

Scikit-Learn Seminar

1.9.4 Gradient Tree Boosting

菊池裕紀

Gradient Boosted Regression Trees

- GBRT

- 以下のモデル

$$f(x) = \sum_{m=1}^M \gamma_m h_m(x) \quad \text{※ } h(x): \text{決定木}$$

- 任意の微分可能損失関数でブースティングを一般化したもの
- Web検索ランキングや生態学に利用されている
- 回帰、分類問題に利用できる

```
from sklearn.ensemble import GradientBoostingClassifier  
from sklearn.ensemble import GradientBoostingRegressor
```

Mathematical formulation

- 前段階のモデルを利用して次の段階のモデルを形成

$$f_m(x) = f_{m-1}(x) + \gamma_m h_m(x)$$

$$f_m(x) = f_{m-1}(x) + \gamma_m \underbrace{\arg \min_h \sum_{i=1}^n L(y_i, f_{m-1}(x_i) - h(x))}_{h_m(x)}$$

$$f_m(x) = f_{m-1}(x) + \gamma_m \sum_{i=1}^n \nabla_f L(y_i, f_{m-1}(x_i))$$

$$\gamma_m = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, f_{m-1}(x_i) - \gamma \frac{\partial L(y_i, f_{m-1}(x_i))}{\partial f_{m-1}(x_i)})$$

Loss functions

- Regression

 - Least square

 - Least absolute deviation

 - Huber

 - Quantile

- Classifier

 - Binomial deviance

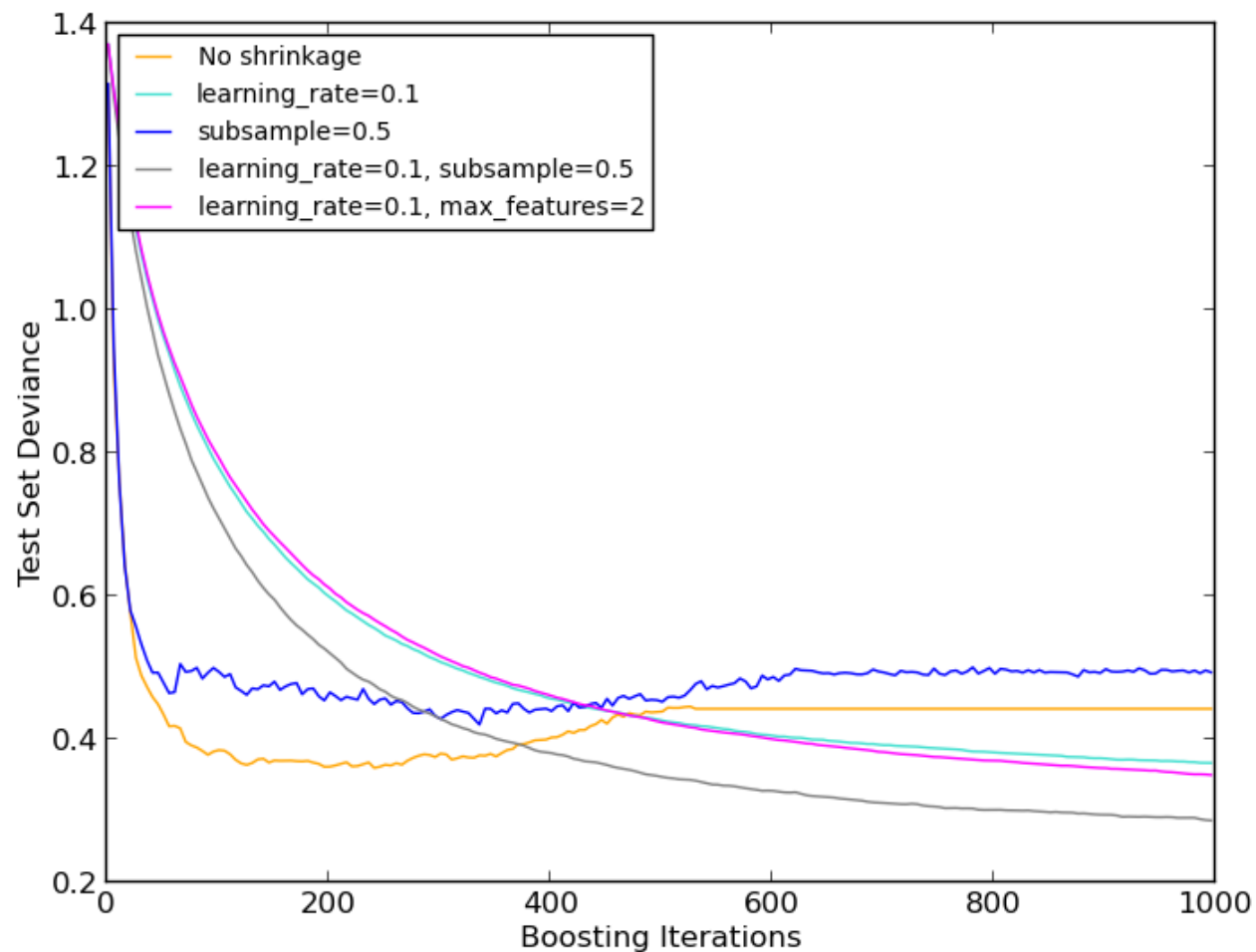
 - Multinomial deviance

Reguralization

- 学習係数で調整

$$f_m(x) = f_{m-1}(x) + v\gamma_m h_m(x)$$

- “sbsample”の設定



Interpretation

- 何百もの決定木を視覚化して分析するのは不可能

- Feature importance

各特徴の重要度をはかることができる

多く分割点に選ばれている特徴はより重要度が高いとみなす

- Partial dependence

各特徴とターゲットの反応との依存関係をしめす

