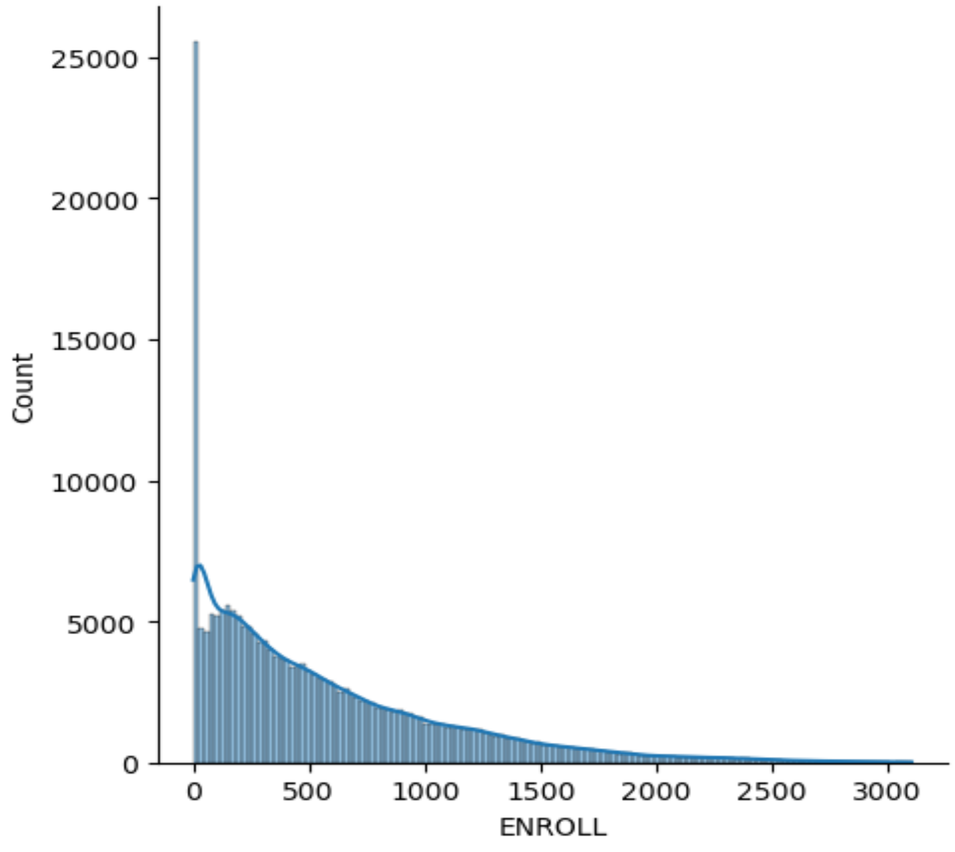
Education AI-ML Case Study

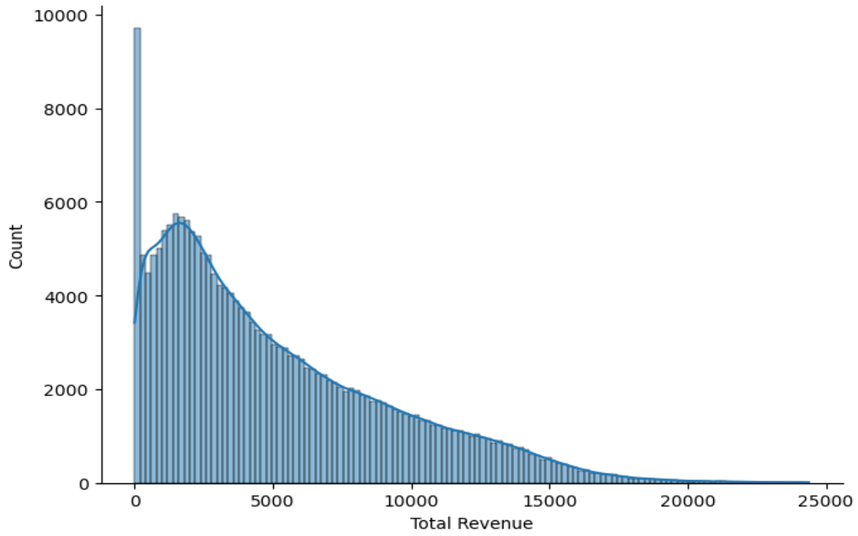
In the education industry, it is important for institutions to accurately predict the number of enrollments for various programs and courses they offer. This information is crucial for planning and allocating resources such as faculty, classrooms, and materials. Auto-ML algorithms can be used to analyze various factors that may impact the number of enrollments and predict the expected number of students.

The objective is to develop a Auto-ML model to predict the number of enrollments in an educational institution for a given semester or academic year. The model should consider various factors that can influence the enrollment numbers, such as the location and reputation of the institution, program offerings, tuition fees, financial aid, student demographics, marketing campaigns, and other external factors like economic trends, population growth, and government policies. The goal is to provide accurate enrollment predictions to help institutions plan their resources and budgets effectively, optimize their student recruitment strategies, and enhance their competitiveness in the education market.

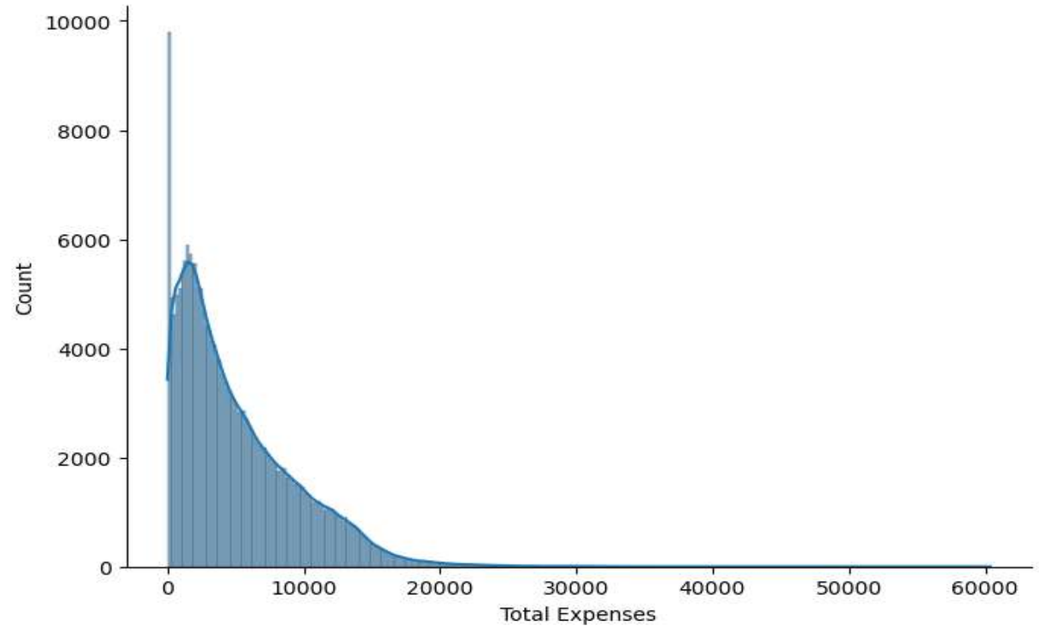


The **number of enrollments** refers to the total number of students who have registered and enrolled in a particular educational institution or program. This metric is important to educational institutions as it directly impacts their funding and revenue, as well as their ability to attract and retain students. The number of enrollments can be influenced by various factors such as the reputation of the institution, the quality of the educational programs offered, the cost of tuition, and the availability of financial aid and scholarships.

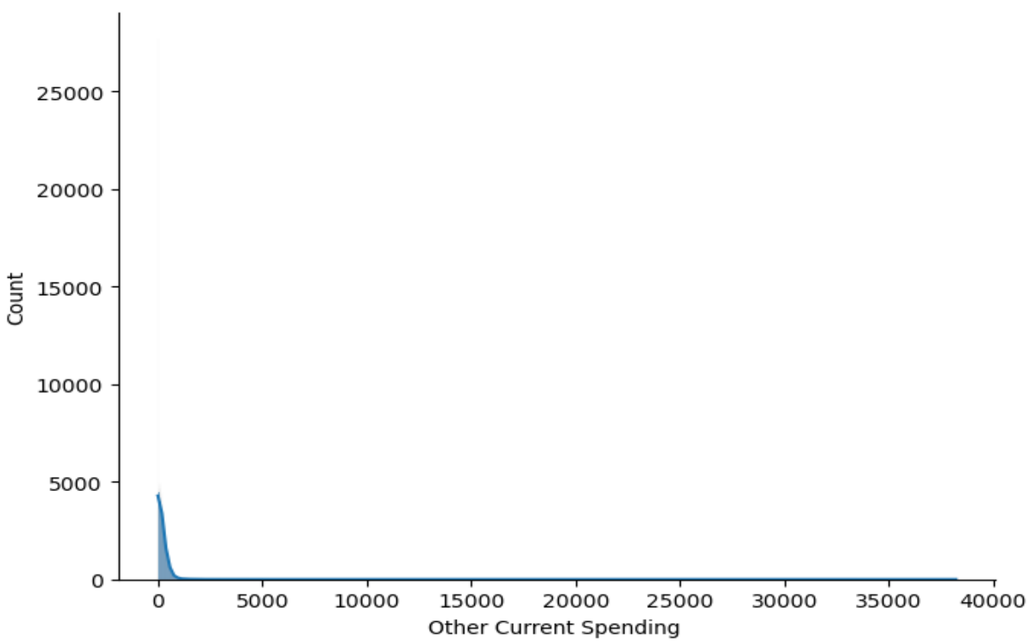
Features Responsible



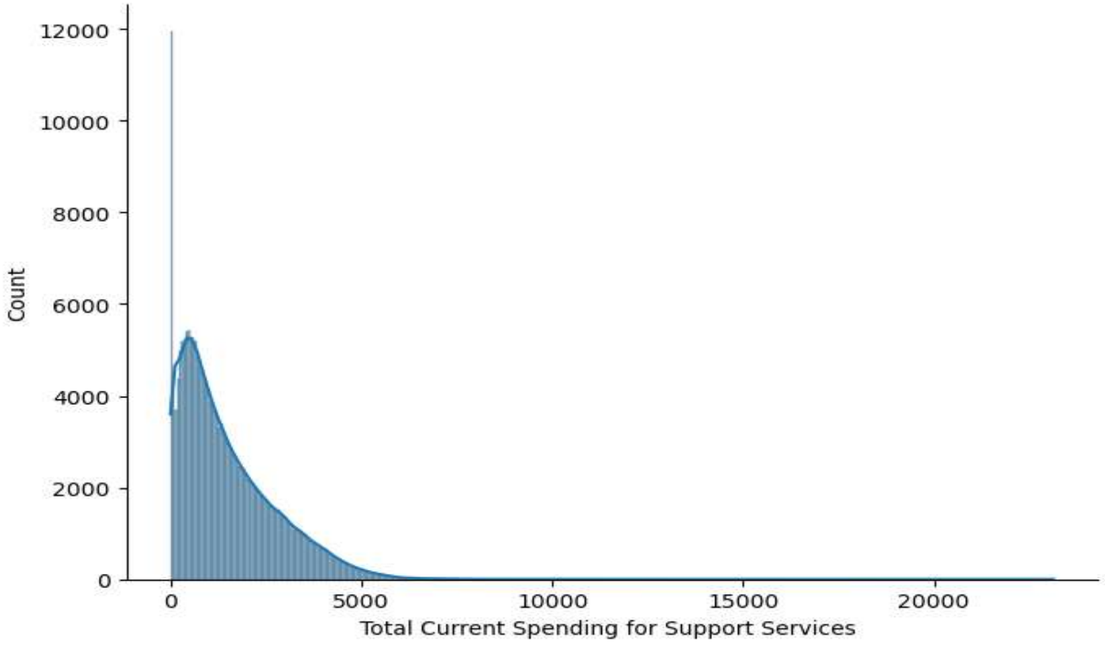
Tracking and analyzing the **Total revenue** can provide valuable insights into the financial health of the institution, and can help in making strategic decisions related to resource allocation, program development, and investment in infrastructure



Analyzing **Total expenses** can provide important insights into the financial health of the institution and its ability to manage its resources effectively.



Other current spendings refers to expenses other than salaries and benefits for staff, and expenditures on buildings and facilities. These expenses could include supplies, equipment, technology, marketing, and other operational costs



Total current spendings for support services refer to the amount of money allocated to the various support services provided in the institution. These may include services such as academic support, counseling, health services, career services, technology support, and student activities, among others.

Auto-ML Methodology Results

Case	Percentile	No. of Features	Random Forest	XGBoost	RNN	MLP	Lasso	Avg. Accuracy
Case 1	25	3	68.6	70.6	70.35	71.1	70	**70.12**
Case 2	50	6	77.9	77.1	69.53	76.7	71.9	**74.62**
Case 3	75	9	81.6	80.1	55.06	78	71.9	**73.32**
Case 4	90	10	81.1	80.8	62.83	78.4	72	**75.02**

- Based on our observation , Random forest was the best performing algorithm with 81.6% accuracy in 75^{th} percentile.
- 90^{th} percentile is the best percentile with an average accuracy of 75.02%.

.

Conclusion

In conclusion, predicting the number of enrollments in an educational institution can be a challenging task due to various factors that can influence student decisions. By analyzing data on factors such as tuition fees, student demographics, marketing expenditure, and other current spendings, machine learning models can be developed to predict the number of enrollments accurately. The dataset has 202775 records with 2 Categorical Features and 11 Numerical Features.

For regression, models were created with algorithms using Auto-ML techniques like Lasso, Recurrent Neural Network, Multilayer Perceptron, Random forest and XGBoost . With these models, performance measurement values were obtained for feature sets of 3,6, 9 and 10. The Auto-ML algorithms were able to predict number of enrollments with an average accuracy between 70% – 75% and helped to identify factors that determine the number of enrollments . The major factors include Total Revenue, Total Expenses , Other current spendings and Total current spendings . The Random forest with 81.6 % accuracy in 75^{th} percentile where tree showed a threshold of Total Current Spending for Instruction >= 5550000 dollars and Other Current Spending >= 407000 dollars which can lead to higher number of enrollments. Based on the performance measurement values obtained, it is possible to say that the study achieved success in predicting in number of enrollments.

.