

NBA Players Value Analysis

Team member: Jianxiao Wang, Henry Hu, Jack Jin

Comments from Claude

- CRISP framework is clearly used on your BA project
- Very sophisticated design of business analytics project by applying CFMC and clustering as unsupervised learning method for clustering NBA players and assigning each player with a class.
- Good use of multiple ETL operators for data preparation, such as dealing missing values, attribute selections and data union.
- Four ML models are clearly illustrated, and explained. It would be perfect to include classification model comparison since it is a necessary step closing up the cycle.
- Good team work and presentation- Henry, JianXiao and Jack!
- Score: 98

Business Understanding

Issue

- Values and classes of NBA players

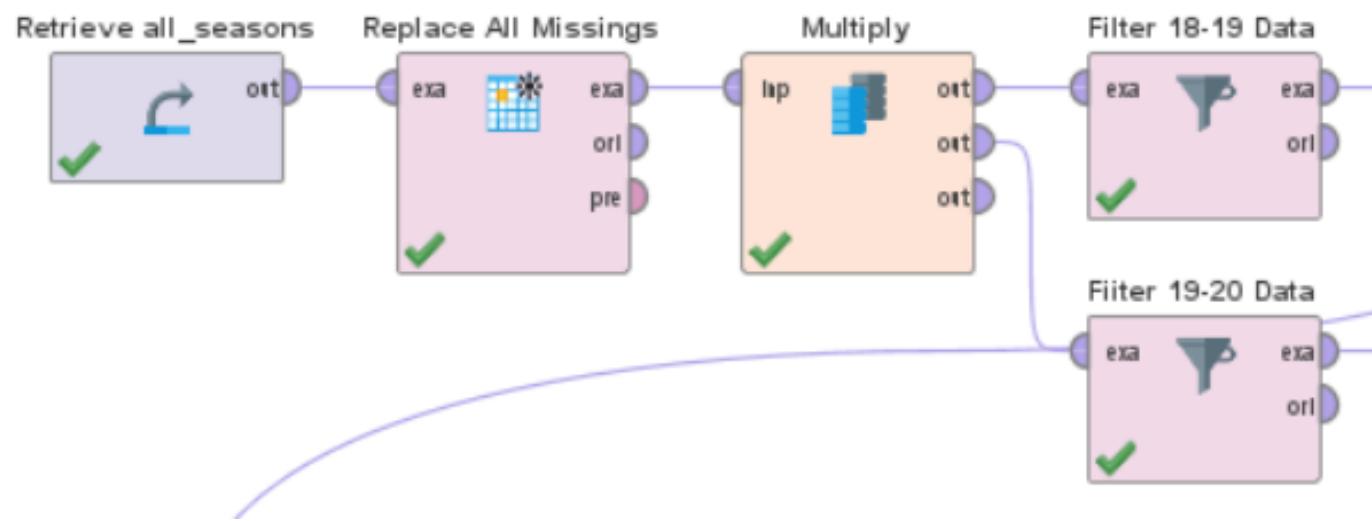
Measurement

- CFMC
 - Competence=net rating
 - Future=age score (50 minus age)
 - Money=salary
 - Contribution=rebounds + assists + points/2

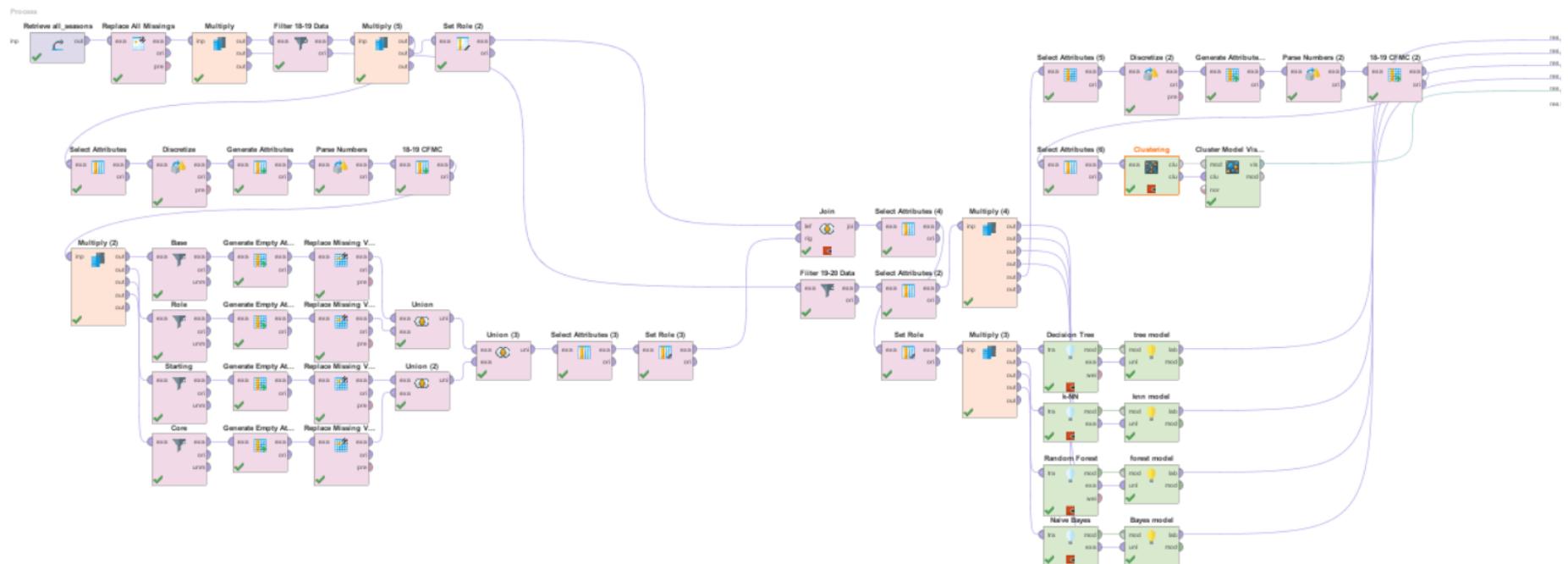
Data Understanding

- Source of data: Kaggle, from NBA Stats API
- Salary: Hoopshype
- Authenticity

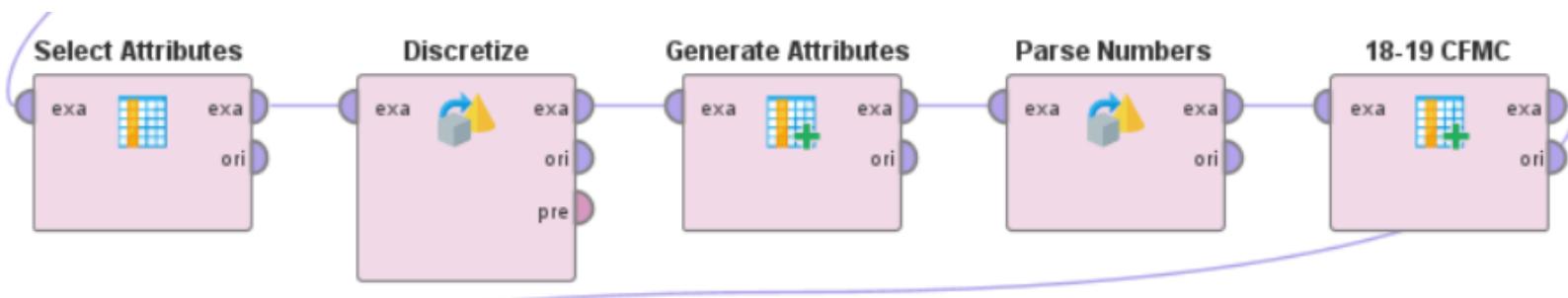
Data Preparation



Modeling



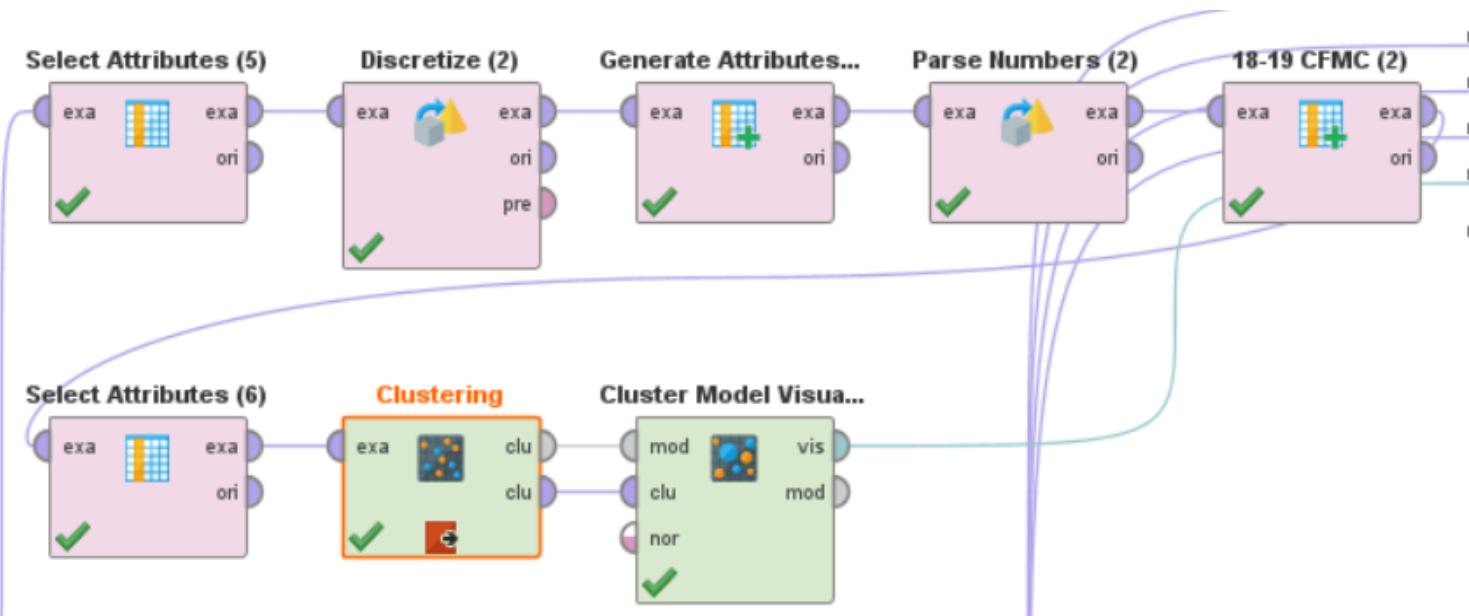
RFM / CFMC



18-19 CFMC Score

Row No.	player_name	C	F	M	Contribution	Score ↓
11	Giannis Antetokounmpo	4	4	4	5	17
85	Joel Embiid	4	4	4	5	17
119	James Harden	4	3	5	5	17
181	Anthony Davis	4	4	4	5	17
275	Russell Westbrook	4	3	5	5	17
478	Nikola Jokic	4	4	4	5	17
161	Blake Griffin	4	3	5	4	16
172	Andre Drummond	4	4	4	4	16
242	Damian Lillard	4	4	4	4	16
282	Stephen Curry	4	3	5	4	16
291	Rudy Gobert	4	4	4	4	16
300	Paul George	4	3	5	4	16
394	Kevin Durant	4	3	5	4	16
420	LeBron James	4	2	5	5	16
449	Karl-Anthony Towns	4	5	2	5	16
450	Kawhi Leonard	4	4	4	4	16
458	Mike Conley	4	3	5	4	16

Clustering



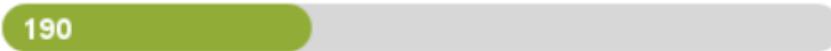
Number of Clusters: 4
Distance Measure: Euclidean Distance
Average Cluster Distance: 1.051
Davies-Bouldin Index: 1.105

Cluster 0  Average Distance: 1.852

M is on average **476.16%** larger, Contribution is on average **157.05%** larger, F is on average **20.07%** smaller

Cluster 1  Average Distance: 1.231

F is on average **42.38%** smaller, M is on average **12.31%** smaller, Contribution is on average **9.18%** smaller

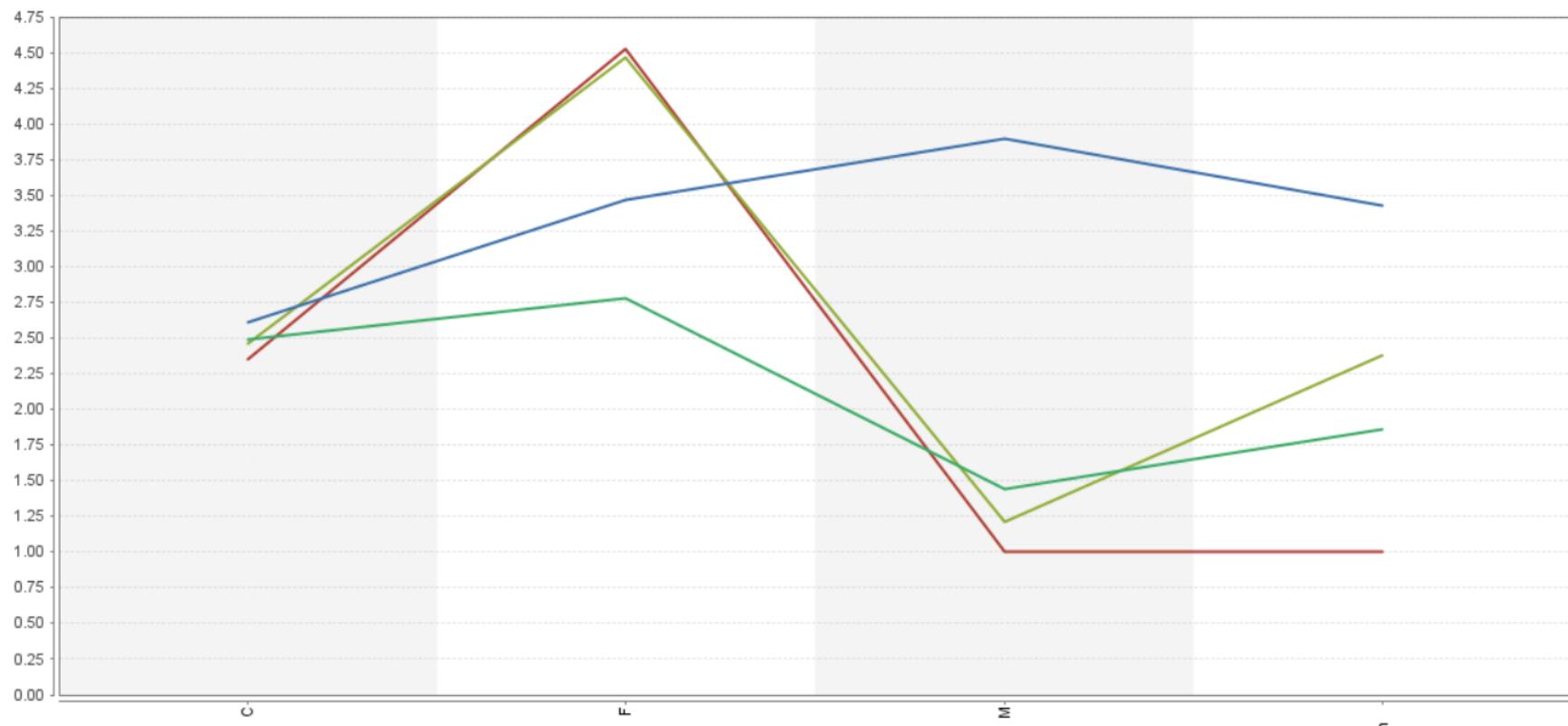
Cluster 2  Average Distance: 1.089

M is on average **58.22%** smaller, Contribution is on average **45.54%** larger, F is on average **12.51%** larger

Cluster 3  Average Distance: 0.642

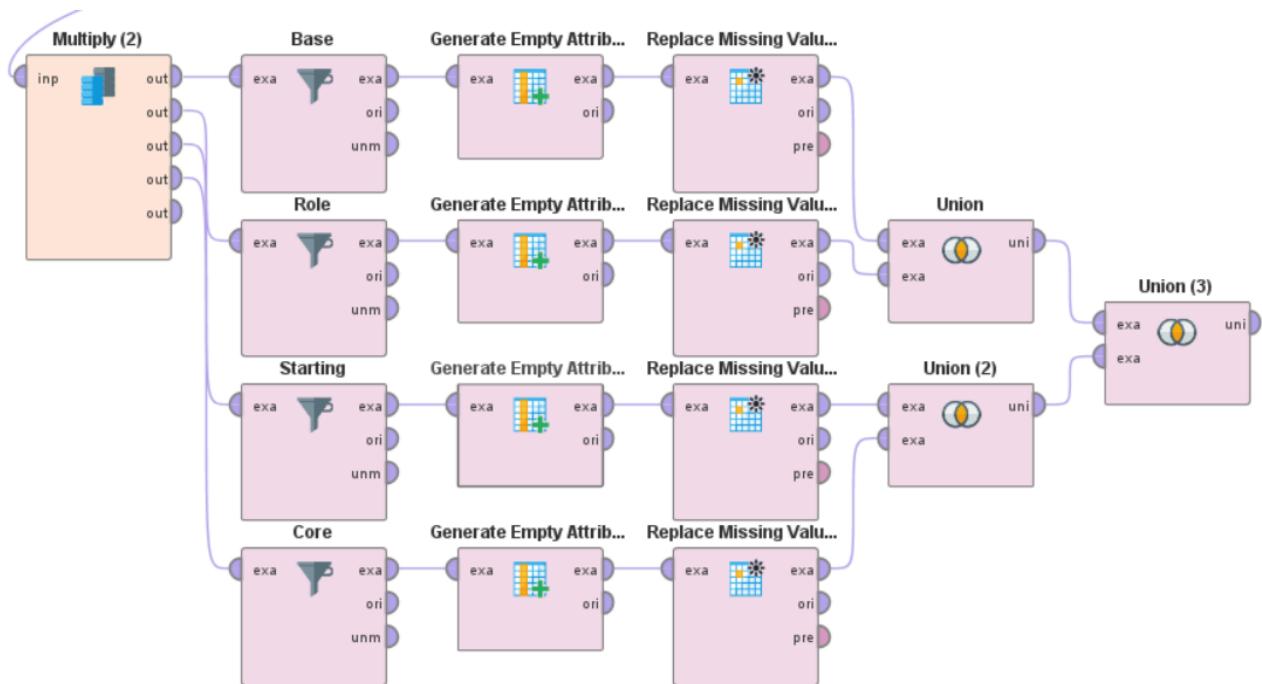
Contribution is on average **100.00%** smaller, M is on average **98.87%** smaller, F is on average **14.28%** larger

Cluster 0 (62) **Cluster 1** (86) **Cluster 2** (190) **Cluster 3** (176)



Continued

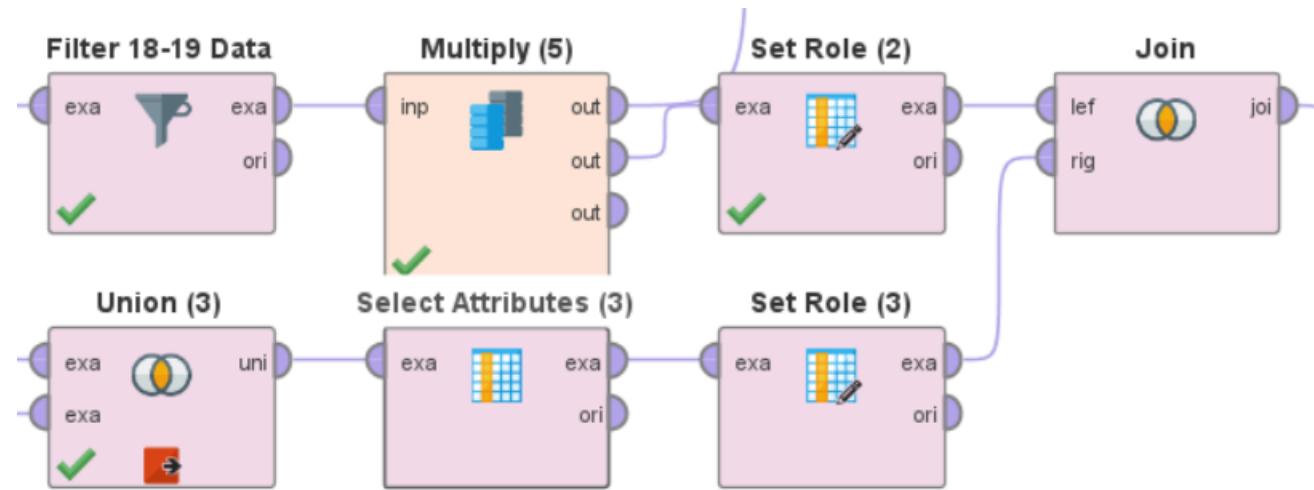
Classification



Class assigned to each player

Row No.	player_name	class	C	F	M	Contribution	Score ↓
504	Giannis Antetokounmpo	core players	4	4	4	5	17
505	Joel Embiid	core players	4	4	4	5	17
506	James Harden	core players	4	3	5	5	17
511	Anthony Davis	core players	4	4	4	5	17
515	Russell Westbrook	core players	4	3	5	5	17
529	Nikola Jokic	core players	4	4	4	5	17
509	Blake Griffin	core players	4	3	5	4	16
510	Andre Drummond	core players	4	4	4	4	16
514	Damian Lillard	core players	4	4	4	4	16
516	Stephen Curry	core players	4	3	5	4	16
518	Rudy Gobert	core players	4	4	4	4	16
519	Paul George	core players	4	3	5	4	16
521	Kevin Durant	core players	4	3	5	4	16
524	LeBron James	core players	4	2	5	5	16
526	Karl-Anthony Towns	core players	4	5	2	5	16
527	Kawhi Leonard	core players	4	4	4	4	16
528	Mike Conley	core players	4	3	5	4	16

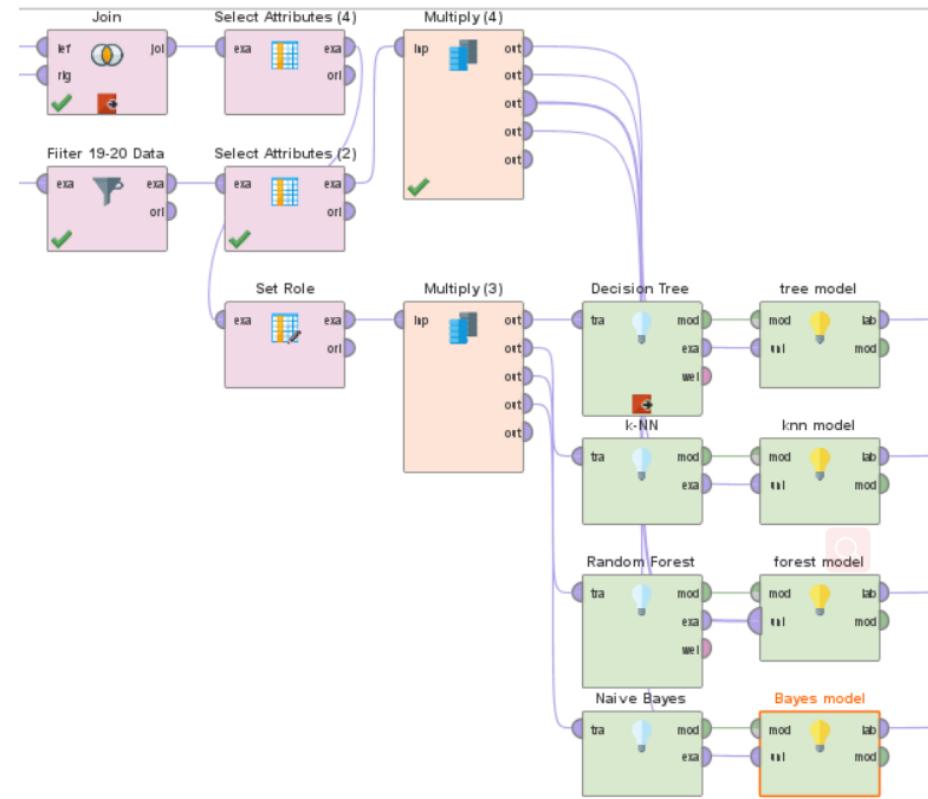
Join Data



Join Result

Row No.	player_name	class	att1	team_abbre...	age	player_height	player_weight	college	country	draft_ye
1	Gary Payton II	role players	10101	WAS	26	190.500	86.182	Oregon State	USA	Undrafte
2	Isaiah Briscoe	role players	10102	ORL	23	190.500	97.522	None	USA	Undrafte
3	Goran Dragic	starting players	10103	MIA	33	190.500	86.182	None	Slovenia	2008
4	Glenn Robinson III	role players	10104	DET	25	198.120	100.697	Michigan	USA	2014
5	Gerald Green	role players	10105	HOU	33	200.660	92.986	None	USA	2005
6	Georges Niang	role players	10106	UTA	26	203.200	104.326	Iowa State	USA	2016
7	George King	minimum sal...	10107	PHX	25	198.120	99.790	University of ...	USA	2018
8	Gary Trent Jr.	role players	10108	POR	20	198.120	92.986	None	USA	Undrafte
9	Gordon Hayward	core players	10109	BOS	29	203.200	102.058	Butler	USA	2010
10	George Hill	role players	10110	MIL	33	190.500	85.275	None	USA	2008
11	Giannis Antetokounmpo	core players	10111	MIL	24	210.820	109.769	None	Greece	2013
12	Gorgui Dieng	role players	10112	MIN	29	210.820	114.305	Louisville	Senegal	2013
13	Haywood Highsmith	role players	10113	PHI	22	200.660	99.790	None	USA	Undrafte
14	Greg Monroe	role players	10114	PHI	29	210.820	120.202	Georgetown	USA	2010
15	Guerschon Yabusele	role players	10115	BOS	23	203.200	117.934	None	France	2016
16	Hamidou Diallo	role players	10116	OKC	20	195.580	89.811	Kentucky	USA	2018
17	Harrison Barnes	starting players	10117	SAC	27	203.200	102.058	North Carolina	USA	2012

Machine Learning Prediction



Row No.	player_name	prediction(↑)	confidence(minimum salary / rookies)	confidence(role players)	confidence(starting players)	confidence(core players)
8	Giannis Antetokounmpo	core players	0	0	0.000	1.000
23	Hassan Whiteside	core players	0	0	0.000	1.000
75	Jimmy Butler	core players	0	0	0.001	0.999
79	Joel Embiid	core players	0	0	0.000	1.000
116	James Harden	core players	0	0	0.000	1.000
136	Bradley Beal	core players	0	0	0.000	1.000
148	CJ McCollum	core players	0	0.000	0.258	0.742
178	Andre Drummond	core players	0	0	0.001	0.999
183	Anthony Davis	core players	0	0	0.000	1.000
203	DeMar DeRozan	core players	0	0	0.022	0.978
218	Devin Booker	core players	0	0	0.002	0.998
232	Chris Paul	core players	0	0	0.000	1.000
248	Damian Lillard	core players	0	0	0.000	1.000
261	Rudy Gobert	core players	0	0	0.003	0.997
263	Russell Westbrook	core players	0	0	0.000	1.000
302	Paul George	core players	0	0	0.070	0.930
371	Tobias Harris	core players	0	0.000	0.201	0.799

Row No.	player_name	prediction(↑)	confidence(minimum salary / rookies)	confidence(role players)	confidence(starting players)	confidence(core play...
8	Giannis Antetokounmpo	core players	0.010	0	0.030	0.960
14	Gordon Hayward	core players	0.001	0.035	0.161	0.803
23	Hassan Whiteside	core players	0	0.006	0.278	0.716
75	Jimmy Butler	core players	0.000	0.021	0.093	0.886
79	Joel Embiid	core players	0	0.015	0.187	0.798
97	Domantas Sabonis	core players	0.001	0.013	0.410	0.577
116	James Harden	core players	0.010	0	0	0.990
130	Blake Griffin	core players	0.001	0.106	0.201	0.692
136	Bradley Beal	core players	0.010	0.005	0.218	0.766
178	Andre Drummond	core players	0	0.001	0.208	0.792
183	Anthony Davis	core players	0	0.003	0.090	0.907
218	Devin Booker	core players	0.000	0.042	0.281	0.677
232	Chris Paul	core players	0.000	0.032	0.077	0.891
235	Clint Capela	core players	0.002	0.076	0.391	0.531
248	Damian Lillard	core players	0	0.001	0.108	0.891
261	Rudy Gobert	core players	0	0.005	0.225	0.770
263	Russell Westbrook	core players	0.010	0	0.020	0.970

Row No.	player_name	prediction(cl... ↑)	confidence(minimum salary / rookies)	confidence(role players)	confidence(starting players)	confidence(core players)
8	Giannis Antetokounmpo	core players	0	0	0.401	0.599
14	Gordon Hayward	core players	0	0	0	1.000
23	Hassan Whiteside	core players	0	0	0.402	0.598
75	Jimmy Butler	core players	0	0	0	1
116	James Harden	core players	0	0	0	1
130	Blake Griffin	core players	0	0	0	1
136	Bradley Beal	core players	0	0	0.402	0.598
178	Andre Drummond	core players	0	0	0.402	0.598
183	Anthony Davis	core players	0	0	0.402	0.598
218	Devin Booker	core players	0	0	0.406	0.594
232	Chris Paul	core players	0	0	0	1.000
243	D'Angelo Russell	core players	0	0	0.406	0.594
248	Damian Lillard	core players	0	0	0.203	0.797
261	Rudy Gobert	core players	0	0	0.195	0.805
263	Russell Westbrook	core players	0	0	0	1
281	Stephen Curry	core players	0	0	0	1
283	Steven Adams	core players	0	0	0.401	0.599

Row No.	player_name	prediction(cl... ↑)	confidence(minimum salary / rookies)	confidence(role players)	confidence(starting player...)	confidence(core players)
8	Giannis Antetokounmpo	core players	0	0	0	1
14	Gordon Hayward	core players	0	0	0	1
23	Hassan Whiteside	core players	0	0	0	1
75	Jimmy Butler	core players	0	0	0	1
79	Joel Embiid	core players	0	0	0	1
116	James Harden	core players	0	0	0	1
130	Blake Griffin	core players	0	0	0	1
136	Bradley Beal	core players	0	0	0	1
178	Andre Drummond	core players	0	0	0	1
183	Anthony Davis	core players	0	0	0	1
218	Devin Booker	core players	0	0	0	1
232	Chris Paul	core players	0	0	0	1
248	Damian Lillard	core players	0	0	0	1
261	Rudy Gobert	core players	0	0	0	1
263	Russell Westbrook	core players	0	0	0	1
281	Stephen Curry	core players	0	0	0	1
302	Paul George	core players	0	0	0	1

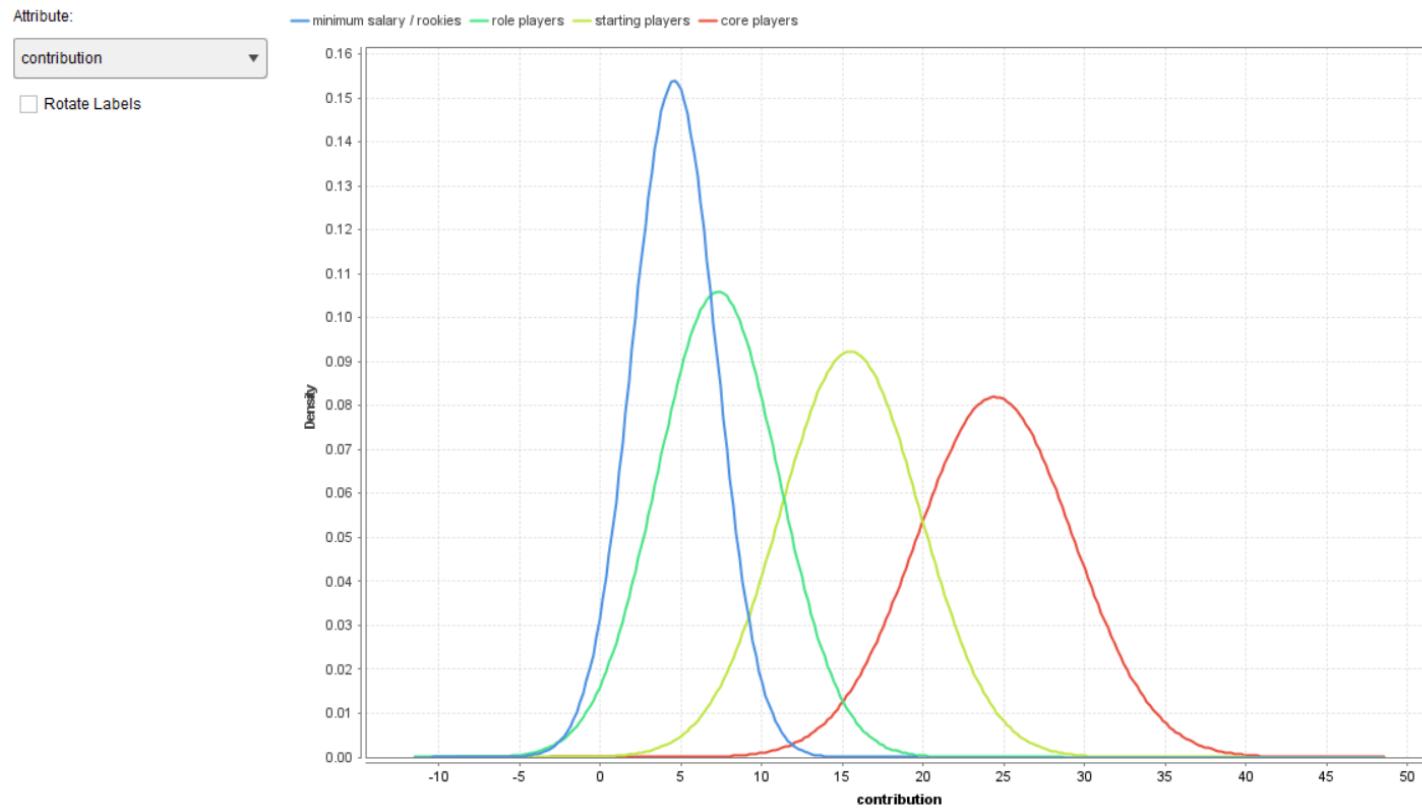
Evaluation

- Naïve Bayes
- Random Forest
- k-NN
- Decision Tree

Naïve Bayes

Row No.	player_name	predictio... ↑	confidence(minimum salary / rookies)	confidence(role players)	confidence(starting players)	confidence(core players)
8	Giannis Antetokounmpo	core players	0	0	0.000	1.000
23	Hassan Whiteside	core players	0	0	0.000	1.000
75	Jimmy Butler	core players	0	0	0.001	0.999
79	Joel Embiid	core players	0	0	0.000	1.000
116	James Harden	core players	0	0	0.000	1.000
136	Bradley Beal	core players	0	0	0.000	1.000
148	CJ McCollum	core players	0	0.000	0.258	0.742
178	Andre Drummond	core players	0	0	0.001	0.999
183	Anthony Davis	core players	0	0	0.000	1.000
203	DeMar DeRozan	core players	0	0	0.022	0.978
218	Devin Booker	core players	0	0	0.002	0.998
232	Chris Paul	core players	0	0	0.000	1.000
248	Damian Lillard	core players	0	0	0.000	1.000
261	Rudy Gobert	core players	0	0	0.003	0.997
263	Russell Westbrook	core players	0	0	0.000	1.000
302	Paul George	core players	0	0	0.070	0.930
371	Tobias Harris	core players	0	0.000	0.201	0.799

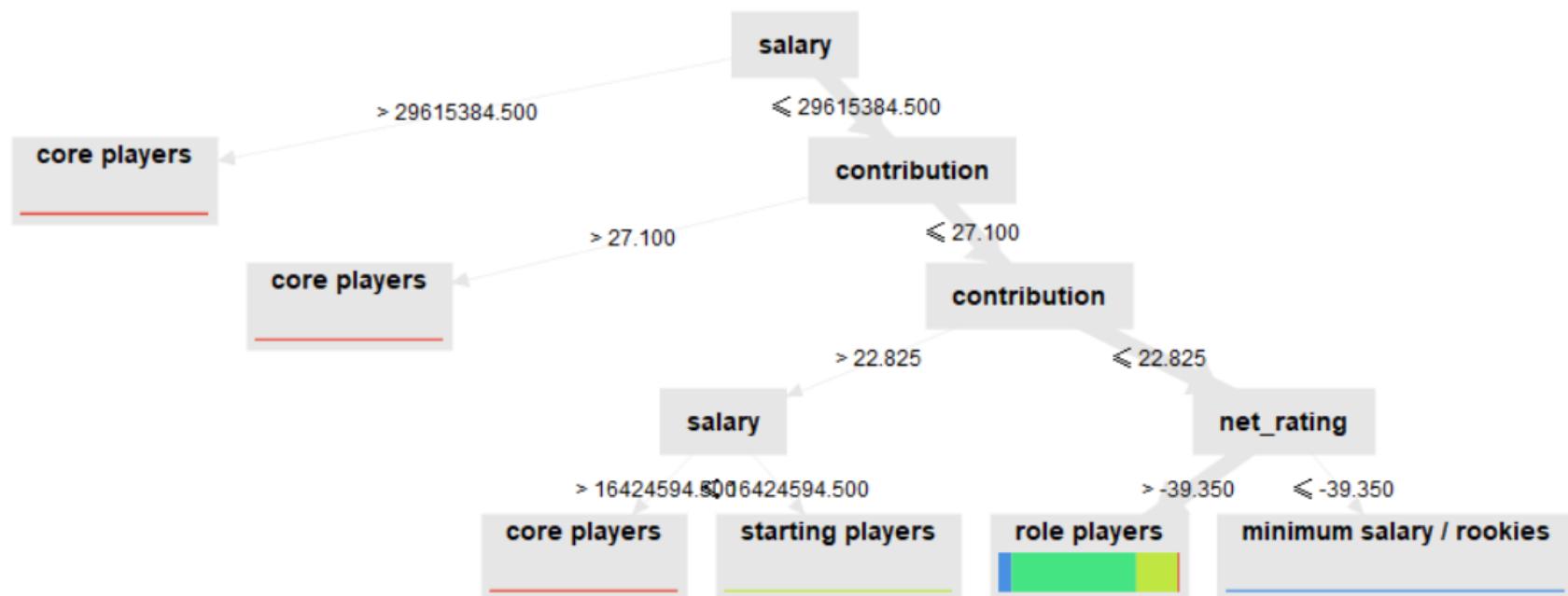
Contribution density curve



Decision Tree

Row No.	player_name	prediction(class ↑)	confidence(minimum salary / rookies)	confidence(role players)	confidence(starting player...)	confidence(core players)
8	Giannis Antetokounmpo	core players	0	0	0	1
14	Gordon Hayward	core players	0	0	0	1
23	Hassan Whiteside	core players	0	0	0	1
75	Jimmy Butler	core players	0	0	0	1
79	Joel Embiid	core players	0	0	0	1
116	James Harden	core players	0	0	0	1
130	Blake Griffin	core players	0	0	0	1
136	Bradley Beal	core players	0	0	0	1
178	Andre Drummond	core players	0	0	0	1
183	Anthony Davis	core players	0	0	0	1
218	Devin Booker	core players	0	0	0	1
232	Chris Paul	core players	0	0	0	1
248	Damian Lillard	core players	0	0	0	1
261	Rudy Gobert	core players	0	0	0	1
263	Russell Westbrook	core players	0	0	0	1
281	Stephen Curry	core players	0	0	0	1
302	Paul George	core players	0	0	0	1

Tree:



Random Forest

Row No.	player_name	prediction(class)	confidence(minimum salary / rookies)	confidence(role players)	confidence(starting players)	confidence(core play...
8	Giannis Antetokounmpo	core players	0.010	0	0.030	0.960
14	Gordon Hayward	core players	0.001	0.035	0.161	0.803
23	Hassan Whiteside	core players	0	0.006	0.278	0.716
75	Jimmy Butler	core players	0.000	0.021	0.093	0.886
79	Joel Embiid	core players	0	0.015	0.187	0.798
97	Domantas Sabonis	core players	0.001	0.013	0.410	0.577
116	James Harden	core players	0.010	0	0	0.990
130	Blake Griffin	core players	0.001	0.106	0.201	0.692
136	Bradley Beal	core players	0.010	0.005	0.218	0.766
178	Andre Drummond	core players	0	0.001	0.208	0.792
183	Anthony Davis	core players	0	0.003	0.090	0.907
218	Devin Booker	core players	0.000	0.042	0.281	0.677
232	Chris Paul	core players	0.000	0.032	0.077	0.891
235	Clint Capela	core players	0.002	0.076	0.391	0.531
248	Damian Lillard	core players	0	0.001	0.108	0.891
261	Rudy Gobert	core players	0	0.005	0.225	0.770
263	Russell Westbrook	core players	0.010	0	0.020	0.970

One example of trees



k-NN

Row No.	player_name	prediction(class)	confidence(minimum salary / rookies)	confidence(role players)	confidence(starting players)	confidence(core players)
8	Giannis Antetokounmpo	core players	0	0	0.401	0.599
14	Gordon Hayward	core players	0	0	0	1.000
23	Hassan Whiteside	core players	0	0	0.402	0.598
75	Jimmy Butler	core players	0	0	0	1
116	James Harden	core players	0	0	0	1
130	Blake Griffin	core players	0	0	0	1
136	Bradley Beal	core players	0	0	0.402	0.598
178	Andre Drummond	core players	0	0	0.402	0.598
183	Anthony Davis	core players	0	0	0.402	0.598
218	Devin Booker	core players	0	0	0.406	0.594
232	Chris Paul	core players	0	0	0	1.000
243	D'Angelo Russell	core players	0	0	0.406	0.594
248	Damian Lillard	core players	0	0	0.203	0.797
261	Rudy Gobert	core players	0	0	0.195	0.805
263	Russell Westbrook	core players	0	0	0	1
281	Stephen Curry	core players	0	0	0	1
283	Steven Adams	core players	0	0	0.401	0.599

k-NN Description

KNNClassification

Weighted 5-Nearest Neighbour model for classification.

The model contains 532 examples with 14 dimensions of the following classes:

minimum salary / rookies

role players

starting players

core players

Example

8	core players	Giannis Antetokounmpo	0	0	0.000	1.000
13	role players	Glenn Robinson III	0.000	0.822	0.178	0.000
28	role players	Goga Bitadze	0.001	0.999	0.000	0
15	starting players	Goran Dragic	0	0.000	0.928	0.072
14	starting players	Gordon Hayward	0	0.000	0.681	0.319
17	starting players	Gorgui Dieng	0.000	0.046	0.954	0.000
18	role players	Grant Williams	0.001	0.999	0.000	0.000
19	role players	Grayson Allen	0.001	0.999	0.000	0

Deployment

- Provide this model to managers
- They only have to input the new season data to generate new class predictions for their trading purposes.



Thank you for listening!

Any questions?