SOCCER PLAYER'S MARKET VALUE PREDICTION & TRANSFER STRATEGY PRESCRIPTION

(15.095) Machine Learning Under a Modern Optimization Lens

Oscar COURBIT, Andrea ZANON

MIT SLOAN SCHOOL OF MANAGEMENT

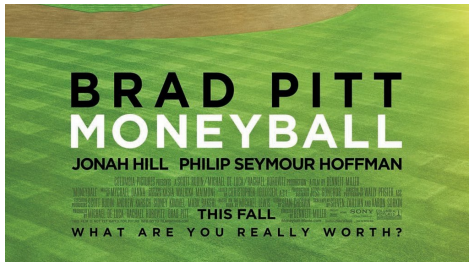
MIT MANAGEMENT SLOAN SCHOOL

PROBLEM AND RELEVANCE

Transfer market

- €41 billion spent over last decade
- 26% average annual inflation growth

- FC Barcelona, €700 M shortfall in 2022
- Neymar, €222 M (2017)



Find cheap, **underrated**, promising players

Improve team **optimally**, avoid useless spending

MIT MANAGEMENT SLOAN SCHOOL

DATA

Dataset

- FIFA players' data
  - 8 years
  - ~20k observations each, 110 features
- Transfers data
  - 30 years
  - ~2k rows each

Train-test split

Training set:

- 2019/20 and 2020/21
- ~ 40k observations

Test set:

- 2021/2022
- ~20k observation

Missing data imputation

Goalkeepers have missing data

⇩

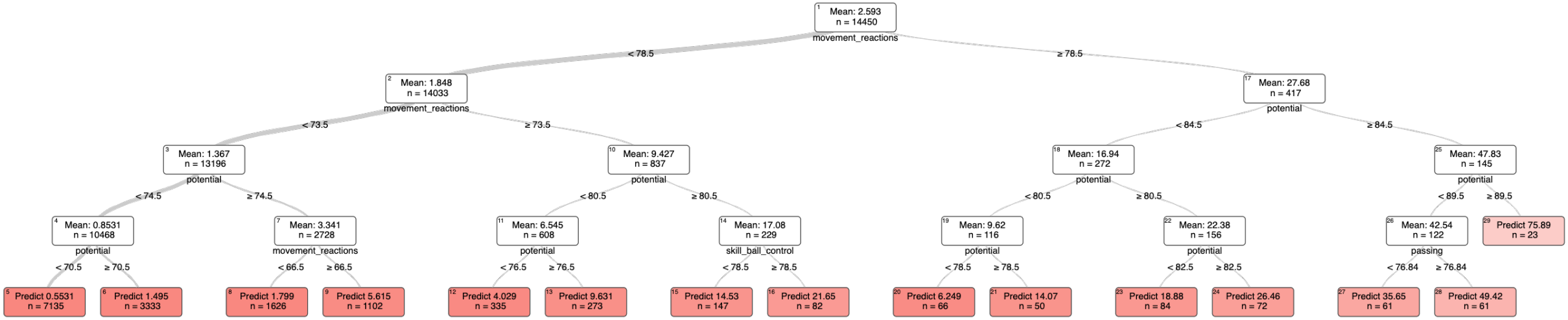
IAI - Optimal Tree Imputation

MIT MANAGEMENT SLOAN SCHOOL

PREDICTION MODEL, INTERPRETABILITY

Regression	Linear	Robust	Holistic
Ensemble	XGBoost	Random Forest	
Trees	ORT	ORT-H	ORT-L

XGBoost
best performing
$R^2 = 0.86$



ORT fitted on XGBoost predictions, $R^2 = 0.86$

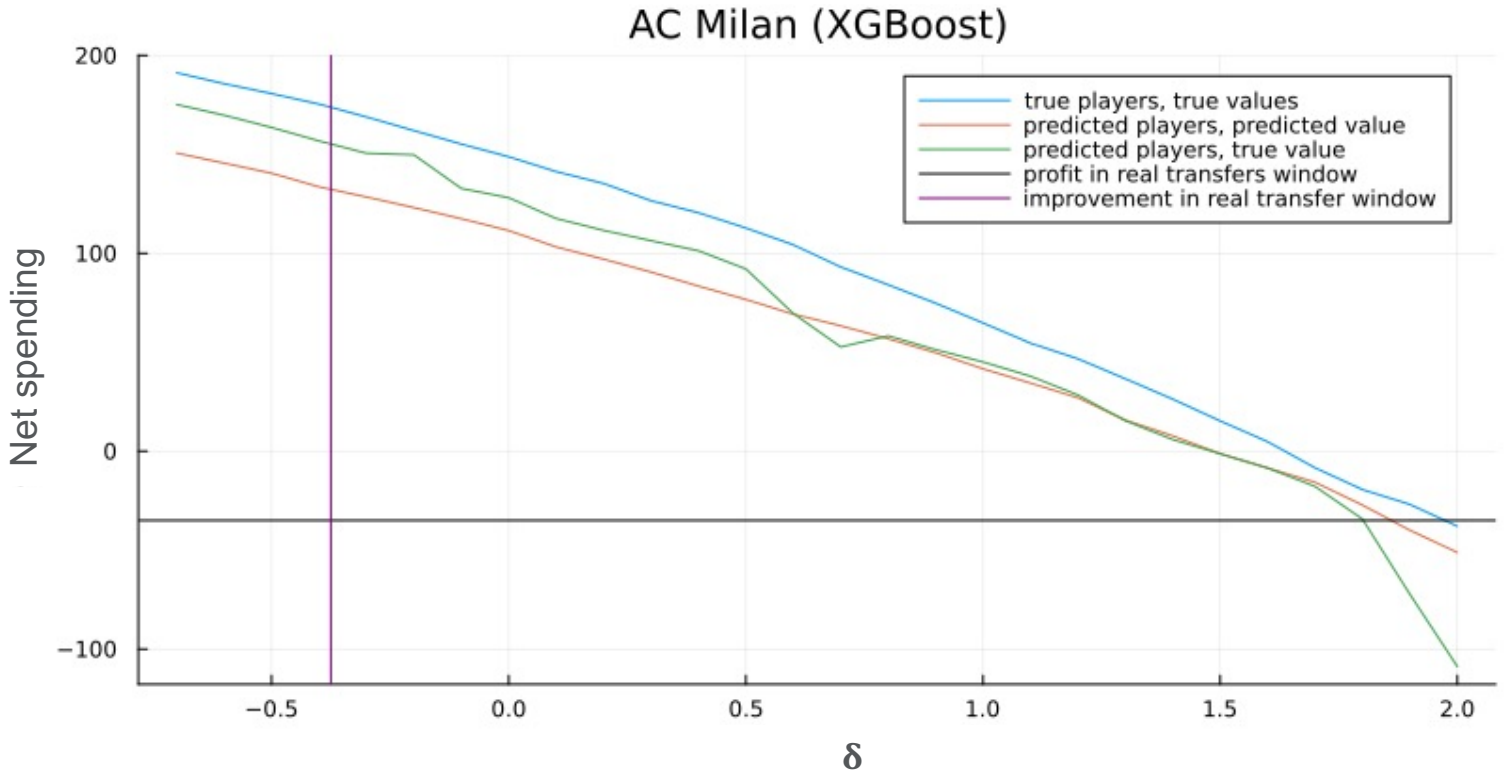
MIT MANAGEMENT SLOAN SCHOOL

PRESCRIPTION AND KEY INSIGHTS

Formulation:

$$\begin{aligned} \min \quad & \textbf{net_spend} \ (\Longleftrightarrow \ \max \quad \textbf{profit}) \\ \text{s.t.} \quad & \textbf{improvement} \geq \delta \end{aligned}$$

Linear objective function:
Regress and Compare



Objective value v. minimum delta improvement

MIT MANAGEMENT SLOAN SCHOOL

SUMMARY, FINDINGS AND NEXT STEPS

Comparison **reality** vs **prescription**	AC Milan		FC Barcelona		PSG		AVERAGE	
	Δ(Profit)	Δ(δ)	Δ(Profit)	Δ(δ)	Δ(Profit)	Δ(δ)	Δ(Profit)	Δ(δ)
For equal Profit	0	**+2.18**	0	**+0.18**	0	**+1.49**	0	**+1.41**
For equal δ	**+185 M**	0	**+85M**	0	**+354 M**	0	**+196M**	0

Improvement achieved by prescription based on XGBoost predictions.
More teams are considered in the average

- Clubs can purposedly adopt a different strategy (PSG)
- Prescription close to reality for teams that truly adopt a “parsimonious” strategy (Barcelona)

- For a **more realistic** model:
  - Consider release clause, wages, free transfers, loans, contract duration
  - **Robust Optimization** model

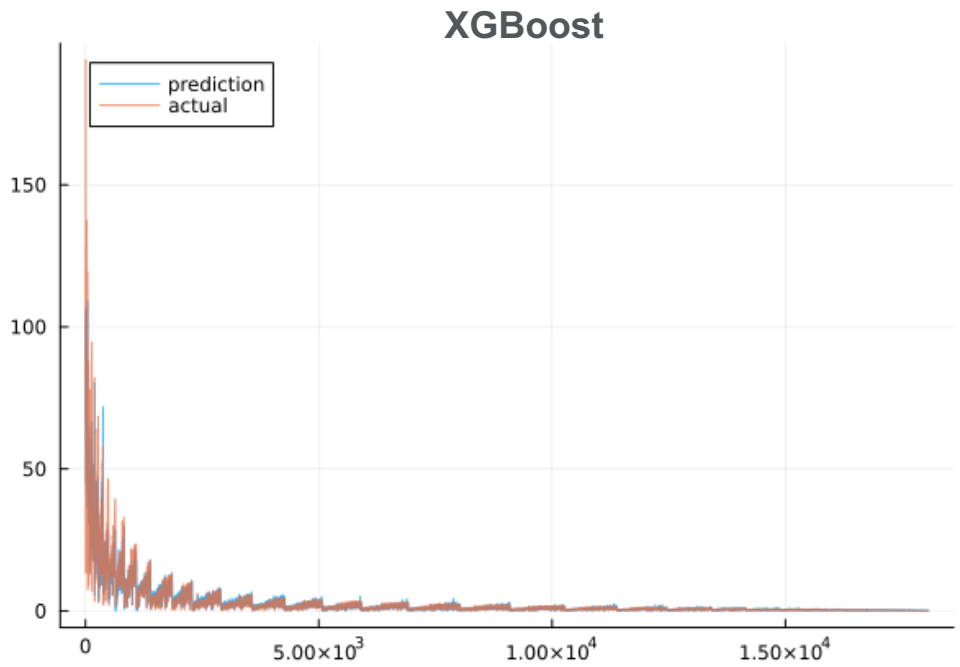
MIT MANAGEMENT SLOAN SCHOOL

THANK YOU

MIT MANAGEMENT
SLOAN SCHOOL

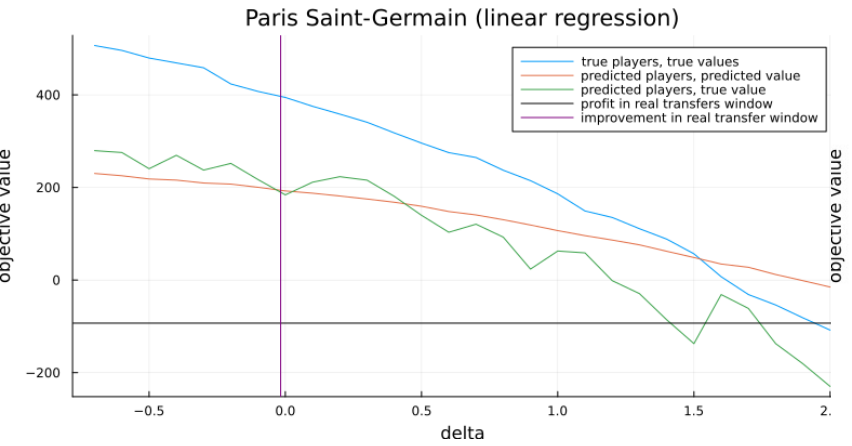
APPENDIX 1: PREDICTION RESULTS

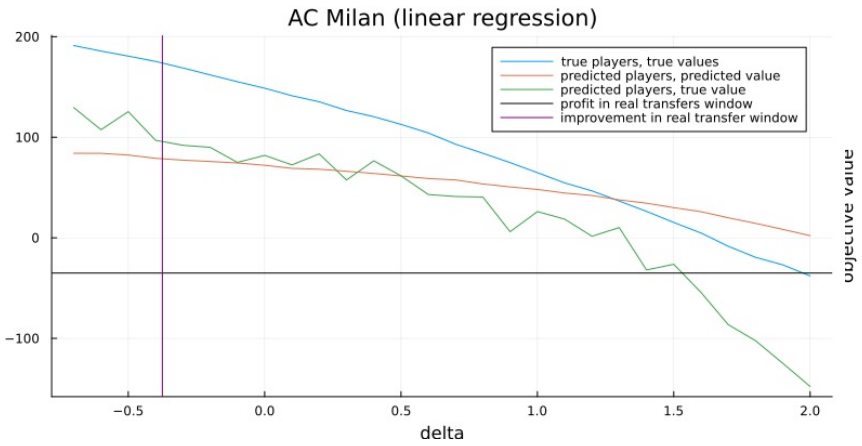
Model	Time (s)	R^2	RMSE	MAE	MAPE
Linear Regression	**0.5**	70.2	4.16	1.91	126.1
LASSO	0.8	66.4	4.54	1.57	89.7
Random Forest	86	87.5	2.75	0.58	21.2
XGBoost	128	**88.7**	**2.60**	**0.51**	**19.2**
Holistic Regression	4800	71.5	4.14	1.73	114.7
ORT	373	80.4	3.37	0.96	67.1
ORT-H	49380	71.7	3.86	0.85	54.8
ORT-L	3260	73.3	3.99	1.58	91.3

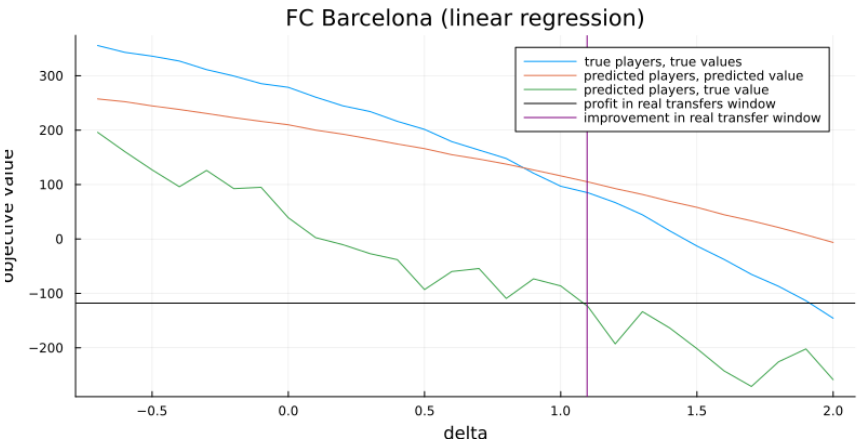


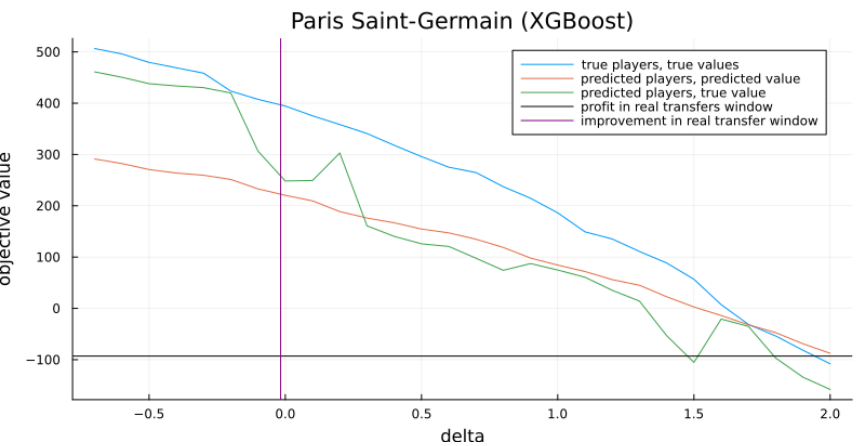
MIT MANAGEMENT SLOAN SCHOOL

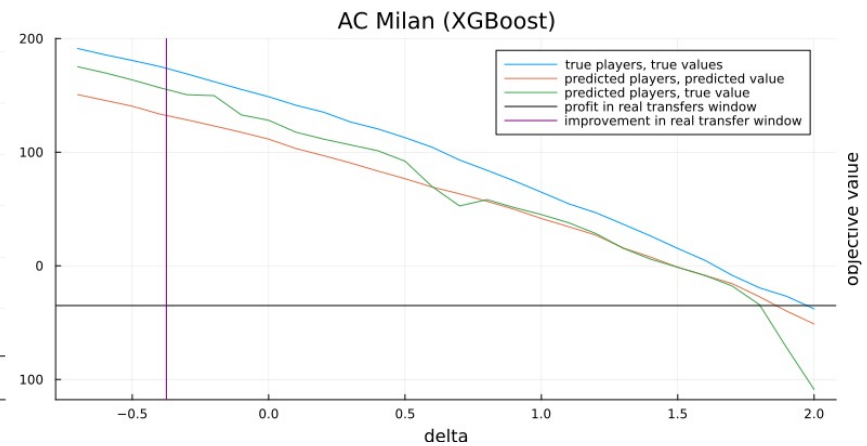
APPENDIX 2: PRESCRIPTION RESULTS

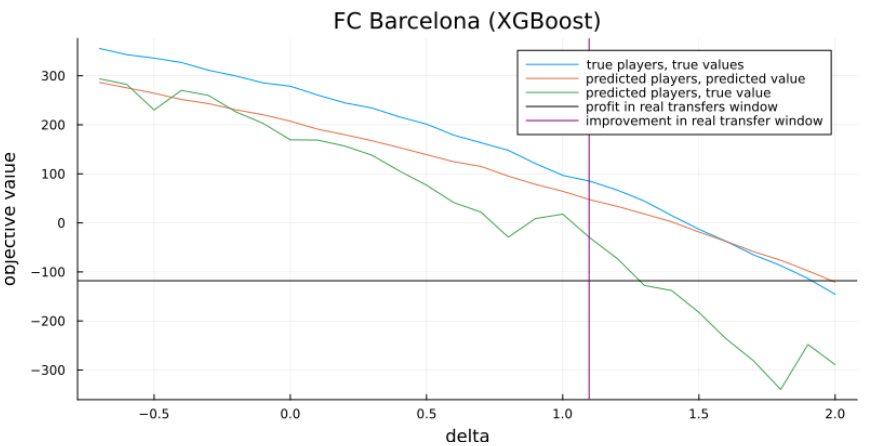






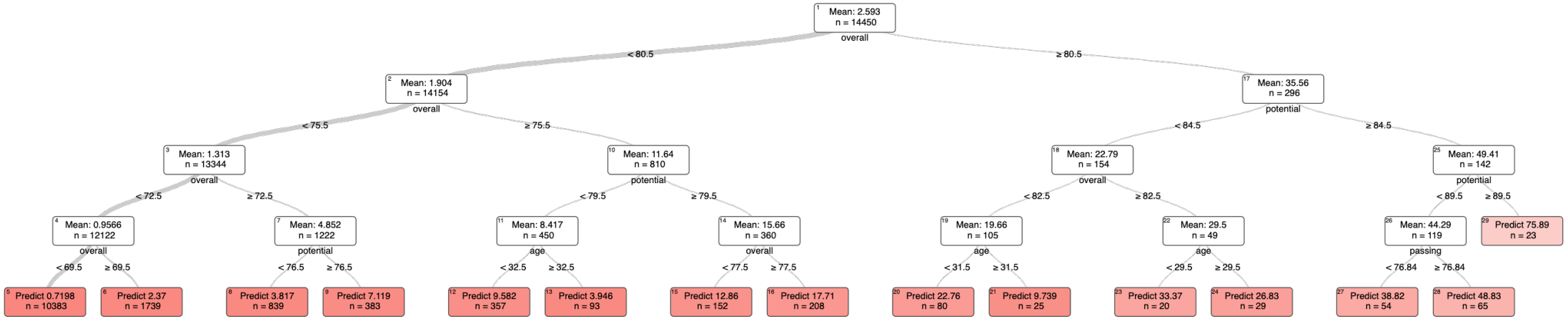




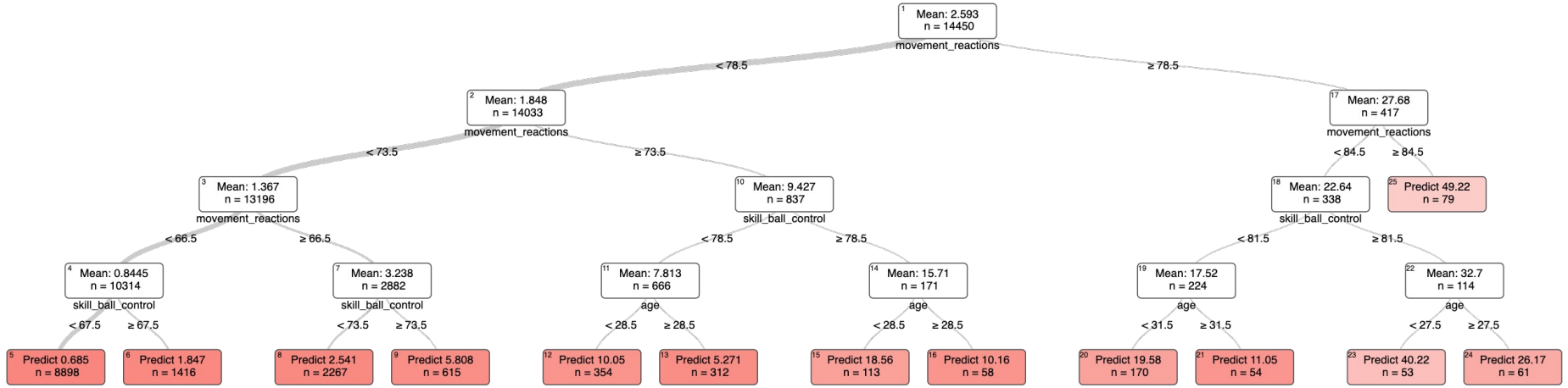


MIT MANAGEMENT SLOAN SCHOOL

APPENDIX 3: INTERPRETABILITY



$R^2 = 0.92$



$R^2 = 0.71$

MIT MANAGEMENT SLOAN SCHOOL