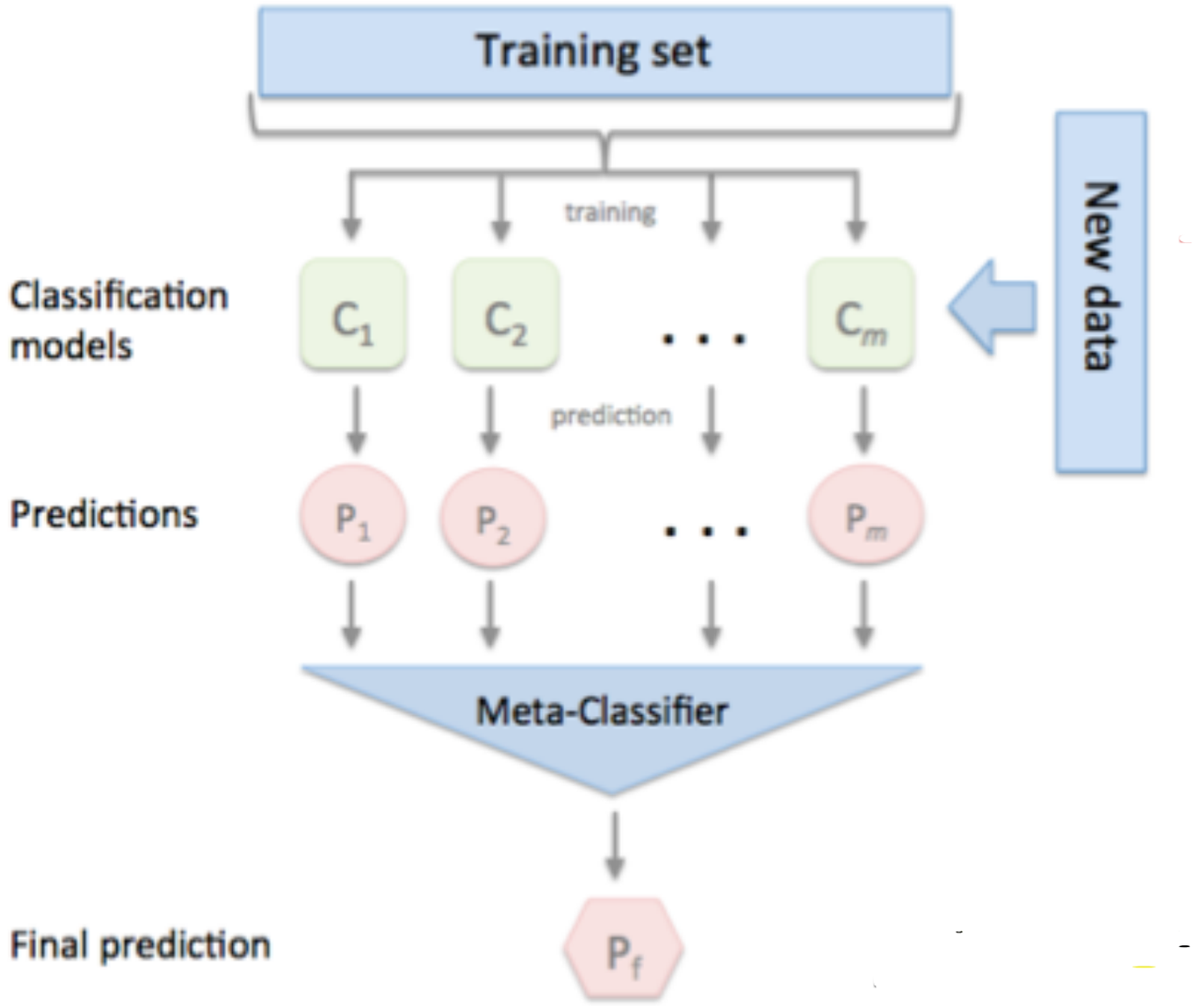
集成学习

Stacking

Stacking


Training set
training
New data
Classification models
C_1
C_2
C_m
prediction
Predictions
P_1
P_2
P_m
Meta-Classifier
Final prediction
P_f

Stacking

输入： 训练集 $D = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots, (\boldsymbol{x}_m, y_m)\}$;
初级学习算法 $\mathfrak{L}_1, \mathfrak{L}_2, \ldots, \mathfrak{L}_T$;
次级学习算法 $\mathfrak{L}$.

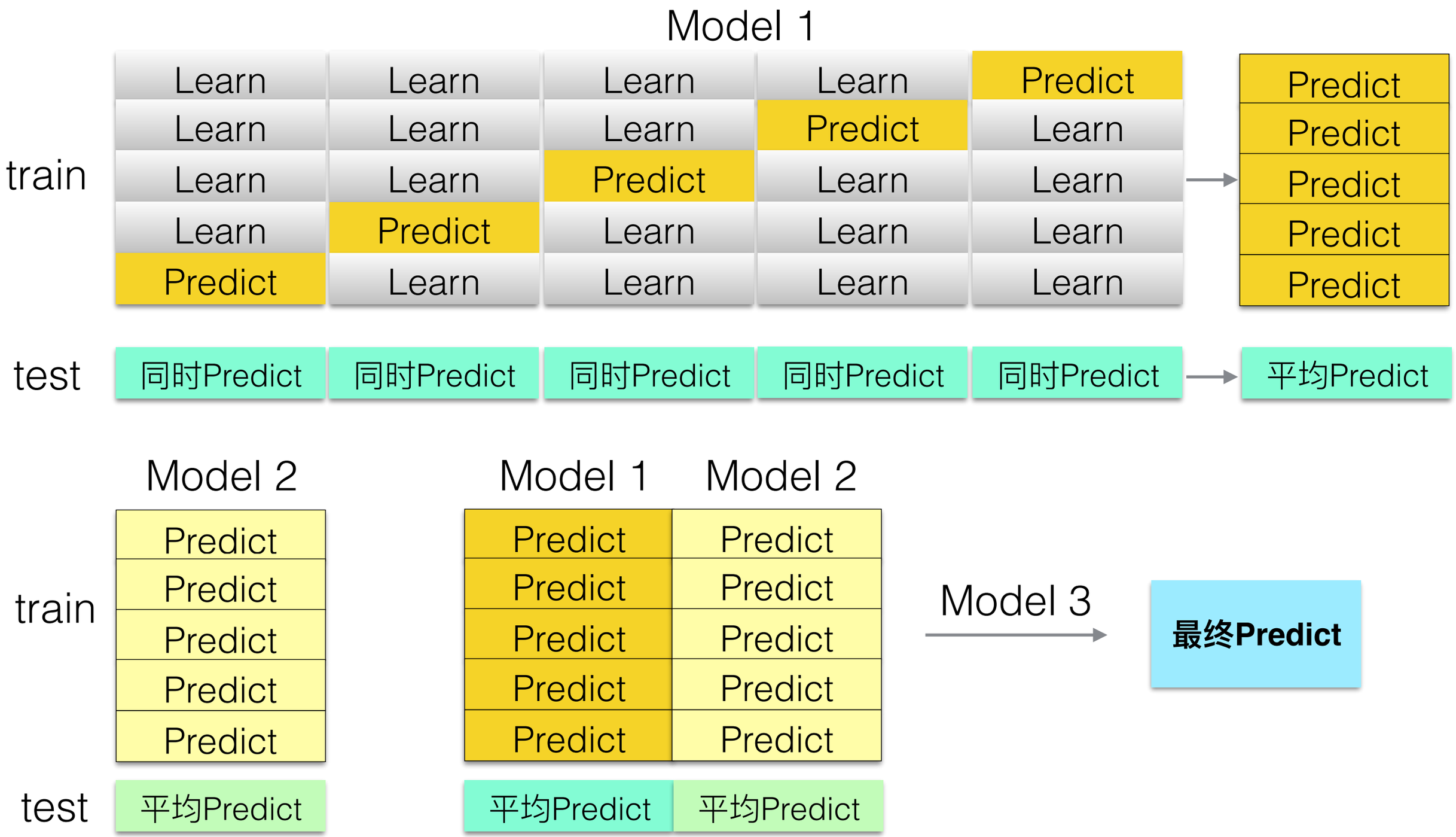
过程：

```
for t = 1,2,...,T do
    h_t = L_t(D);
end for
D' = ∅;
for i = 1,2,...,m do
    for t = 1,2,...,T do
        z_it = h_t(x_i);
    end for
    D' = D' ∪ ((z_i1, z_i2, ..., z_iT), y_i);
end for
h' = L(D');
```

输出： $H(\boldsymbol{x}) = h'(h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \ldots, h_T(\boldsymbol{x}))$

图 8.9 Stacking 算法

Stacking


Model 1
train
Learn Learn Learn Learn Predict
Learn Learn Learn Predict Learn
Learn Learn Predict Learn Learn
Learn Predict Learn Learn Learn
Predict Learn Learn Learn Learn
Predict
Predict
Predict
Predict
Predict
test
同时Predict 同时Predict 同时Predict 同时Predict 同时Predict
平均Predict
Model 2
train
Predict
Predict
Predict
Predict
Predict
test
平均Predict
Model 1 Model 2
Predict Predict
Predict Predict
Predict Predict
Predict Predict
Predict Predict
平均Predict 平均Predict
Model 3
最终Predict

Stacking

安装：

pip install mlxtend

conda install mlxtend