

odfmark=/PUT,Raw={ThisPage}«/Trans « /S /Replace » »

ack



CDMC2020 Task Presentation

Rongxin Xu, Xiaoyan Wang

2022-04-27

pdfmark=PUT,Raw=(ThisPage)/Trans/S/Replace
remark.eps
logos/tulip.eps

Overview

pdfmark=PUT,Raw=(ThisPage)/Trans/S/Resampling method

remark.eps

logos/tulip.eps

Resampling method

pdfmark-PUT,Raw-(ThisPage)/Trans/S/Replace-Remark.eps
logos/tulip.eps

Data set processing

Defn

According to the label distribution, stratified sampling is performed on the original training set to form a new training set and a validation set. After this, comprehensive sampling is performed on the new training set, and the obtained data set is used as the training set for model training.

slides-20201018-pics.pdf

Figure 1: split and resampling

pdfmark=/PUT,Raw=(ThisPage)/Trans/S/Replace/Remark.
Remark.
logos/tulip.eps

Cross-validation
Defn Use 5-fold cross-validation to evaluate the effect of the model, and get the average of accuracy and f1_macro.

Table 1: Mean value of accuracy and f1_macro under 5-fold cross validation

	mean_acc	mean_f1_macro
XGBoost	96.10%	96.14%
LightGBM	95.89%	95.94%
Randomforest	96.17%	96.21%

pdfmark=/PUT,Raw=(ThisPage)/Trans/S/Replace=

remark.eps

logos/tulip.eps

Defn

Use three models to predict the validation set and observe the effect.

Table 2: The accuracy and f1_macro on the validation set.

	accuracy	f1_macro
XGBoost	91%	77%
LightGBM	91%	78%
Randomforest	92%	79%

pdfmark-PUT,Raw-(ThisPage)/Trans/S/Replace	One-Class SVM
remark.eps	
logos/tulip.eps	Defn Use the data of each category to train a single classifier and then integrate it with XGBoost.

slides-20201018-pics.pdf

pdfmark-PUT,Raw-(ThisPage)/Trans/S/Replace	result
remark.eps	
logos/tulip.eps	

Defn	The accuracy rate is 67%, and f1_macro is 37%.
------	--

pdfmark=/PUT,Raw=(ThisPage)/Tr/S/Replaces
remark.eps

logos/tulip.eps

More work and question

pdfmark=/PUT,Raw=(ThisPage)/Trans/S/Replace=question

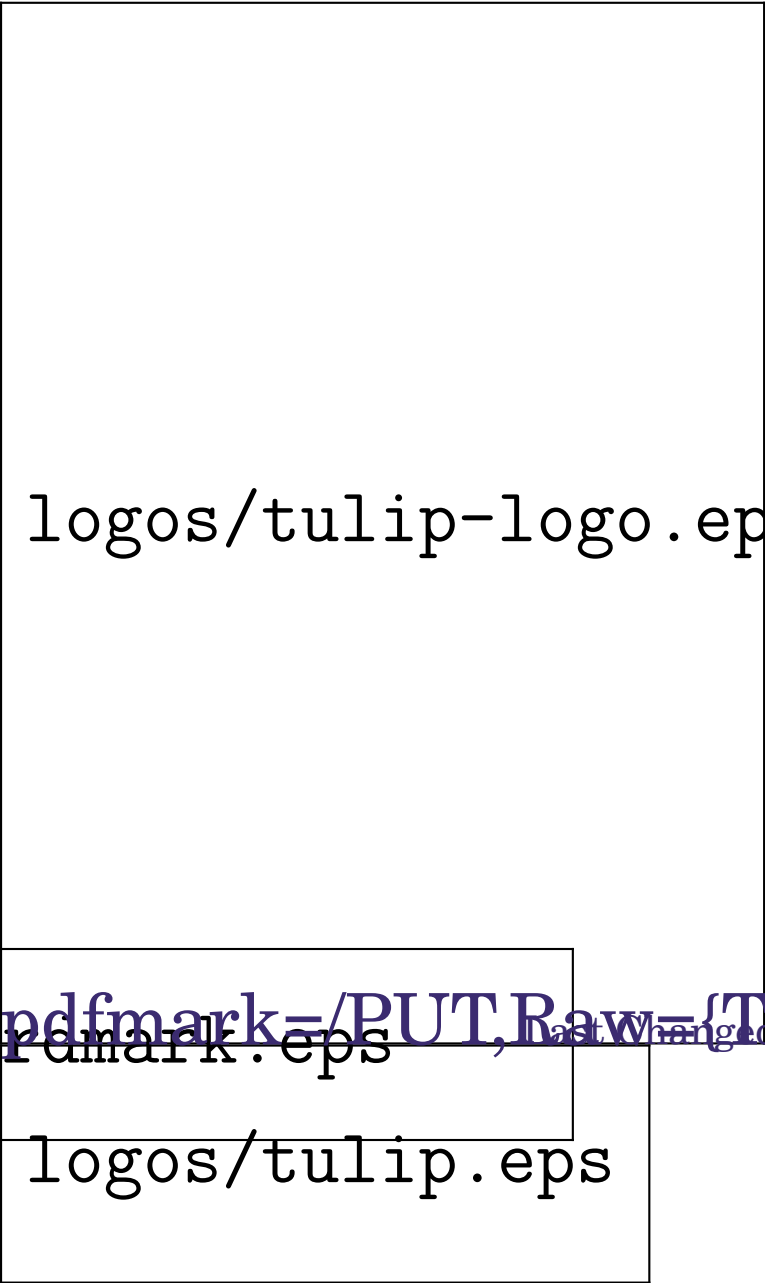
remark.eps

logos/tulip.eps

■ Hierarchial VAE

■ Whether task1 wants to predict two target values?

slides-20201018-pics.pdf



pdfmark=PUT Raw=(This Page)/Trans/S/Replace/End
remark.eps, % Warning: If you use the file, please cite it.

logos/tulip.eps

Temporary page!

L^AT_EX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it. If you rerun the document (without altering it) this surplus page will go away, because L^AT_EX now knows how many pages to expect for this document.