

This app can automatically analyze big data classification, regression, and time series.

Version 20230313

# Applications

• It can be used in big data analysis, stock price inference, financial index inference, student grade forecast, air pollution forecast, Prediction of survival rate of critically ill patients, etc.

- It has three modes available:
  - Inference : The inference of a single algorithm.
  - Auto Analysis : Select multiple algorithms for automatic training and inference.
  - Optimize X : Find the set of X factors that are close to the target result Y from the set X factor value interval.

						C:/Uters/ie/Desktop/V	W-8022ATA-erc-20220313;8gDeta-22;8eta;Caserfoxfor-Titaric;Inairing_input.cnv
u	ers'ai Desktop/J	PP-BIODATA-	ec-20230313/Big	Data-22 idata Class	ification-Titanic In	sining input.csv	
	Passengerid	Pclass	Age	Fare	Sex	Survived	
		3	22	7.25	0	0	
		1	38	71.2833	1	1	
		3	26	7.925	1	1	
ļ		1	35	53.1	1	1	
		3	35	8.05	0	0	
1		3	29.69911765	8.4583	0	0	
		1	54	\$1.8625	0	0	
1		3	2	21.075	0	0	
1		3	27	11.1333	1	1	
ł	0	2	14	30.0708	1	1	
	1	3	4	16.7	1	1	
ł	2	1	58	26.55	1	1	
ł	3	3	20	8.05	0	0	
ĺ	4	3	39	31.275	0	0	
	5	3	14	7.8542	1	0	
ł	6	2	55	16	1	1	
ł	7	3	2	29.125	0	0	
ŀ	8	2	29.69911765	13	0	1	
ł	9	3	31	18	1	0	
ł	0	3	29.69911765	7.225	1	1	



#### • File

- Open : open file.
- Save : save the current window.
- Save As : save the current window as a new file.
- Close File : close file.
- Close All Files : close fll files.
- Browse Data : browse data folder.
- Close App : close App.

BGD L	EADERG BigData	a - Version 2	3		
File	Auto Analysis	Inference	Optimize X	Document	LibreOffice Calc
	Open				
	Save				
	Save As				
	Close File				
	Close All Files				
	Browse Data				
	Close App				

- Inference : The inference of a single algorithm.
- Auto Analysis : Select multiple algorithms for automatic training and inference.
- Optimize X : Find the set of X factors that are close to the target result Y from the set X factor value interval.
- The analysis includes the following three items:
  - Classification :
    - XGBoost ` LightGBM ` Gradient Boosting.
  - Regression :
    - XGBoost `LightGBM `Conv1D `DNN `Gradient Boosting `Linear `Polynomial `PCA `SVM.
  - TimeSeries:
    - XGBoost , LSTM.





- Document : Open Document.
- Spreadsheet : Open Office.

#### LEADERG BigData - Version 23

File Auto Analysis Inference Optimize X Document LibreOffice Calc

🖬 ປ	Intitled 1	- LibreOf	fice Calc								- C		×
<u>F</u> ile	Edit \	<u>/</u> iew <u>I</u> n	sert F <u>o</u> rmat	Styles Shee	et <u>D</u> ata <u>T</u> oo	ls <u>W</u> indow	<u>H</u> elp						×
₿.	- 📄 - 🖥	□	l 🖨 🔁	X 🖻 🔂 •	🎸 👍   🗧	• ~   A	abc   📰 🗸 🚺	- <u>A</u>	Z 🕹 🏹   🎦	Ω 🖾 🖬	• 🔮 🖵 📋		>>
Arial			~ 10	opt ∨ B	I <u>U</u> - A	• 🕫 • 📄			≟   ጬ - %	6 0.0 🔁 🛛 .	ΩΩ .0Q   🔁 🤅		>>
A1			$\sim   f_X \sum -$	=								$\bullet$	=
	А		в	С	D	E	F	G	н		J		=
1													Δ.
2												- 11	8
3													
5													
6													c
7													Jx
8												_	
9												- 1	
10												- 1	
12												- 1	
13													
14													
15												- 1	
16												~	
<												>	
- M		+ 9	Sheet1									-	
Shee	et 1 of 1			Defau	ilt	English (US/	A)	-I G	Average:	; Sum: 0		-+ 1	00%

• Select an item to be automatically analyzed.

 LEADERG BigData - Version 23
 File Auto Analysis Inference Optimize X Document LibreOffice Calc
 Classification Regression Time Series

• Select the training and inference file.

BGD [Classifi	cation] Please select a train file.		×	[Classification] Please select an inference file.		×
$\leftarrow  \rightarrow$	Y ↑	ن المحمد المعنون المحمد الم المحمد المحمد المح المحمد المحمد المحمم المحمد المحمد المحمد المحمد المحمد المحمد المحمد المحمد المحمد المحم المحمد المحمد المحمد المحمد المحمم محمد محمم محمد م	ation-Titanic	$\leftarrow$ $\rightarrow$ $\checkmark$ $\uparrow$ $\blacksquare$ « data > Classification $\checkmark$	Search Classifica	ation-Titanic
Organize	<ul> <li>New folder</li> </ul>	*==- *= <b>*</b>		Organize - New folder		□ ?
^	Name	Date modified	Туре	Name	Date modified	Туре
	inference_input.csv	4/12/2021 1:48 PM	Microsoft Exce	inference_input.csv	4/12/2021 1:48 PM	Microsoft Exce
	🛂 train_input.csv	7/18/2022 1:45 PM	Microsoft Exce	Train_input.csv	7/18/2022 1:45 PM	Microsoft Exce
	<u>x</u>		>			>
	File name: train_input.csv	<ul> <li>Files(*.csv *.CSV *.XLS)</li> <li>Open</li> </ul>	X *.xlsx .txt ~ Cancel	File name: inference_input.csv	V Files(*.csv *.CSV *.XLS)	<pre>K*.xlsx .txt ~ Cancel</pre>

• Select the algorithm for analysis. After selecting, please click this item and set the parameters in the Setting area on the right side of the window.

EADERG BigData - Version 23 File Auto Analysis Inference Optimize X Document LibreOffice Calc		- 0	×
C:/Users/al/Desktop/APP-BIGDATA-src-20230313/BigData-22/data/Classification-Titanic/training_input.csv 🔀 C:/Users/al/Desktop/APP-BIGDATA-src-20230313/BigData-22/data/Classification-	Titanic/inference_input.csv	Auto Analysis - Classification	×
select       training input : C:/Users/ai/Desktop/APP-BIGDATA-src-20230313/BigData-22/data/Classification-Titan         select       inference input : C:/Users/ai/Desktop/APP-BIGDATA-src-20230313/BigData-22/data/Classification-Titan         select       inference input : C:/Users/ai/Desktop/APP-BIGDATA-src-20230313/BigData-22/data/Classification-Titan         start       start	ic/training_input.csv tanic/inference_input.c	·sv	
□ XGBoost_Classification	GradientBoosting_Classification		
GradientBoosting_Classification	SDK		
Select algorithm	numEstimator Setting	parameters	

• After selecting the algorithm and setting the parameters, press Start to start the analysis. Messages during analysis will be displayed in the Log window below.

C:/Users/ai/Desktop/APP-BIGDATA-src-20230313/BigData-22/data/Classification-Titanic/inference_input.csv 🔟 Auto Analysis - Classification 🔀 train_output.	.csv 🔀 training_result 🔀 inference_output.csv 🔀 inference_result 🔀
select training input : C:/Users/ai/Desktop/APP-BIGDATA-src-20230313/BigData-22/data/Classification	on-Titanic/training input.csv
cite information impute ()///Loars/si/Dealston/ADD DICDATA and 20220212/DiaData 22/data/Classifian	c
miletence input : C//Users/al/Desktop/AFF-BIODATA-src-20250515/BigData-22/data/Classifica	ation-Thanic/Interence_Input.csv
start stop Start and Stop	
XGBoost Classification [inference completed]	Setting
Abboost addition (interface completed)	YGBoost Classification
LightGBM Classification	XODOUSC_Classification
LightGBM_Llassification GradientBoosting_Classification	SDK
LightGBM_Classification GradientBoosting_Classification	SDK
LightGBM_Classification GradientBoosting_Classification	SDK
LightGBM_Classification         GradientBoosting_Classification         Image: State of the	SDK
LightGBM_Classification         GradientBoosting_Classification         ====== XGBoost_Classification training ======         ccuracy = 98.75 %         IPython.core.display.HTML object>         ====== XGBoost_Classification inference ======         ccuracy = 77.19 %         IPython.core.display.HTML object>         IPython.core.display.HTML object>	SDK

• The training and inference are completed, and the training and inference results are output.

train_output.csv 🗵	train_result 🗵	inference_output.csv 🗵	inference_result 🗵
ELEDERG BigBass - Vencion 19 - Analyzing     File Training Inference Audo Analyzin Optimize X Document LikerOffice Calc     CharacterizeDeskop/IPP 00224 - 00222222 desexa/bigBass - Mitlen/CalminStation Transchara - ges.ct.er      CharacterizeDeskop/IPP 00224	TA 412 2022 1230 Mekana (NgSuka -15) Kata, Chanafaston Tainc, Inference, yout, an 💿 Auto. And you's - Classification 📄 trans.	- O - X bany_cav i bany_taken_capacity i herence_taken	
XGBoost_Classification result : Accuracy = 98.75 % Feature Importance		^	
Pare	Toature importances		
Pclass -6 -4 -2 0 2 4 6 8 Low SHAP value (impact on model output)			
Sex 42.11	+2.69		
Age +1.58 Pclass +1.2 00 0/5 1/0 1/5 2/0 2/5		×	

Sort b	train_output.csv inference_output.c	sv 🔝 interence_result 🔝 train_output.csv	train_resuit		
			Window Snip		
		train accuracy(%)	inference accuracy(%)	average accuracy(%)	
	XGBoost_Classification	98.75	77.19	87.97	
	GradientBoosting_Classification	98.75	76.78	87.765	1
	LightGBM_Classification	97.5	74.34	85.92	1
					-

#### • Open the inference file. Open -> File.

EADERG BigData - Version 23

– 0 ×

File Auto Analysis Inference Optimize X Document LibreOffice Calc

	Passengerld	Pclass	Age	Fare	Sex	Survived	
4	401	3	39	7.925	0	1	
2	402	3	26	8.05	0	0	
3 4	403	3	21	9.825	1	0	
4	404	3	28	15.85	0	0	
5 4	405	3	20	8.6625	1	0	
5 4	406	2	34	21	0	0	
7 4	407	3	51	7.75	0	0	
8 4	408	2	3	18.75	0	1	
9 4	409	3	21	7.775	0	0	
10 4	410	3	29.69911765	25.4667	1	0	

• Select the inference algorithm.

	ERG BigData - Versior	23					
File	Auto Analysis	Inference	Optimize	Х	Document	LibreOffice	Calc
		Classifie	cation 🔹 🕨		XGBoost		
		Regress	sion 🕨		LightGBM		
		Time Se	eries 🕨 🕨		Gradient Bo	posting	

• Select the model and pkl file for inference.

Please select XGBoost_Classification model	×	Please select XGBoost_Classification scaler	×
$\leftarrow \rightarrow \checkmark \uparrow$ . XGBoost_Clas > model $\checkmark$ O	✓ Search model	$\leftarrow \rightarrow \checkmark \uparrow$ ] « XGBoost_Class > model	✓ ひ Search model
Organize   New folder	:==	Organize   New folder	III <b>▼</b> []
Name Date modified	d Type Si	Name	Date modified Type Si
<b>model.sav</b> 7/22/2022 3:5	2 PM SAV File	scaler.pkl	7/22/2022 3:52 PM PKL File
File name	Filos(ž sou ž oth)	File namer scalar pkl	
	Open Cancel	File name: scaler.pki	Open Cancel

- Get training results, including:
  - result csv : predicted results.
  - result : Display inference results such as accuracy, curve graph, etc.

Auto Analysis	Inference Opt	mize X Document	LibreOffice Cald	:				nie AutoAnalysis interence Optimie X. Document: Lancottice Calc C. (JusephiloPedetau/PHC) (Statistica) (Stati	inference_result
rs/ai/Desktop/APP-BIGDA	TA-src-20230313/BigData-22/data	Classification-Titanic/inference_inp	ut.csv 🗵	inference_ou	utput.csv		inference_result		
Boost_Classif	ication inference res	alt :						XGBoost Classification result :	
Passenger	rld Pclass	Age	Fare	Sex	Survived	prediction	^		
401	3	39	7.925	0	1	0		Accuracy = 77.19 %	
402	3	26	8.05	0	0	0		High	
403	3	21	9.825	1	0	0		Sex B	
404	3	28	15.85	0	0	0			
405	3	20	8.6625	1	0	0		Pclass	
406	2	34	21	0	0	0		-6 -4 -2 0 2 4 6	
407	3	51	7.75	0	0	0		SHAP value (impact on model output)	
408	2	3	18.75	0	1	1	Vindow Snip	Sex 12.63	
409	3	21	7.775	0	0	0		Fare +1.98	
0 410	3	29.69911765	25.4667	1	0	0		Age +1.59	
		<< <	Page	1	of 25 >	>>	v 	Polass +1.14 00 0.5 10 rean(ISHP value)	

- Select an item to optimize X.
- EADERG BigData Version 23

File	Auto Analysis	Inference	Optimize X Document		LibreOffice Calc	
			Classificat	tion	•	XGBoost
			Regression	n	•	LightGBM
						Gradient Boosting

- Select file, file can be
  - train\_input, inference\_input with headers.
  - the optimizeX file set by the previous optimize X.
- Select the model file.
- Select the pkl file.

👳 Please se	lect csv		×	🔤 Pleas	e select XGBoost_Classification [model]		>	×	🔤 Please	e select XGBoost_Classification [scaler]		×
$\leftarrow \rightarrow$ ,	↑ 📜 « data > Classification	ע גע אין	rch Classification-Titanic	$\leftarrow \rightarrow$	✓ ↑	el ∨ Ŭ ∕ Sea	rch model		$\leftarrow \rightarrow$	✓ ↑	iel 🗸 🖸 🔎 Sear	rch model
Organize 🔻	New folder		:== - (?)	Organiz	ze 🔻 New folder		::: • 🔟 ?		Organiz	ze 🔻 New folder		:= • 🔳 ?
^	Name	Date modified	Type Si	^	Name	Date modified	Туре	Si	^	Name	Date modified	Type Si
*	XGBoost_Classification	8/24/2022 04:37 PM	File folder	*	📄 model.sav	8/24/2022 04:37 PM	SAV File		<b>*</b>	scaler.pkl	8/24/2022 04:37 PM	PKL File
÷	inference_input.csv	4/12/2021 01:48 PM	Microsoft Excel Co	<b>6</b> 7					<b>*</b>			
	🛂 optimizeX.csv	8/23/2022 09:51 AM	Microsoft Excel Co									
	uin_input.csv	7/18/2022 01:45 PM	Microsoft Excel Co									
~ <			>	~	<			>	~	<		>
	File name: optimizeX.csv	<ul> <li>✓ Files(*.csv</li> </ul>	*.CSV *.XLSX *.xlsx .txt $\vee$		File name: model.sav	✓ Files(*.sav	) ~			File name: scaler.pkl	<ul> <li>✓ Files(*.pkl</li> </ul>	) ~
		Oper	n Cancel			Ope	n Cancel				Oper	n Cancel

• Y target is the target value of the output.

- Sets the interval for each Xfactor value.
  - min is the minimum value of X.
  - max is the maximum value of X.
  - step is the interval.
  - For example, when Pcalss sets min=1, max=3, step=1, then Pclass = [1, 2, 3].



• Press Start to begin. Messages will be displayed in the Log window during the process.

LEADERG BigData - Version 2	3				CEADERG BigData - Version 23 – 🗇							
File Auto Analysis	Inference Optimi	ze X Document	LibreOffice Calc		ssification-Titanic/inference_input.csv	OptimizeX - XGBoost_Classification	optimizeX.log 🗵	inference_input_optimize_x.csv 🗵	inference_ouput_optimize_x.csv 🗵 4			
C:/Users/ai/Desktop/APP Start St Y target : 1 X range :	-BIGDATA-src-20230	313/BigData-22/data,	/Classification-Titanic/ii	ence_input.csv	Pclass : 3 Age : 3 Fare : 3 Sex : 3 Number of combinations : 81 generate optimize X generate optimize X end Open inference_input_optimize_x end ======= XGBoost_Classification inferen Accuracy = 39.51 % Open inference_ouput_optimize_x end	ce ======						
Pclass	Age	Fare	Sex		LUg							
min 1	0.42	0.0	0									
max 3	80.0	512.3292	1									
step 1.0	39.79	256.1646	0.5									

- Optimize X completion and output results.
- The data closest to the target Y value is ranked at the top.

cation-Titanic/inference_input.csv 🗵			OptimizeX - XGBoos	t_Classification 🔝	optimizeX.log 🔝	inference_input_	optimize_x.csv 🗵	inference_ouput_optimize_x.csv 🗵		
	Passengerld	Pclass	Age	Fare	Sex	targetY	prediction	prediction - targetY		
1	1	1.0	0.42	0.0	0.0	1	1	0		
2	23	1.0	80.0	256.1646	0.5	1	1	0		
3	26	1.0	80.0	512.3292	0.5	1	1	0		
4	27	1.0	80.0	512.3292	1.0	1	1	0		
5	28	2.0	0.42	0.0	0.0	1	1	0		
6	29	2.0	0.42	0.0	0.5	1	1	0		
7	30	2.0	0.42	0.0	1.0	1	1	0		
8	31	2.0	0.42	256.1646	0.0	1	1	O Mindow Snip		
9	34	2.0	0.42	512.3292	0.0	1	1	0		
10	2	1.0	0.42	0.0	0.5	1	1	0		
	42	2.0	40.04		1.0	4	4	^		

#### Browse the APP folder

• From the App4AI interface, find the target icon, right-click, and then click Browse to browse the APP folder.



### Reference

- Please refer to the readme.txt in the APP folder.
- LEADERG AppForAI : <u>https://www.leaderg.com/appforai-windows</u>
- Copyright © LEADERG INC. All rights reserved.