ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

ELPM: EDUPRENEURIAL LEADERSHIP PREDICTION MODEL

Umamaheswari.D

Assistant Professor
Department of Computer Science, CHRIST University, Bangalore, India

Jaisree M K

II MSc. Data Science,
Department of Computer Science, CHRIST University, Bangalore, India

Madhumitha P

II MSc. Data Science
Department of Computer Science, CHRIST University, Bangalore, India

ABSTRACT

This study presents a novel approach Edupreneurial Leadership Prediction Model that uses a machine learning based framework which has been specifically designed for the identification of the leadership potential in the educational entrepreneurship. Most of the traditional methods use the concept of subjective judgment and prolonged observation. The proposed model leverages the data-driven techniques which can be used to enhance and improve the accuracy and the performance efficiency of the model. This model uses clustering algorithms to group similar candidate profiles and then applies a Random Forest classifier to predict the leadership qualities of an individual. To uncover the hidden patterns in the data association rule mining is used which also improves the feature selection and the predictive analytics of the model. The incorporation of VADER sentiment analysis to access the written reports leverages deep insights into a candidate's attitude and communication styles. This framework is built on Python and scikit-learn which includes modules for evaluating the educational backgrounds, entrepreneurial experience, personality traits and the contextual factors. This allows the advantage of customization across various educational institutions and universities thereby demonstrating a higher accuracy. The future enhancements include connecting the professional networking platforms with adaptive learning mechanisms for continuous improvement. This model supports informed decision making in the educational sector by transforming the leadership identification into a fast data-driven process.

Keywords— Edupreneurial Leadership, Machine Learning, Predictive Analytics, Random Forest, Clustering Algorithms, Data-Driven Techniques, Association Rule Mining, VADER Sentiment Analysis, Entrepreneurial Experience, Informed Decision Making, Personality Trait, Adaptive Learning, Leadership Identification.

I. INTRODUCTION

Education plays a pivotal role in shaping the minds of individuals and communities by fostering knowledge, skills, values and attitude which are necessary for both personal and societal advancements.

Student data can be useful for various educators, administration and the policymakers which can help in the formulation of new curriculum designs, resource distribution and personalized support thereby offering valuable insights. The usage of predictive analysis helps in enhancing this process detecting and forecasting the academic trends, finding the student at risk stage and also enabling proactive interventions which helps in improving the learning

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

outcomes. The proposed Edupreneurial Leadership Prediction Model (ELPM) web application helps in simplifying the decision-making process using applications such as Python and scikit-learn[1] based platform for the data analysis. This helps in accelerating the identification of leadership potential in educational domain, and also it supports data-informed strategies for innovation and investment.

Traditionally methods were mostly based on manual effort, but our ELPM automates the complex analytical tasks. Using the advanced algorithms and the efficient workflow, this system processes weeks of data and gives output in minutes. Data cleaning, statistical modelling, user-friendly visualizations are the key features of this module, this eliminates the need for technical expertise. This platform also generates detailed reports, growth opportunities, pinpointing leadership candidates and customized intervention strategies.

The ELPM enables the educators to focus on strategic initiatives rather than the data management process, thereby minimizing the administrative workloads. The real-time processing of this ensures the up-to-date insights for the timely decision-making process.

It employes various methods and enhances the learning outcomes through the usage of predictive leadership analytics and identifying the high-potential students in the educational entrepreneurship. Based on this information the institutions can tailor the mentorship, coursework and experiential learning to nurture and help grow these talents. This system also helps in identifying the systemic trends which affect the leadership development that allow curriculum improvements which benefit all the students.

The presence of continuous feedback mechanism measures the impact of the leadership program thereby facilitating the evidence-based adjustments needed for improvement. This dynamic approach helps in creating an adaptive learning environment and equipping the future educational entrepreneurs with the necessary skills and techniques to tackle and overcome the sector-based challenges. Ultimately, the ELPM focuses on optimizing the resource allocation, improving leadership competencies and the entrepreneurial preparedness.

II. RELATED WORKS

The Edupreneurial Leadership Prediction Model (ELPM) uses machine learning to find students with leadership potential in educational entrepreneurship. Unlike subjective methods, ELPM takes a data-driven approach. It analyzes data using techniques like sentiment analysis, classification, and clustering [1]. Built on scikit-learn, a Python machine learning library, it employs K-means clustering to group similar student profiles [3, 11] and uses Random Forest to predict leadership based on factors like academic performance and personality traits [5].

Using the Apriori algorithm for Association Rule Mining, ELPM reveals connections between traits. For example, "high self-confidence and entrepreneurial experience indicate leadership potential" [6]. This insight helps educators design mentorship programs. ELPM also applies VADER sentiment analysis to gauge students' views on entrepreneurship [7, 8]. This is supported by research on hybrid models that combine VADER with logistic regression [9].

ELPM reflects studies that highlight self-efficacy, creativity, and mindset in leadership development [12, 18]. Its flexible, reusable design satisfies the need for modern tools in higher education [15, 16] while following data privacy standards [12]. Future enhancements could involve real-time data from professional networks [15] or hybrid models that combine text and numerical data for better accuracy [9]. By reducing bias, ELPM helps institutions effectively identify and nurture future leaders.

III. PROPOSED METHODOLOGY

This section discusses the entire work process behind ELPM. It clearly explains how the tool works and integrates the data and specifies basic prerequisites for the application to perform.

A. System Requirements - Functional Specification

This section deals with all the actions the user should take for ELPM to perform.

a) User Authentication and Authorization

In the developed ELPM system, initially the users are required to create accounts, but none of the stored data is stored permanently in the database. The usage of this design is mainly used for prioritizing the security and the privacy by avoiding the long-term data storage in the databases. This helps in minimizing the risk of the associated data breaches. This system does not use any traditional user authentication or credential verification methods. Instead, the entire access is controlled through predefined user roles, such as Faculty users and General users. This ensures that each user can interact and access only the features which are relevant to their roles. This process simplifies and makes it user friendly while maintaining the accessibility and security.

b) Data Upload

Users can upload student performance data in .csv format into the system. The system must validate the submitted data; it does so by presenting the user with an error message if any other format data is uploaded and asking the user to ensure that the data uploaded complies with predetermined format and schema requirements.

Example Scenario:

Suppose a professor can upload data of different universities' student feedback regarding their interest in becoming business leaders or entrepreneurs in .csv format that includes student data, like age, gender, mindset, and interest. That application will process the data in seconds, while no sensitive information is stored for analysis.

Benefits: The Edupreneurial Leadership Model (ELPM) adheres strictly to the GDPR and the data privacy standards thereby using only non-sensitive data and anonymized data without retaining personal details. This makes sure that it uses secure data processing and limited access. The built system is well suitable for various applications such as academic research, educational policy development and institutional decision making where the privacy and the ethical handling of data is a must.

c) Data Visualization

The system will provide interactive graphical representations of student mindset data, like bar charts, etc., allowing users to customize their view of the data because it analyzes only categorical data.

Example Use Cases: A financial analyst utilizes bar charts to contrast the distribution of credit ratings (e.g., AAA, AA, A, BBB) across different industry sectors such as Technology, Healthcare, Energy, and Consumer Goods. Through these visualizations, the analysts can easily evaluate the risk profile and the creditworthiness of firms in various industries. For instance, take a sector having a large number of AAA-rated firms, e.g., healthcare, it can reflect a reduced investment risk, while in the other situation involving BBB-rater firms it can imply higher volatility. Through visualization and interpretation of the categorical data, the analysts will be able to make better decisions in regards to portfolio diversification, risk management tactics and sector allocation.

Future Enhancements: The application can be further improved to provide an export and save facility for visualizations in different formats like PDF, PNG or SVG. To enhance the analytical depth as well as usability of the application we can incorporate some more features like legends, axis labels, color schemes and comparing multiple data sets.

D) VADER Sentiment Analysis:

The system employs the VADER (Valence Aware Dictionary and eSentiment Reasoner) algorithm to scan text data from questionnaires and students' feedback. It identifies sentiment polarity (positive, neutral, negative) and emotional intensity, allowing users to grasp the general mindset and motivational level of students towards leadership and entrepreneurship.

Example Scenario: An education researcher applies the VADER sentiment analysis module to review open-ended survey feedback from students on whether or not they would be interested in a business startup. By detecting underlying sentiment trends by demographic, the researchers can better customize entrepreneurship programs and optimize student engagement activities.

e) Association Rule Mining:

The system employs association rule mining to reveal hidden associations between categorical variables in the data. It identifies strong rules and patterns, such as which sets of characteristics are likely to occur together with high potential for leadership, providing interpretability and feature selection.

Example Scenario: A college administrator employs association rule mining to identify patterns of traits (such as high self-confidence, involvement in groups and prior entrepreneurial experience) that professional staff accompanying students rated highly on leadership potential will have in cahoots with student performance and personality data. This will assist the professional staff in improving selection criteria for leadership development programs and in developing mentorship programs.

B. System Requirements - Non Functional Specification

This section deals with the system's capabilities and conditions that enable ELPM to function efficiently and hence perform effectively.

a) System Constraints:

To be fully accessible, ELPM must work with Windows, macOS, Linux, and on Chrome, Firefox, Safari, and Edge. Managing more data volume and more users can only be enhanced through scalability. Regular updates from Python and Streamlit will be immutable down the road. Keeping Streamlit and Python libraries updated is important to success, so as managing dependencies through virtual environments. Compatibility is also improved through consistent testing.

b) Performance:

The current data load times are less than three seconds for the initial data load, and the design is intended to return product quickly. Most standard analysis require less than 5–10 sec. after the data load, so a fast turnaround should never be an issue. The system can accommodate at least 100 concurrent users without any performance issues, assuring a seamless user experience. Although we currently have no contingency plan for high user loads (more than 100 users at the same time), we will implement one in the near future.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

c) Reusability

The modular design of the ELPM allows for separate development, and the components are reused. Documentation of APIs makes the use of extensions and integration easier. Reusable and well-documented code guarantees scalability and maintenance. Git makes functionality formal by allowing for module replacement.

D) Modifiability

Configuration files, interface-based programming, and a modular architecture are all essential in ensuring ELPM can be extended. Version control through Git and documentation will all work toward the implementation of new features as well as development and collaboration on changes to the platform.

e) Usability

The user interface we have developed has been based on accessibility to ensure controls are user-friendly, navigation is clear and seamless as well as content discovery is simple. We have large buttons, enabled keyboard navigation, and high-contrast colours to make opinions easy for all users - including disabled users. Where problems arise, our app will provide guiding error messages, in which we will get feedback from the user when we include a feedback section in the app so we can continually improve the app based on this feedback to ensure it continues to be a great experience for all users.

e) Reliability

Given the result of a redundant infrastructure design with multiple servers and load balancers. Our objective is to achieve high availability with 99.9% uptime. We will also utilize auto-scaling features so instances can be added and removed automatically according to traffic. We'll also implement proper error recovery features to recover gracefully where things do fail.

f) Maintainability

Use design patterns, clean code and coding standards to produce quality code. Make sure testing (unit testing, integration testing and end-to-end testing) is done to ensure code reliability. Documentation helps keep the codebase recognizable. CI/CD helps keep code reliable through automated testing and deployments, which reduces mistakes. All of this means forming a reliable and reusable codebase that can be used again and again.

C. Dataset Description

This section gives an overview of the primary dataset used in the ELPM system. The data is collected through a structured Google Form filled out by the university students. It includes details about their demographic profile, their exposure to entrepreneurial education, their views on edupreneurial leadership, institutional culture, individual innovativeness, and entrepreneurial mindset.

Table 1 Dataset attributes with their description

Section	**Variable Name**	**Type**	**Description**
Demographic & Background	Name	Text	Respondent's name (to be anonymized during analysis)
Demographic & Background	Age	Categorical	Age group: Below 18, 18–20, 21–23, 24–26, Above 26

Demographic & Background	Gender	Categorical	Gender identity
Demographic & Background	Program_of_Study	Categorical	Undergraduate / Postgraduate / M.Phil./Ph.D. / Other
Demographic & Background	Department_School	Text	Academic department or school
Demographic & Background	Year_of_Study	Categorical	First, Second, Third, Final Year, Other
Demographic & Background	Campus	Categorical	Campus location (e.g., Central, Kengeri, Pune Lavasa, etc.)
Demographic & Background	Entrepreneur_Courses	Categorical	Whether the student has taken entrepreneurship/innovation courses
Demographic & Background	Entrepreneur_Activities	Categorical	Involvement in entrepreneurial activities (Yes / No / Planning)
Demographic & Background	Leadership_Position	Categorical	Holds leadership roles at university (Yes / No)
Perception of Edupreneurial Leadership	Promotion_of_Entrepreneurial_Thinking	Likert (1–5)	Perceived promotion of entrepreneurial thinking by university
Perception of Edupreneurial Leadership	Frequency_of_Leader_Entrepreneurial_Actions	Likert (1–5)	How often faculty/admins engage in entrepreneurial initiatives
Perception of Edupreneurial Leadership	University_Support_for_Initiatives	Likert (1–5)	Perceived support for student-led projects
Perception of Edupreneurial Leadership	Leadership_Role_Expectations	Multi-choice	Roles university leadership should play: mentoring, resources, exposure, etc.
Impact on Educational Culture	Culture_Encourages_Innovation	Likert (1–5)	Whether university culture encourages creativity and innovation
Impact on Educational Culture	Leadership_Culture_Impact_Factors	Multi-choice	Actions by leadership affecting culture: seminars, funding, R&D, etc.
Impact on Educational Culture	Resource_Accessibility	Likert (1–5)	Accessibility to innovation resources (labs, incubators, funding, etc.)
Influence on Student Innovativeness	Feel_Inspired_to_Innovate	Likert (1–5)	Whether student feels inspired to create based on university experience
Influence on Student Innovativeness	Resources_Helping_Innovation	Multi-choice	Resources that aided innovation (mentors, labs, events, etc.)
Influence on Student Innovativeness	Faculty_Mentorship_Impact	Likert (1–5)	Impact of mentorship on student's ability to innovate
Entrepreneurial Mindset	Encouraged_Entrepreneurial_Thin	Likert (1–5)	Whether university helped shape entrepreneurial thinking

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

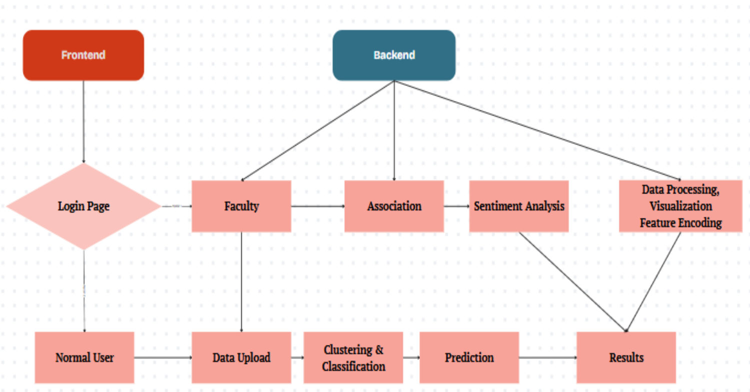
Development	king		
Entrepreneurial Mindset Development	Confidence_in_Pursuing_Entrepreneurship	Likert (1–5)	Student's self-confidence in starting an entrepreneurial career
Entrepreneurial Mindset Development	Influential_Programs_on_Mindset	Multi-choice	University programs shaping mindset: labs, courses, events, competitions
Entrepreneurial Mindset Development	Skills_Gained_From_Experience	Multi-choice	Entrepreneurship-related skills gained: problem-solving, networking, etc.
Open-Ended Responses	Suggestions_for_Improvement	Text	Student's input on how the university can enhance edupreneurial support
Open-Ended Responses	Personal_Entrepreneurial_Experience	Text	Descriptions of personal experiences shaping entrepreneurial mindset
Open-Ended Responses	Desired_Resources_or_Opportunities	Text	Student's suggestions for additional support in entrepreneurship

D. Block Diagram

Figure 1 provides a visual depiction of the block diagram showing the flow of the processing of the input data or uploaded data via ELPM. System has two essential components, Frontend and Backend.

A strong backend and an easy-to-use frontend build system architecture. The Frontend allows users to access the site by signing into a Login Page which authenticates them and directs them to interfaces suitable for their respective roles. Ordinary Users can use prediction tools in a User interface to assess leadership readiness, be able to see the cluster and classification results, upload data from instruments to and about students and leader, and review results, including downloadable reports and visualizations.

The functioning of the system is supported through links between modules connected to the user. The Association module maintains links between students, classes, and assessments; the Faculty Module manages instructor accounts and class construction. The Data Processing, Visualization, and Feature encoding Module represents the analytical heart of the system, taking raw data, preparing it for machine learning and generating insights, and ensuring that all components of the system are able to work seamlessly together for full end-to-end functionality.



ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

Fig. 1 ELPM Block Diagram

IV. RESULTS AND DISCUSSIONS

The section deals with a demonstration of ELPM's functionalities, mentioning all the features the tool has.

A. ELPM Sign Up Page

A complete and safe user registration process is the first point of access to functionality via the ELPM application. Users can create accounts on the Sign-Up Page while specifying their role (either Normal User or Faculty), entering a unique username, and establishing a password. Role-based access control, allows access to features permitting an application to provide differentiated functionality by providing users access to functionality according to their responsibilities.

Fig. 2 Sign Up Page of ELPM Application

B. Login Page

Once a user has successfully created an account, they will use the Login Page to gain access to the system securely and are authorized to access the ELPM platform through the Review and validation of the Users username, password, and role. The reason for only allowing verified Users with approved credentials to access the ELPM platform is to protect sensitive analysis modules, and the integrity of the data.

Fig. 3 Login Page of ELPM Application

C. ELPM Web Application Interface

The ELPM interface is intended to be easy to use. After login, the user will be taken to the dashboard, which has a number of tabs: Home, Clustering, Association Rule Mining, Sentiment Analysis, Classification, and Prediction. The navigation is intuitive, allowing users to move between tasks and more easily access the various analytical modules. This nurturing of the user experience and analytical efficiency worked in concert with each analytical activity's similarities and differences.

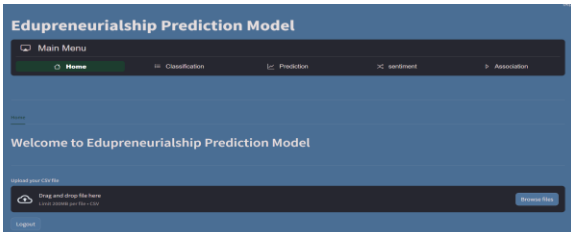


Fig. 4 Main Interface Page of ELPM Application

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

D. Data Upload module

The Data Upload module represents an important feature that plays a critical role in ensuring that only valid structured input is available to the analysis algorithms. It is noteworthy that the platform does not facilitate unsupported formats; for instance, it can only process CSV files containing categorical data. Before conducting an analysis, users are shown a preview of the data they uploaded and must accept them as valid for processing before headings and data are validated. Furthermore, schema validation verifies that the data being processed is in a provided categorical format (e.g., planned for the resulting aggregation) so subsequent analysis can operate on valid output.

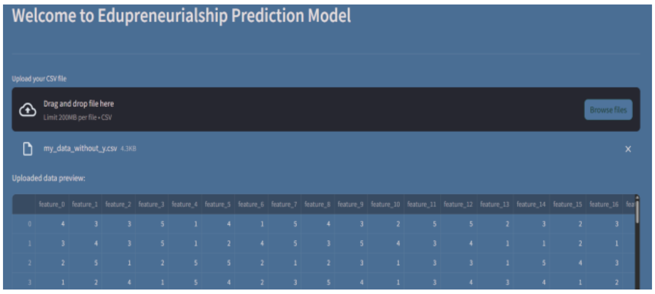


Fig. 5 Data Upload

E. Data Processing Module

When data are uploaded, the Data Processing Module will execute multiple preprocessing operations. One of the first operations will be to familiarize to users with the data by providing descriptive statistics which will include the mean, median, mode, variance, and standard deviation. The module will also consult pertaining to missing values, detecting and eliminating outliers, and encoding of categorical variables through proper transformation processes. Within the module's process there are also tools for feature selection, to assist with which categorical attributes are best options to ultimately investigate.

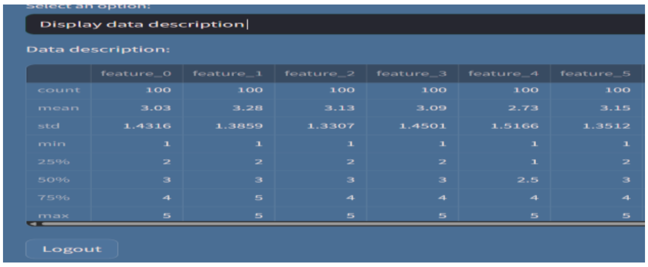


Fig. 6 Data Processing in ELPM Application

F. Clustering Module

The Clustering Module is the initial step in revealing the structure present in unlabeled categorical datasets. Categorical variables are have to be converted to numerical data for clustering via one-hot encoding. Potential methods such as the elbow method can be used to the decide the ideal number of clusters. Data points are then assigned to clusters iteratively, using K-means method of assigning clusters based upon distance to centroids; it will not stop assigning until the centroids of clusters converge. Although the clusters exist in high-dimensional space, they can be displayed in two-dimensional space by the transforms offered by techniques such as Principal Component Analysis-PCA and full cluster labels were applied to the original data for further use after interpretation of each cluster by analysis of the summary profiles of the most dominant categorical variables in each cluster.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

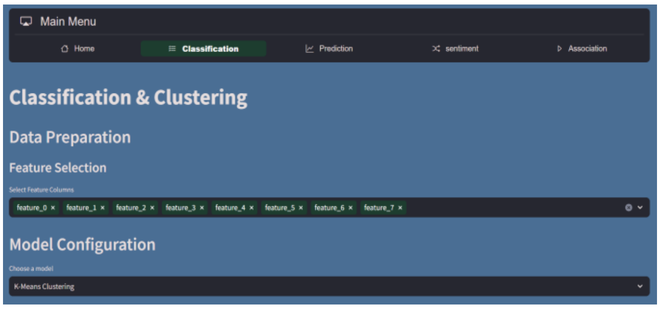


Fig. 7 Clustering Module in ELPM Application

G. Classification and Prediction Module

Next, after clustering, the resultant cluster labels are used as target variables for supervised learning by the ELPM system. The Classification [3] and Prediction Module includes feature engineering techniques such as interaction terms development, and convoluted category transformations. A set of decision trees, or Random Forest classifier[5], is created, each trained on different bootstrap sample and random subset of features as a model. Each model combined predictions together by majority vote, while the output of each model also includes meaningful decision rules that outline the traits of clusters. Furthermore, this module creates the opportunity for scaling and repurposing the trained model, producing predictions with unobserved data to predict cluster assignments.

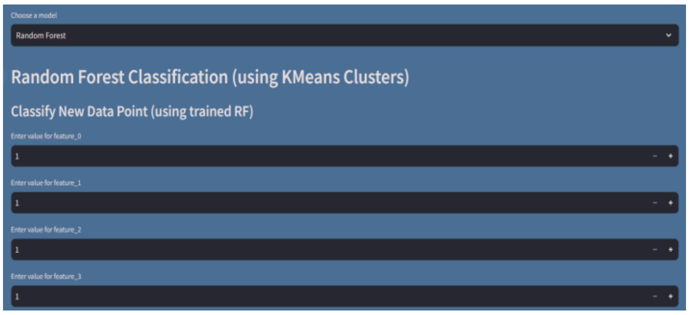


Fig. 8 Classification and Prediction Module in ELPM Application

H. Association Rule Mining

The Association Rule Mining module [6], which enhances the level of pattern discovery in categorical datasets is only available for faculty members. Access to the Association Rule Mining module is based on verifying credentials and restricting access. The Association Rule Mining module discovers frequently occurring itemsets and creates association rules utilizing the Apriori algorithm. Users can modify attributes such as lift, confidence, and minimum support to improve the rule-generating process.

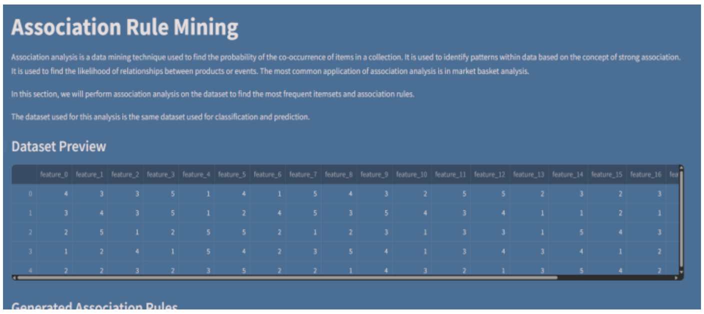


Fig. 9 Association Rule Mining in ELPM Application

Rule Filtering: Options to sort and filter rules based on metrics (support, confidence, lift)

	antecedents	consequents	support	confidence	lift
0	frozenset({'feature_17_4'})	frozenset({'feature_0_2'})	0.05	0.5556	3.268
1	frozenset({'feature_7_1'})	frozenset({'feature_0_4'})	0.11	0.5	2.2727
2	frozenset({'feature_0_4'})	frozenset({'feature_7_1'})	0.11	0.5	2.2727
3	frozenset({'feature_15_4'})	frozenset({'feature_0_4'})	0.07	0.5	2.2727
4	frozenset({'feature_1_1'})	frozenset({'feature_7_5'})	0.07	0.5385	1.7949
5	frozenset({'feature_1_1'})	frozenset({'feature_11_4'})	0.07	0.5385	1.9231
6	frozenset({'feature_14_4'})	frozenset({'feature_1_3'})	0.07	0.5	2.2727
7	frozenset({'feature_2_3'})	frozenset({'feature_4_1'})	0.11	0.5789	1.8676
8	frozenset({'feature_2_3'})	frozenset({'feature_15_2'})	0.1	0.5263	2.0243
9	frozenset({'feature_3_1'})	frozenset({'feature_2_4'})	0.11	0.5238	2.0952

Fig. 10 Image of Association Rule Mining Analysis

Users can export discovered rules in a number of formats for external review, and the module provides filtering rules in part based on these metrics. Because of this element, faculty are able to draw insights while maintaining analytic control.

I. Sentiment Analysis:

The primary function of the ELPM Sentiment Analysis[7] module is to assess the qualitative characteristics of student comments. The module is displayed in Fig. 3.9 which uses the VADER (Valence Aware Dictionary and Sentiment Reasoner)[8] sentiment analysis program that is vocabulary and rule based and was developed for short text types and social media.

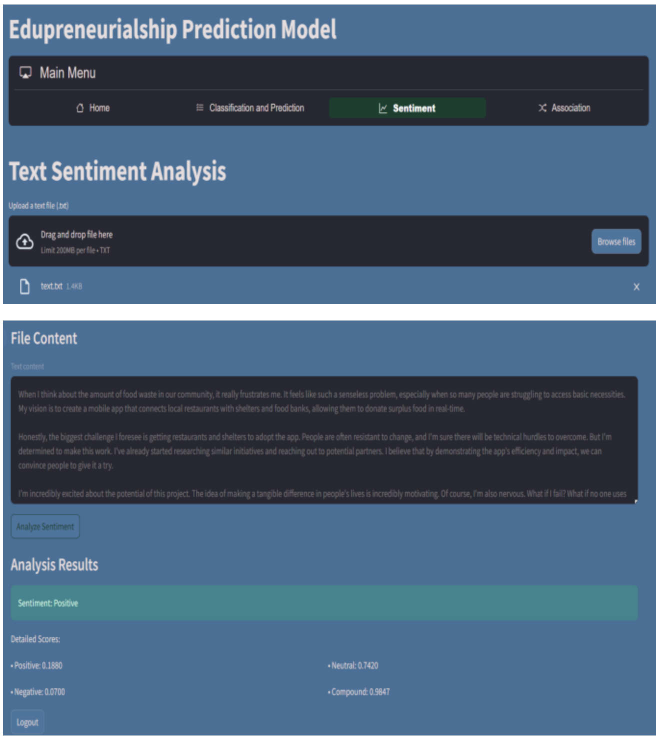


Fig. 11 Image of Sentiment Analysis Result

This module uses text data collected from open-ended comments in the leadership surveys and will categorize student sentiment as positive, neutral, or negative. The use of technology to analyze sentiment orientation allows researchers and teachers to assess students' positive or negative feeling toward a leadership topic, adding another variable to the prediction structure.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

J. Performance Metrics

In the ELPM application's VADER Sentiment Analysis[9], the textual responses are evaluated with a rule-based model that provides sentiment polarity scores based on lexical choices and the intensity of context. VADER uses four performance measures to evaluate its scores: positive, negative, neutral, and compound scores. The compound score is the only measure used to assess overall sentiment and is calculated based on the normalized sum of valence scores of each word in the input text. The Compound Score is between -1 (extremely negative) to +1 (extremely positive).

The formula used for calculating the compound score is:

compound = normalized(sum of valence scores of each token) in Fig. 12.

Normalization is done using a standard logistic function.

$$\frac{x}{\sqrt{x^2 + \alpha}}$$

Fig. 12 Formula of compound score

This score is then interpreted using standard thresholds:

Metric	Value Range	Interpretation
Positive Score	0 to 1	Proportion of text with positive sentiment
Negative Score	0 to 1	Proportion of text with negative sentiment
Neutral Score	0 to 1	Proportion of text that is neutral in tone
Compound Score	-1.0 to +1.0	Aggregated normalized score reflecting overall sentiment
Threshold Rule	≥ 0.05	Positive sentiment
	≤ -0.05	Negative sentiment
	Between -0.05 and 0.05	Neutral sentiment

The evaluation framework in ELPM for evaluating student input on leadership and entrepreneurial mindsets consists of these thresholds plus the compound score, which provides a measurable indicator of sentiment orientation in textual data.

V. CONCLUSION

The Entrepreneurship Leadership Prediction Model (ELPM) shows how systematically evaluating leadership potential through categorical data-based machine learning techniques[10] can be utilized to evaluate and understand leadership in educational contexts. The model's system has been validated and is a good system to comprehensively assess, with leadership models made up of subjective and perception-based assessments as categorical variables. All of the modules in the system have passed all key test cases!

The ELPM penchant for unsupervised learning to identify hidden structures is a unique feature. K-means clustering[11] identifies naturally occurring clusters of leadership attributes and does so without a labelled and classified data set as a priori. In an educational context, the landmark classifications made in a variety of formats sometimes seem arbitrary or complicated. The combination of unsupervised machine learning techniques is exceptional for institutions that wish to assess the leadership attributes and mindsets of students uniformly and non-subjectively[[12]].

The analytical process begins with both clustering and classification as two prominent phases and builds a model that can identify patterns and make effective predictions. ELPM provides a powerful predictive engine by assigning new student records into similar leadership groups and translating those clustering assignments to target labels for use within the Random Forest classifier. The programmed, tiered design of the model also provided a more exploitable model that could generalize and adapt within different institutional contexts.

Although it can be difficult to assess distinct leadership qualities in educational administrations, ELPM allows only categorical data analyses. The ELPM allows administrators to measure a variety of leadership qualities within one categorical scale (1-5), and subsequently performed the arithmetic automatically. ELPM's capacity to allow for category variation in the areas where performance indicators could have only a tenuous definition also has the advantage of allowing the system to work without needing a target column.

ELPM also adds an extra level of depth through the use of its Association Rule Mining module (for faculty only); this module reveals complex relationships among varying levels of different variables. This will unpack the conditional dependencies and co-occurring features that will influence the leadership development strategy most effectively. Faculty and researchers can do more than think in terms of classification and prediction. ELPM has evolved into a tenant capable of qualitatively predicting educational leadership by adopting sentiment analysis through VADER being the newest edition. ELPM as a platform has now developed into a fully scalable, domain-sensitive approach to predicting educational leadership.

REFERENCES

1. Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." Journal of Machine Learning Research, vol. 12, 2011, pp. 2825–2830.
2. Hennig, Christian, et al., editors. Handbook of Cluster Analysis. CRC Press, 2015.
3. MacQueen, J. "Some methods for classification and analysis of multivariate observations." Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, no. 14, 1967, pp. 281–297.K. Elissa, “Title of paper if known,” unpublished.
4. Breiman, Leo, et al. Classification and Regression Trees. Wadsworth International Group, 1984.
5. Breiman, Leo. "Random forests." Machine Learning, vol. 45, no. 1, 2001, pp. 5–32. Springer, https://doi.org/10.1023/A:1010933404324.
6. Agrawal, Rakesh, and Ramakrishnan Srikant. "Fast algorithms for mining association rules in large databases." Proceedings of the 20th International Conference on Very Large Data Bases (VLDB), 1994, pp. 487–499.
7. Vonitsanos, Gerasimos, et al. "Social Media Insights into Climate Change: Sentiment Analysis Using VADER and RoBERTa." Proceedings of the 19th International Workshop on Semantic and Social Media Adaptation & Personalization (SMAP), Nov. 2024, DOI:10.1109/SMAP63474.2024.00036.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

8. Hutto, Clayton J., and Eric Gilbert. "VADER: A parsimonious rule-based model for sentiment analysis of social media text." Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM), 2014.
9. Arief, Murahartawaty, and Noor Azah Samsudin. "Hybrid Approach with VADER and Multinomial Logistic Regression for Multiclass Sentiment Analysis in Online Customer Review." International Journal of Advanced Computer Science and Applications, vol. 14, no. 12, 2023, pp. 1–8. DOI:10.14569/IJACSA.2023.0141232.
10. Agresti, Alan. Categorical Data Analysis. 3rd ed., Wiley-Interscience, 2012.
11. Jain, Anil K. "Data clustering: 50 years beyond K-means." Pattern Recognition Letters, vol. 31, no. 8, 2010, pp. 651–666. Elsevier, https://doi.org/10.1016/j.patrec.2009.09.011.
12. Wang, Zihan, and Geovanny Genaro Reivan Ortiz. "Assessing the management student's entrepreneurial intentions: Role of entrepreneurship education and technology transfer." Frontiers in Psychology 13 (2022): 953324.
13. Bernardes, Oscar, Vanessa Amorim, and Bárbara Santos. "Influence and Perception of Innovation and Creativity in Higher Education." EdTech Economy and the Transformation of Education. IGI Global, 2022. 36-52.
14. Fabeil, Noor Fzlinda, Entrepreneurship Education and Entrepreneurial Behaviour Among Undergraduate Students in Sabah, Malaysia (October 16, 2019). Journal of Economics and Business, Vol. 2 No.4, 2019, DOI: 10.31014/aior.1992.02.04.151, Available at SSRN: https://ssrn.com/abstract=3470250
15. Bezanilla, María José, et al. "Developing the entrepreneurial university: Factors of influence." Sustainability 12.3 (2020): 842.
16. Bizri, Rima, et al. "The entrepreneurial university: a proposed model for developing nations." Journal of Management Development 38.5 (2019): 383-404.
17. Saputra, Yori Andes, Asep Herry Hernawan, and Laksmi Dewi. "Entrepreneurship Curriculum in Higher Education." West Science Social and Humanities Studies 2.01 (2024): 106-112.
18. Li, Yue, Keyan Cao, and Hashem Salarzadeh Jenatabadi. "Effect of entrepreneurial education and creativity on entrepreneurial intention in college students: mediating entrepreneurial inspiration, mindset, and self-efficiency." Frontiers in Psychology 14 (2023): 1240910.
19. Lukito, Daniel. "Transformation of Student Leadership: Adopting Entrepreneurship to Face Global Challenges."
20. Lin, Minyi. "Research on Higher Education Management and the Cultivation of Innovation and Entrepreneurship Abilities in University Students." Academic Journal of Management and Social Sciences 5.3: 2023.