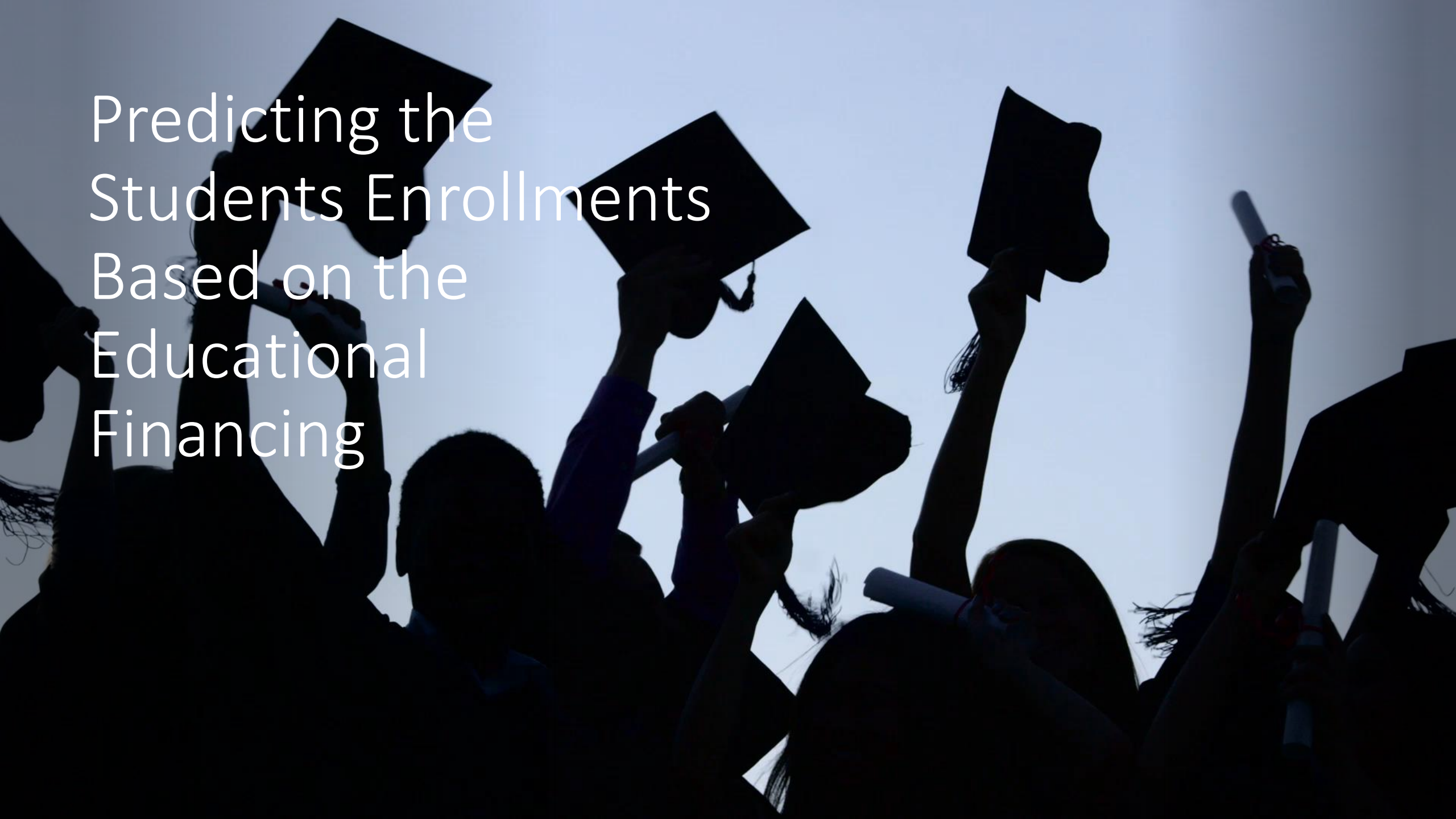


Predicting the Students Enrollments Based on the Educational Financing

Statement of the Problems

- School district organization by strategic planning concern the elementary student attrition and retention, and student enrollment are highly related to the financing management activities;
- The institutional support on the enrollment concerns that the secondary school student graduation college choice, the transition to colleges are related to the allocation of financial aid and large expenditure of awards.
- Considering the state, federal revenue and spending on instruction, this research explores different factors that presumably influence the enrollment amount of the school districts.
- This work intends to provide decision-makers in the enrollment management administration a better understanding of the factors that are highly correlated to the enrollment capacity.
- We validate our methods using real data of NAEP (National Assessment of Educational Progress) data and Educational financing data from the National Finance Census (National Center for Educational Statistics).
- Linear regression (LR) and predict the enrollment of district and K-nearest Neighbor (KNN) on cohort levels.

Research Questions

What are the correlations between the features of educational financing and the elementary and secondary students' enrollments of local school districts?

What is the model fit for the relevant educational financing features to predict the likelihood of enrollment in the next school year?

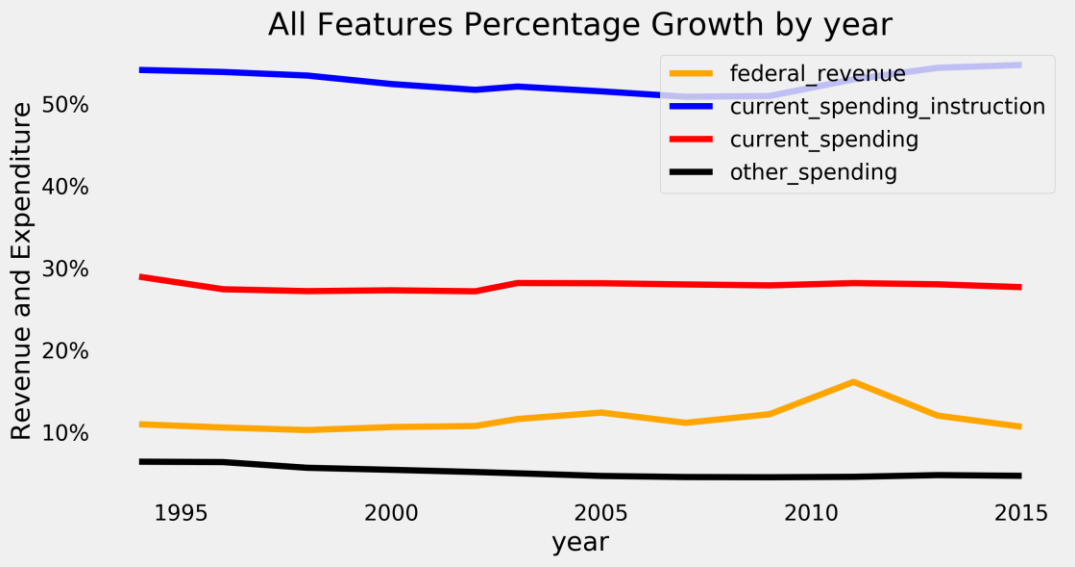
What is the recommendation to the stakeholders on financing management to increase the enrollment of elementary and secondary students?

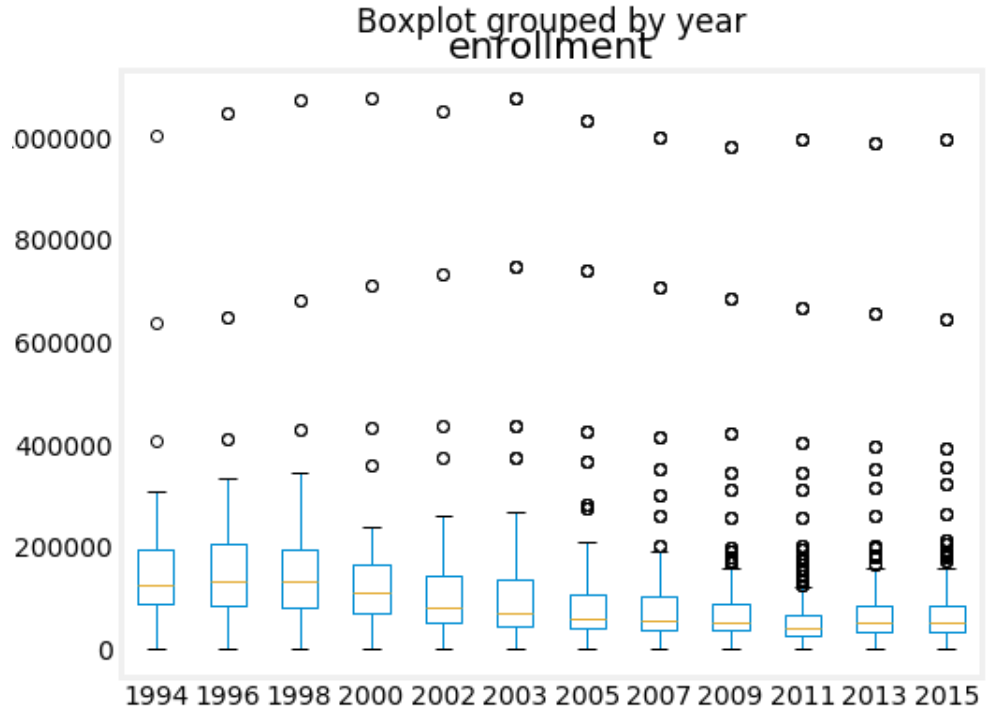
Metrics and Feature Description

District Enrollment: overall enrollment number of elementary and secondary students aggregated in districts
Continuous: number count
Categorical: number count scale: < 50000, 50000 ~ 70000, > 70000

Features	Description	Feature Selection
Total elementary-secondary revenue	Sum of Revenue from Federal, State, Local Sources	*
Total Revenue from Federal Sources	• Compensatory (Title I) • federal aid	*
Total Revenue from State Sources	• General formula assistance • Special education programs • Transportation programs ……	
Total Revenue from Local Sources	• All taxes • Property taxes • Parent government contributions • Revenue from other school systems • Charges	*
Total elementary-secondary expenditure	Sum of current spending, payment to the government, total capitals.	*
Total Current Spending for Instruction	• Total salaries and wages • Total employee benefit payments	*
Other Current Spending	• Salaries and wages for instruction • Employee benefits for instruction	*
Total Capital Outlay Expenditure		

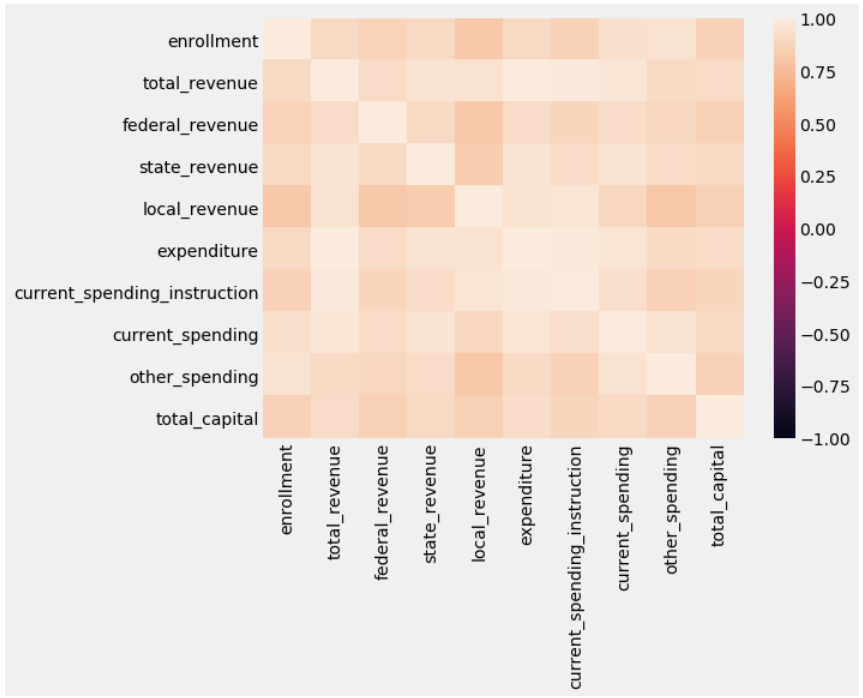
Education Financing Data Explorative Analysis

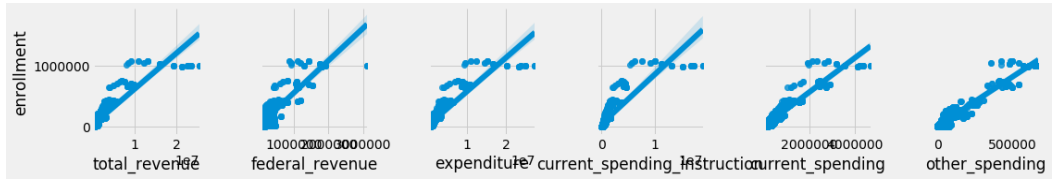




Features Correlation

- Features of revenue and spending have highly correlation to the enrollments on district levels.





Significant Predictors

OLS Regression Results

Dep. Variable:	enrollment	R-squared:	0.924
Model:	OLS	Adj. R-squared:	0.924
Method:	Least Squares	F-statistic:	7784.
Date:	Sat, 13 Feb 2021	Prob (F-statistic):	0.00
Time:	18:41:33	Log-Likelihood:	-45890.
No. Observations:	3855	AIC:	9.179e+04
Df Residuals:	3848	BIC:	9.184e+04
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	4230.3263	814.701	5.192	0.000	2633.039	5827.613
total_revenue	-0.0724	0.005	-13.166	0.000	-0.083	-0.062
federal_revenue	-0.0502	0.008	-6.654	0.000	-0.065	-0.035
expenditure	0.0493	0.006	8.693	0.000	0.038	0.060
current_spending_instruction	0.0159	0.006	2.781	0.005	0.005	0.027
current_spending	0.1727	0.010	16.700	0.000	0.152	0.193
other_spending	1.0255	0.027	38.685	0.000	0.974	1.077

Omnibus:	2230.006	Durbin-Watson:	0.582
Prob(Omnibus):	0.000	Jarque-Bera (JB):	136071.575
Skew:	1.988	Prob(JB):	0.00
Kurtosis:	31.833	Cond. No.	5.17e+06

Warnings:

- Multiple Regression indicate the statistically significant predictors of the Features.
- The selected features have main effect on enrollment,
- p <.001

Performance Evaluation

LR Algorithm with Feature Selection

- *Enrollments* = 4230.32 -0.1 * total_revenue - 0.1 *federal_revenue

  + 0.1 *expenditure+ 0.02 *current_spending_instruction

  + 0.2 *current_spending + 1.0 *other_spending

- MAE: 19949.20
- RMSE: 30726.0

Regularized Regression Algorithm after compare the features

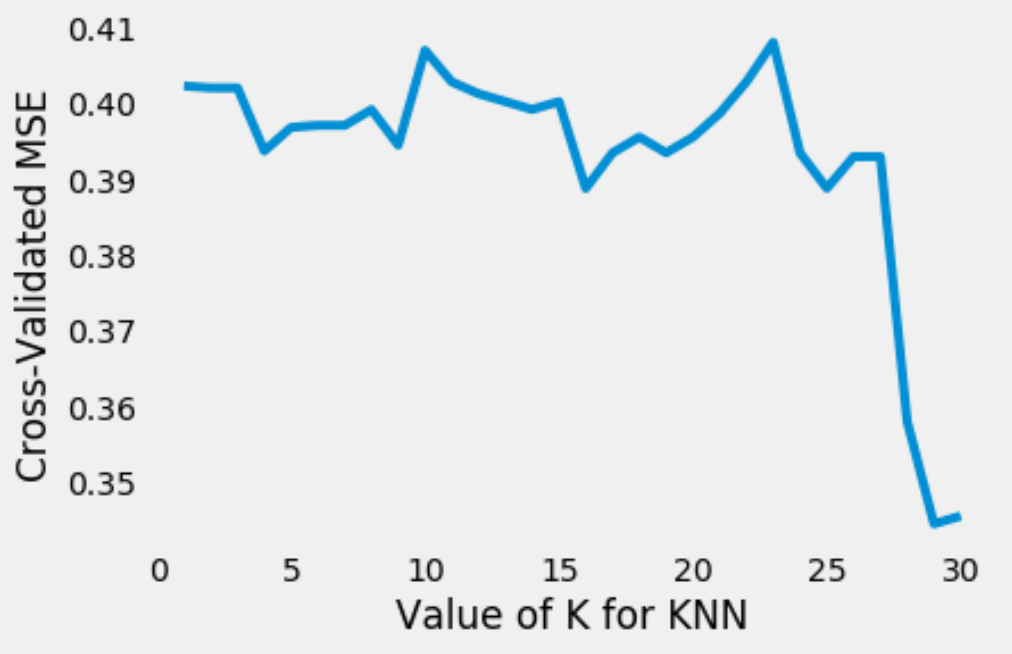
- *Enrollments* = 4124 -0.1 * total_revenue - 0.1*federal_revenue

  + 0.16 *current_spending + .02 *current_spending_instruction + 1.1 *other_spending

- RMSE: 30554.39 (alpha=0.0)

K Nearest Neighbors

- KNN is used to create the classifier to predict the enrollment with a result or roughly 93.14 % accuracy.
- 5-folds Cross validation was used to evaluate the KNN algorithm
- Both the accuracy and MSE indicated the best K is 28 ~ 29.



Comparison of LR and KNN

- The proposed models of LR algorithm and KNN can predict enrollment with reliable accuracy.
- LR Algorithm can be use for the projection of next year enrollment.

	LR	KNN
Train Data	2874	2891
Test Data	958	964
Accuracy	88.6%	93.14%

Findings & Implication

- The results indicated the spending on the instruction, payments to the government, and institutional supports are more likely to increase the district enrollments.
- Expanding the federal and total revenue such as allocation of financial aid and scholarship is not the best solution to maximize the enrollments for elementary and secondary students .
- Using the projection of LR algorithm and KNN by applying the enrollment from last year to project the enrollment of next year can help the management to assign money rewards efficiently.

Recommendation for Stakeholders

- The elementary and secondary students' cohorts on the district level have large variation for the management of educational financing and budget the next year compared to the school year.
- In terms of planning and evaluating the enrollments, using K-nearest neighbor would directly predict the enrollment of the cohort without identifying the variation on the school district's levels.
- For the research purpose, the institutional strategic planning would approach the enrollment by implement LR projection.

Acknowledgement

General Assembly
Data Scientist
Part-time Course

Questions & Answers

yuzhang1@seattleschools.org