CASE STUDY

EDUCATION INDUSTRY

Sriya.AI
AI ANYWHERE ANYTIME

Objective

- To get the prediction accuracy of Institutions with High Enrollment by Auto-AI and SXI and compare.
- Precision AI using Target SXI based Random Forest trees. Target increase in No. of Institutions with High Enrollment is **20%** up from current levels.

SXI Hypothesis

- SXI is a proxy/surrogate for all features responsible for ensuring institutions with high or low Enrollment rate. The higher the SXI, the better the institutions with High Enrollment and hence increasing SXI score should lead to increase in No. of institutions with high Enrollments.

SXI Definition

- **Sriya Expert Index (SXI):** Dynamic score/index obtained from a proprietary formula consisting of weights from 10 ML algorithms. SXI is a super feature and is a true weighted representative of all important features. Converts a multi-dimensional hard to solve problem into a simpler 2-dimensional solution (problem solved).
- **SCORE + CORRELATE = IMPROVE**

Discussion & Results

1. Exploratory Data Analysis

202775 enrollments were distributed to **78516** Good and **124259** bad. Good are the number of institutions with high enrollments and Bad are the number of institutions with low enrollments. So, **38.78%** is the current institutions with High Enrollments and **79.39%** is institutions with Low Enrollments.

2. <u>SXI - Exploratory Data Analysis</u>

Current Average SXI is **7.32**. Number of institutions Enrollment above 7.32 is **91552** and of these **63960** are number of institutions with high enrollment and **27592** are number of institutions with low enrollment. So High institutions Enrollment (%) is **69.86**% and low are institutions Enrollment **30.14**%.

Correspondingly No. of total enrollments below 7.32 is **111223** and of these **14556** are high and **96667** are low. So High institutions Enrollment (%) is **13.09**% and low institutions Enrollment is **86.91**%.

So SXI is a perfect proxy/surrogate for High Enrollment and above average SXI the ratio of good outcome is **1.8x** overall average and below average SXI this ratio of good outcome is **0.33** overall average. So, increase in SXI leads to increase in No. of Institutions with High Enrollment.

3. <u>Predictive AI</u>

- Auto-AI Prediction accuracy is **80.8%** and best performing algorithm is **XGBoost**.
- SXI Prediction accuracy of Institutions with High Enrollment is **100%**.
- Ratio of SXI/Auto-AI prediction accuracy is **1.24.**

4. <u>Precision AI</u>

Desired increase in target outcome which is No. of Institutions with high enrollments rate is 20%. Original No. of Institutions with high enrollments rate is **38.72**% so a **20%** increase should lead to a **46.5**% overall Number of Institutions with High Enrollment (**38.72*1.2**). Which means **94219** institutions from 202775 would become institutions with High enrollment rather than current **78516.**

The correlation between SXI and Institutions with High Enrollment is **0.99**. This implies that SXI and Institutions with High Enrollment are highly, positively correlated to each other. Hence, increase in SXI will definitely result in increase in Institutions with High Enrollments.

Correlation graph: Institutions with High Enrollments (%) & SXI
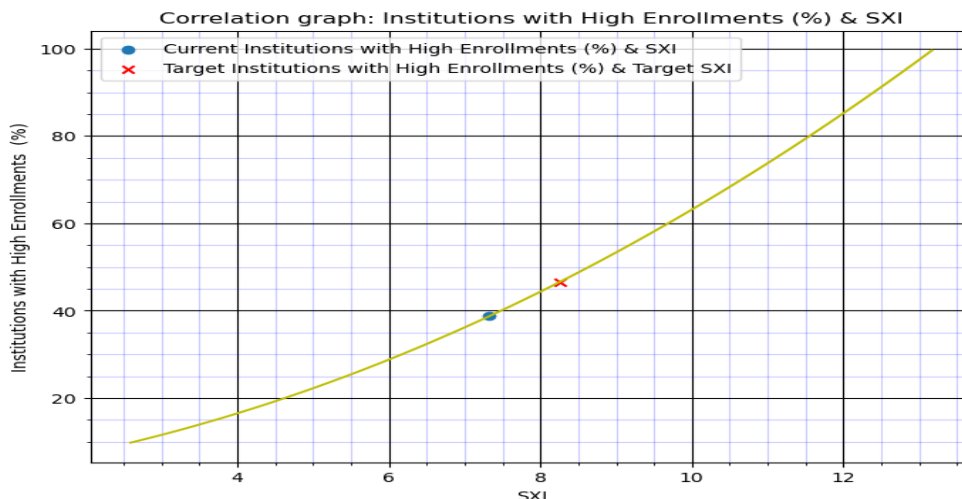
Current SXI and Target SXI Decision Trees

a. Current SXI Decision Tree
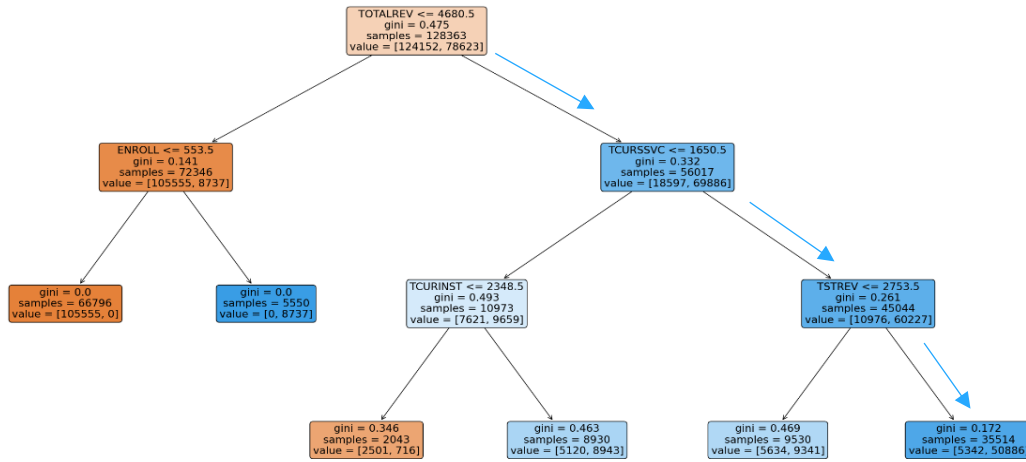
Interpretation: -

Node 1: Total Revenue >= $ 4,680,500 (No. of High Enrollment in parent node: 78623) (Right split so it is False symbol changes from < to >)

Left split: 8737 gini:0.14, **Right Split**: 69886- majority positive class; gini:0.3

(Total value for the next split: 69886)

Node 2: Total Current Spending for Support Services >= $1,650,500 (Right split so it is False symbol changes from < to >)

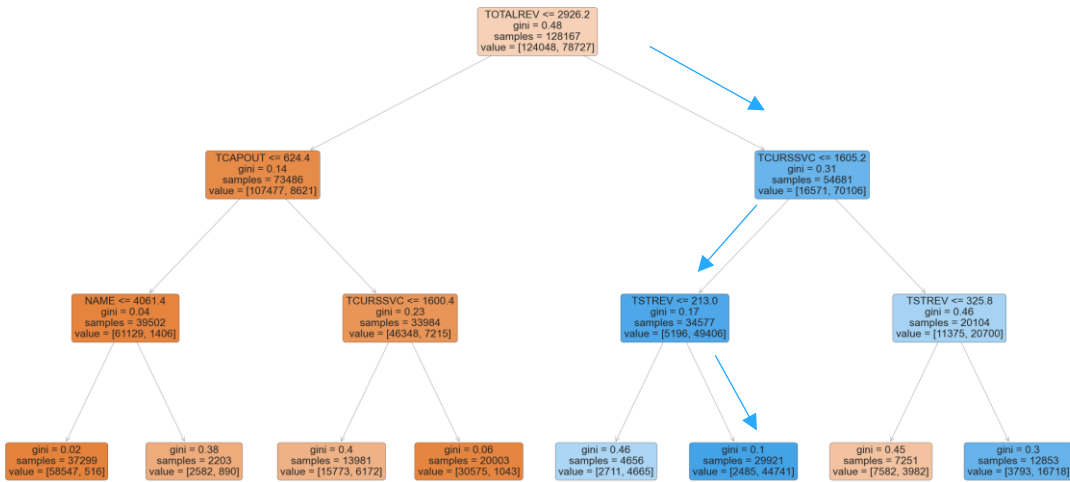**Left split**: 9659; gini:0.49, **Right Split**: 60227- majority positive class; gini:0.26

(Total value for the next split: 60227)

Node 3: Total Revenue from State Sources >= $2,753,500 (Right split so it is False symbol changes from < to >)

Left split: 9341; gini:0.46, **Right Split**: 50886 - majority positive class; gini:0.17 – Final Leaf Node.

✓ High/Low Enrollment Ratio is **9.5**

b. Target SXI Decision Tree

Target SXI from correlation curve for 20% increase in target outcome of Institutions with High Enrollment is **8.25**.

Interpretation: -

Node 1: Total Revenue >= $ 2,926,200 (No. of High Enrollments in parent node: 78727) (Right split so it is False symbol changes from < to >)

Left split: 8621 gini:0.14, **Right Split**: 70106- majority positive class; gini:0.31

(Total value for the next split: 70106)

Node 2: Total Current Spending for Support Services <= $ 1,605,200

Left split: 49406- majority positive class; gini:0.17, **Right Split**: 20700; gini:0.46

(Total value for the next split: 49406)

Node 3: Total Revenue from State Source >= $ 213,000 (Right split so it is False symbol changes from < to >)

Left split: 4665; gini:0.46, **Right Split**: 44741 - majority positive class; gini:0.1 – Final Leaf Node.

- ✓ High/Low Enrollment Ratio is **18**

Conclusion 5

1. SXI Prediction accuracy is **1.24** times Auto AI prediction accuracy and hence is **24%** superior.

2. Institutions, whose SXI score is higher than current average SXI score of **7.32** have **80%** of Institutions with High enrollment than overall of average Enrollment rates.

3. Target **20**% increase in number of Institutions with higher Enrollments is achievable by increasing target SXI to **8.25** from current **7.32** levels. This would result in **94219** Institutions with higher Enrollments up from current **78516** levels.

Initial Increase from current levels: 20% or 15,703.	SXI Impact Potential

4. Target High/Low Enrolment ratio is **18** while the current ratio is **9.5.** This represents a **potential 89.47% compounded increase** if all recommendations in target SXI are completely implemented**.**

Compounding Increase from current levels: 89.47% or 70,248.	SXI Impact Potential