

# AI Credit Scoring System For University Students

Advisor : Prof . YoungJoo Shin / Team : PassAbility / Members : HanByeol Jang  
YoungSang Kwon  
YongRak Choi

## Overview

- Needs to check credit score automatically
- For university students who can't be checked official credit score because of the loss of details.

## SW Architecture

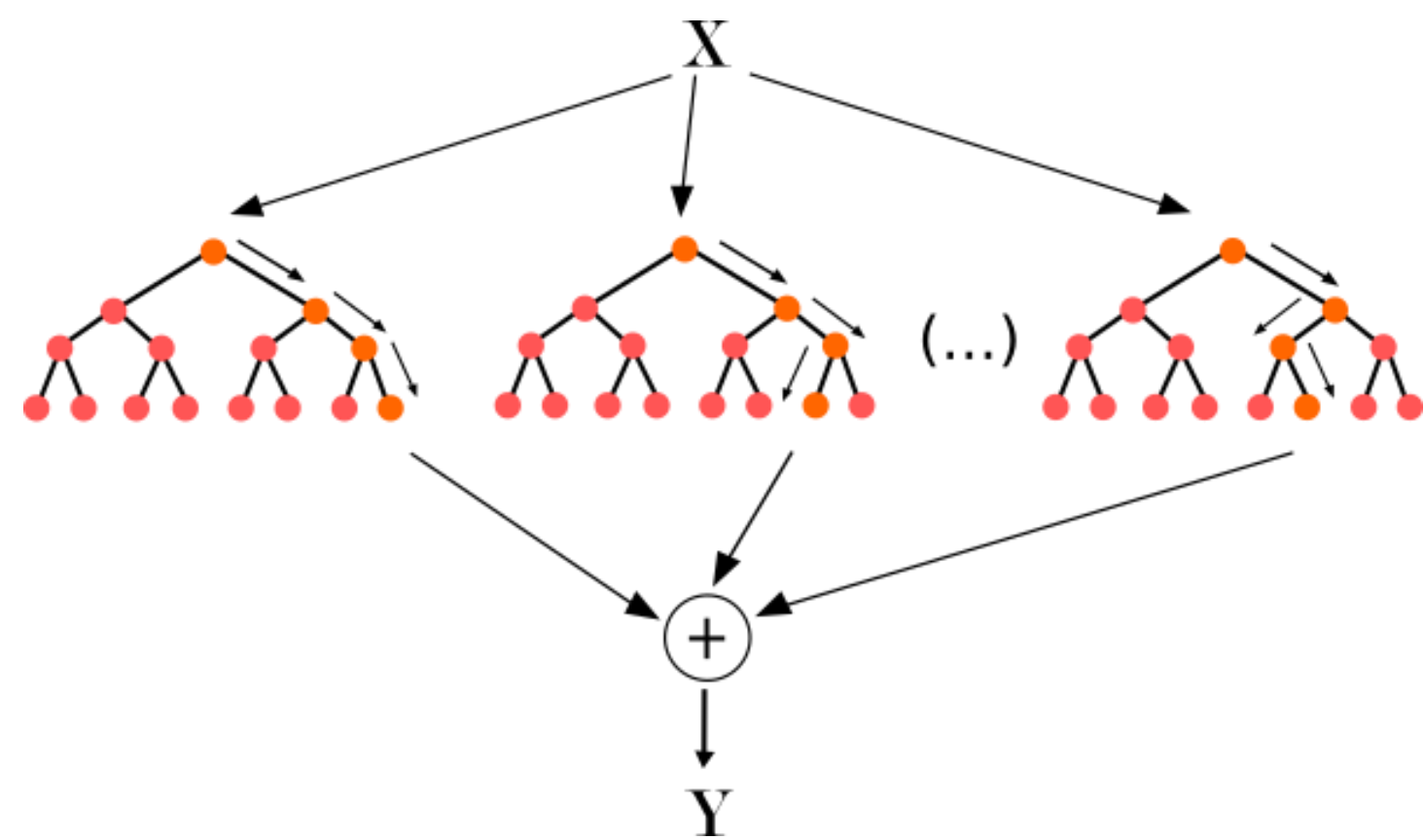
- Machine learning for learning model of financial data with nonlinear characteristics affected by external environment variables.
- Credit scoring with RandomForest, LighGBM and XGBoost that are classification algorithms
  - After that, cross validation for each models and ensemble them to improve the accuracy

## Random Forest

- The disadvantage of the decision tree is large variation in performance, which can be overcome by using randomization technologies such as RandomForest Algorithm's Bagging or Randomized node optimization
- Validation methods : Cross validation, Oob(Out of Bag) validation

Parameter	Value
n_estimators	300
max_features	auto
random_state	84
oob_score	True

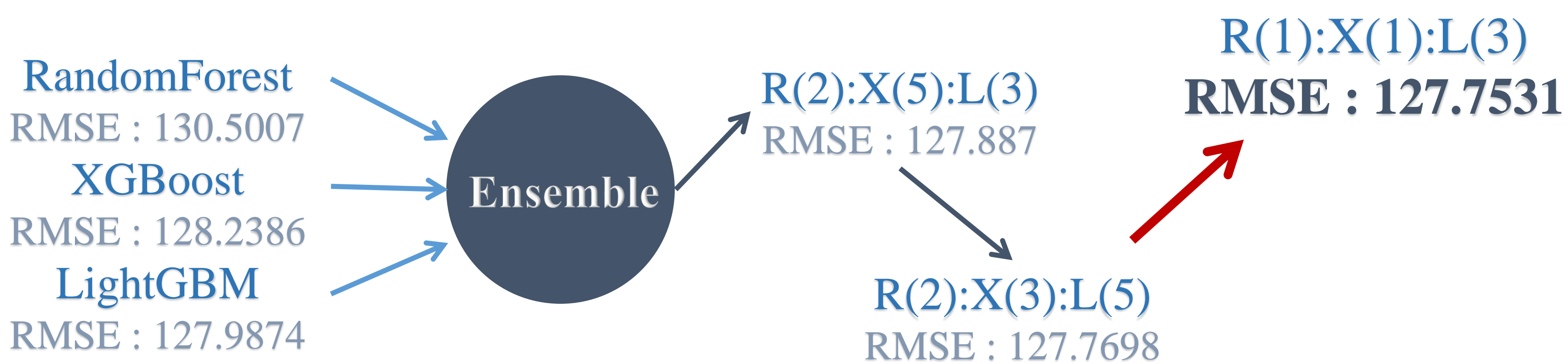
- Parameter values -



- Tree Ensemble process -

## Results of Ensemble

- Ensemble Random Forest, XGBoost, LightGBM to derive results
- Accuracy indicator criteria  
RMSE (Root Mean Squared Error)
- Ensemble  
That combine several Machine Learning techniques in order to improve Machine Learning results



## Conclusion

- LightGBM is faster compared to any other method. Besides, better results. And more exactly after ensemble
- Utilization Plans
  - Using the AI Credit Scoring System, we can simulate the credit evaluation of university students in the bank and classify them to some degrees
  - As a result, It makes more efficient to attract and manage potential customers who can't be checked credit score like university students

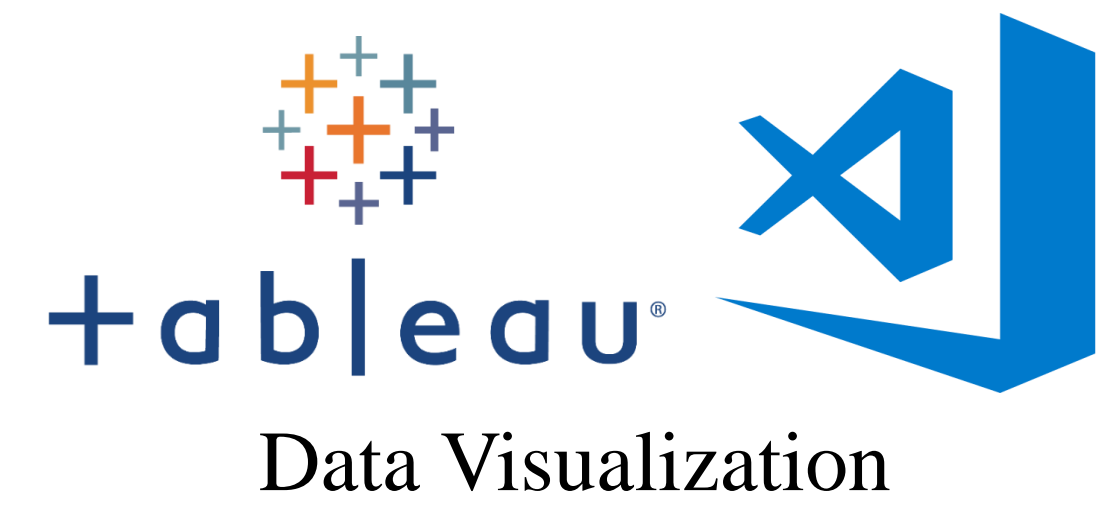
## IDE & Editor



For several machine learning algorithms



Data Preprocessing  
& Data Augmentation



Data Visualization  
& UI Design

## Data Composition



To get the datasets  
of Economics and Finance



For Categorization

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	15	disbursl	asset	value	term	type	disbursl_date	education	transport	travel	mobile	medical	count_loan	count_loan	count_loan	count_loan	count_loan	count_loan	count_loan	count_loan	count_loan	count_loan
2	1	10000	10000	10000	10000	student	2000-01-01	6	1	0	0	1	0	0	0	0	0	0	0	0	0	0
3	2	47450	15500	7123	12-07-25	student	2000-09-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	3	12170	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
5	4	57113	16111	8848	20-12-23	student	2002-10-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
6	5	12170	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
7	6	54513	13300	8848	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
8	7	48480	13300	7843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
9	8	48984	13300	7189	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
10	9	57113	13300	8843	20-12-23	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
11	10	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
12	11	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
13	12	48700	74100	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
14	13	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
15	14	48700	74100	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
16	15	48700	74100	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
17	16	48700	74100	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
18	17	57113	13300	8843	20-12-23	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
19	18	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
20	19	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
21	20	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
22	21	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
23	22	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
24	23	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
25	24	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
26	25	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
27	26	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
28	27	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
29	28	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
30	29	52003	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1
31	30	54513	13300	8843	20-08-05	student	2001-08-18	6	1	0	0	1	1	1	1	1	1	1	1	1	1	1

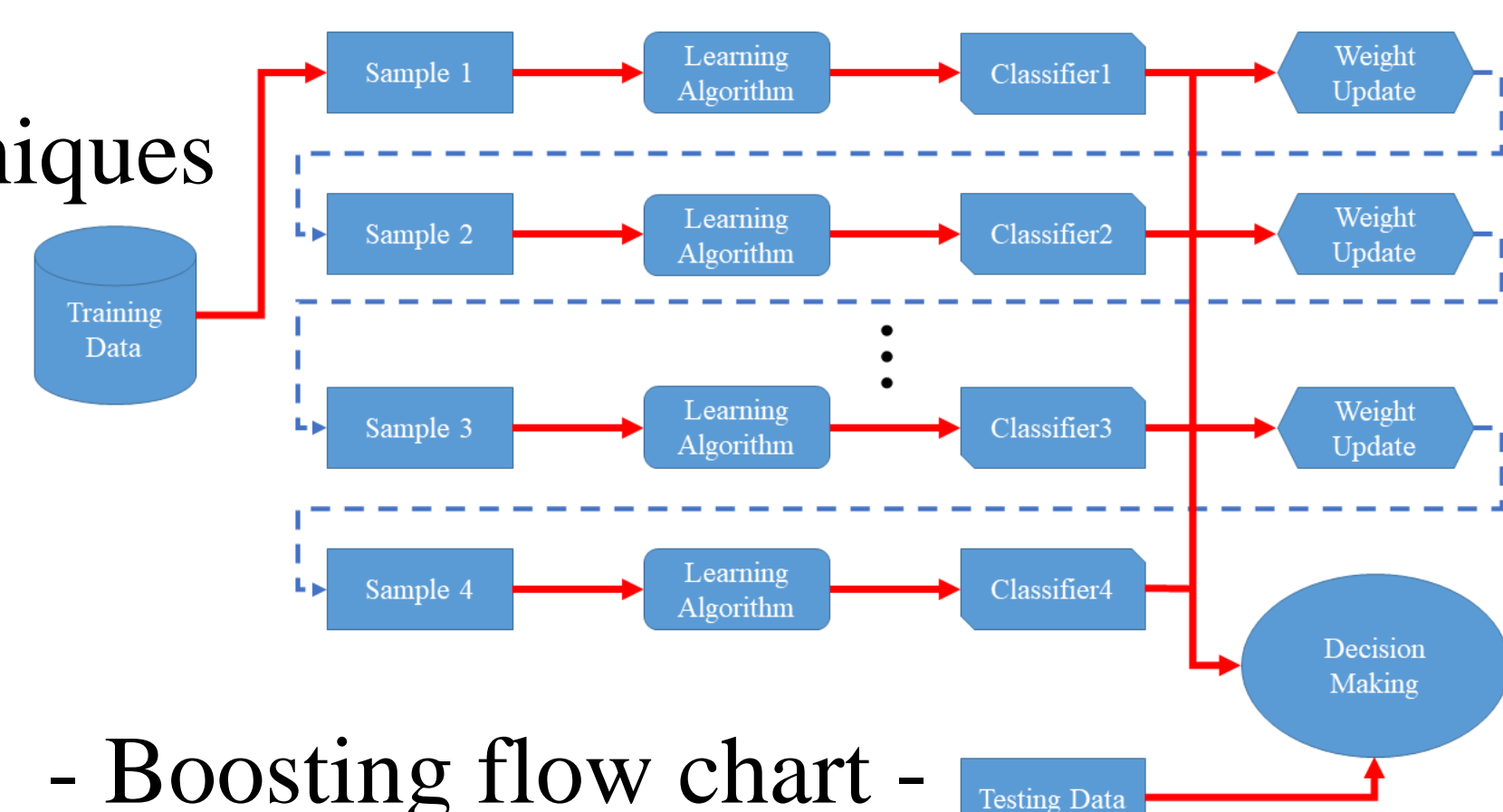
After that,  
Data Preprocessing  
& Data Augmentation in R

Microsoft  
SQL Server  
Database

## Gradient Boosting

- Unlike Random Forest, which scores as an ensemble among independent trees, the Gradient Boosting complements the residual of the previous classifier to create a strong classifier
- Gradient Boosting used in model implementation
  - LightGBM, XGBoost
- XGBoost
  - Using parallel processing techniques in decision tree configuration
    - ✓ Increase performance
- Performance comparison
  - LightGBM > XGBoost

- Boosting flow chart -



## UI Design

