J. Empir. Soc. Sci. Stud. 7(1)

A Natural Language Processing Approach to Grouping Students by Shared Interests

Aravind Sasidharan Pillai

Principal Data Architect, Data Engineering, Cox Automotive Inc., USA, Master in Data Science, University of Illinois Urbana-Champaign.

Abstract

This research introduces an automated, Natural Language Processing (NLP)-based method for assembling students into groups based on shared interests, extracted from personal narratives. For the experiment, each student in the class was required to compose several stories, ranging from 300 to 400 words, to facilitate the extraction of common phrases. These phrases were then used to cluster students according to shared interests revealed in their personal stories. The study applied the Rapid Automatic Keyword Extraction (RAKE) algorithm, an unsupervised and language-agnostic technique for extracting keywords. This method is distinguished by its independence from specific linguistic structures, rendering it broadly applicable across various types of documents and fields. The RAKE algorithm operates through several distinct phases: The first phase involves removing all stopwords and phrase delimiters from the text. This step isolates potential key phrases within the narrative text. Contrary to the traditional use of TF-IDF (Term Frequency-Inverse Document Frequency) metrics, RAKE employs a keyword score-matrix based on Word Frequency, Word Degree, and the Degree to Frequency Ratio. In the final phase, RAKE identifies the highest-scoring phrases among the phrase candidates. These phrases, representing the document's most significant themes or topics, are then used as the basis for student grouping, capturing the core interests manifest in the narratives.

Keywords: keyword extraction, natural language processing, personal narratives, RAKE algorithm, shared interests, student clustering

Introduction

Interest-based grouping students by topics and subjects of interest, can be a powerful strategy in personalized learning. This approach aligns educational content and activities with the personal interests of each student, which can significantly enhance engagement, motivation, and ultimately, learning outcomes.

Personalized Learning presents a progressive shift from traditional models of education, which often cater to the average, rather than the individual. This modern approach tailors education to meet each student's unique needs, abilities, and interests at their own pace[1,2]. Rather than adhering to a rigid, uniform curriculum, personalized learning aims to provide customized educational experiences that adapt to learning styles and preferences. This method allows students to explore subjects in ways that are most effective for them, thereby enhancing engagement and enabling deeper understanding of material [3,4].

Learning strategies play a crucial role in this educational model, involving deliberate actions and mental efforts by learners to comprehend and retain information. These strategies are categorized into three types: metacognitive, cognitive, and socially affective. Metacognitive strategies involve planning, organizing, and evaluating one's own learning processes. Cognitive strategies, on the other hand, include active techniques such as taking notes, using academic resources effectively, and employing repetition for memorization. Socially affective strategies focus on collaborative learning, where students engage with peers or teachers to solve problems and share knowledge, fostering a communal learning environment [5].

The combination of personalized learning and strategic educational methods encourages students to take charge of their learning journey. By engaging with material that resonates with their personal interests and using strategies that match their learning style, students are more likely to achieve mastery in various subjects. This approach not only makes learning more relevant and exciting but also equips students with the skills necessary to thrive in an ever-changing world. It represents a shift towards a more inclusive, learner-centered paradigm in education that recognizes and respects individual differences.

Table 1. **key components of personalized learning:**	
Component	**Description**
Learner Profiles	Detailed records of students' strengths, weaknesses, preferences, and goals.
Personalized Learning Paths	Tailored educational paths that align with individual student profiles.
Competency-Based Progression	Students progress by demonstrating mastery of a subject, rather than time spent on content.
Flexible Learning Environments	Adapting physical or digital spaces to suit individual learning needs.
Adaptive Teaching Practices	Teachers use varied instructional strategies based on student needs.
Robust Data Systems	Data from assessments and activities is used to refine learning experiences.
Technology Integration	Use of educational technologies like LMS and adaptive learning software to support personalization.
Student Autonomy and Ownership	Students take control of their learning process by setting goals and choosing their activities.

Learning strategies are specialized approaches that students adopt to accomplish academic tasks and enhance their capabilities. These strategies help transform learners into independent, proactive participants within their educational environments, equipping them with valuable skills that are transferable to future professional contexts. Such strategic learning fosters a sense of autonomy and responsibility, enabling students to navigate and adapt to various learning challenges effectively. By integrating these strategies into daily study habits, students not only improve their immediate academic outcomes but also lay a strong foundation for lifelong learning and adaptability in their careers [6,7].

Personalized learning tailors educational experiences to align with individual student needs and prior experiences, significantly boosting motivation and engagement. This tailored approach optimizes learning opportunities, allowing each student to explore and excel at their own pace and in ways that resonate with their unique learning styles. As a result, personalized learning not only enhances academic performance but also ensures a more inclusive and adaptable educational system. Such a system acknowledges diverse learning needs and trajectories, thereby promoting an educational environment where every student has the opportunity to achieve their maximum potential. This concept of personalized learning encompasses various facets, including individual developmental trajectories and learning pathways, reflecting a holistic approach to education.

Implementing personalized learning in a school district requires significant changes in the classroom environment, with outcomes that can vary widely depending on numerous factors, particularly the attributes of the teachers involved. Characteristics such as a teacher’s experience, the extent and quality of their professional training,

their personal beliefs and background, as well as their knowledge of educational pedagogy and their satisfaction with their current job roles, all play critical roles in the success of this educational approach. These elements combine to either facilitate or hinder the adaptation to personalized learning models, highlighting the importance of teacher support and continuous professional development in the transition process.

In some schools, technology is employed as a tool to facilitate personalized learning, although it is not strictly necessary for such educational models. For instance, schools that follow the Montessori Method already incorporate key principles similar to personalized learning, such as promoting individualized learning paths and deeper cognitive skills without relying on digital tools. Where technology is used, it serves as a bridge to allow students to engage at their own pace and level, enhancing their learning through additional practice and mastery demonstrations. This not only makes use of students' familiarity with digital environments but also minimizes the learning curve for integrating technology into educational practices, enriching the learning experience and providing opportunities for more personalized interaction between students and teachers [8].

The integration of technology into personalized learning environments enables a richer, more diverse approach to education [9]. Students have access to various digital resources, including audiobooks, multilingual texts, writing aids, and educational games, all tailored to their learning needs. This method encourages students to collaborate, share ideas, and seek help from peers and online resources, fostering a sense of community and collective problem-solving. By blending traditional teaching methods with digital tools, personalized learning becomes a comprehensive approach that not only addresses diverse educational needs but also enhances the overall delivery and effectiveness of content, ensuring a more inclusive and adaptable educational system.

Proposed framework

Interest-Based Clustering, through grouping students by topics and subjects of interest, is a strategy in personalized learning. This approach aligns educational content and activities with the personal interests of each student, which can significantly enhance engagement, motivation, and ultimately, learning outcomes.

Table 2. **key aspects of implementing Interest-Based Clustering:**	
Step	**Description**
Identify Student Interests	Use surveys, interviews, and informal conversations to gather data on students' favorite subjects, hobbies, career goals, and areas of curiosity. This forms the basis for tailoring the curriculum.
Curriculum Design	Develop educational units that are adaptable to include elements catering to various interests. For instance, a project on environmental science could incorporate perspectives related to technology, politics, art, or economics.
Flexible Grouping	Organize students into groups based on their interests, with the flexibility to adjust these groups as interests change over time. This approach prevents stagnation and promotes continuous engagement.
Incorporate Technology	Use educational technology tools like online platforms, digital libraries, and apps to provide personalized learning experiences and access to a diverse range of resources.
Facilitate Peer Learning	Encourage collaboration within interest groups to enhance learning through the sharing of ideas and perspectives, deepening understanding of the subject matter.
Assessment and Reflection	Include assessments and reflective activities that relate the educational material back to students' interests, helping to reinforce the relevance and encouraging self-assessment and goal setting.

Why Interest-Based Clustering Works

When students are learning about topics they're genuinely interested in, their engagement levels naturally rise. Engaged students are more likely to participate actively in class discussions, delve deeper into subject matter, and pursue additional learning opportunities outside the classroom. Learning about preferred subjects can serve as a strong motivational factor. Students often exhibit a greater willingness to tackle challenging tasks and persist through difficulties when they're motivated by interest. Interest-driven learning has been linked to improved comprehension, retention, and application of knowledge. Students are more likely to integrate new information with their existing knowledge when they find the material relevant and engaging. Allowing students to choose learning paths based on their interests supports the development of autonomy and self-directed learning skills. This can lead to lifelong learning habits, as students become more adept at seeking out information and learning opportunities on their own.

Interest-Based Clustering is a teaching strategy designed to enhance student engagement and personalize learning by aligning educational activities with individual student interests. The first step in implementing this approach involves identifying these interests. Teachers can utilize various methods such as surveys,

interviews, and casual conversations to gather comprehensive information about what students are passionate about. These may range from favorite academic subjects to hobbies and future career aspirations. Understanding these interests helps in creating a curriculum that resonates more deeply with students, making learning more relevant and engaging.

Once student interests are identified, the next step is curriculum design. This involves developing educational units, projects, and assignments that can be adapted to include elements related to different student interests. For instance, a unit on environmental science could be designed to appeal to students with diverse inclinations—such as those interested in technology, who might focus on innovative conservation techniques, or those drawn to politics, who could explore environmental law and policy. This customization ensures that students are not just passive recipients of information, but active participants who see the value of their learning in real-world contexts.

The third component of Interest-Based Clustering is flexible grouping. This strategy involves organizing students into groups based on their current interests, with the flexibility to reorganize these groups as interests evolve or new ones emerge. Such flexibility helps prevent the stagnation that can occur in static groups and encourages continuous exploration and engagement among students. It also facilitates the dynamic adaptation of group activities to better meet the changing needs and interests of the student body.

Incorporating technology plays a crucial role in Interest-Based Clustering. Educational technologies, such as online platforms, digital libraries, and specialized apps, can provide access to a wide array of resources tailored to different interests. These tools enable students to pursue personalized learning pathways at their own pace, extending learning opportunities beyond the traditional classroom setting. Technology not only diversifies the educational resources available but also supports the varied learning styles and preferences that exist within a classroom. Interest-Based Clustering involves facilitating peer learning and incorporating assessment and reflection practices that tie learning back to student interests. Encouraging students to work together within interest groups fosters an environment where they can share knowledge and perspectives, enriching the learning experience. Meanwhile, using interest-aligned assessments and reflection activities helps students connect their academic achievements to personal goals and interests, reinforcing the relevance of what they learn and promoting self-motivation. Despite

its challenges, such as the need for flexible curriculum design and the management of diverse educational activities, the benefits of increased engagement and personalized learning are substantial, making Interest-Based Clustering a valuable strategy in modern education.

Methods

This research proposed an automated NLP based approach to group students based on common interests extracted personal stories. For experiment each student in a class were tasked to write several 300-400 words stories in order to extract common phrases from the stories to group the students based on the common phrases/interests in their personal stories.

The study employed the RAKE (Rapid Automatic Keyword Extraction) algorithm, an unsupervised, language-neutral method for keyword extraction [10,11]. This technique does not rely on any specific linguistic structure and is universally applicable across various document types and domains. The RAKE algorithm was operated through a series of distinct steps [10,12,13]:

- **Selection of Candidate Phrases**: The initial step involved stripping the text of all stopwords and phrase delimiters. This process isolates potential key phrases from the stories text, preparing the data for further analysis.
- **Computation of Feature Values**: Instead of leveraging the commonly used TF-IDF (Term Frequency-Inverse Document Frequency) metrics, RAKE calculates a keyword score-matrix. This matrix is built upon **Word Frequency**, **Word Degree**, and **Degree to Frequency Ratio**.
- **Keyword Identification**: The concluding phase of the RAKE algorithm involves selecting the highest scoring phrases from the list of candidate phrases. These phrases are deemed the document's keywords, encapsulating the core topics or themes represented in the text.

Once the grouping is complete, the list of clusters/groups with their member's IDs will be stored in database. The following diagram shows the architecture of the proposed model in this study.

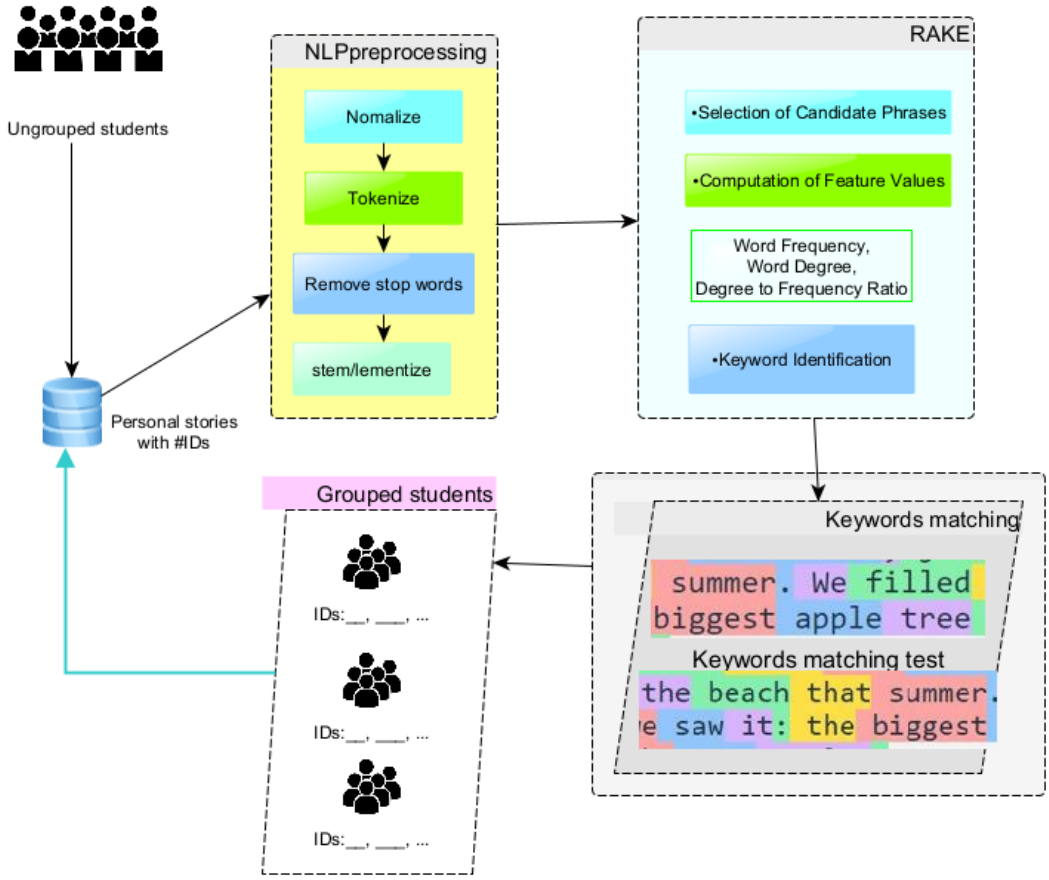
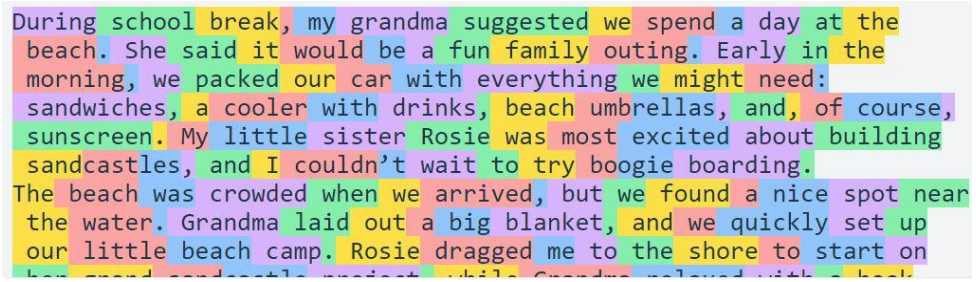


Figure 1. Architecture of the proposed framework (source: Author)

Preprocessing and results



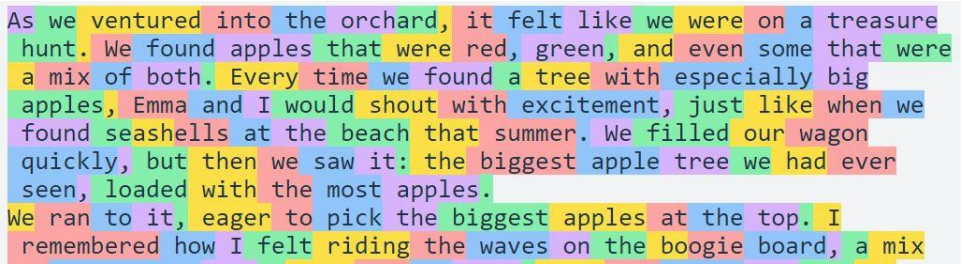
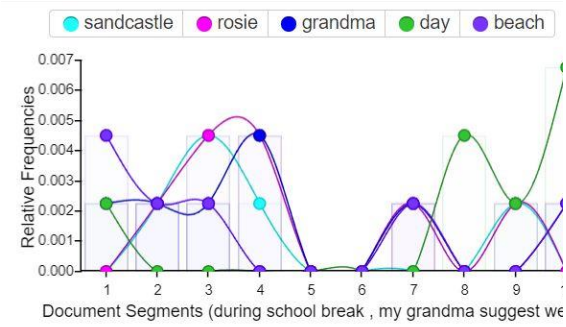


Figure 2. Tokenization

Table 3. After Text Cleaning and spaCy Word Lemmatize	
Story ID@A01. After Text Cleaning and spaCy Word Lemmatize	Story ID@B03 After Text Cleaning and spaCy Word Lemmatize
during school break , my grandma suggest we spend a day at the beach . she say it would be a fun family outing . early in the morning , we pack our car with everything we might need : sandwich , a cooler with drink , beach umbrella , and , of course , sunscreen . my little sister Rosie be most excited about build sandcastle , and I could not wait to try boogie boarding . the beach be crowd when we arrive , but we find a nice spot near the water . Grandma lay out a big blanket , and we quickly set up our little beach camp . Rosie drag I to the shore to start on her grand sandcastle project , while Grandma relax with a book under the	last fall , my uncle decide we should all go apple pick at a farm outside the city . he say it be go to be an adventure , and he be not wrong . my cousin Emma and I be excited because we 'd never go apple pick before . we imagine come home with bag full of apple and then bake pie and make cider . the farm be huge , with row upon row of apple tree . the air be crisp , and you could smell the apple . it remind I of the fresh , clean smell of the beach in the morning , just without the sea . we get a little wagon to haul our apple , and the farm owner give we a map , show we where the different kind

umbrella . build sandcastle with Rosie be more fun than I expect . we build a huge castle with wall and tower . Rosie even find some seashell to decorate it . she be so proud of our creation ; she make Grandma come and see it . Grandma say it be the good sandcastle she 'd ever see , and Rosie beam with pride . after a while , I decide to try boogie boarding . the wave be perfect , not too big and not too small . at first , I keep fall off , and I swallow a lot of seawater , which be gross . but after a few try , I start to get the hang of it . ride a wave all the way to the shore be the good feeling . I feel like I be fly . Lunchtime come , and we be all starve . the sandwich taste extra delicious with the salty air and the sound of the wave . after lunch , Grandma surprise we with her famous chocolate chip cookie . they be a bit melt from the heat , but that make they even well . in the afternoon , Rosie and I explore the beach , collect the pretty seashell we could find . we even spot a crab scuttle along the shore and follow it until it disappear into a hole in the sand . we decide that be enough excitement for one day . as the day come to an end , we pack up our thing , tired but happy . before leave , we take a family picture with the sunset behind we , a beautiful end to a perfect day . the drive home be quiet . Rosie fall asleep almost immediately , clutch a seashell in her hand . I look out the window , think about the wave and the sandcastle . it be one of those day you wish could last forever . Grandma say we should do it again soon , and I could not agree more . it be the good day of the summer , a simple	of apple be . Emma and I decide we would find the big , juiciest apple on the farm . as we venture into the orchard , it feel like we be on a treasure hunt . we find apple that be red , green , and even some that be a mix of both . every time we find a tree with especially big apple , Emma and I would shout with excitement , just like when we find seashell at the beach that summer . we fill our wagon quickly , but then we see it : the big apple tree we have ever see , load with the most apple . we run to it , eager to pick the big apple at the top . I remember how I feel ride the wave on the boogie board , a mix of excitement and a tiny bit of fear . climbing tree be not my strong suit , but Emma be fearless . she remind I of Rosie with her sandcastle , determined and proud . Emma climb up , reach for the big apple , while I hold onto the ladder , my heart racing . suddenly , Emma shout , " I get it ! " but as she do , the branch she be on give way , and down she come , apple and all . it be a scary moment , and it feel like everything be happen in slow motion . but thankfully , she land on some hay the farm owner have put around the tree . she be a bit shaken but laugh it off , hold the big apple triumphantly above her head . we make our way back , our wagon now overflow with apple . that evening , our family make apple pie , apple sauce , and even try our hand at make cider . it be a day full of adventure , laughter , and , most importantly , family . just like our day at the beach , apple picking become a cherished memory , a story we 'd tell

day at the beach that turn into a treasure trove of memory .	over and over , especially the part about Emma 's big fall . it be a reminder that sometimes the good adventure be the simple one you share with people you love
Readability Index: 2.621 Average Words Per Sentence: 44.5 Most frequent words in the corpus: rosie (7); grandma (7); day (7); beach (6); sandcastle (5)	Readability Index: 2.053 Average Words Per Sentence: 91.2 Most frequent words in the corpus: apple (20); emma (7); big (7); tree (5); farm (5)



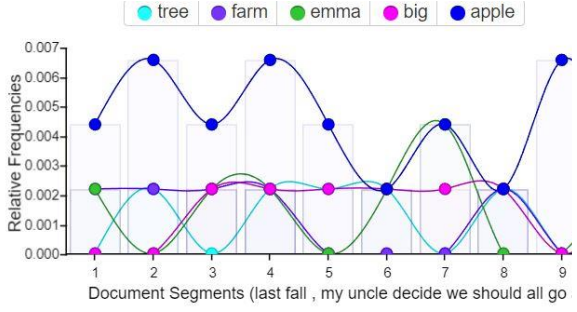


Figure 3(a). Frequent words in story 1.

Figure 3(b). Frequent words in story 1.

The shared keywords between this story and the previous ones are "**beach**," "**adventure**," "**excitement**," and "**family**." These themes are woven throughout both narratives, creating connections through the joy of discovery and the value of spending time with loved ones.

During school break, my grandma suggested we spend a day at the beach. She said it would be a fun family outing. Early in the morning, we packed our car with everything we might need: sandwiches, a cooler with drinks, beach umbrellas, and, of course, sunscreen. My little sister Rosie was most excited about building sandcastles, and I couldn't wait to try boogie boarding.

The beach was crowded when we arrived, but we found a nice spot near the water. Grandma laid out a big blanket, and we quickly set up our little beach camp. Rosie dragged me to the shore to start on her grand sandcastle project, while Grandma relaxed with a book under the umbrella.

Building sandcastles with Rosie was more fun than I expected. We built a huge castle with walls and towers. Rosie even found some seashells to decorate it. She was so proud of our creation; she made Grandma come and see it. Grandma said it was the best sandcastle she'd ever seen, and Rosie beamed with pride.

After a while, I decided to try boogie boarding. The waves were perfect, not too big and not too small. At first, I kept falling off, and I swallowed a lot of seawater, which was gross. But after a few tries, I started to get the hang of it. Riding a wave all the way to the shore was the best feeling. I felt like I was flying.

Lunchtime came, and we were all starving. The sandwiches tasted extra delicious with the salty air and the sound of the waves. After lunch, Grandma surprised us with her famous chocolate chip cookies. They were a bit melted from the heat, but that made them even better.

In the afternoon, Rosie and I explored the beach, collecting the prettiest seashells we could find. We even spotted a crab scuttling along the shore and followed it until it disappeared into a hole in the sand.

We decided that was enough excitement for one day. As the day came to an end, we packed up our things, tired but happy. Before leaving, we took a family picture with the sunset behind us, a beautiful end to a perfect day. The drive home was quiet. Rosie fell asleep almost immediately, clutching a seashell in her hand. I looked out the window, thinking about the waves and the sandcastles. It was one of those days you wish could last forever. Grandma said we should do it again soon, and I couldn't agree more. It was the best day of the summer, a simple day at the beach that turned into a treasure trove of memories.

Figure 4(a). Story ID@A01 keypharses

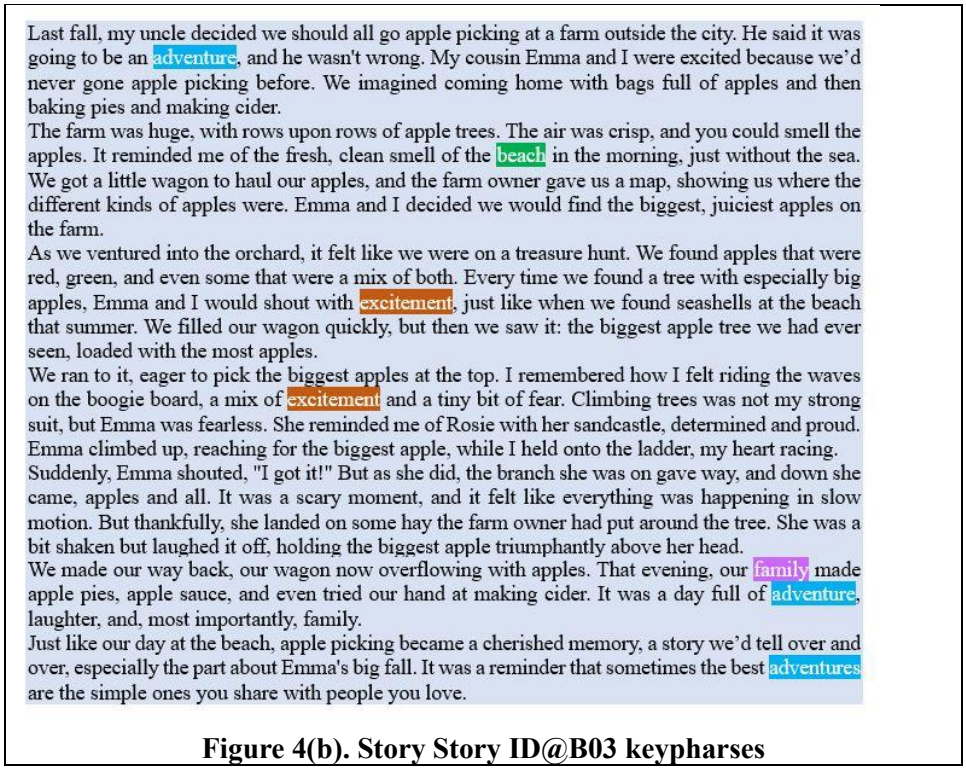

Last fall, my uncle decided we should all go apple picking at a farm outside the city. He said it was going to be an adventure, and he wasn't wrong. My cousin Emma and I were excited because we'd never gone apple picking before. We imagined coming home with bags full of apples and then baking pies and making cider.
The farm was huge, with rows upon rows of apple trees. The air was crisp, and you could smell the apples. It reminded me of the fresh, clean smell of the beach in the morning, just without the sea. We got a little wagon to haul our apples, and the farm owner gave us a map, showing us where the different kinds of apples were. Emma and I decided we would find the biggest, juiciest apples on the farm.
As we ventured into the orchard, it felt like we were on a treasure hunt. We found apples that were red, green, and even some that were a mix of both. Every time we found a tree with especially big apples, Emma and I would shout with excitement, just like when we found seashells at the beach that summer. We filled our wagon quickly, but then we saw it: the biggest apple tree we had ever seen, loaded with the most apples.
We ran to it, eager to pick the biggest apples at the top. I remembered how I felt riding the waves on the boogie board, a mix of excitement and a tiny bit of fear. Climbing trees was not my strong suit, but Emma was fearless. She reminded me of Rosie with her sandcastle, determined and proud. Emma climbed up, reaching for the biggest apple, while I held onto the ladder, my heart racing.
Suddenly, Emma shouted, "I got it!" But as she did, the branch she was on gave way, and down she came, apples and all. It was a scary moment, and it felt like everything was happening in slow motion. But thankfully, she landed on some hay the farm owner had put around the tree. She was a bit shaken but laughed it off, holding the biggest apple triumphantly above her head.
We made our way back, our wagon now overflowing with apples. That evening, our family made apple pies, apple sauce, and even tried our hand at making cider. It was a day full of adventure, laughter, and, most importantly, family.
Just like our day at the beach, apple picking became a cherished memory, a story we'd tell over and over, especially the part about Emma's big fall. It was a reminder that sometimes the best adventures are the simple ones you share with people you love.

Figure 4(b). Story Story ID@B03 keypharses

Conclusion

The study adapts the Rapid Automatic Keyword Extraction (RAKE) algorithm for a novel application—grouping students based on the thematic content of their personal narratives. The research demonstrates the feasibility of using an unsupervised, language-agnostic method in educational settings. This is particularly valuable as it suggests that the technique can be applied without the need for extensive training data or language-specific adjustments, making it adaptable to diverse educational contexts and student populations.

The study contributes to the pedagogical methods by potentially increasing engagement in collaborative learning environments by grouping students based on shared interests identified through their own writings. This approach ensures that student groupings are organically aligned with their interests, which can improve collaboration and outcome in group projects.

The research highlights a method that could streamline the administrative aspects of forming student groups in educational programs. Educational institutions can save time and resources, reducing the subjective bias of manual group assignments by automating the grouping process.

The effectiveness of the grouping relies heavily on the quality and depth of the personal narratives provided by the students. If students write superficially or off-topic, the extracted keywords might not accurately represent their true interests, leading to less effective groupings. Although RAKE is described as language-agnostic, its performance can still be influenced by cultural and linguistic nuances in the text. The algorithm's ability to extract meaningful keywords might vary significantly across narratives written in different languages or by individuals from diverse cultural backgrounds, potentially affecting the homogeneity of the groupings. The unsupervised nature of the RAKE algorithm might lead to biases based on the frequency and co-occurrence of words, which do not necessarily correlate with the importance or relevance of the themes to the students' actual interests. This can skew the grouping process, favoring dominant but potentially less relevant themes. While the study demonstrates the application of RAKE in a classroom setting, scalability to larger groups or across multiple classes and institutions might present challenges. The computational resources and time required to process a large volume of narratives efficiently can be significant.

The unsupervised approach does not provide clear interpretative frameworks for why certain students are grouped together beyond the algorithmic output. This can make it difficult for educators to manually adjust or understand the basis of group formations without a deeper analysis of the underlying keyword connections. Students' interests can evolve over time, and a static grouping based on a single set of narratives might not reflect these changes. The algorithm's static snapshot approach does not account for the dynamic nature of student interests, potentially leading to outdated or irrelevant groupings as the course progresses.

References

1. Campbell, R. J., Robinson, W., Neelands, J., Hewston, R. & Mazzoli, L. PERSONALISED LEARNING: AMBIGUITIES IN THEORY AND PRACTICE. *British Journal of Educational Studies* **55**, 135–154 (2007).
2. Jéirveléi'k, S. Personalised learning? New insights into fostering learning capacity. *Schooling for Tomorrow Personalising Education* (2006).
3. Tolmie, E. Implementing personalised learning in New Zealand primary schools innovative learning environments. (2016).
4. Keamy, (ron) Kim & Nicholas, H. Personalised learning: Can governments guarantee diversity for individuals? *Int. J. Divers. Organ. Communities Nations* **7**, 137–146 (2007).
5. Holmes, W., Anastopoulou, S., Schaumburg, H. & Mavrikis, M. Technology-enhanced Personalised Learning: Untangling the Evidence. 112 (2018).
6. Miliband, D. Choice and voice in personalised learning. *Schooling for tomorrow: Personalising education* 21–30 (2006).
7. Prain, V. *et al.* Personalised learning: lessons to be learnt. *Br. Educ. Res. J.* **39**, 654–676 (2013).
8. Sebba, J., Brown, N., Steward, S., Galton, M. & James, M. An investigation of personalised learning approaches used by schools. *Nottingham: DfES* (2007).

9. Saxena, A. K. Evaluating the Regulatory and Policy Recommendations for Promoting Information Diversity in the Digital Age. *International Journal of Responsible Artificial Intelligence* **11**, 33–42 (2021).
10. Pay, T., Lucci, S. & Cox, J. L. An ensemble of automatic keyword extractors: TextRank, RAKE and TAKE. *Computación y Sistemas* (2019).
11. Anjali, S., Meera, N. M. & Thushara, M. G. A Graph based Approach for Keyword Extraction from Documents. in *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)* 1–4 (IEEE, 2019).
12. Thushara, M. G., Mownika, T. & Mangamuru, R. A Comparative Study on different Keyword Extraction Algorithms. in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* 969–973 (IEEE, 2019).
13. Rose, S., Engel, D., Cramer, N. & Cowley, W. Automatic keyword extraction from individual documents. in *Text Mining* 1–20 (John Wiley & Sons, Ltd, Chichester, UK, 2010).