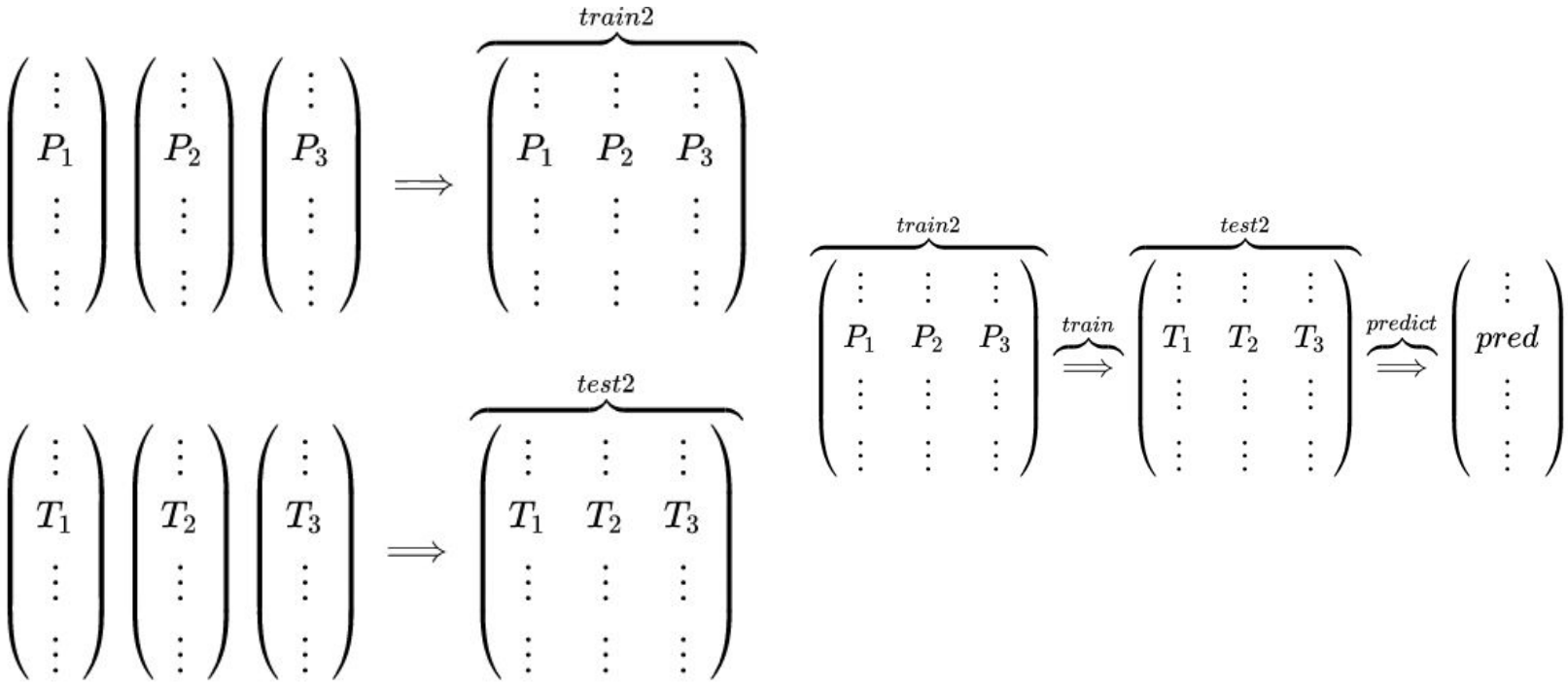
模型融合—Stacking

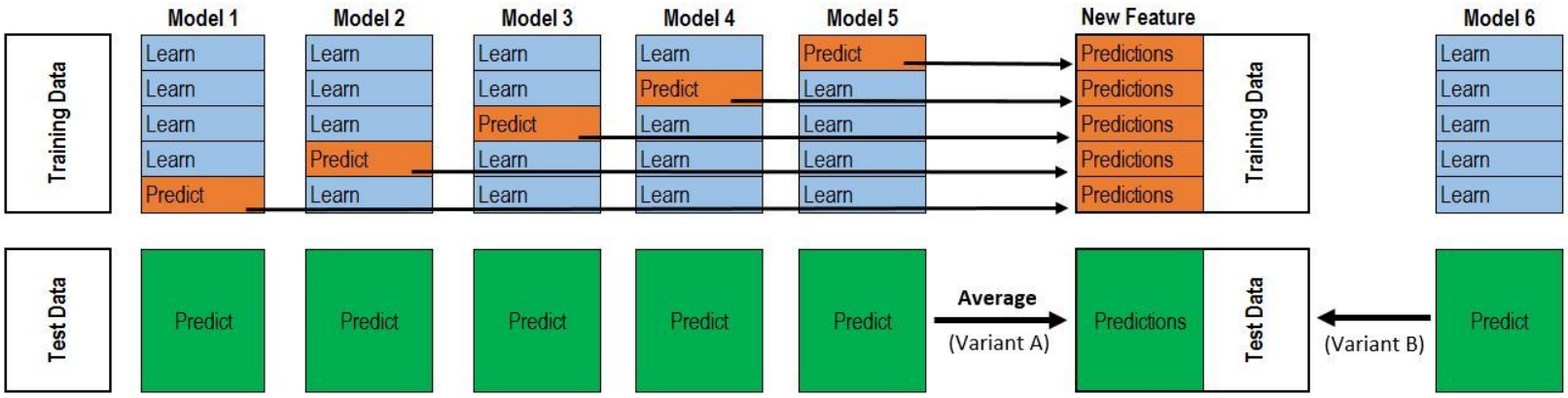
方敏
驰声研发

摘要

- 基本思想及方法
- 实验结果

1、基本思想及方法




Model 1
Model 2
Model 3
Model 4
Model 5
New Feature
Model 6
Training Data
Learn
Predict
Predictions
Test Data
Average
(Variant A)
(Variant B)

2、实验结果

- 分类任务

```
||
|| combine1          ||  98.29%   || 70.71%   ||
|| combine2          ||  98.07%   || 69.64%   ||
|| combine3          ||  100%     || 69.96%   ||
|| combine4          ||  100%     || 65.67%   ||
|| stacking1         ||  68.92%   || 70.92%   ||
说明：
combine1: unigram-TF + unigram-normTF + bigram-TF + bigram-one-hot -> svm-gauss
combine2: unigram-TF + bigram-one-hot -> svm-gauss
combine3: unigram-TF + bigram-one-hot + trigram-one-hot -> svm-gauss
combine4: P_svm-gauss_unigram-TF + P_svm-gauss_unigram-normTF + P_svm-gauss_bigram-TF + P_svm-gauss_bigram-one-hot -> svm-gauss
stacking1: unigram-TF + unigram-normTF + bigram-TF + bigram-one-hot + trigram-one-hot > svm-gauss -> svm-gauss
```

- 回归任务

```
||-----------------------
|| baseline            || 0.8253/0.9095 || 0.8090/0.9032    ||
|| combine             || 0.8317/0.9113 || 0.8047/0.9054    ||
||stacking1            || 0.7856/0.8935 || 0.8176/0.9070    ||
说明：
combine:unigram-TF-disc + bigram-TF-disc + bigram-one-hot-cont + trigram-tf-cont + trigram-one-hot-disc + basefeature
stacking1：base-fea +  unigram-TF + unigram-normTF + bigram-TF + bigram-one-hot + trigram-one-hot -> stacking -> svr-gauss
```

END