

Board Game Data

Group 2

STOR 320 - Final Project



1.

Introduction + Data

The Data

boardgamegeek.com

Scraped and uploaded to Kaggle
by users Lime and ZakWhite in
January 13, 2020

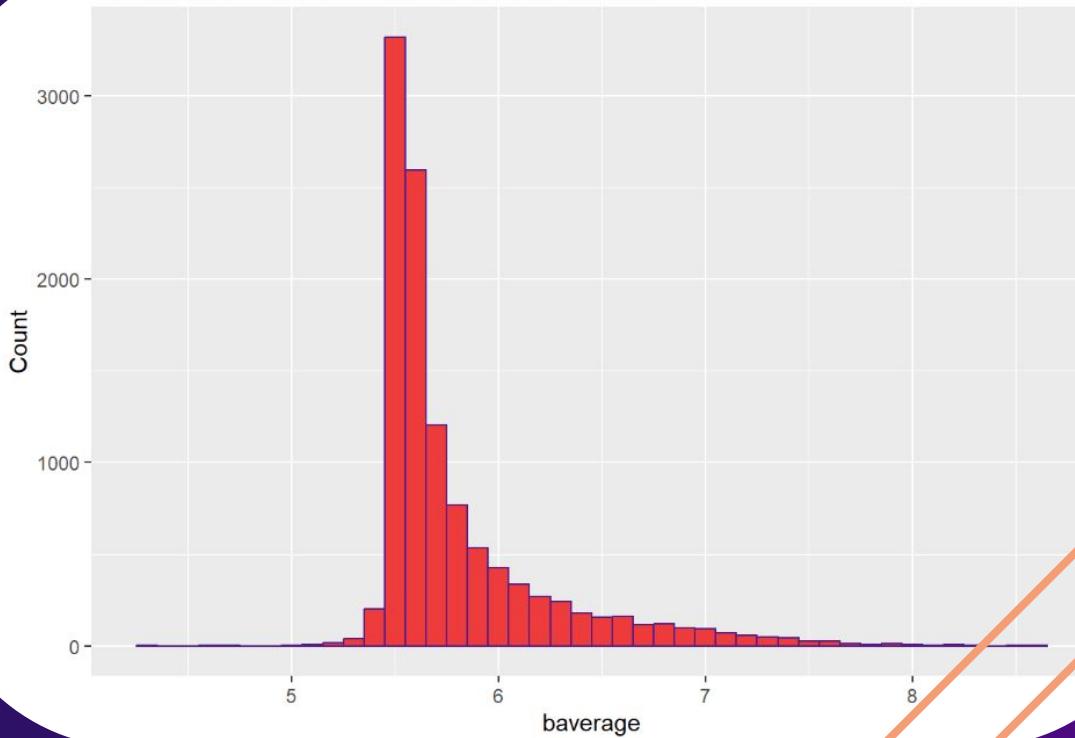


BOARD
GAME
GEEK

52 columns
20,000 rows

Response Variable

Distribution of Board Game B-average Ratings



Top Ten Games

Top Ten Board Games (2020)

Board Game	Weighted User Rating (baverage)	(Log) Total Media	Mechanic Count	Average Weight	Minimum Players	Median Play Time
Gloomhaven	8.58424	6.473891	12	3.8078	3	90.0
Pandemic Legacy: Season 1	8.47159	6.732211	8	2.8301	4	60.0
Terraforming Mars	8.26781	7.229839	9	3.2313	3	120.0
Through the Ages: A New Story of Civilization	8.23513	5.568345	5	4.3850	3	120.0
Brass: Birmingham	8.20459	5.081404	8	3.9122	3	90.0
Twilight Imperium (Fourth Edition)	8.17076	4.890349	11	4.1966	6	360.0
Twilight Struggle	8.16138	6.196444	9	3.5724	2	150.0
Star Wars: Rebellion	8.15575	5.720312	6	3.6939	2	210.0
Gaia Project	8.14076	5.241747	7	4.3131	3	105.0
Scythe	8.09991	6.997596	10	3.3810	4	102.5

Gloomhaven

Top Ten Board Games (2020)

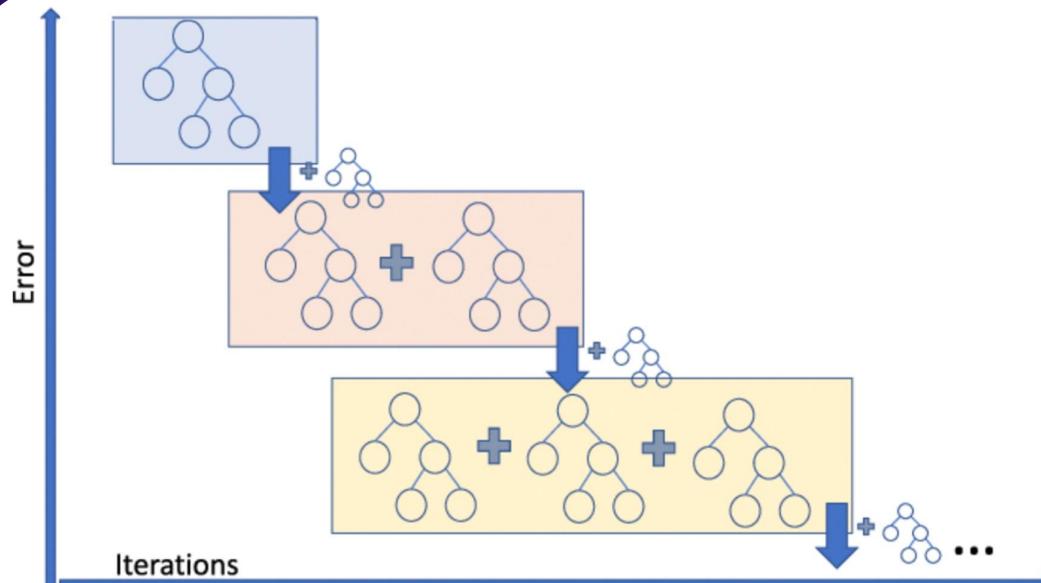
Board Game	Weighted User Rating (baverage)	(Log) Total Media	Mechanic Count	Average Weight	Minimum Players	Median Play Time
Gloomhaven	8.58424	6.473891	12	3.8078	3	90.0
Settlers of Catan Legacy: Season 1	8.47159	6.732211	8	2.8301	4	60.0
Exploding Mars	8.26781	7.229839	9	3.2313	3	120.0
New Story	8.23513	5.568345	5	4.3850	3	120.0
Ally (Second Edition)	8.20459	5.081404	8	3.9122	3	90.0
Scythe (Second Edition)	8.17076	4.890349	11	4.1966	6	360.0
Graveyard	8.16138	6.196444	9	3.5724	2	150.0
Scrabble	8.15575	5.720312	6	3.6939	2	210.0
Ally	8.14076	5.241747	7	4.3131	3	105.0
Scrabble (Second Edition)	8.09991	6.997596	10	3.3810	4	102.5



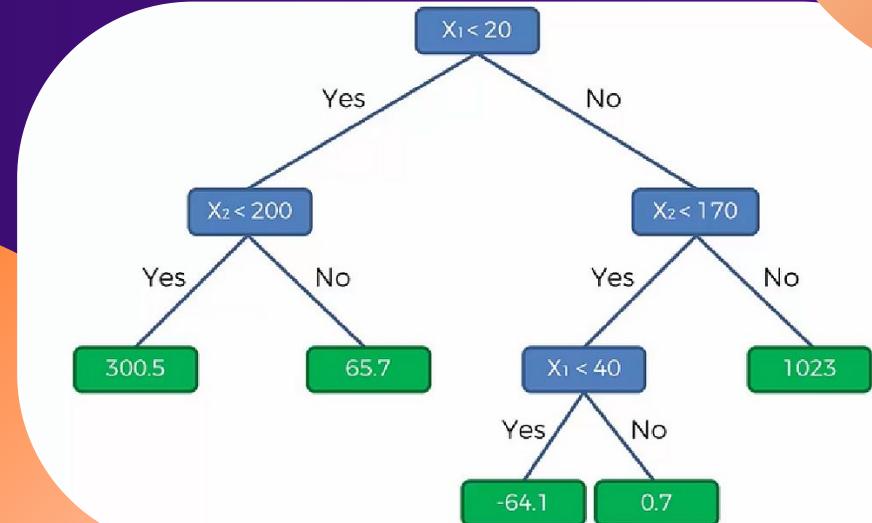
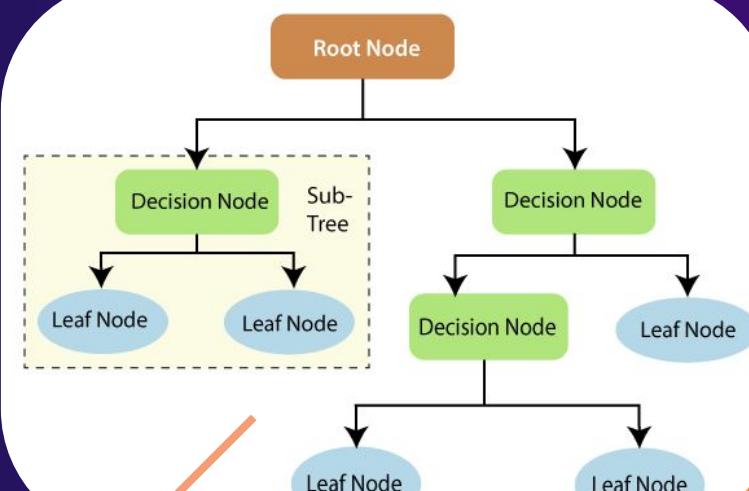
2. Results

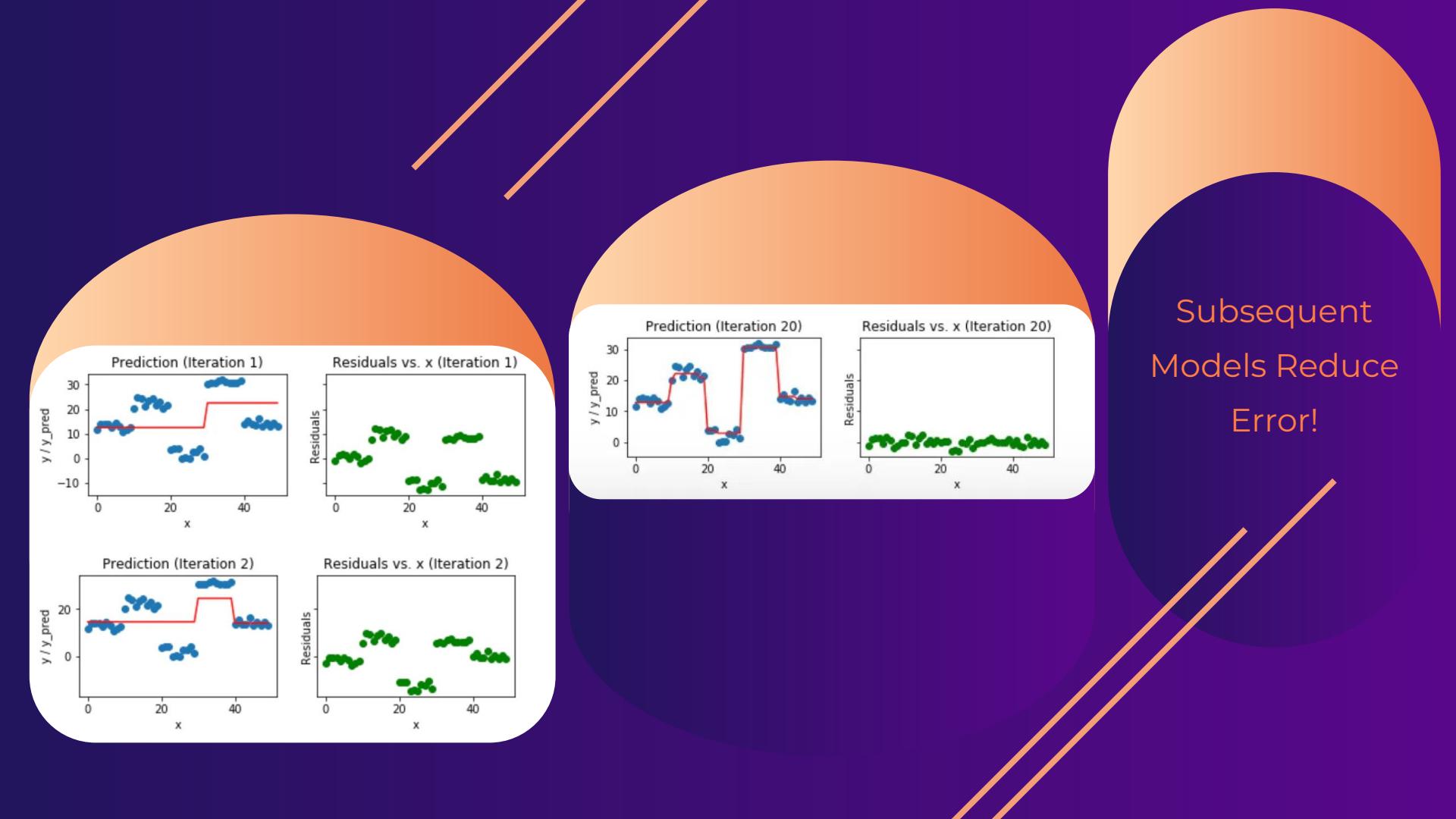


Gradient Boosting

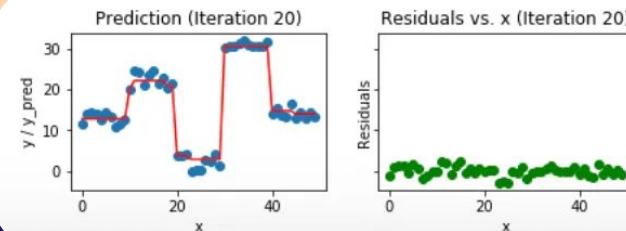
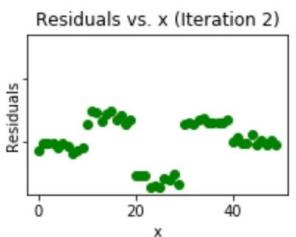
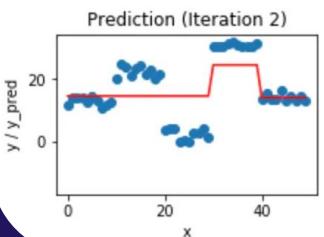
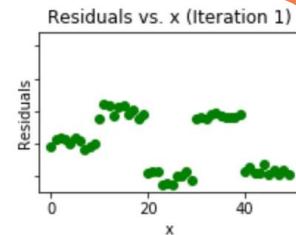
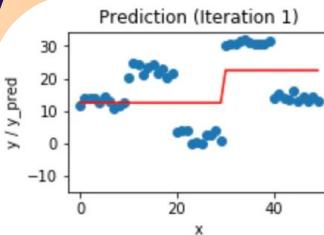


Decision Trees





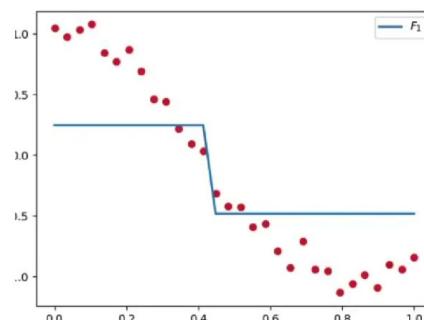
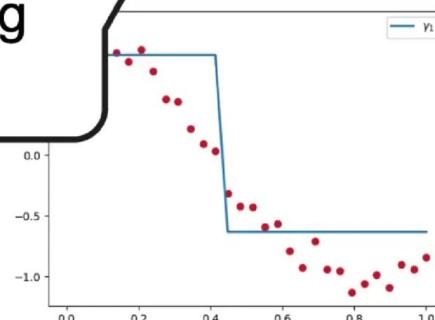
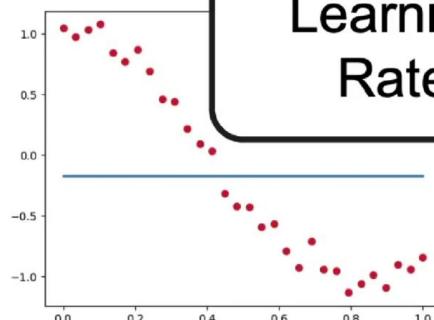
Subsequent
Models Reduce
Error!



Learning Rate

$$F_0 + \nu \gamma_1 = F_1$$

Learning
Rate



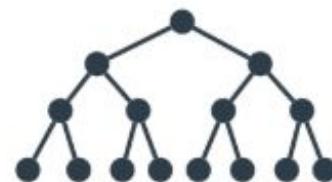
Tree Depth



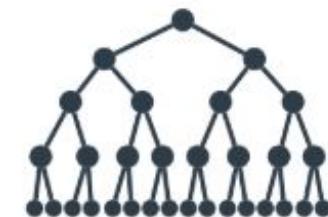
Depth = 1



Depth = 2



Depth = 3

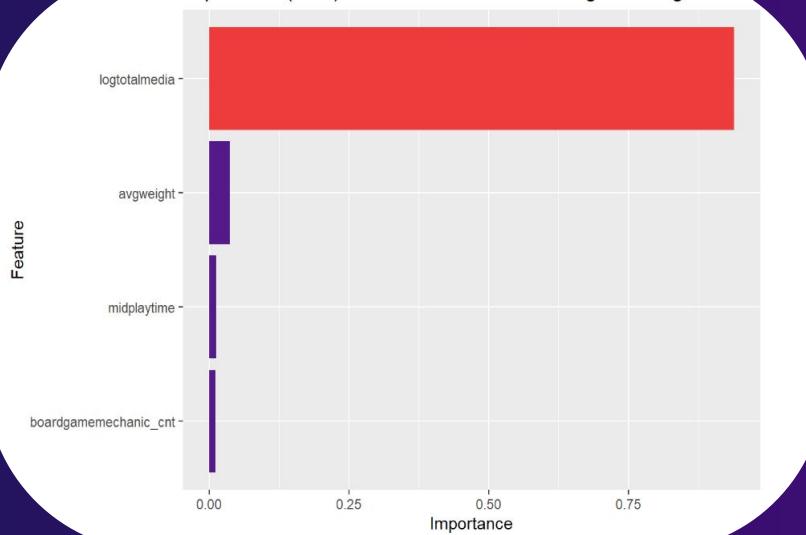


Depth = 4

Maximum depth of a decision tree

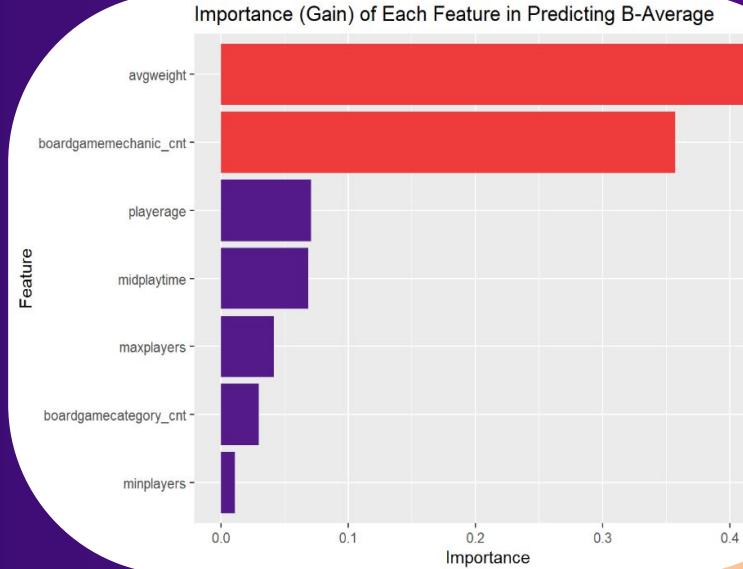
Variable Importance in Model

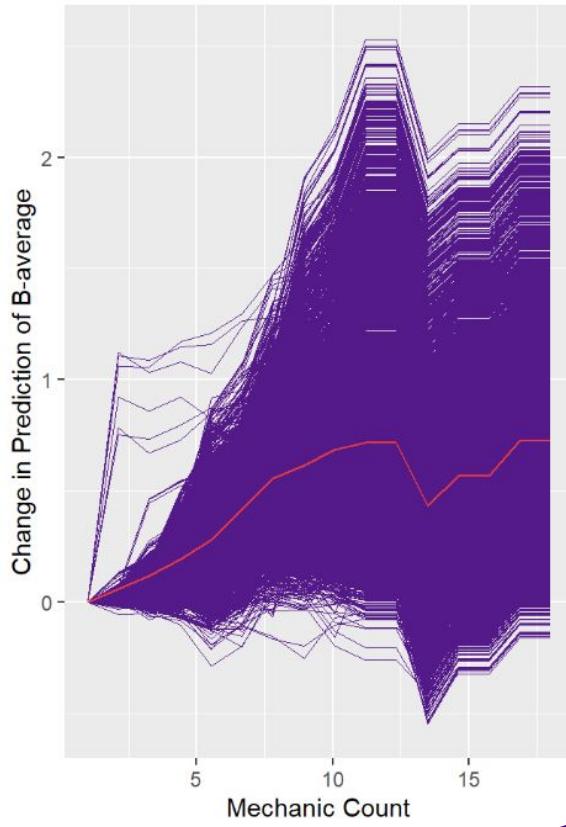
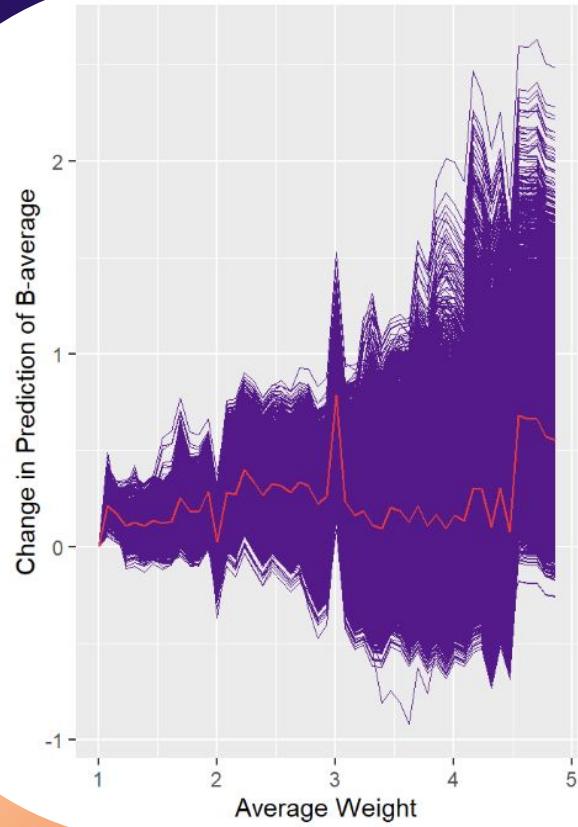
Importance (Gain) of Each Feature in Predicting B-Average



With *logtotalmedia*

Without *logtotalmedia*





PDPs

(Partial
Dependence
Plots)

3. Conclusion



