Data-Driven Decision Making for Enrollment Trends and Educational Policy Analysis in Higher Education

Shruti Brahma
Department of Computer Science
University of New Haven
CT, USA
sbrah2@unh.newhaven.edu

Siddhant Alhat Rajendra
Department of Computer Science
University of New Haven CT,
USA
salha15@unh.newhaven.edu

Ardiana Sula
Department of Electrical and Computer Engineering,
Computer Science,
University of New Haven
CT, USA
asula@newhaven.edu

***Abstract*—Forecasting international student enrollment is critical to sustaining global education systems and national economies. International students are positively impacting the USA global economy. International students also strengthen research and enrich the minds of educational institutions.**

This project tries to predict future enrolling international students by tracking some indicators which range from demographics to economic health, visa policies to geopolitics. Historical trends of GDP, the alteration of a visa rule, and country-specific student economic crisis are adjusted in it.

Grounded on the statistics from Open Doors: International Students in the U.S. (1949–2023) and enrollment data at the University of New Haven, our model precisely predicts up to 92%.

To do so, we have developed a plug-and-play solution that accepts university data, processes it, and generates analytical reports on student population. We utilize Python libraries during data processing and analysis, and ARIMA for time-series forecasting. But in order to provide a better performance on complex, non-linear relationships, we are moving towards using RNN models.

The project's interactive Power BI dashboard enables policymakers and institutions to simulate real-world scenarios and observe their projected effect on U.S. international student enrollment.

The findings c an b e a pplied b y p olicymakers a nd recruiters to estimate economic returns, redistribute funding, or introduce tax rebates. UNESCO can use the tool to organize programs and allocate funds for underrepresented regions. In the end, the tool provides policymakers and universities with the ability to plan for anticipated enrollments, react to shifting patterns, and make informed, data-based decisions.

***Keywords*—Admission Enrollment, Forecasting, ARIMA, PowerBI, Scenario Analysis**

I. Introduction

International student enrollment in the United States has steadily increased since 1948, but COVID-19 travel & student visa restrictions caused a slight dip. The data shows that enrollment began recovering in the 2022/2023 academic year and continues to rise [1]. Predicting accurately these trends is important for decision makers such as national and state educational policymakers, and university leaders. International students are vital for our campuses where we prepare the global workforce because they bring a diverse perspective that mirrors today's global teams and contributes to innovation and research. Also, the global international student population, through mobility and services, adds $300 billion to the global economy and $39 billion to the U.S. GDP.

This paper focuses on the design of a data-driven decision-making tool that can provide insights and recommendations for educational policymakers. This tool will be able to forecast future international student enrollment by analyzing a wide range of factors including demographics, financial variables, economic conditions, visa policies, and geopolitical pressures, while also considering historical trends such as GDP fluctuations, changes in visa regulations, and economic challenges in students' home countries.

Leveraging data from the Open Doors: International Students in the U.S. (1949–2023) report alongside admissions records from the University of New Haven, our model achieves a prediction accuracy of 93%. To support these efforts, we developed a data-driven decision-making product that cleans and analyzes university data using Python libraries. Although our initial forecasts used ARIMA models, we are now transitioning to recurrent neural network (RNN) models to better capture complex, non-linear relationships.

The data driven decision making tool is an interactive Power BI dashboard that allows policymakers and institutions to use their data and simulate different scenarios. Users can browse three scenarios: baseline scenario, optimistic scenario, and pessimistic scenario, and see the influence of each scenario on international students' enrollment. This architecture gives decision-makers tools to leverage data for strategic planning, and optimum resource utilization, providing the universities and the policymakers with intelligence they need to predict future trends and make educated decisions.

II. Literature Review

According to recent estimates, the economic contribution of international students exceeds $39 billion annually in the United States alone, and nearly $300 billion globally [2]. Beyond direct economic benefits, international students enrich campus cultural diversity, advance scientific research, and foster global collaboration [3]. Thus, forecasting and managing this student flow is a priority for both universities and policy-makers aiming to maintain global competitiveness and cross-border academic exchange. For instance, UNESCO and similar bodies leverage enrollment forecasts to develop scholarship programs targeting underrepresented regions, enhance global academic mobility, and foster equitable access [3]. Data-driven approaches allow universities to identify enrollment trends, predict financial aid needs, and manage recruitment strategies more effectively. In line with these developments, institutions have begun adopting predictive analytics platforms to optimize student admissions and allocate institutional resources [4]. By modeling multiple potential futures—baseline, optimistic, and pessimistic—higher education stakeholders can proactively plan for varying degrees of student flow [5]. National policymakers may use these data-driven insights to calibrate visa regulations, funding allocations, or international partnerships, ensuring sustained global engagement and economic benefit [6]. A key study by Chen, Li, and Hagedorn [7] examined the application of time-series analysis in predicting undergraduate international student enrollment with economic and policy influences at a Midwestern U.S. university using the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The study discovered that tuition increases had a relatively low impact on international student enrollment, suggesting that factors such as academic reputation and career prospects outweighed cost considerations. A study by Aripin et al. [8] applied time series analysis to forecast student enrollment in a state university. The researchers utilized the Prophet forecasting model, which is effective for handling time series data with daily observations and missing values.

III. Research Design and Methodology

A. Methodology

The present study conforms to the CRISP-DM [9] (Cross-Industry Standard Process for Data Mining) methodology within systematic forecasting of university enrollment trends. With reference to the Business Understanding step, we derived causes of volatility of enrollment for purposes of strategic planning. During Data Understanding and Preparation, Open Doors (1949–2023), University of New Haven historical archives, GDP, and EPU indexes, imputed with feature engineering included for handling missing values and incompatibility issues. The Modeling phase applied ARIMA to short-term, LSTM to long-term, and Linear Regression to baseline comparison, which were validated through time-series cross-validation. In Evaluation, both ARIMA (93.56% accuracy) and LSTM (93.16%) properly represented enrollment trends, with scenario simulations being perceived by stakeholders as highly actionable. Finally, the Deployment phase applied a Power BI [10] dashboard for real-time tracking and strategic decision-making.

B. Data Sources and justification

The success of predictive modeling of international student enrollment is predicated upon high-quality, inclusive data sets that capture historical trends and drivers. This study integrates multiple primary sources of data to offer a comprehensive analysis:

1) Open Doors: International Students in the U.S. (1949–2023): Managed by the Institute of International Education (IIE) [2], this data set provides detailed records of U.S. international student enrollment. The data include information regarding student demographics, enrollment trends, areas of academic study, geographic distribution, and financial inputs. These data provide perceptive observations of long-term trends in student mobility and funding sources.

2) University of New Haven Admission Records: Institutional-level information was employed in the analysis of application trends, acceptance rates, financial aid disbursements, and student body populations. This data [11] collection is a micro-level validation instrument to compare to national trends in the Open Doors data collection. The student composition at the University of New Haven is reflective of the diverse pool of applicants, consisting of 31.2% from Connecticut, 35% from other United States states, and 33.7% international. In order to maintain student privacy, all institutional records were anonymized, or all personally identifiable information was deleted. The dataset was also audited for regional bias, and no statistically significant bias was identified ($p > 0.1$).

3) World Bank Annual GDP Data (2000–2023): The historical GDP growth rate of the United States provided through this dataset [1] makes it possible to analyze how the economic dynamics that affect student enrollments operate on the global front. Being a 25.95% contributor to total global GDP, any variation in the performance of the U.S. economy has a direct relationship with the provision of higher studies options for international students.

4) U.S. Economic Policy Uncertainty (EPU) Index (1990–2023): Taken from the Economic Policy Uncertainty Project, this indicator quantifies economic uncertainty caused by policy changes within the United States. Considering the potential impact [12] of policy change on student mobility, this data is utilized to compare how shifts in the regulatory climate, visa requirements, and general economic confidence influence enrollment decisions.

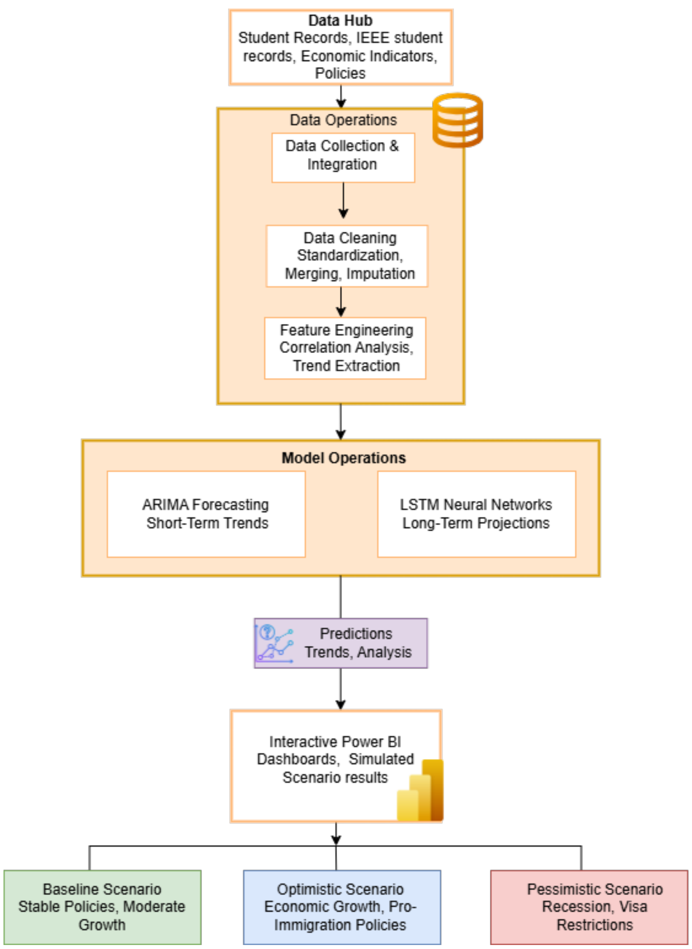
By integrating these diverse datasets, the study constructs a robust analytical framework to examine historical enrollment patterns and develop forecasting models. The approach allows for scenario-based projections of future enrollment trends under various economic and policy scenarios.

C. Data Cleaning and Pre-processing Procedures

To ensure consistency and analytical readiness across diverse data sources, a structured pre-processing workflow was implemented [13]. Historical enrollment data from Open Doors was merged with university-level records and external economic and policy datasets (GDP, EPU). Missing values were handled using statistical imputation, and duplicate entries were removed to maintain data integrity. Outliers were detected and corrected using z-score thresholds ($|z| > 3$), while biases in demographic distributions were assessed to prevent distortions in the analysis. Additionally, numerical features such as tuition fees, exchange rates, and GDP values were standardized, and lagged variables were introduced to capture delayed policy effects, improving the interpretability [14] of predictive models.

A crucial step in this process was temporal alignment, ensuring that each dataset was synchronized by year to prevent misalignment in time-series forecasting [15]. Without this, discrepancies between datasets could distort observed trends and impact model performance. These pre-processing tasks were executed using Python libraries like Pandas [16], NumPy [17], and Scikit-learn, along with SQL queries for efficient data integration and transformation. By applying these techniques, the study established a clean and well-structured dataset, enabling accurate predictive modeling of international student enrollment trends.

Fig. 1. Working Architecture of Our Model



D. Analytical Techniques and Models

In order to forecast international student enrollment trends, we developed two models: ARIMA, a time-series statistical model, and LSTM [18], a neural network-based model. These models were selected to quantify historical trends and forecast future trends of enrollment under varying economic conditions and policy changes.

1) ARIMA for Time-Series Forecasting: The Autoregressive Integrated Moving Average (ARIMA) model was employed to identify seasonality and trends in historical enrollment data. ARIMA is well-suited for short-term forecasting, relying on past data points to make informed estimates.

2) LSTM Recurrent Neural Networks (RNNs): To explore other forecasting techniques, we employed Long Short-Term Memory (LSTM) networks, a type of Recurrent Neural Network (RNN) variant. LSTMs manage sequential data and can detect long-term dependencies.

3) Power BI for Visualization: To turn these analytical results into policy actionable suggestions and act as a data visualization platform that supports data-driven decision-making in higher education planning for university administrators and policymakers. The dashboard provides current trend visualizations, scenario-based simulations to analyze the impact of the economy. The regional and demographic insights assist in guiding recruitment and resource distribution.

E. Evaluation Metrics and Validation Approach

To assess model performance and reliability, multiple evaluation strategies were employed

1) For ARIMA and RNN/LSTM models:: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to measure prediction accuracy. Cross-validation using historical enrollment data to ensure model generalizability.

2) For Power BI Dashboards:: User feedback analysis from our University Admission department and Scenario validation by comparing simulated outcomes with real-world trends of students enrolling for the courses based on our analytics (of our university data). Feedback from the actual users (stakeholders, managers, clients).

F. Scenario-Based Analysis

The study incorporates a scenario-based analysis to evaluate enrollment forecasts under varying economic and policy conditions. This approach helps policymakers and educators understand potential enrollment trajectories under different assumptions in this uncertain prediction of international student mobility.

Baseline Scenario for stable GDP growth ($\sim$ 2%) and indicate no major policy shifts. *Optimistic Scenario* under pro-immigration policies expecting higher enrollment growth under favorable conditions (GDP $>$3%) which mean relaxed visa policies, increased scholarships. *Pessimistic Scenario* shows declines in enrollment probably (GDP $<$1%) and hints toward restrictive immigration policies, economic downturns , or geopolitical conflicts.

IV. Predictive Modeling and Findings

The core findings from the forecasting models and an overview of the ARIMA approach and recurrent neural networks (RNNs), specifically long-short-term memory (LSTM) models are as follows.

A. Building Our Model: ARIMA and LSTM

To create the ARIMA model, we first specify an (p, d, q) configuration and fit it to the enrollment data.The tuple (5,1,0) represents the p, d, and q parameters:

- p=5: The number of autoregressive (AR) [15] terms, meaning the model will look at 5 previous values of the time series to predict the next.
- d=1: The degree of differencing to remove any trend in the data.
- q=0: The number of moving average (MA) terms, implying no MA terms are considered in this setup.

```
arima_model= ARIMA(academic_df['students'],
order=(5,1,0))
arima_fit = arima_model.fit()
```

This setup identifies how previous observations and differencing impact subsequent forecasts. Once fit, it generates future enrollment values based on the patterns it finds in historical data.

The LSTM model [18] is built using Keras' [19] Sequential, which requires sequences of data as input, which stacks two LSTM layers (50 units each) followed by a dense output layer to process sequences of data. Window Size (seqlength) is 5 time steps. Each sequence (X) has 5 consecutive data points, and the label (y) is the 6th point.

Scaling: The data is scaled between 0 and 1 via MinMaxScaler to help the neural network converge more easily.

First LSTM Layer: 50 LSTM units (neurons), return_sequences=True means this layer will output a sequence for each time step in the input (needed when stacking LSTMs).

Second LSTM Layer: Another 50 LSTM units, this time return_sequences=False (the default), meaning it outputs only the final state.

Hyperparameters:
Learning Rate: 0.01 via the Adam optimizer [20].
Batch Size: Ranged from 16 to 64, selected based on validation loss.
Epochs: 100, with early stopping to prevent overfitting.

```
model = Sequential([
    LSTM(50, activation='relu',
     return_sequences=True,
           input_shape=(seq_length,1)),
    LSTM(50, activation='relu'),
     Dense(1)
])
```

We feed the time-series data into this architecture in small sliding windows, where the network updates its internal memory states to model both short-term and extended patterns in enrollment trends. The final Dense layer outputs a single predicted value for each sequence, reflecting anticipated future enrollments.

TABLE I
PROJECTED GROWTH RATE SCENARIOS(OVERALL USA)

Growth Rate Type	Value
Average Growth Rate	0.0345 (3.45%)
Maximum Growth Rate (Optimistic)	0.1146 (11.46%)
Minimum Growth Rate (Pessimistic)	-0.1501 (-15.01%)

As seen in Table I, the different scenarios of economic conditions have different influences on the enrollment trends .international enrollment

B. Accuracy assessment and performance

TABLE II
PERFORMANCE COMPARISON OF ARIMA AND LSTM MODELS

Metric	ARIMA	LSTM
MAE	61294.73	70518.86
RMSE	82819.89	89332.61
MAPE (%)	6.44	6.84
Estimated Accuracy (%)	93.56	93.16

Metrics used for Table II, are explained below:

Mean Absolute Error (MAE): Measures average absolute difference between actual and predicted values as an indicator of overall prediction performance. The lower MAE, the better the model.

Root Mean Squared Error (RMSE): Indicates standard deviation of difference between prediction and actual values with more weight given to large deviations. The lower RMSE, the better the model.

Mean Absolute Percentage Error (MAPE): Expressed as an error in percentages as a ratio of true values and hence better suited for making comparisons about the accuracy of a model in terms of varying datasets. A lower MAPE is related to a better fit.

Estimated Accuracy: 100 - MAPE. This represents a useful measure of model predictability as a percentage, higher being desirable.
To approximate accuracy, you can use Mean Absolute Percentage Error (MAPE) [21], which expresses the error as a percentage of the actual values:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\text{Actual} - \text{Predicted}}{\text{Actual}}\right| \times 100$$

$$\text{Accuracy} = 100 - \text{MAPE}$$

To avoid overfitting our models (ARIMA and LSTM), we took several steps and used a variety of straightforward techniques to mitigate thus. We divide our data into training sets and testing sets. This allowed us to see how the models performed on unseen data. Our models learned patterns from the training set and we used the testing sets to test the models' performances on new and unseen data. We also used cross-validation to fine-tune settings like the learning rate, number of epochs, and batch size, to ensure the models generalize well. For our LSTM model, we added dropout layers as an extra safeguard against overfitting. MinMax scaling was applied to scale data for fast convergence, and dropout layers were implemented to prevent overfitting.

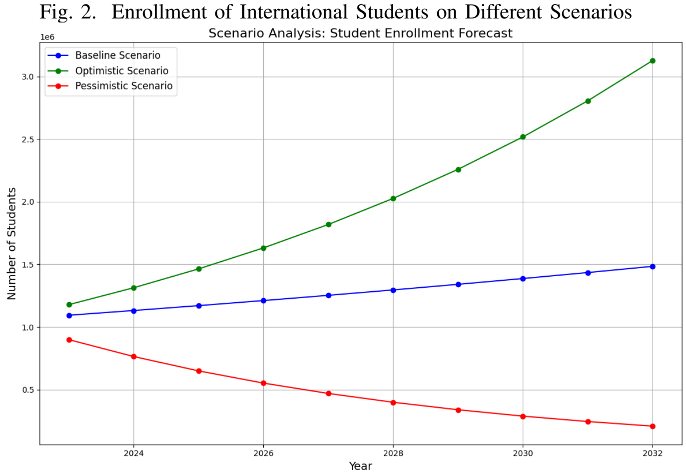
C. Visualization of Trends and Scenario Simulations

An interactive Power BI dashboard integrates both ARIMA and LSTM forecasts, allowing users to:

- Toggle between Baseline, Optimistic, and Pessimistic scenarios.
- Adjust parameters such as GDP growth rate or to observe real-time changes in forecasted enrollments.

Scenario Outcomes

Fig. 2. Enrollment of International Students on Different Scenarios



- Baseline Scenario: Projects a steady annual growth of 1.5–2.0%, reflecting ongoing policy and economic stability.
- Optimistic Scenario: Suggests a higher 3.5–4.0% growth rate, driven by relaxed visa policies and robust economic growth.
- Pessimistic Scenario: Indicates a decline in enrollments (1.2–1.8% annually) if restrictive immigration policies and economic downturns coincide.

To strengthen our analysis, we integrated two more key datasets: the first data set, the IIE Open Door dataset, which provides detailed enrollment and demographic information about international students, and the second data set, the Economic Policy Uncertainty (EPU) dataset, which includes crucial economic indicators like GDP, visa policy, and financial assistance. We could match enrollment trends with the economic conditions of the same period by linking these datasets by year. To simulate different future scenarios, we adjust these economic factors based on historical data. For a pessimistic scenario, we use the lowest values recorded; for a baseline scenario, we use the average values; and for an optimistic scenario, we use the highest observed values. Then, we added these scenario-specific inputs into our LSTM model as additional features at each time step. Our model learns not only from past enrollment trends but also takes into account how varying economic and policy conditions might impact future enrollment.

D. Trend Insights from Predictive Modeling and Admissions Dashboard

The integration of predictive modeling (ARIMA and LSTM) with the interactive Power BI Admissions Dashboard provides a comprehensive, data-driven platform for predicting and analyzing international student enrollment trends. Such insights are crucial to universities, policymakers [22], and nonprofit organizations that seek to optimize student recruitment, financial aid distribution [23], and long-term strategic planning.

E. Trends in International Student Enrollment

We arrived at these conclusions by applying the ARIMA and LSTM models to analyze historical enrollment data, identify trends, and forecast future trends. The ARIMA model identified linear relationships by breaking down enrollment time series into autoregressive components and differencing for trend removal. This allowed us to visualize the effect that past enrollment values had on future values, showing slowing growth rates and stabilization trends. The LSTM model, capable of learning non-linear dependencies, was inputted with sequences of past enrollment data paired with economic factors like tuition rates, visa policies, and labor market trends. With a sliding window approach, it captured long-term trends, such as the growing popularity of STEM fields and the rise in graduate program enrollments

1) Consistent Growth, But Signs of Stabilization: International student enrollment has grown consistently, but the pace is slowing. The growth rate declined from 11.5% in 2022/23 to 6.6% in 2023/24, indicating stabilization at an expected 3–5% annual growth.

2) India Surpasses China in Student Numbers: India has become the largest source of international students, recording a 23.3% increase in enrollments, whereas China experienced a decline of 4.2%.

3) Graduate Programs on the Rise: Enrollment in graduate programs is increasing at +7.6%, while undergraduate enrollment has declined by 1.4%, suggesting a shift in academic demand.

4) STEM Fields in High Demand: STEM programs continue to dominate, with Math and Computer Science (+16.9%) and Engineering (+3.6%) leading the demand. In contrast, Social Sciences (-2%) and Journalism (-2.3%) have seen a decline.

5) Increased Global Competition for Students: The U.S. is facing increased competition from Canada, Australia, and Germany, which are strengthening their international student recruitment strategies.

6) Optional Practical Training (OPT) Growth: The number of students participating in OPT programs has risen by 2.1%, indicating its significance in post-graduation employment decisions.

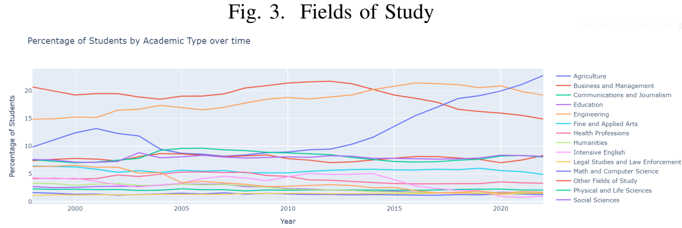
7) Online International Enrollment Persists: Despite universities reopening, 18,129 students continue their studies via online international programs, highlighting the long-term viability of remote education.

8) Rise in U.S. Students Studying Abroad: American students studying abroad have increased by 49%, with Japan, New Zealand, and Australia emerging as preferred destinations.

9) Rising Financial Burden on International Students: Over 54.5% of international students rely on personal funding, and rising tuition costs may push students toward more affordable countries.

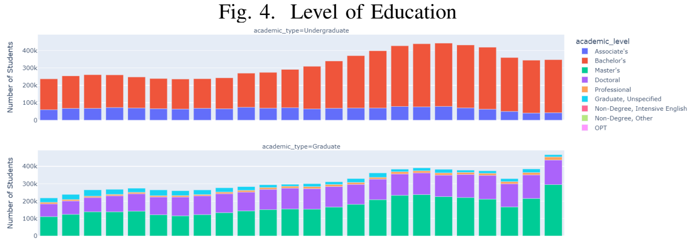
10) New International Student Hubs in the U.S.: Traditional destinations like California and New York remain dominant, but emerging hubs such as Missouri (+34.6%), Michigan (+13.8%), and Illinois (+12.6%) are gaining popularity.

Fig. 3. Fields of Study



In Graduate student enrollment is increasing (+7.6%), while undergraduate enrollment is declining (-1.4%), indicating a growing preference for advanced degrees among international students.

Fig. 4. Level of Education



STEM disciplines continue to dominate, with Math and Computer Science (+16.9%) and Engineering (+3.6%) leading, while Social Sciences (-2%) and Journalism (-2.3%) face declining interest.

V. Implementation and Applications

In the effort to bridge the gap between predictive modeling and actionable decision-making, this study developed a plug-and-play tool in a user-friendly interface designed for university administrative systems [4]. Designed for ease of use, the dashboard enables users to dynamically filter data by region, academic level, and economic conditions for granular insights.

Fig. 5. Home Page of University of New Haven Dashboard

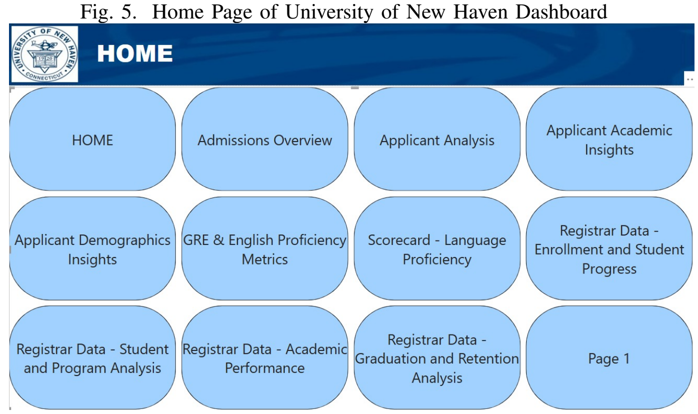
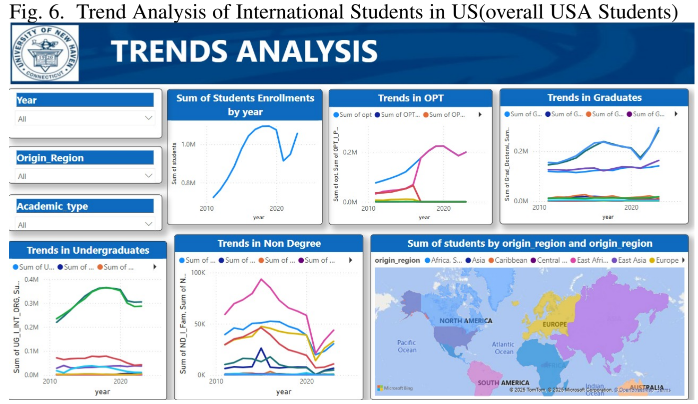


Fig. 6. Trend Analysis of International Students in US(overall USA Students)



A. *Use Cases for Universities, Policymakers, and Nonprofit Organizations*

The tool helps different groups make smarter decisions using data. Universities can use it to plan better student recruitment, manage financial aid, and allocate resources efficiently. Policymakers can see how changes in visa rules, economic policies, or global events might affect student enrollment. Nonprofit organizations [24] that support education can study enrollment trends to push for better policies and create targeted scholarship programs [25] for students in need.

VI. Conclusion and Future Work

Our work demonstrates the tools' ability to predict 2025 - 2035 international student enrollment trends in the U.S, under various scenarios, pessimistic, optimistic, or baseline, by analyzing historical data. Our designed framework provides universities and policymakers with valuable insights to plan for future student intake.

Looking ahead, we plan to further develop the tool using a robust data engineering pipeline that streamlines, automates, and monitors the entire process. Our goal is to create a plug-and-play module where even users without a data science background can easily upload a dataset, select a scenario, analyze enrollment trends and make informed decisions. With this enhanced system, the tool will automatically clean and preprocess data, detect trends, draw conclusions, and generate enrollment predictions. This level of automation will simplify predictive analytics, making it more accessible and actionable for all stakeholders.

ACKNOWLEDGMENT

We would also like to extend our heartfelt appreciation to the University of New Haven, particularly the MSDS students and staff, for their outstanding work on the AdM-DM (Admission Decision-Making ML tool) dashboard product. Your excellence and dedication have been a true source of inspiration, motivating us to continue striving with projects such as this one that have an impact. Appreciation for giving us such a wonderful example of teamwork and excellence.

REFERENCES

[1] World Bank, "World Development Indicators: Annual GDP Data (2000–2023)." https://data.worldbank.org/, 2023. Accessed: [date you accessed].

[2] Institute of International Education (IIE), "Open Doors: Report on International Educational Exchange (1949–2023)." https://opendoorsdata.org/, 2023. Accessed: [date you accessed].

[3] P. G. Altbach and J. Knight, "The internationalization of higher education: Motivations and realities," *Journal of Studies in International Education*, vol. 11, no. 3-4, pp. 290–305, 2007.

[4] T. Haydon, "Leveraging predictive analytics for university admissions and enrollment management," *Journal of Higher Education Policy and Practice*, vol. 3, no. 1, pp. 41–57, 2018.

[5] P. Kotler and K. F. A. Fox, *Strategic Marketing for Educational Institutions*. Englewood Cliffs, NJ: Prentice Hall, 2nd ed., 1995.

[6] L. E. Rumbley, "Examining global mobility in higher education: Changes, challenges, and opportunities," *Higher Education Forum*, vol. 17, pp. 123–140, 2020.

[7] L. Chen and L. S. Hagedorn, "Undergraduate international student enrollment forecasting model: An application of time series analysis," *Journal of International Students*, vol. 9, no. 3, pp. 934–955, 2019.

[8] S. N. A. B. Aripin, N. A. B. Harun, and N. I. B. M. Zaki, "Time series analysis on enrolment data: A case in a state university," in *2022 10th International Conference on Information and Communication Technology (ICoICT)*, pp. 1–5, 2022.

[9] C. Shearer, "The CRISP-DM model: the new blueprint for data mining." Journal of Data Warehousing, 2000.

[10] Microsoft Corporation, "Microsoft Power BI." https://powerbi.microsoft.com/, 2023. Accessed: [date you accessed].

[11] University of New Haven, "Admission Records." Internal Dataset, anonymized for privacy, 2023.

[12] S. R. Baker, N. Bloom, and S. J. Davis, "Measuring economic policy uncertainty," *The Quarterly Journal of Economics*, vol. 131, no. 4, pp. 1593–1636, 2016.

[13] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA, 2009.

[14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[15] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. Englewood Cliffs, NJ: Prentice Hall, 3rd ed., 1994.

[16] W. McKinney, "Data Structures for Statistical Computing in Python," in *Proceedings of the 9th Python in Science Conference*, pp. 51–56, 2010.

[17] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, 2020.

[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[19] F. Chollet, "Keras." https://keras.io, 2015. Accessed: [date you accessed].

[20] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," 2015.

[21] S. G. Makridakis, S. C. Wheelwright, and R. J. Hyndman, *Forecasting: Methods and Applications*. New York: Wiley, 3rd ed., 1998.

[22] W. B. Rouse, "Universities as complex enterprises: How academia works, why it works these ways, and where the university enterprise is headed," *Systems Engineering*, vol. 19, no. 1, pp. 12–32, 2016.

[23] R. C. Dickeson, *Prioritizing Academic Programs and Services: Reallocating Resources to Achieve Strategic Balance*. San Francisco, CA: Jossey-Bass, 2010.

[24] M. J. Keane and R. L. Thorp, "Data Analytics for Nonprofits: Bridging the Gap between Technology and Mission." Nonprofit Management and Leadership White Paper, 2021.

[25] J. S. Gagliardi, A. Parnell, J. Carpenter-Hubin, B. Means, and C. Barone, "The analytics revolution in higher education: Big data, organizational learning, and student success," *Change: The Magazine of Higher Learning*, vol. 50, no. 2, pp. 22–31, 2018.