JORDAN VS LEBRON

Aprenentatge Computacional: Cas Kaggle
Matheus Mingorance Maciel, 1562738

CONTINGUT DE LA PRESENTACIÓ

01

Base de dades i anàlisi d'atributs

Presentació del dataset i dels objectius de l'estudi

02

Primer model: Regressió

Predir els PPG de Lebron a partir dels 33 anys

03

Segon model: Classificació

Model per classificar partits de Lebron i Jordan

04

Conclusió final

Falsos positius, negatius, i reflexió final

01

BASE DE DADES I ANÀLISI D'ATRIBUTS

=== Primeres files del dataset de Michael Jordan ===

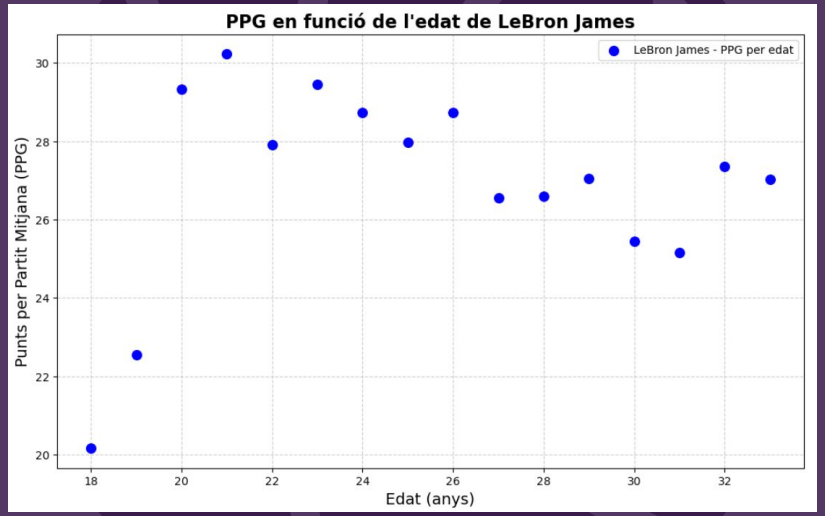
	game	date	age	team	opp	result	mp	fg	fga	fgp	three	threeatt	threep	ft	fta	ftp	orb	drb	trb	ast	stl	blk	tov	pts	game_score	minus_plus
0	1	1984-10-26	21-252	CHI	WSB	W (+16)	40:00	5	16	0.313	0	0	NaN	6	7	0.857	1	5	6	7	2	4	5	16	12.5	NaN
1	2	1984-10-27	21-253	CHI	MIL	L (-2)	34:00	8	13	0.615	0	0	NaN	5	5	1.000	3	2	5	5	2	1	3	21	19.4	NaN
2	3	1984-10-29	21-255	CHI	MIL	W (+6)	34:00	13	24	0.542	0	0	NaN	11	13	0.846	2	2	4	5	6	2	3	37	32.9	NaN
3	4	1984-10-30	21-256	CHI	KCK	W (+5)	36:00	8	21	0.381	0	0	NaN	9	9	1.000	2	2	4	5	3	1	6	25	14.7	NaN
4	5	1984-11-01	21-258	CHI	DEN	L (-16)	33:00	7	15	0.467	0	0	NaN	3	4	0.750	3	2	5	5	1	1	2	17	13.2	NaN

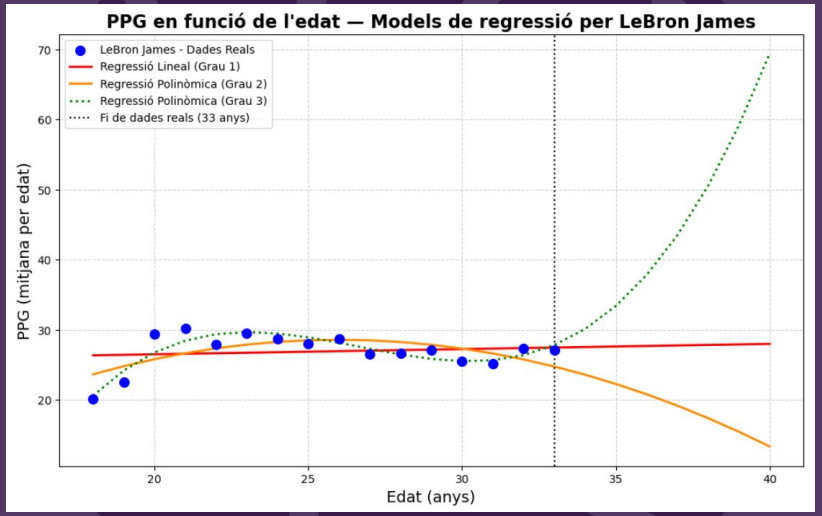
=== Primeres files del dataset de LeBron James ===

	game	date	age	team	opp	result	mp	fg	fga	fgp	three	threeatt	threep	ft	fta	ftp	orb	drb	trb	ast	stl	blk	tov	pts	game_score	minus_plus
0	1	2003-10-29	18-303	CLE	SAC	L (-14)	42:00	12	20	0.600	0	2	0.0	1	3	0.333	2	4	6	9	4	0	2	25	24.7	-9
1	2	2003-10-30	18-304	CLE	PHO	L (-9)	41:00	8	17	0.471	1	5	0.2	4	7	0.571	2	10	12	8	1	0	7	21	14.7	-3
2	3	2003-11-01	18-306	CLE	POR	L (-19)	39:00	3	12	0.250	0	1	0.0	2	2	1.000	0	4	4	6	2	0	2	8	5.0	-21
3	4	2003-11-05	18-310	CLE	DEN	L (-4)	41:00	3	11	0.273	0	2	0.0	1	1	1.000	2	9	11	7	2	3	2	7	11.2	-3
4	5	2003-11-07	18-312	CLE	IND	L (-1)	44:00	8	18	0.444	1	2	0.5	6	7	0.857	0	5	5	3	0	0	7	23	9.0	-7

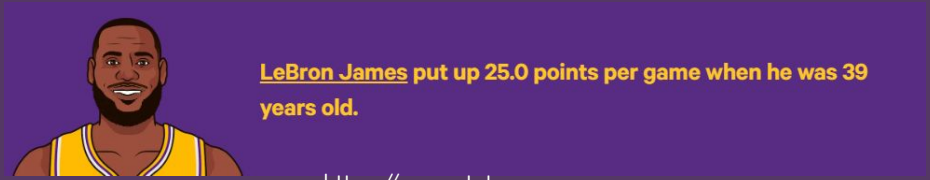
02 REGRESSIÓ

Predir els PPG de Lebron a partir dels 33 anys

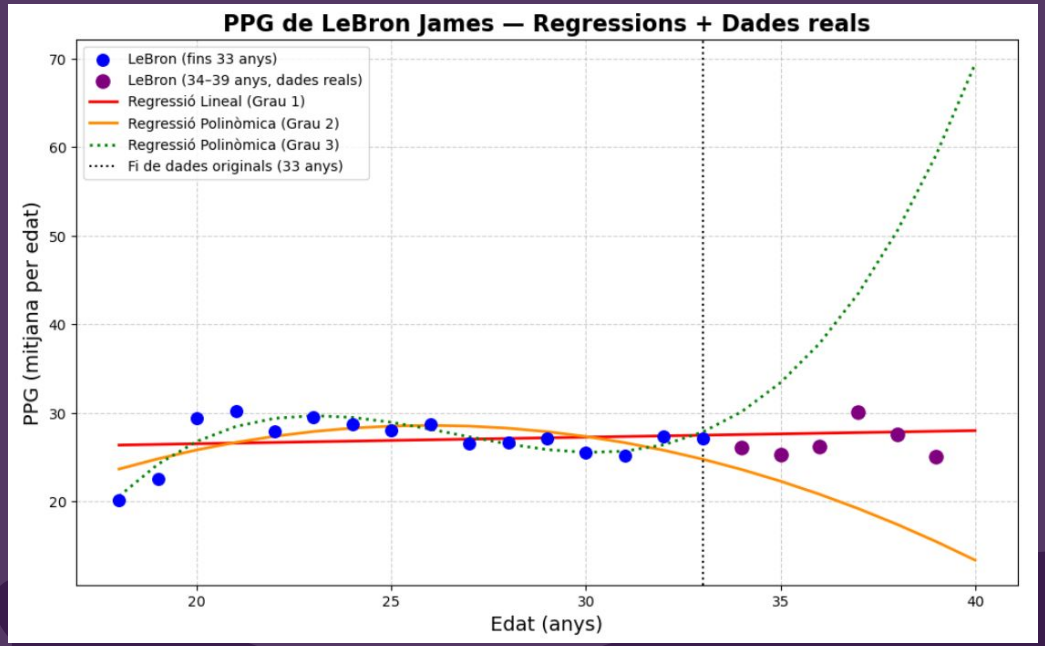

PPG en funció de l'edat de LeBron James
LeBron James - PPG per edat
Punts per Partit Mitjana (PPG)
Edat (anys)
20
22
24
26
28
30
18
20
22
24
26
28
30
32


PPG en funció de l'edat — Models de regressió per LeBron James
LeBron James - Dades Reals
Regressió Lineal (Grau 1)
Regressió Polinòmica (Grau 2)
Regressió Polinòmica (Grau 3)
Fi de dades reals (33 anys)
PPG (mitjana per edat)
Edat (anys)
20
30
40
50
60
70
20
25
30
35
40

DADES REALS



https://www.statmuse.com



===== COMPARACIÓ DE MODELS PER EDAT (34–39) =====

Edat	PPG Real	Pred Lineal	Err Lin	Acc Lin (%)	Pred Grau 2	Err G2	Acc G2 (%)	Pred Grau 3	Err G3	Acc G3 (%)
34.00	26.10	27.52	1.42	94.56	23.57	2.53	90.31	30.10	4.00	84.69
35.00	25.30	27.59	2.29	90.93	22.25	3.05	87.96	33.38	8.08	68.05
36.00	26.20	27.67	1.47	94.40	20.78	5.42	79.31	37.79	11.59	55.75
37.00	30.10	27.74	2.36	92.17	19.15	10.95	63.62	43.46	13.36	55.60
38.00	27.50	27.82	0.32	98.85	17.37	10.13	63.15	50.53	23.03	16.27
39.00	25.00	27.89	2.89	88.44	15.43	9.57	61.71	59.12	34.12	-36.47

LeBron James' longevity makes no sense once you confront the forgotten truth

Year 23 should not look this good for the Los Angeles Lakers superstar.

By Svyatoslav Rovenchuk | Nov 26, 2025

How LeBron James' unprecedented longevity factors into the GOAT debate

Kevin Pelton
Dec 30, 2020, 07:00 AM ET

LeBron James: The GOAT of longevity

Jim Johnson · March 14, 2025 · Lead by Example

03

CLASSIFICACIÓ

Model per classificar partits de Lebron i Jordan

NOU DATASET: CAREERS

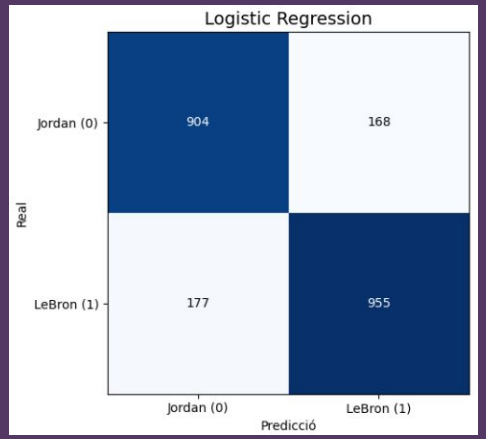
```
=== Primeres files del dataset de classificació ===
```

	mp	fg	fga	fgp	three	threeatt	threep	ft	fta	ftp	orb	drb	trb	ast	stl	blk	tov	pts	player
0	40:00:00	5	16	0.313	0	0	NaN	6	7	0.857	1	5	6	7	2	4	5	16	0
1	34:00:00	8	13	0.615	0	0	NaN	5	5	1.000	3	2	5	5	2	1	3	21	0
2	34:00:00	13	24	0.542	0	0	NaN	11	13	0.846	2	2	4	5	6	2	3	37	0
3	36:00:00	8	21	0.381	0	0	NaN	9	9	1.000	2	2	4	5	3	1	6	25	0
4	33:00:00	7	15	0.467	0	0	NaN	3	4	0.750	3	2	5	5	1	1	2	17	0

```
Percentatge de NaNs a threep: 16.79%
Percentatge de NaNs a ftp: 1.91%
```

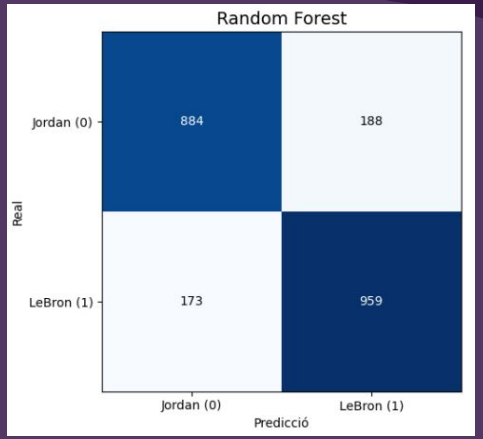
- AST_per_PTS
- 3PT_Attempt_Ratio
- TRB_per_FGA

GRÀFIQUES: TRAIN



Regressió Logística

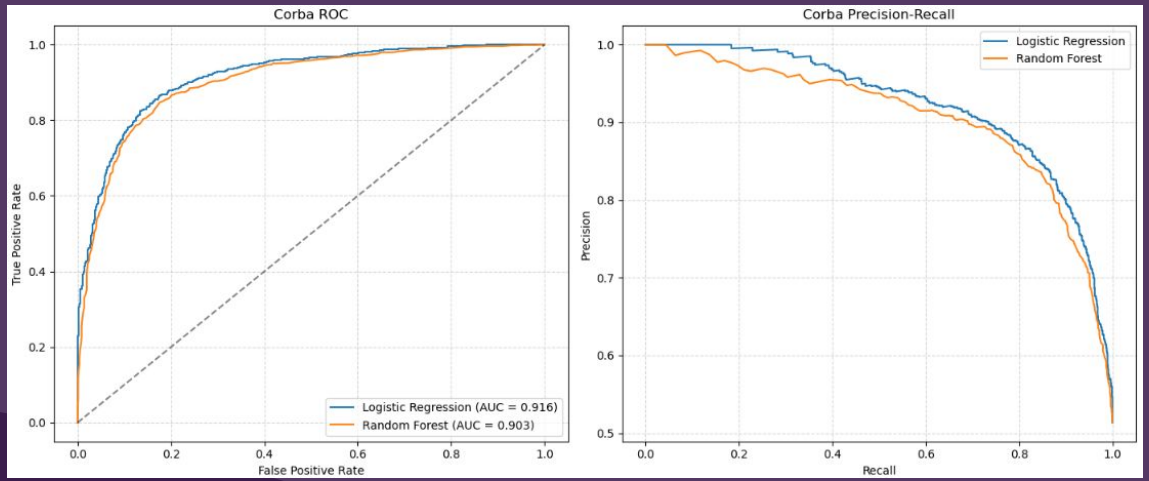
84,35%



Random Forest

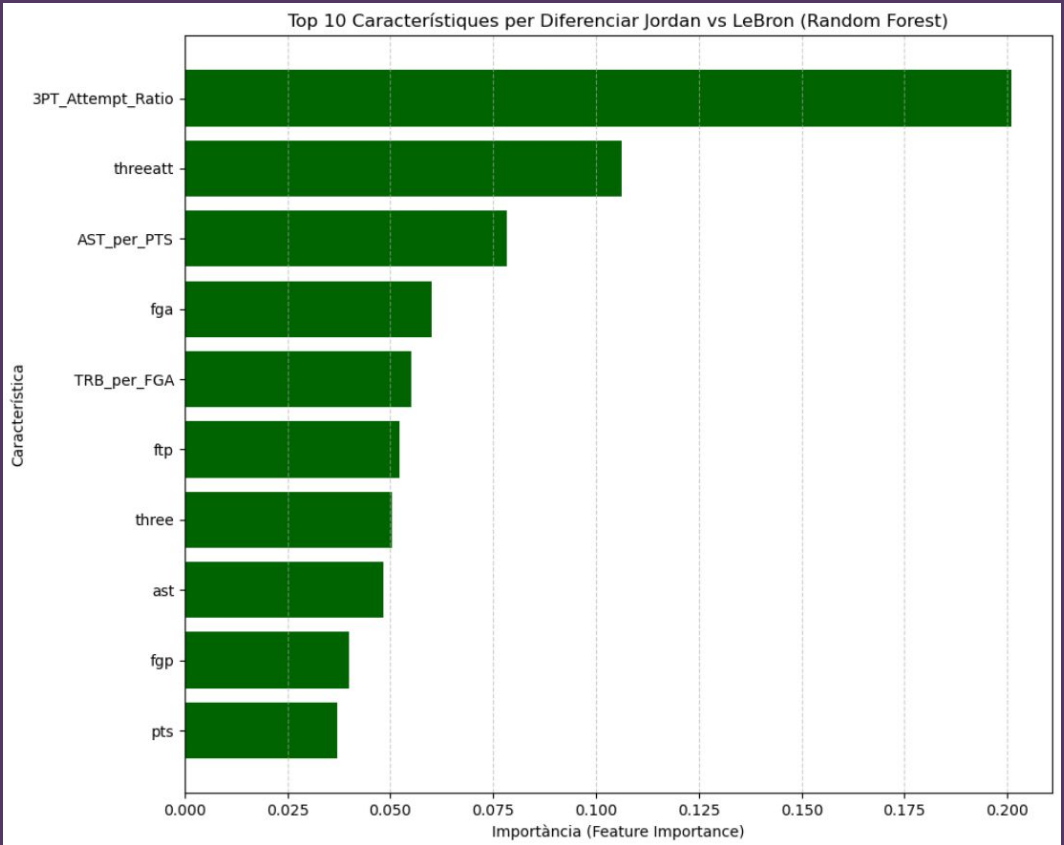
83,62%

ROC



PR

IMPORTÀNCIA DE CARACTERÍSTIQUES



CONJUNT DE TEST: PLAYOFFS

Michael Jordan Playoffs Game Log Menu

Michael Jordan Overview

Game Logs

High School & Amateur 1984-85
1985-86 1986-87 1987-88
1988-89 1989-90 1990-91
1991-92 1992-93 1994-95
1995-96 1996-97 1997-98
2001-02 2002-03
Career Playoffs

Splits

Shooting

Lineups

On/Off

More

On this page

Game Log Summary
Playoffs
Full Site Menu

179 Starts, 119 Wins, 60 Losses, 39 Double-Doubles, 2 Triple-Doubles

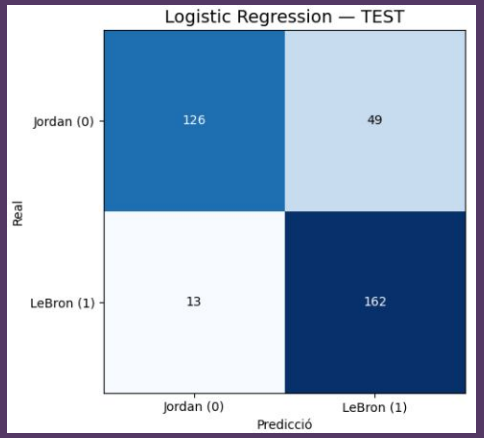
Playoffs

✨Upgraded✨ » Switch to Advanced Game Log Share & Export ▾ Glossary

Rk	Gcar	Gtm	Date	Team		Opp	Result	GS	MP	FG	FGA	FG%	3P	3PA	3P%	2P	2PA	2P%	eFG%	FT	FTA	FT%	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS	GmSc	+/-
1	1	1	1985-04-19	CHI	@	MIL	L, 100-109	*	44:00	7	19	.368	0	1	.000	7	18	.389	.368	9	10	.900	1	3	4	10	3	0	1	3	23	21.5	
2	2	2	1985-04-21	CHI	@	MIL	L, 115-122	*	43:00	9	17	.529	1	4	.250	8	13	.615	.559	11	12	.917	1	3	4	12	2	0	7	5	30	24.3	
3	3	3	1985-04-24	CHI		MIL	W, 109-107	*	42:00	12	26	.462	0	1	.000	12	25	.480	.462	11	16	.688	3	5	8	7	4	1	3	4	35	28.2	
4	4	4	1985-04-26	CHI		MIL	L, 97-105	*	42:00	6	16	.375	0	2	.000	6	14	.429	.375	17	20	.850	2	5	7	5	2	3	4	3	29	24.3	
Rk	Gcar	Gtm	Date	Team		Opp	Result	GS	MP	FG	FGA	FG%	3P	3PA	3P%	2P	2PA	2P%	eFG%	FT	FTA	FT%	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS	GmSc	+/-
5	5	1	1986-04-17	CHI	@	BOS	L, 104-123	*	43:00	18	36	.500	0	0		18	36	.500	.500	13	15	.867	0	4	4	2	2	1	5	3	49	29.3	
6	6	2	1986-04-20	CHI	@	BOS	L, 131-135 (2OT)	*	53:00	22	41	.537	0	0		22	41	.537	.537	19	21	.905	1	4	5	6	3	2	4	4	63	47.2	
7	7	3	1986-04-22	CHI		BOS	L, 104-122	*	39:00	8	18	.444	1	1	1.000	7	17	.412	.472	2	3	.667	4	6	10	9	2	1	5	6	19	15.4	
Rk	Gcar	Gtm	Date	Team		Opp	Result	GS	MP	FG	FGA	FG%	3P	3PA	3P%	2P	2PA	2P%	eFG%	FT	FTA	FT%	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS	GmSc	+/-
8	8	1	1987-04-23	CHI	@	BOS	L, 104-108	*	44:00	12	25	.480	0	1	.000	12	24	.500	.480	11	11	1.000	0	6	6	7	4	2	2	4	35	30.8	
9	9	2	1987-04-26	CHI	@	BOS	L, 96-105	*	45:00	14	29	.483	2	3	.667	12	26	.462	.517	12	14	.857	1	3	4	4	0	4	4	3	42	28.5	
10	10	3	1987-04-28	CHI		BOS	L, 94-105	*	39:00	9	30	.300	0	1	.000	9	29	.310	.300	12	14	.857	6	5	11	7	2	1	2	4	30	21.5	
Rk	Gcar	Gtm	Date	Team		Opp	Result	GS	MP	FG	FGA	FG%	3P	3PA	3P%	2P	2PA	2P%	eFG%	FT	FTA	FT%	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS	GmSc	+/-
11	11	1	1988-04-28	CHI		CLE	W, 104-93	*	44:00	19	35	.543	0	0		19	35	.543	.543	12	12	1.000	1	6	7	2	2	2	5	3	50	34.2	
12	12	2	1988-05-01	CHI		CLE	W, 106-101	*	44:00	24	45	.533	0	0		24	45	.533	.533	7	7	1.000	3	3	6	3	4	1	1	4	55	40.3	
13	13	3	1988-05-03	CHI	@	CLE	L, 102-110	*	44:00	12	25	.480	0	0		12	25	.480	.480	14	15	.933	2	3	5	9	2	1	4	5	38	30.2	
14	14	4	1988-05-05	CHI	@	CLE	L, 91-97	*	42:00	18	25	.720	0	0		18	25	.720	.720	8	9	.889	2	3	5	4	4	2	4	6	44	37.4	
15	15	5	1988-05-08	CHI		CLE	W, 107-101	*	43:00	12	22	.545	0	0		12	22	.545	.545	15	18	.833	1	3	4	6	2	2	7	4	39	27.8	
16	16	6	1988-05-10	CHI	@	DET	L, 82-93	*	40:00	10	22	.455	0	1	.000	10	21	.476	.455	9	10	.900	5	6	11	6	0	1	5	4	29	20.8	
17	17	7	1988-05-12	CHI	@	DET	W, 105-95	*	44:00	14	22	.636	0	0		14	22	.636	.636	8	10	.800	4	7	11	1	1	1	2	5	36	28.7	
18	18	8	1988-05-14	CHI		DET	L, 79-101	*	38:00	8	20	.400	0	0		8	20	.400	.400	8	9	.889	1	6	7	3	1	1	5	3	24	12.9	
19	19	9	1988-05-15	CHI		DET	L, 77-96	*	45:00	11	22	.500	0	0		11	22	.500	.500	1	2	.500	2	5	7	5	6	0	3	0	23	21.0	
20	20	10	1988-05-18	CHI	@	DET	L, 95-102	*	43:00	10	22	.455	1	2	.500	9	20	.450	.477	4	7	.571	2	6	8	8	2	0	3	4	25	18.6	
Rk	Gcar	Gtm	Date	Team		Opp	Result	GS	MP	FG	FGA	FG%	3P	3PA	3P%	2P	2PA	2P%	eFG%	FT	FTA	FT%	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS	GmSc	+/-
21	21	1	1989-04-28	CHI	@	CLE	W, 95-88	*	40:00	12	21	.571	0	1	.000	12	20	.600	.571	7	8	.875	0	5	5	11	4	1	6	5	31	26.6	
22	22	2	1989-04-30	CHI	@	CLE	L, 88-96	*	44:00	10	22	.455	0	1	.000	10	21	.476	.455	10	10	1.000	3	2	5	10	2	0	7	2	30	22.5	
23	23	3	1989-05-03	CHI		CLE	W, 101-94	*	39:00	18	34	.529	1	1	1.000	17	33	.515	.544	7	9	.778	2	5	7	10	5	1	5	4	44	35.6	
24	24	4	1989-05-05	CHI		CLE	L, 105-108 (OT)	*	43:00	14	28	.500	0	1	.000	14	27	.519	.500	22	27	.815	0	3	3	4	3	0	1	6	50	37.3	
25	25	5	1989-05-07	CHI	@	CLE	W, 101-100	*	44:00	17	32	.531	1	1	1.000	16	31	.516	.547	9	13	.692	3	6	9	6	1	0	2	3	44	32.7	
26	26	6	1989-05-09	CHI	@	NYK	W, 120-109 (OT)	*	46:00	11	17	.647	1	2	.500	10	15	.667	.676	11	11	1.000	2	8	10	12	2	0	6	5	34	32.7	
27	27	7	1989-05-11	CHI	@	NYK	L, 97-114	*	34:00	7	17	.412	0	2	.000	7	15	.467	.412	1	4	.250	2	6	8	3	1	0	1	4	15	8.4	
28	28	8	1989-05-13	CHI		NYK	W, 111-88	*	39:00	14	25	.560	1	2	.500	13	23	.565	.580	11	13	.846	2	13	15	9	6	1	3	3	40	41.4	
29	29	9	1989-05-14	CHI		NYK	W, 106-93	*	42:00	12	18	.667	0	1	.000	12	17	.706	.667	23	28	.821	0	11	11	6	1	2	6	3	47	39.9	
30	30	10	1989-05-16	CHI	@	NYK	L, 114-121	*	43:00	13	30	.433	1	3	.333	12	27	.444	.450	11	15	.733	5	3	8	10	5	1	3	2	38	33.9	
31	31	11	1989-05-19	CHI		NYK	W, 113-111	*	45:00	14	22	.636	1	3	.333	13	19	.684	.659	11	12	.917	1	4	5	10	0	4	6	3	40	34.3	

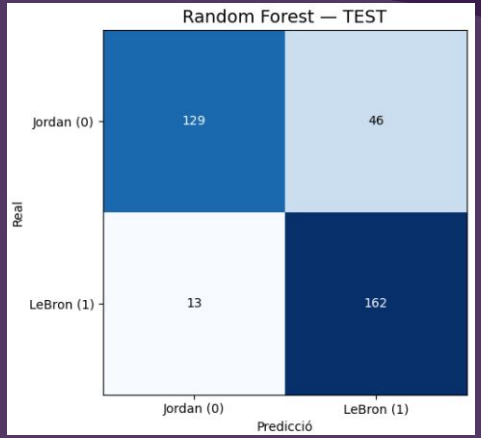
https://www.basketball-reference.com

GRÀFIQUES: TEST



Regressió Logística

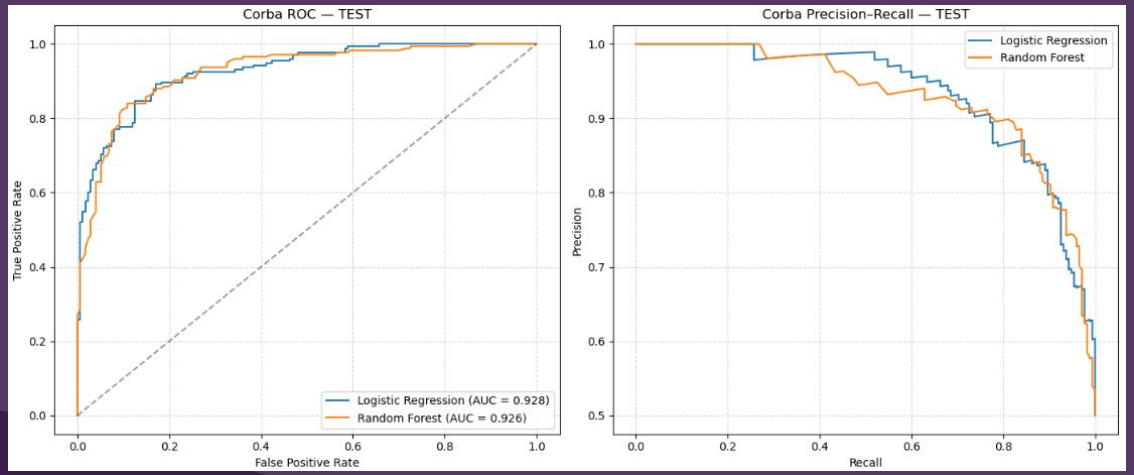
82,29%



Random Forest

83,14%

ROC



PR

04 CONCLUSIÓ

Falsos positius, negatius, i reflexió final

===== LOGISTIC REGRESSION — FALSE POSITIVES (Jordan → LeBron) =====

	mp	fg	fga	fgp	three	threeatt	threep	ft	fta	ftp	orb	drb	trb	ast	stl	blk	tov	pts	player	AST_per_PTS	3PT_Attempt_Ratio	TRB_per_FGA	pred_logreg	pred_rf
15	18.975595	10	22	0.455	1	2	0.5	4	7	0.571	2	6	8	8	2	0	3	25	0	0.320000	0.090909	0.363636	1	1
21	18.975595	11	17	0.647	1	2	0.5	11	11	1.000	2	8	10	12	2	0	6	34	0	0.352941	0.117647	0.588235	1	1
22	18.975595	7	17	0.412	0	2	0.0	1	4	0.250	2	6	8	3	1	0	1	15	0	0.200000	0.117647	0.470588	1	1

===== LOGISTIC REGRESSION — FALSE NEGATIVES (LeBron → Jordan) =====

	mp	fg	fga	fgp	three	threeatt	threep	ft	fta	ftp	orb	drb	trb	ast	stl	blk	tov	pts	player	AST_per_PTS	3PT_Attempt_Ratio	TRB_per_FGA	pred_logreg	pred_rf
184	18.975595	8	17	0.471	0	1	0.0	6	6	1.000	0	4	4	3	0	0	0	22	1	0.136364	0.058824	0.235294	0	0
188	18.975595	8	17	0.471	1	2	0.5	15	16	0.938	2	6	8	4	2	1	2	32	1	0.125000	0.117647	0.470588	0	0
192	18.975595	15	25	0.600	1	2	0.5	2	4	0.500	3	8	11	4	2	2	1	33	1	0.121212	0.080000	0.440000	0	0

GRÀCIES

Aprenentatge Computacional Curs 2025/2026