ASEC 2025 Conference

The 6th International Electronic Conference on Applied Sciences

09–11 December 2025 | Online

MDPI

Clustering Student Profiles with Parental Responsibilities Using Unsupervised Learning Algoritnms

Cutipa-Santi Michael Newton[1], QuelcaVelasquez Daisy Crishyely[2], Cutipa-Santi Yeni[3]
Universidad Nacional del Altiplano Puno

INTRODUCTION & AIM

University students with parental responsibilities face additional challenges that impact their academic performance, well-being, and persistence in their studies. This study analyzes 206 records from the 2024 University Census to identify patterns and profiles using unsupervised learning algorithms. It considers academic, family, socioeconomic, and time-management variables to gain insights into these challenges.

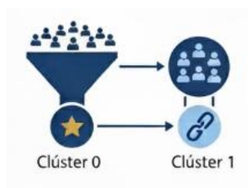
The Challenge: Balance of Roles

Institutional Support Gap

Academic & Time Overload

Institutional Economic Gap

Objective: Profile Identification



Objective

To identify and characterize profiles of students with parental responsibilities using K-Means, DBSCAN, and Agglomerative Clustering, in order to guide personalized academic interventions.

METHOD

The methodology of this study employed machine learning techniques, specifically clustering algorithms, to analyze the profiles of students with parental responsibilities. The methodology can be divided into several key stages:

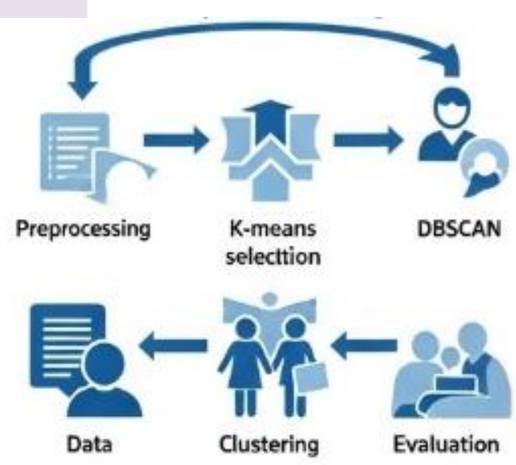
Data Selection and Cleaning

236 records were collected, and 206 were analyzed after a cleaning process.

Duplicates, null values, and irrelevant columns were removed, leaving 39 relevant variables for analysis.

Dimensionality Reduction (PCA)

PCA was used to reduce the dataset to 10 components, explaining nearly 100% of the variance.



Clustering Algorithms

Three algorithms were implemented:

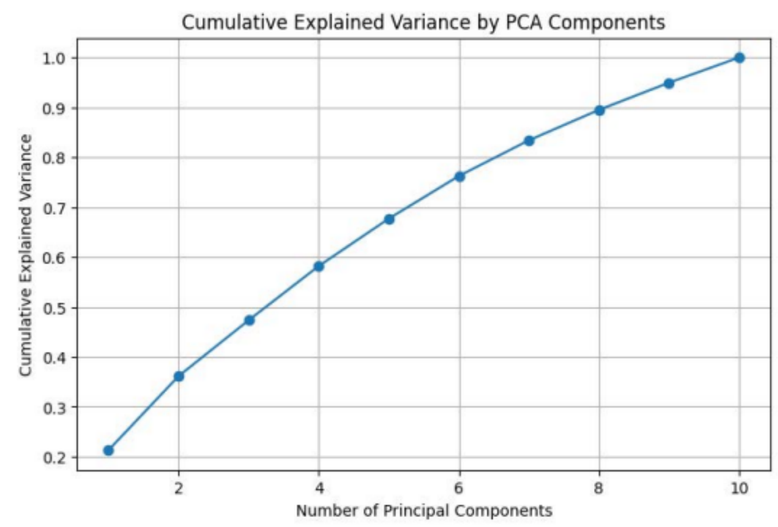
- ✓ K-Means with k=2 (optimized using the Elbow Method and the Silhouette Index).
- ✓ DBSCAN with parameters eps=1.5 and min_samples=5.
- ✓ Agglomerative Clustering with k=2 and Ward linkage.

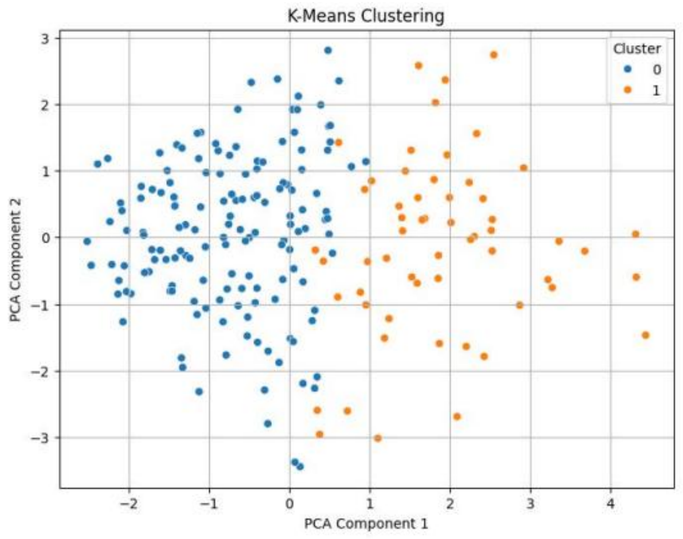
RESULTS & DISCUSSION

This study applied clustering algorithms to segment students with parental responsibilities. The following key findings were observed:

Principal Component Analysis (PCA):

shows that with 10 principal components, nearly 100% of the variance in the dataset was explained, ensuring efficient dimensionality reduction while retaining essential information for clustering analysis.





K-Means Clustering:

K-Means successfully identified two well-defined clusters. Cluster 0 consisted of older students (average age: 28), with better academic performance and lower academic overload, while Cluster 1 included younger students (average age: 24.6), who faced more study pauses and had a greater impact on their academic performance.

These results highlight the effectiveness of K-Means in identifying distinct student profiles and underscore the importance of using dimensionality reduction to improve clustering performance.

CONCLUSION

This study identified two profiles of students with parental responsibilities using the K-Means algorithm, which proved to be the most effective. Cluster 0 included more stable students with better academic performance and lower overload, while Cluster 1 grouped younger students with more pauses in their studies and a negative impact on their performance.

These results highlight the usefulness of K-Means in understanding the academic dynamics of students with parental responsibilities, providing a foundation for designing personalized interventions to support their academic success and well-being.

FUTURE WORK / REFERENCES

Incorporate more variables, such as socioeconomic and geographic factors, that could enrich the analysis. It is also suggested to explore the use of advanced clustering techniques and the integration of more emotional and social variables to improve the effectiveness of personalized interventions

•Alalawi, S. J. S., Shaharanee, I. N. M., & Jamil, J. M. (2023). Clustering student performance data using K-Means algorithms. *Journal of Computational Innovation and Analytics (JCIA), 2*(1), Artículo 1.