



Wage Against the Machine: A Generalized Deep-Learning Market Test of Dataset Value

Philip Z. Maymin Trefz School of Business University of Bridgeport, Bridgeport, CT

How can you tell if a particular sports dataset really adds value?

(In other words: does measuring whether a hand was up on defense *matter*?)
If you ask one genius to extract all possible insights from dataset X, and another genius to extract all possible insights from datasets X+Y, if the first genius is smarter or luckier or both, he may get more insights from less data. What's the solution?
Use deep-learning on both datasets to try to outperform betting markets.

8000

6000 -				
4000 -	Some additional metric	s available from Vant	age Sports	
2000 -	Scoring: Contest+ FG%, Op Shot Defense: Block-to-Pos Movement and Involvemen Disruptions: In-Air TO%,U Passing: Indirect Pass Rate, Rebounding: O/DBlockouts Screens: Screens Received/S	en+ Freq.,Open+ FG%,Po session,Points Against per nt: Activity Rate,Cut Effici nforced TO%,Effective Bu Assist+ Screen%,Deflected s per 100 Opps,O/DReb Pu Set per Chance,Split%,Sol	oints per Chance, Shot,Shots per Chance, ency,Touches per Chance, ump%,Front Post D%,Pre d-Pass Rate,True Facilitations ursuit Rate, id Screen%,KIF%,Hedge?	, ssure Rate, ion,
0	Jan	Feb	Mar	Apr

Starting with an initial bankroll of \$5,000, the daily rolling deep learning algorithm **using the standard NBA dataset** is correct on 49% of wagers and ends the season with \$1,700. This means that the standard dataset combined with deep learning is unable to do better than a coin toss. (Indeed, the 49% is not statistically significantly different from 50%.) This should not be surprising, as the betting markets are indeed quite efficient, and we should expect that they incorporate all standard publicly available information. **Using additional Vantage Sports data**, the algorithm is correct on 54% of wagers and ends the season with \$6,500. The difference is highly statistically significant (p-value < 0.01). And it exceeds the breakeven probability.