On the Novel Use of the Welch-Young Inverse Bio-elasticity Coefficient when Building Binary Classification Trees

Stephen Welch[1], Alison Young[1,2], Gregory McScienceface[3]

[1]stephen@superfancyuniversity.edu, [2]ayoung@someinstututeyouveneverheardof.gov, [3]mcscienceface@anevenfancieruniversitythatyoucouldnevergetinto.edu

Abstract

The use of information gain and gini impurity as split criteria when growing decision trees is shown to be critically-hyper -sub-optimal by extension of the Welch-Takai sampling theorem. A novel split criteria, the Welch-Young Inverse Bio-elasticity Coefficient is introduced and shown to be hyper-optimal by reverse inverse L-norm sampling. Decision tree performance using our novel split criteria is evaluated on a wide range of datasets, showing that we are the fucking shit, and everyone else's shit is bullshit.

1. Introduction

As everyone knows, because it's obvious as shit,[1] misclassification error rate, gini impurity, and information gain are non-optimal criteria for growing decision trees.

(1) $$I(p) = min(p, 1-p)$$

(2) $$Gini(p) = p(1-p)$$

(3) $$Entropy(p) = p \cdot log\left(\frac{1}{p}\right) + (1-p) \cdot log\left(\frac{1}{1-p}\right)$$

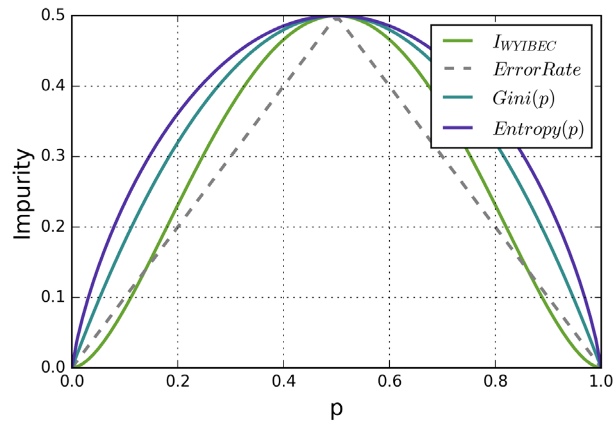

Figure 1. The Young-Welch Inverse Bio-elascticity Coefficient. Our decision tree is class two hyper-optimal, and guarantees performance above the theoretical Welch-Waka-Flocka upper bound.

Below, drawing from information theory, rare bird mating patterns, and quantum sub-particle modeling, we derive the Welch-Young Inverse Bio-eleasticity coefficient, a vastly superior split criteria for growing decision trees. Rather than risk insulting the reader by explaining what th fuck is actually going on, we've chosen to just write down a bunch of convoluted equations.

$$\frac{D}{Dt}\overline{w'^i w'^j} + \overline{w'^i w'^\alpha}\nabla_\alpha \bar{u}^j + \overline{w'^j w'^\alpha}\nabla_\alpha \bar{u}^i - \alpha\left(g^{i\alpha}\overline{w'^j\frac{T'}{\bar{T}}} + g^{j\alpha}\overline{w'^i\frac{T'}{\bar{T}}}\right)\left(\nabla_\alpha\bar{\Phi} + \frac{D\bar{u}_\alpha}{Dt}\right)$$

$$+ \frac{1}{\bar{\rho}}\nabla_\alpha[\overline{\bar{\rho} u'^\alpha w'^i w'^j} + \overline{(g^{i\alpha}w'^j + g^{j\alpha}w'^i)P'} - \overline{w'^i\sigma^{j\alpha}(u')} - \overline{w'^j\sigma^{i\alpha}(u')}]$$

$$+ \frac{1}{\bar{\rho}}\overline{w'^i w'^j\nabla_\alpha(\bar{\rho} u'^\alpha)} - \overline{P'(g^{i\alpha}\nabla_\alpha w'^j + g^{j\alpha}\nabla_\alpha w'^i)} = -\frac{1}{\bar{\rho}}[\overline{\sigma^{i\alpha}(u')\nabla_\alpha w'^j} + \overline{\sigma^{j\alpha}(u')\nabla_\alpha w'^i}] = -\epsilon_2^{ij}\,,$$

$$(1+e_4)\frac{D}{Dt}\overline{\left(\frac{T'}{\bar{T}}\right)^2} - 2f(t)\overline{\left(\frac{T'}{\bar{T}}\right)^2} - 2\overline{w'^\alpha\frac{T'}{\bar{T}}}D_\alpha + \frac{1}{(1+e_4)\bar{\rho} C_p^2}\nabla_\alpha\left[(1+e_4)^2 C_p^2\bar{\rho}\overline{w'^\alpha\left(\frac{T'}{\bar{T}}\right)^2}\right] + \frac{1+e_4}{\bar{\rho}}\overline{\left(\frac{T'}{\bar{T}}\right)^2\nabla_\alpha(\rho u'^\alpha)}$$

$$+ \frac{2}{\bar{\rho}\bar{T}C_p}\overline{\frac{T'}{\bar{T}}\left[P'\nabla_\alpha w'^\alpha - \nabla_\alpha(P'_g w'^\alpha) - \frac{DP'_g}{Dt}\right]} = \frac{2}{\bar{\rho}\bar{T}C_p}\overline{\frac{T'}{\bar{T}}[\sigma^{\alpha\beta}(u')\nabla_\alpha u'_\beta - \nabla_\alpha F_r'^\alpha]} = -\epsilon_2\,,$$

$$(1+e_4)\left[\frac{D}{Dt}\left(\overline{w'^i\frac{T'}{\bar{T}}}\right) + \overline{w'^\alpha\frac{T'}{\bar{T}}}\nabla_\alpha\bar{u}^i - \alpha\overline{\left(\frac{T'}{\bar{T}}\right)^2}g^{i\alpha}\left(\nabla_\alpha\bar{\Phi} + \frac{D\bar{u}_\alpha}{Dt}\right)\right] - f(t)\overline{w'^i\frac{T'}{\bar{T}}} - \overline{w'^i w'^\alpha}D_\alpha$$

$$+ \frac{1}{\bar{\rho} C_p}\nabla_\alpha\left[(1+e_4)C_p\bar{\rho}\overline{w'^i w'^\alpha\frac{T'}{\bar{T}}}\right] + \frac{1+e_4}{\bar{\rho}}\overline{w'^i\frac{T'}{\bar{T}}\nabla_\alpha(\rho u'^\alpha)} + \frac{1}{\bar{\rho}\bar{T}C_p}\overline{w'^i\left[P'\nabla_\alpha w'^\alpha - \nabla_\alpha(P'_g w'^\alpha) - \frac{DP'_g}{Dt}\right]}$$

$$= \frac{1+e_4}{\bar{\rho}}\overline{\frac{T'}{\bar{T}}\nabla_\alpha\sigma^{i\alpha}(u')} + \frac{1}{\bar{\rho}\bar{T}C}\overline{w'^i[\sigma^{\alpha\beta}(u')\nabla_\alpha u'_\beta - \nabla_\alpha F_r'^\alpha]} = -\epsilon_2^i\,,$$

$$1 = 1$$

$$I_{WYIBEC} = 4\left(p \cdot log\left(\frac{1}{p}\right) + (1-p) \cdot log\left(\frac{1}{1-p}\right)\right)p(1-p)$$

---

1 Obviously