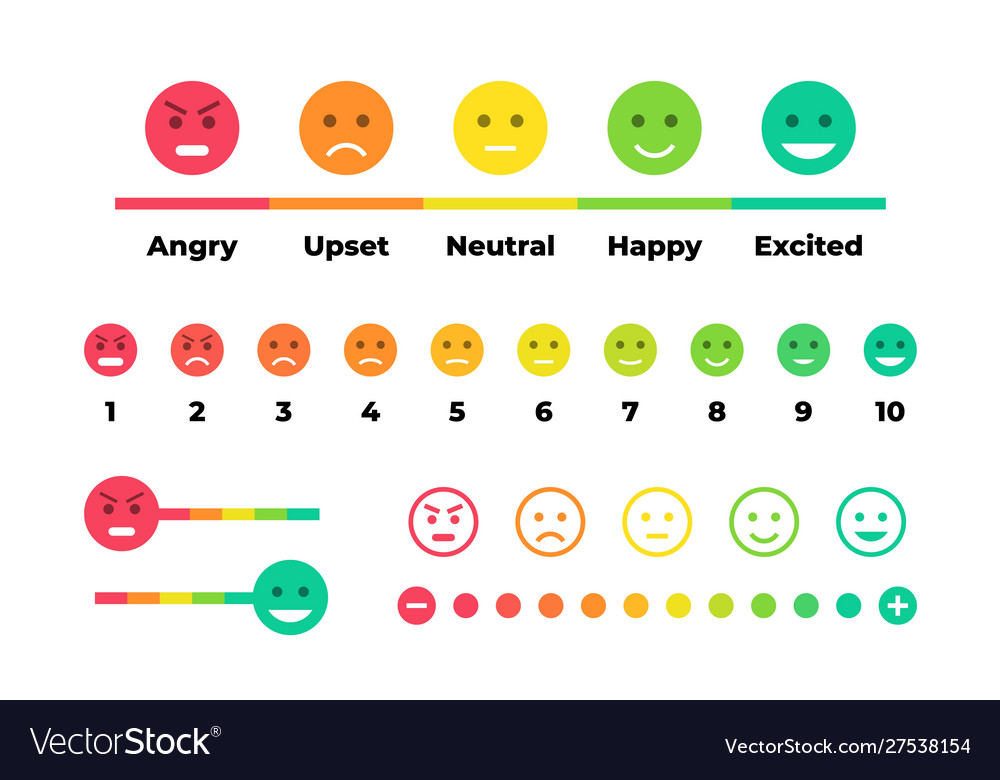


Figure 11 Numeric review scale description

This numeric review score will be compared with sentiment analysis paths positive sentiment score, negative sentiment score, and compound score to identify the role models of free-riding effect combined with group psychology which will be elaborate on results and discussion. The main algorithm design to divide its functionality after and within the pre-processing stage. Even after the separation, both algorithm paths will follow the basic functionality process of a machine learning environment as shown in Figure 12.

Text

Description automatically generated

Figure 12 Seven steps of machine learning approach

Diagram

Description automatically generated

Figure 13 Two major paths of the main algorithm

Before move on to the steps of the each path’s machine learning stages of the algorithm, we will go through the main algorithm’s paths structure. The main algorithm divides into two paths to process the textual data achieve from the prior pre-processing stage of the algorithm. the first path leads the algorithm to the topic extraction component which has dominant topic coverage and coherence drop detection graphs. The second path leads the algorithm to sentiment score calculation to compare the numeric review scores get from the project group members.

## **Data collection**

Data collection to the algorithm will input will be conducted through similar existing practices on project modules on the modules of the undergraduate project which is cloud space. beginning of the project module CSV template to review process will be published and share with project module students.

After that students of the project module can maintain the template each month by providing the proper reviews to the peer-review component. Students of project modules are responsible for maintaining every month's reviews of every team member of the project group.

These review sources will go through processing steps in various stages to provide the optimal visualization to the algorithm to process the information on those stages. some of the stages will have to concatenate processes and some of the will have data frame reorder and redesign to provide more or less information to the algorithm.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Review Id** | **Reviewee Id** | **Month** | **Review Score** | **Textual Review** |
| Student 1 | Student 2 | Month 1 |  |  |
| Student 1 | Student 2 | Month 2 |  |  |
| Student 1 | Student 2 | Month 3 |  |  |
| Student 1 | Student 2 | Month 4 |  |  |
| Student 1 | Student 2 | Month 5 |  |  |
| Student 1 | Student 2 | Month 6 |  |  |

Table 2 CSV template of the reviews

**algorithm access range**

**student access range**

Project Team 001

Review Folder

Student 1

Student Reviews

Student1\_Student2.CSV

Student1\_Student3.CSV

Student 2

Student Reviews

Student2\_Student1.CSV

Student2\_Student3.CSV

Pre-processed Reviews

Figure 14 Access range comparison in the cloud space

As in figure 13 the maintenance of the review data, it will need a proper naming convention and access management to provide a positive feedback culture to the project module. properly named the CSV templates will lead to efficient data collection and productive review data maintenance in the cloud space for the algorithm's current and future usage. Also, the access range for the supervisor will be defined in the cloud space to maintain a positive feedback culture.

**algorithm access range**

**Supervisor access range**

Project Team 001

Review Folder

Student 1

Review Report

Student1\_Student2.pdf

Student1\_Student3.pdf

Student 2

Review Report

Figure 15 Access range comparison in the cloud space

## **Data preparation**

In data, preparation algorithm will add and remove necessary rows and columns depend on its functionality of the stage to the above mention table 2 which is the basic input feedback collecting by peer-reviews in data collection. to visualize a theory or compare these reviews alongside with algorithm result from raw data won’t be enough that’s why there are data preparation stages include in most of the stages.

­­

Figure 16 Sentiment analysis data preparation

Numeric review Score

Positive Sentiment Score

Neutral Sentiment

Score

Score

Negative Sentiment

Score

Compound Sentiment

Score

With this Sentiment score and Numeric review score comparison, we can demonstrate the team member's performance with both result-oriented and process-oriented indicators to achieve the expected learning outcomes.

The numeric review score is the only raw data material on above figure 15. all the other mentioned sentimental scores come from the sentiment analysis path outcomes. Those can be helper scores to the numeric review score most of the time. But if it's not hate speech or review bomb situation can be occur on the review inputs to the algorithm.

## **Data pre-processing**

In data-preprocessing there are several methodologies that exist in the research studies related to the sentiment analysis field with lots of approaches. in this research study, we are using regular expressions to go through the data-preprocessing stage of the algorithm.

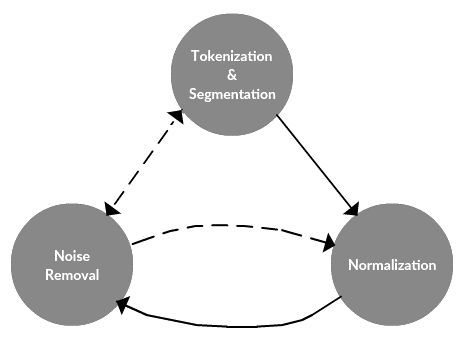
****

Figure 17 The text data preprocessing framework

### **data preprocessing framework**

There are three major sections that follow in this data-preprocessing in the algorithm as shown in figure 17. some of them are light processes and some of them are heavy processes depend on the tasks and approaches with the algorithm. The noise removal is task-specific and tokenization and normalization is task-independent.

### **Noise removal**

This algorithm's pre-processing pipeline goes through some light noise removal stages because we get data on specific CSV templates. If we get the data on HTML, XML, or JSON types of templates, we will need heavy noise removal with their headers, footers, and meta data.

Noise removal follows before any segmentation or tokenizing stage because of the contractions. The contraction will lead the tokenization process to split the without expanding which is not ideal for a sentimental analysis process. As shown in the figure 18 it can lead the word to invert the values and meanings.

Also putting this step after tokenization will generate the same result because of the segmentation already happens. We can take bigram functions to combined segmented negative words but it's highly unlikely to expect the same outcome we get from the noise removal process.

Figure 18 Importance of noise removal process

Contractions word

Standard Normalization

Tokenization

Advanced Normalization

Didn’t

Didnt

Did

nt

Did

### **Normalization**

The normalization process comes after the noise removal process in the pre-processing pipeline. normalization will divide into two categories based on the functionality. The first one is the standard normalization process. as shown in the figure 16 it has a few sub-processes.

Remove new line characters

distracting single quotes

Remove punctuation

Lower casting Texts

Remove non-ASCII characters

Replace all integer occurrences

Figure 19 Standard Normalization

#### **Standard normalization**

Standard normalization is about setting up the fair ground for all the texts in the context. some of the steps taken in standard normalization can be re-use in upcoming stages or steps to validate or for the confirmation purposes such as removing punctuations as shown in figure 17.

Standard normalization

Actual Step

Remove punctuations (RE)

Tokenization

Validation Step

Remove punctuations (gensim)

Figure 20 Duplicate processes to validations

The validation process or the confirmation process will come as a build-in function on the some of the stages and steps depend on the resources usage. For example in the figure 17 remove punctuations basically initiated in standard normalization as a regular expression function with the given data frames. But after that validation process comes with the tokenization stage with gensim utills which is a complete different approach.

#### **Advanced normalization**

Advanced normalization will come after the standard normalization which will set the fair ground for all the textual data. Advanced normalization again divides into three steps based on functionality. There stop word removal, stemming, and lemmatization.

##### **Stop word removal process**

Stop words can be identified as the common words used in textual data which has minimum impact on its overall meaning.in our research study, we are using nltk library to remove the stop words. Nltk is identified as The Natural Language Toolkit which is one of the well-recognized and recommended NLP libraries in the Python ecosystem which is used for tasks like tokenization, stemming, part of speech tagging and etc.

##### **Stemming process**

The stemming process is about removing affixes which including suffixes, prefixes, infixes, circumfixes from the given textual data to get the stem word of it.

Useful

Suffix

Stemming

Stem word

Use

-ful

Figure 21 Stemming -suffix removal

-Un

ited

United

Stem word

Suffix

Stemming

Figure 22 Stemming -prefix removal

-s

passerby

passersby

Stem word

Suffix

Stemming

Figure 23 Stemming -infix removal

Suffix

Stem word

Immaturity

matur

-Im -ity

Stemming

Figure 24 Stemming -circumfix removal

##### **Phrase detection - Bigram and Trigram models**

Bigram and trigram model building and generation is a helper function to the lemmatization process. Bigrams can be described as two words that score together in the textual data context. Trigrams are three words score frequently in the textual data context. in this research study, gensim phrases model is building bigrams and trigram for the next lemmatization process. There are two core arguments in the gensim phrases model which are min\_count and threshold. higher argument settings can be lead to low bigram and trigram count. After defining min\_count and threshold for the bigram and trigram models process will move on to the function calling for the models before the lemmatization process.

|  |  |
| --- | --- |
| Sentence | His time management is brilliant |
| Bigrams | His Time |
| Time Management |
| Management is |
| Is brilliant |

Table 3 Bigram generation

|  |  |
| --- | --- |
| Sentence | His time management is brilliant |
| Trigrams | His Time Management |
| Time Management is |
| Management is brilliant |

Table 4 Trigram generation

##### **Lemmatization process**

Lemmatization is bind to the stemming process but it is different because what is capturing is citation forms which are based on word’s lemma state. Lemma is refer to the head words which incudes in the dictionary.

Sometimes lexeme form which is the variation of the lemma or head word can differ from its lemma which can lead to change the impact it can put on the textual context.

But the sentimental value will be the same as the variation of the lemma. because of that lemma variations will not affect the sentimental value of the lemma even if it differs from the large margin to the head word chosen by the defined dictionary on the algorithm.

Also, the other processes like the stemming process or the tokenization process won't be affected to the word sentiment score based on if it comes prior to the lemmatization or after lemmatization because of the sentimental value won’t affect by the order of those processes.

Lexeme forms

Lemma

Breaking

matur

Figure 25 Lemmatization process

Lemmatization

# **Topic** **Model**

## **Pre-requisites**

To initiate any topic model it needs a proper input. in our research before move on to the topic model process we need two main inputs for our LDA topic model. There are the dictionary and corpus.

Corpus is a bag of words with a unique id and a frequency label created by the gensim library with the defined dictionary using the pre-processed data we given in earlier stages. when converting textual data to the bag of words library consider all the words are tokenized and normalized In the early stages. It can be exist in Unicode and utf8-encoded strings.no further pre-processing processes will be conducted after the tokenization, stemming, and lemmatizing stages. Dictionary is the module that encapsulates the mapping between pre-processed textual data with their integer ids. without a dictionary, corpus cant be processed so corpus depends on the dictionary.

TOPIC

MODEL

BOW

Figure 26 LDA topic model process

## **Model build**

Now the algorithm has the required input for the model but the algorithm needs a topic to count to the train the module. Depend on the model training environment we can decide the LDA model type. If the training environment has a low specification algorithm can run with gensim LDA model but if the training environment is capable of providing multiple cores algorithm can run with gensim LDA Multicore model.

Master process

Update

Result

Queue

Model

process

Input

Queue

Chunks

Input corpus iterator

Worker 1

Chunk

process

Figure 27 LDA Multicore process

In the multicore model master process consume one core of the training environment and all the worker consume given cores to their processes accordingly. in the master process chunks of the documents are divided and sent to the worker process and after worker processes are completed it master process collects them and updates the model process . LDA multicore module is relatively quick than LDA general model.

## **Model training**

### **Base model training**

There are several optional parameters that need to be set on the model when to initiate with the given topic count. There is the random state, update every, chunk size, passes, alpha, per word topics.

Chunk size is the number of documents used in the training chunk. chunk size is an integer and it's optional in the base model. With a multicore model, we can customize the chunk size for faster implementations. Passes also an optional integer type parameter which is the number of passes through the corpus in the training phase of the model. Update every is an integer and an optional parameter that defines the number of documents that will iterate on each update phase. If it is set to zero it will set to batch learning and if it set to one it will set to online iterative learning.

The random state is optional integer parameter use in the reproducibility which can be an object or a seed to the model. Per word topics are the Boolean valued parameter. When its true model computes the list of topics and sorted it based on the likelihood of the topics for each word in the text.it will also multiply the phi values by word count.

Alpha is an optional float type parameter which most of the time is set to auto state.it has two other states named as symmetric and asymmetric. Alpha is a hyperparameter that will be tune in post stages. Alpha controls the mixture of the topic in the document. If we increase the alpha document will have a minimum mixture of topics. If we decreased the document will have a maximum number of topics.

## **Model evaluation**

### **Perplexity Score**

Perplexity score is a statical score of the model’s prediction on the given sample. It evaluates the model’s log-likelihood on the given test sample.to get the best outcome of the model predictions perplexity score has to be the lowest. Accurate probabilistic models have a high likelihood with low perplexity. But predictive likelihood scores and human judgments are least co-related, in some scenarios it can be even anti-correlated. [36]

Shape

Description automatically generated with low confidence

Figure 28 Perplexity score

As figure 25 perplexity can be defined as the inverse probability of the given sample set which is normalized by the word count.[36]

### **Coherence Score**

Topic coherence score monitors the list of topics generated by the model and the informative rate of the topics. There are multiple methods to calculate the coherence score such as C\_v, C\_p, C\_uci, C\_umass, C\_npmi, C\_a. in this research study we are we are using C-v and C\_umasss.

C\_v score or the coefficient of variance is based on single set segmentation of the top words with measures that have an indirect confirmation approach to normalized pointwise mutual information (NPMI) and cosine similarity.[37][38]

In this research study except for the coefficient of variance score, we are calculating the UMass score to validate the accuracy of the topic model.it measures the words only with properties with preceding and succeeding types and required an order list of words. In other words, its confirmation measures are based on document co-occurrences units and single preceding segmentation with logarithmic conditional probability.[38]

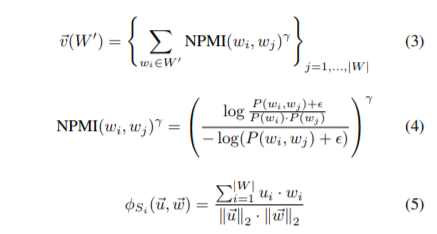
****

Figure 29 Umass calculation

## **Parameter tuning**

### **Model hyper parameters**

There are two types of parameters called model parameter and model hyper parameters. Model hyper parameters can be considered as machine learning algorithm settings that developers can tune before training. In our research study, it will be the number o topics in the model.

### **Model parameters**

The model parameter can be considered as the model learning factor in the real-time training environment. in our research studies, it can be weight assignments to the given topics.

## **Sensitivity tests**

In this research study, we develop the base model and train the base model using default parameters and default hyperparameter. in this sensitivity tests process conducting the test to determine the optimal hyper parameters.

We are conducting a sensitivity test as a series for the determination of the number of topics, Dirichlet alpha hyperparameter which is document topic density, and Dirichlet beta hyperparameter which is the word topic density. Sensitivity tests running with two different validation corpus sets and the metric to calculate the performance comparison of the test sets will be the coefficient of the variance.

The best coherence score is the last increment point of the graph before graph flattening or major drop. Alphas and beta values will be set according to those effects.

# **Commercialization**

This research component is functioning independently with other components of the learning platform which means the algorithm runs on the given backend and the cloud space file directory which will maintain the reviews templates, updated reviews, and review reports.

## **Multi-language support**

This research component has defined review language in every major step with nltk libraries, spacy libraries even In the pre-processing stages support both Unicode and utf8-encoding types which makes the main algorithm and helper algorithm to adapt to any defined language besides its default English language.

## **Multicore environment support**

With multicore model adaptability algorithm can switch between multicore high-performance environment or single-core low-performance environment. With this support algorithm can swift between gensim LDA model and gensim LDA multicore model to build and train the base model in the preferred environment.

## **Scalability and efficiency**

Besides the adaptability to performance orientation algorithm can swift between defining hyperparameter and parameter to make the multicore environments more productive with faster implementations. The algorithm offers defining workers according to existing cores of the training environment.

# **Testing & Implementation**

In this section, we will go through all algorithm implementations with an in-depth explanation.

### **Data collection**

As described in the earlier section data collection happens on cloud space through CSV templates which are located and maintain dedicated to individual students on their folders. every review folder and review template are following a general naming convention which will maintain the positive feedback culture.

If the months of the review collected are fixed with the module outline algorithm set to that naming pattern or months cant be fixed to an exact timeline algorithm will follow the numeric naming convention.



Table 5 Stable timeline month naming convention



Table 6 Unstable timeline month convention

### **Reading data**

To read data from specific URLs and defined locations we are using the panda library and io library in python and Jupyter notebook environment.

# Importing modules

import pandas as pd

import os

#Read data into papers

papers = pd.read\_csv(‘specified review cloud space url’)

papers.shape

# Print head

papers.head()

### **Data processing**

Since each iteration belongs to an individual student algorithm will drop columns and rows which are not associated with the review data.

# Remove the columns

papers = papers.drop(columns=['review id', 'target id', 'month'], axis=1)

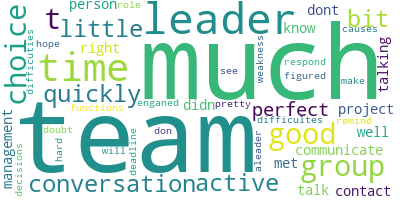
### **Pre-processing stage**

After pre-processing stage completes as described in the methodology we are comparing the processed feedback data with a word cloud to demonstrate algorithm is following best practices to achieve optimal pre-processed data.

Text

Description automatically generated

Figure 30 pre-processed sample with word cloud comparison



### **Sensitivity Tests and hyper parameter tuning**

In this section, we will go through the hyper-parameter tuning functions. As described in the methodology hyper parameter tuning is conducted by series of sensitivity tests. With sensitivity tests, we are setting up the optimal alpha value, beta value, coherence, and the topic count. This table 7 is belongs to the sample of the first validation set which is 75% of the corpus. topics count and alpha values are stable to two and 0.01 but both alpha value and coherence values are incrementing.



Table 7 Sensitivity test results part 1

In this table 8, its alpha value change to symmetric type from the auto but topics count are still stable. Corpus is the same as the previous validation set. But we can see the coherence drop with beta increment.

Table 8 Sensitivity test results part 2

In this table 9, we can see the alpha value again shift from symmetric type to asymmetric type. Still, the same topics count with the same validation set. The beta increment is the same as the previous. But we can see the coherence is unstable in this asymmetric alpha status. it has two coherence drops.

Table 9 Sensitivity test results part 3

Now we move on to the next validation set which is the full corpus sample. in this table 10, we can see symmetric alpha values with beta increment and coherence drop to the stable topic count.



Table 10 Sensitivity test results part 4

This second iteration of the full corpus sample has shifted its alpha values to asymmetric from the symmetric type and we can see a significant coherence drop to a beta increment and a stable topic count.



Table 11 Sensitivity test results part 5

In the previous iteration, we see topic counts stable with the symmetric and asymmetric count. But here its shifts to default alpha values with topic increment by one level and still beta is incrementing. In the full corpus sample, we can see significant coherence drops more often than in the first validation set.



Table 12 Sensitivity test results part 6

## **Results**

In this result section we are going through results related to all the models in the algorithm in this first segment we can see the topic generation f the LDA base model to given topic count, in this base model alpha values are in auto status.

|  |  |
| --- | --- |
| corpus=corpus, | chunksize=100, |
| id2word=id2word, | passes=10, |
| num\_topics=num\_topics, | alpha='auto', |
| random\_state=100, | per\_word\_topics=True |
| update\_every=1, |  |

Table 13 LDA base model parameters

[(0,

'0.024\*"much" + 0.024\*"leader" + 0.024\*"team" + 0.024\*"time" + '

'0.024\*"choice" + 0.024\*"talk" + 0.024\*"perfect" + 0.024\*"good" + '

'0.024\*"right" + 0.024\*"well"'),

(1,

'0.159\*"leader" + 0.159\*"much" + 0.083\*"role" + 0.083\*"perfect" + '

'0.083\*"pretty" + 0.083\*"active" + 0.083\*"good" + 0.008\*"team" + '

'0.008\*"choice" + 0.008\*"talk"'),

(2,

'0.090\*"weakness" + 0.090\*"aleader" + 0.090\*"perfect" + 0.090\*"see" + '

'0.090\*"doubt" + 0.090\*"choice" + 0.090\*"team" + 0.090\*"leader" + '

'0.008\*"much" + 0.008\*"talk"'),

(3,

'0.025\*"time" + 0.025\*"team" + 0.024\*"conversation" + 0.024\*"bit" + '

'0.024\*"hard" + 0.024\*"group" + 0.024\*"deadline" + 0.024\*"make" + '

'0.024\*"cause" + 0.024\*"decision"'),

(4,

'0.123\*"much" + 0.083\*"team" + 0.083\*"leader" + 0.083\*"talk" + 0.044\*"time" '

'+ 0.044\*"choice" + 0.044\*"good" + 0.044\*"group" + 0.044\*"know" + '

'0.044\*"person"'),

(5,

'0.024\*"leader" + 0.024\*"much" + 0.024\*"team" + 0.024\*"talk" + 0.024\*"time" '

'+ 0.024\*"perfect" + 0.024\*"right" + 0.024\*"choice" + 0.024\*"project" + '

'0.024\*"group"'),

(6,

'0.104\*"team" + 0.104\*"time" + 0.054\*"active" + 0.054\*"much" + '

'0.054\*"little" + 0.054\*"remind" + 0.054\*"decision" + 0.054\*"cause" + '

'0.054\*"make" + 0.054\*"deadline"'),

(7,

'0.109\*"quickly" + 0.057\*"hope" + 0.057\*"function" + 0.057\*"difficulte" + '

'0.057\*"conversation" + 0.057\*"contact" + 0.057\*"bit" + 0.057\*"difficutie" + '

'0.057\*"little" + 0.057\*"figure"'),

(8,

'0.024\*"leader" + 0.024\*"much" + 0.024\*"team" + 0.024\*"talk" + 0.024\*"good" '

'+ 0.024\*"choice" + 0.024\*"perfect" + 0.024\*"time" + 0.024\*"right" + '

'0.024\*"well"'),

(9,

'0.024\*"team" + 0.024\*"leader" + 0.024\*"much" + 0.024\*"talk" + '

'0.024\*"choice" + 0.024\*"time" + 0.024\*"perfect" + 0.024\*"good" + '

'0.024\*"little" + 0.024\*"right"')]

We are also calculating each model training duration.in this base model its training duration is :

--- 0.04599738121032715 seconds ---

After calculating the training duration we are calculating the coherence scores with cv and UMass metrics.

Graphical user interface, application, Word

Description automatically generated

Figure 31 Cv score calculation of the LDA base model

**Graphical user interface, text, application

Description automatically generated**

Figure 32 Umass calculation of the base model

As mention in the methodology, lower perplexity scores demonstrate the best algorithms in language models. After training time and coherence score calculations algorithm moves on to the data visualization stage with the PyLDAvis library. We will go through each topic counts intertopic distance maps with multidimensional scaling and the relevant token frequency maps with saliency and relevance metrics.

Saliency is defined as below. First We need computational probability. For a word w its computational probability which is P(T|w), the likelihood of the word w is generated by T latent topic. We also need marginal probability. Which is the likelihood of random-word(w`) generated by T latent topic.[39]

Distinctiveness is kullback-leibler divergence[40].to calculate the saliency we need distinctiveness measurements with computational and marginal probability values.

distinctiveness(w) = X T P(T|w) log P(T|w)/ P(T)

in information theories, we can describe it as generated topics information contribution to the document which it exists with its frequency. Saliency is needed to generate the term-topic matrix which we can see in the below figure.

saliency(w) = P(w) × distinctiveness(w)

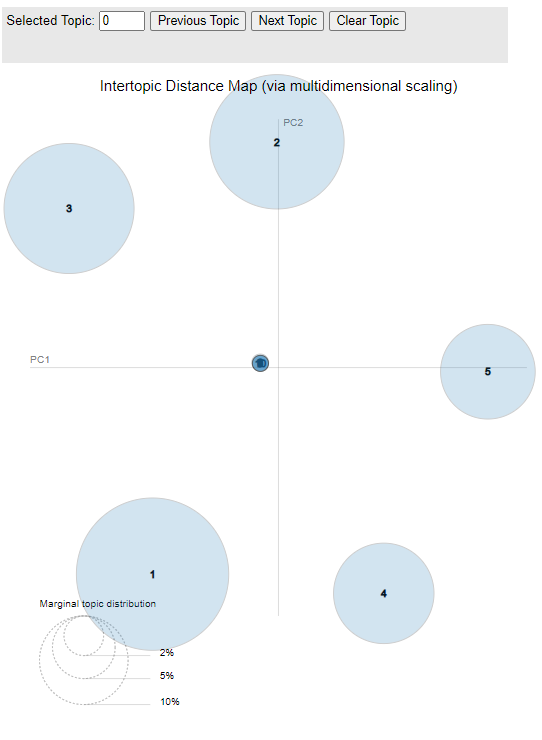
****In this figure 33, we have the initial intertopic map with none of the topic selected. Its

Figure 33 Initial intertopic distance map

demonstrate the intertopic dimensions before topics go through the map. As we can see estimated term frequency does not exist in the initial map.

Figure 34 Term-topic distribution

**Graphical user interface, application

Description automatically generated**

In this figure we can also see the estimated term frequency is not exist like on the map. Also term related to overall frequency only exist with its defined relevance score.

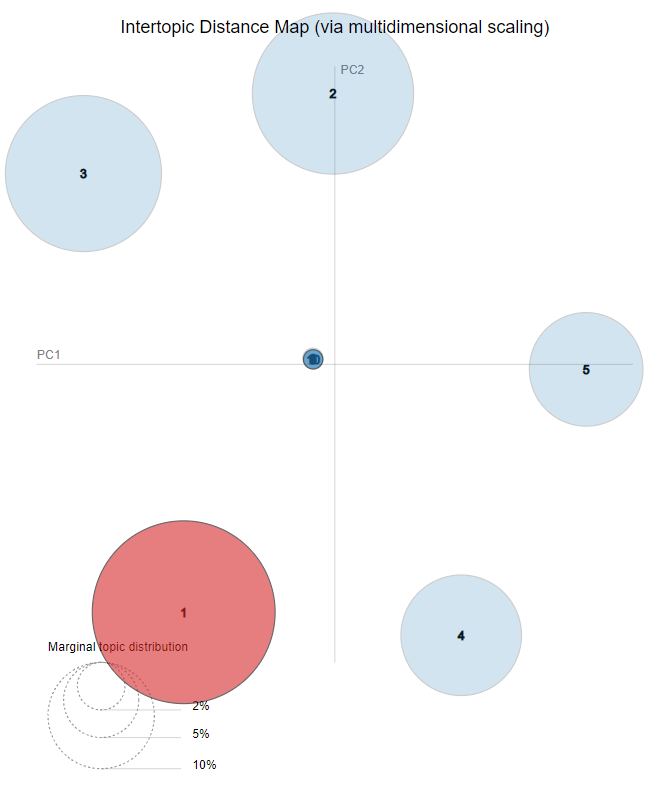
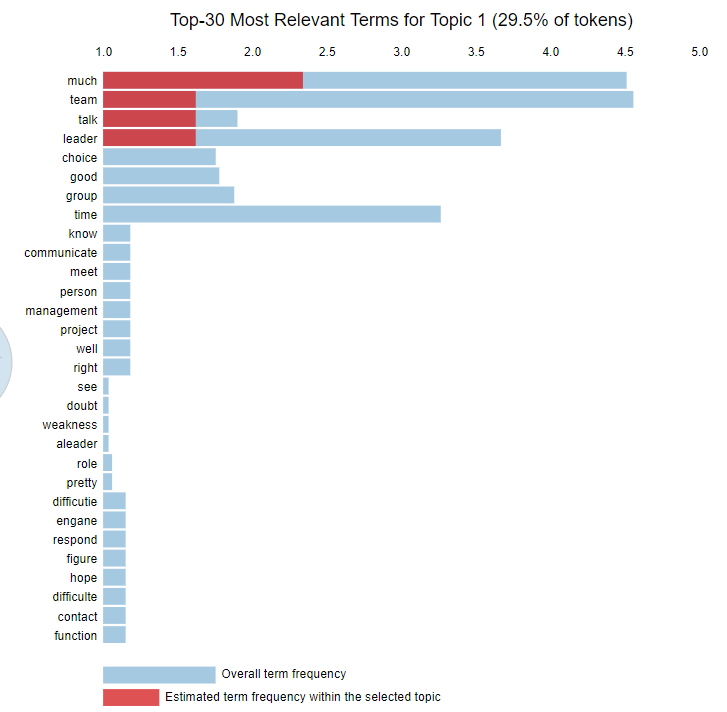
****

Figure 35 Intertopic distance map first iteration

In this figure 35, we can see the first iteration of the term-topic distribution. in this map, we have estimated term frequency indicates in the first word topic cluster. Like that term frequency will goes through each cluster according to their relevance and saliency scores.

****

*Figure 36 Term-topic distribution first iteration*

In this figure 36, this term frequency allocation it indicates the first topic estimated term frequency with the overall term frequency illustrations. This is a direct adaption from the map which both use pyLDAvis library.

**Text, letter

Description automatically generated**

This figure demonstrates the gensim LDA mallet topic generation pattern up to selected 3 topic variations it can set to prefer count .in the each topic cluster most probability calculated word will be represented as the topic according to its term-topic frequencies.

Its training time calculation is demonstrated below figure as same as the base model training time. It spends 3 more seconds than the base model which is not bg different when it comes to individual review processes.

--- 48.57359981536865 seconds ---

## **Research Findings**

In the research finding, we compare the base model and modified model to demonstrate setting up the optimal values makes the coherence more accurate. in the below calculations are testing results collected from the testing phase of the algorithm which proves mallet has lower coherence with gensim base model.

Mallet Coherence Score: 0.40945393726671275

Gensim Coherence Score: 0.44023133334642955

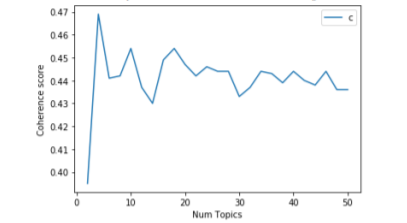
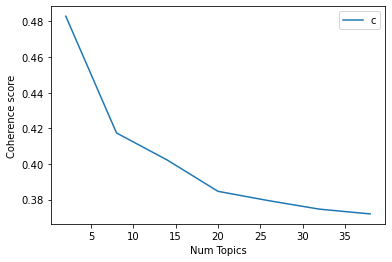
In data collection we get the reviews from both free-riding effects to maintain students and the active learning students. in the below comparison, we can see the free-riding effect affects to the review, and it inverted the graph which is supposed to increment and be stable within a certain level with iterations of up and down motions in the stabilized level. But in the free-riding graphs never have the up and downs between a stable level it completely decreased with more review context.

Figure 37 Altered coherence progression

Figure 38 Expected coherence progression

In this section, we are demonstrating the process of optimal model topic generation which is after base model and parameter tuning. This leads to getting the exact dominant topics of the review with tuned hyper parameters.

[(0,

'1.000\*"talk" + 0.000\*"decision" + 0.000\*"make" + 0.000\*"hope" + '

'0.000\*"little" + 0.000\*"quickly" + 0.000\*"respond" + 0.000\*"active" + '

'0.000\*"cause" + 0.000\*"deadline"'),

(1,

'0.500\*"decision" + 0.500\*"function" + 0.000\*"hope" + 0.000\*"little" + '

'0.000\*"quickly" + 0.000\*"respond" + 0.000\*"active" + 0.000\*"cause" + '

'0.000\*"deadline" + 0.000\*"figure"'),

(2,

'0.667\*"quickly" + 0.333\*"person" + 0.000\*"decision" + 0.000\*"make" + '

'0.000\*"hope" + 0.000\*"little" + 0.000\*"respond" + 0.000\*"active" + '

'0.000\*"cause" + 0.000\*"deadline"'),

(3,

'1.000\*"talk" + 0.000\*"decision" + 0.000\*"make" + 0.000\*"hope" + '

'0.000\*"little" + 0.000\*"quickly" + 0.000\*"respond" + 0.000\*"active" + '

'0.000\*"cause" + 0.000\*"deadline"'),

(4,

'1.000\*"team" + 0.000\*"hard" + 0.000\*"hope" + 0.000\*"little" + '

'0.000\*"quickly" + 0.000\*"respond" + 0.000\*"active" + 0.000\*"cause" + '

'0.000\*"deadline" + 0.000\*"decision"'),

(5,

'0.667\*"good" + 0.333\*"make" + 0.000\*"decision" + 0.000\*"hope" + '

'0.000\*"little" + 0.000\*"quickly" + 0.000\*"respond" + 0.000\*"active" + '

'0.000\*"cause" + 0.000\*"deadline"'),

'0.000\*"make" + 0.000\*"hope"'),

Graphical user interface, application

Description automatically generated

Table 14 Dominant topic coverage

After optimal model build and training algorithm generating the dominant topic coverage with its relevant keywords. This table has document numbers provided, dominant topic unit, topic contribution, topic percentage contribution, and the keywords for each contributed topic.

**Diagram

Description automatically generated with low confidence**

Table 15 Topic number coverage

After the dominant topic coverage algorithm demonstrates the topic numbers with topic percentage contribution according to their keyword for each topic contribution.

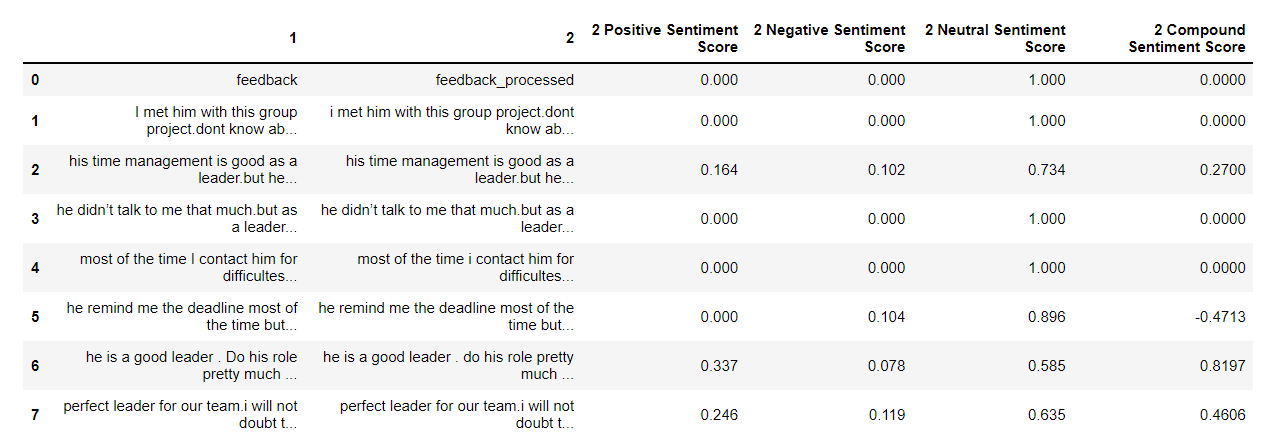
****

Table 16 Sentiment score coverage

After the main algorithm end with its coverage reports we algorithms helper function will generate the sentiment scores according to pre-processed data from the early stages to compare the sentiment scores with given students review scores. Below we listed the score variation that can be occur with the algorithm functionality.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Review score | Positive sentiment score | Neutral sentiment score | Negative sentiment score | Compound sentiment score |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Table 17 Basic review flow

## **Discussion**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Review score | Positive sentiment score | Neutral sentiment score | Negative sentiment score | Compound sentiment score |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Table 18 Intermediate review flow

The mixture of the various types of sentiment score is complex with the review complexity and the topic variation. Color intensity demonstrates the mixture of the each topic contribution in these tables. In some of the scenarios, actual human judgment can be differ from the sentiment values of the review such as neutral review can be biased into negative sentiment value based on the strong sentiments usage.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Review score | Positive sentiment score | Neutral sentiment score | Negative sentiment score | Compound sentiment score |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Table 19 Complex review flow

# **Conclusion**

With the demonstrate reports and results its clear algorithm is achieved the proposed objectives which are establishing the active learning environments in the outcome-based education with peer-review assessments to detect free-riding effects using coherence drop and the dominant topic coverage.

The algorithm proves hyperparameter tuned model is more accurate than a default base model with several calculations. Data collection also encourages the active learning environment with the cloud space engagements and the template and review maintenance.

The algorithm is not a fully automated decision-making process there are humans interrupting with the tuning process and decision-making process which will be the review component maintaining developer and the project team supervisors.

As discussed in the methodology human judgments are varied and differently co-related with the sentimental values which they are expressing. because of that after reports and results project team's active learning experience will depend on the decisions taken by their supervisor depends and based on the algorithm results of the reviews.

Depends on the client's environment we can adjust the performance metrics of the algorithm as discussed in the memorialization parts. with that functionality, we can expand the target audience for this component from primary-scale school to an industrial environment which can provide multicore environments to the algorithm model training.

There are complex backgrounds bind to the algorithm with all these machine learning approaches and the unsupervised learning methodology. Because of that before introducing this algorithm and its outcome to the academic it will need a brief introduction to the academic staff which will engage with the reports and results of the algorithm.

Some of the time students can try to copy-paste the same feedbacks to the all members of their team on every month which will make the algorithm results more biased to be neutral about that team and team members.it can detect through the main algorithm outcomes with topic contribution but like mentioned before there are lots of parameters and knowledge involve with this reports and algorithm results which need proper knowledge to achieve the best results through it.

Without proper knowledge about the outcomes and the results of the algorithm and its helper functions teams can be headed to more passive environments which is the opposite objective of the our research study. Proper knowledge transfer to the supervisor and the students are the most sensitive fact of this algorithmic infrastructure.

Some of the student vocabularies will be misunderstood by the algorithm sometimes because of the heavy sentiments valued word usage. Most of the time algorithm pre-processing will detect it and not going to affect most of the time. But in introvert situations, they will use less topic contribution based on their personality. That can lead to a coherence inversion of the topic contribution graphs and reports which will detect as a passive freeriding effect. The supervisor has the responsibility to understand the student's personalities and work through their review to provide the better feedback decisions based on algorithm results.

# **References**

[1] Vinichenko, M. V., Melnichuk, A. V., Kirillov, A. V., Makushkin, S. A., & Melnichuk, Y. A. (2016). Modern views on the gamification of business.

[2] Spady, W. G. (1994). Outcome-Based Education: Critical Issues and Answers. American Association of School Administrators.

[3] Barron, B. J., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., & Bransford, J. D. (1998). Doing with understanding: Lessons from research on problem-and project-based learning.

[4] Malan, S. P. T. (2000). The 'new paradigm' of outcomes-based education in perspective. Journal of Family Ecology and Consumer Sciences/Tydskrifvir Gesinsekologie en Verbruikerswetenskappe.

[5] Bhat, Shankaranarayana. (2018). Outcome Based Education - An Overview. Introduction to Outcome Based Education.

[6] Biggs, John B. and Catherine So-kum Tang. “Teaching For Quality Learning At University.”

[7] Implementing outcome-based education and student-centered learning in Afghan public universities: the current practices and challenges Rahmatullah Katawazai.

[8] Gandhi, Parag Omar. (2002). Outcomes Based Education.

[9] Macayan, Jonathan Veran. “Implementing Outcome-Based Education ( OBE ) Framework : Implications for Assessment of Students ’ Performance.” (2017).

[10] Sun, P. H., & Lee, S. Y. (2020). The importance and challenges of outcome-based

education – a case study in a private higher education institution. Malaysian Journal

of Learning and Instruction.

[11] Damit, M. A. A., Omar, M. K., & Puad, M. H. M. (2021). Issues and Challenges of Outcomebased Education (OBE) Implementation among Malaysian Vocational College Teachers.

[12] Student Project Performance Management System for Effective Final Year and Dissertation Projects Supervision Imed Romdhani, Martin Tawse, Safa Habibullah Edinburgh Napier University, School of Computing.

[13] Impact of Time Management Behaviors on Undergraduate Engineering Students’ Performance Richelle V. Adams and Erik Blair.

[14] Spremic, Mario, Zlatan Zmirak, and Krunoslav Kraljevic. "IT and business process performance management: Case study of ITIL implementation in finance service industry." ITI 2008-30th International Conference on Information Technology Interfaces. IEEE, 2008.

[15] Day, Jacqueline, and Milena Bobeva. "Applying performance management principles to a learning environment for undergraduate dissertations: A case study." International Journal for Quality and Standards 1.1 (2007).

[16] Kairuz, Therése, et al. "Consequences of KPIs and performance management in higher education." International Journal of Educational Management (2016).

[17] London, Manuel, and James W. Smither. "Feedback orientation, feedback culture, and the longitudinal performance management process." Human Resource Management Review 12.1 (2002).

[18] Dahling, J., & O’Malley, A. (2011). Supportive Feedback Environments Can Mend Broken Performance Management Systems. Industrial and Organizational Psychology, 4(2).

[19] Steelman, Lisa & Levy, Paul & Snell, Andrea. (2004). The Feedback Environment Scale: Construct Definition, Measurement, and Validation. Educational and Psychological Measurement - EDUC PSYCHOL MEAS. 64.

[20] Chauhan, Ganpat Singh, Preksha Agrawal, and Yogesh Kumar Meena. "Aspect-based sentiment analysis of students’ feedback to improve teaching–learning process." Information and Communication Technology for Intelligent Systems. Springer, Singapore, 2019. 259-266.

[21] P. D. Turney, “Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews,” in Proceedings of the 40th annual meeting on association for computational ling

[22] M. Fernandez-Gavilanes, T. ´ Alvarez-L ´ opez, J. Juncal-Mart ´ ´ınez,E. Costa-Montenegro, and F. J. Gonzalez-Casta ´ no, “Unsupervised ˜method for sentiment analysis in online texts,” Expert Systems with Applications, vol. 58, pp. 57–75, 2016.

[23] Nasim, Zarmeen, Quratulain Rajput, and Sajjad Haider. "Sentiment analysis of student feedback using machine learning and lexicon based approaches." 2017 international conference on research and innovation in information systems (ICRIIS). IEEE, 2017.

[24] Q. Rajput, S. Haider, and S. Ghani, “Lexicon-based sentiment analysis of teachers evaluation,” Applied Computational Intelligence and Soft Computing, vol. 2016, p. 1, 2016.

[25] ] L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, “Combining lexicon-based and learning-based methods for twitter sentiment analysis,” Technical Report, 2011.

[26] El Maaddawy, Tamer & Deneen, Christopher. (2017). Outcomes-Based Assessment and Learning: Trialling Change in a Postgraduate Civil Engineering Course. Journal of University Teaching and Learning Practice.

[27] Walaa Medhat, Ahmed Hassan, Hoda Korashy,Sentiment analysis algorithms and applications: A survey,Ain Shams Engineering Journal,Volume 5, Issue 4,2014,

[28] Kaur, Maninder and Akshay Girdhar. “A Framework for the Indirect Assessment Tool for Outcome Based Education Using Data Mining.” 2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC) (2018): 1-5.

[29] Maxey-Harris, Charlene & Dawes, Lorna. (2015). Building A Scholarly Network In Learning Communities at the University of Nebraska-Lincoln.

[30] W. Cheng and M. Warren, “Making a difference: using peers to assess

individual students’ contributions to a group project,” Teaching in

Higher Education, vol. 5

[31] H. H. af Segerstad, M. Helgesson, M. Ringborg, and L. Svedin,

”Problembaserade lärande. Idén, handledaren och gruppen.” Författarna

och Liber AB.

[32] ] Dommeyer, C. J. (2007). Using the diary method to deal with social loafers on the group project: Its effects on peer evaluations, group behavior, and attitudes. Journal of Marketing Education, 29, 175-188

[33] Aggarwal P and O’Brien CL (2008) Social loafing on group projects: Structural antecedents and effect on student satisfaction. Journal of Marketing Education 30(3): 255–64

[34] ] Fiechtner, S. B., & Davis, E. A. (1984-85). Why some groups fail: A survey of students' experiences with learning groups. Organizational Behavior Teaching Review, 9, 58-71.

[35] Bosworth, K., & Hamilton, J. (1994). Collaborative learning: Underlying processes and effective techniques. New Directions for Teaching and Learning, No. 59. San Francisco: Jossey-Bass

[36] Exploring popular topic models, Poonam Tijare and P Jhansi Rani 2020 J. Phys.: Conf. Ser. 1706 012171

[37] S. Syed and M. Spruit, "Full-Text or Abstract? Examining Topic Coherence Scores Using Latent Dirichlet Allocation," 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2017, pp. 165-174, doi: 10.1109/DSAA.2017.61.

[38] Röder, Michael et al. “Exploring the Space of Topic Coherence Measures.” Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (2015): n. pag.

[39] Termite: Visualization Techniques for Assessing Textual Topic ModelsJason Chuang, Christopher D. Manning, Jeffrey Heer Stanford University Computer Science Department

[40] S. Kullback and R. A. Leibler. On information and sufficiency.Annals of Mathematical Statistics, 22(1):79–86

# 9.**Appendices**