

# 自主課題研究概要

-プロ野球選手の年俵を推測する-

情報システム工学科 3年 022 澤田 紘樹

統計処理ソフトウェアRを用いてプロ野球選手の年俵を推測する。Webページ (<http://home.a07.itscom.net/kazoo/pro/pro.htm>) においてあるプロ野球選手それぞれの2008年度の成績と年俵を使用し、そのデータをRに学習させて、2009年度の年俵を予測した。今回は線形回帰のlm関数とksvm関数(サポートベクター)という各変数に重みをつけて推測するという大きく分けて2つの方法を使用して年俵を予測していく。年俵には前年度の成績と今までの実績が大きく関わってくるが、学習させるデータを選抜したり、予測器のカーネルを変更したりして、より正確な予測を目指す。表1は使用した成績、表2はksvm関数(サポートベクター)で予測した1つの例である。

表1: 成績

投打	年齢	年数	年俵	打率	試合	打数	安打	本	打点	盗塁
ラミレス 右右	34	8	50000	0.319	144	548	175	45	125	1
小笠原 道 右左	35	12	38000	0.31	144	520	161	36	96	0
阿部 慎之 右左	29	8	24000	0.271	125	428	116	24	67	1
坂本 勇人 右右	20	2	850	0.257	144	521	134	8	43	10
木村 拓也 右両	36	18	6500	0.293	124	372	109	7	31	1
谷 佳知 右右	35	12	26000	0.295	120	349	103	10	45	5
亀井 義行 右左	26	4	1400	0.268	96	276	74	5	23	7
高橋 由伸 右左	33	11	35000	0.236	91	275	65	17	41	1
鈴木 尚広 右左	30	12	3000	0.304	105	247	75	3	17	30
古城 茂幸 右左	32	11	1700	0.214	83	159	34	3	16	6

表2: 予測

選手名	実年俵	ksvm予測
ラミレス	50000	49510
小笠原 道	38000	37837
阿部 慎之	24000	24222
坂本 勇人	850	1729
木村 拓也	6500	7240
谷 佳知	26000	26158
亀井 義行	1400	2258
高橋 由伸	35000	34892
鈴木 尚広	3000	3805
古城 茂幸	1700	2568

今回の研究の精度の良さを知るためには2008年度の成績によって推測された年俵の値が、実際の2009年の年俵の値と近い値をとっているかということである。今回は年齢を成績の要素から外してみたり、重要とされると考えられる成績の比重を上げてみたりと数十回の実験を繰り返してみたが、選手の特徴によって、適切と思われる予測方法を分類することはできたが、どの方法も全選手の年俵の予測値が平均的に翌年度の年俵と近い値をとることはなかった。

最後に実際に年俵はどのようにして計算されているのかということ調べてみた。すると中国新聞のWebページに以下のような記事があった。これは、広島カープというチームの年俵の査定方法である。

**年俵算出の基準となるのは査定ポイントだ。これはプレーごとに決められている。2人の担当者が、千を超える査定項目に照らして全試合をチェックする。**

Webページ(<http://www.chugoku-np.co.jp/Carp/Cw200912300064.html>より引用)

今回は成績の中で主要な11項目を使用して推測したが、実際に野球の成績の中には、悪い要素を持つものや、少しマニアックと思われるようなものなどなど様々なものがある。また機会があれば、より多くのデータを使用して正確な年俵の予測を可能にしたい。