1. INTRODUCTION

н

$$f(Y_i) = p_i^{Y_i} (1 ! p_i)^{1!Y_i}$$
 $Y_i = 0,1; i = 1,..., n.$

$$\begin{pmatrix} & 1 \end{pmatrix} \quad i \quad \frac{e^{-0-1X_1}}{1-e^{-0-1X_1}}$$

$$AIC_{p} = -2\log L(b) + 2p = -2^{\&} (\prod_{i=1}^{n} Y_{i}(X_{i}^{T})) - (\prod_{i=1}^{n} \log(1 + e^{X_{i}^{T}})) + 2p$$

2.3 GRADIENT BOOSTING

2.3.1 Regression Trees, overview

$$\hat{f}(X) = \underbrace{\stackrel{5}{!}}_{m=1} c_m I\{(X_1, X_2) \mid R_m\}$$

2.3.1 Regression Trees, detail

.



(

T Ú

Т

$$z_{i} = -\frac{\#}{\#f(x_{i})} ! (y_{i}, f(x_{i}))|_{f(x_{i})=\hat{f}(x_{i})}$$

)

$_{k} = \operatorname{argmin}_{x_{i} \ S_{k}} \left(y_{i}, \hat{\mathcal{H}}(x_{i}) + \right)$ Table 1

 $\hat{p}_i = 0.5,!!$

Predictor Variables	Yr1 FA	Yr1 Sshp	Yr1 Grnt	Yr1 Loan	Ann. FA	Ann. Sshp	Ann. Grnt	Ann. Loan
Response Variables								
Grad. in 6 Yrs								
Grad. in Any								
Yrs to Grad.								
Yr1 Ret.								
Yr2 Ret.								
Yr3 Ret.								

Figure 3.

Figure 4.

3.3 LOGISTIC REGRESSION VS. GRADIENT BOOSTING, REAL DATA



5. APPENDIX

5.1. R code to simulate data.

REFERENCES

A Working Guide to Boosted Trees.

The Elements of Statistical Learning: Data Mining,

Inference and Prediction, 2nd Edition.

Applied Linear Statistical Models.