

# Progetto Finale IDS 2020/21

## Regressione di Poisson per predizioni di partite di calcio e analisi in base al modulo di gioco

Rodolfo Centanaro 4719894, Marien Mehilka 4554200

### Contesto e motivazioni

- **Dataset**

FC Internazionale Milano vs AC Milan

Dataset di 200 partite con risultati e moduli di gioco utilizzati

File csv:

	goals_ht	goals_at	home_t	away_t	formation_ht	formation_at
0	1	1	Inter	Milan	3-5-2	4-4-2
1	0	1	Inter	Milan	5-3-2	5-4-1
2	1	2	Inter	Milan	3-5-2	4-3-3
3	4	2	Inter	Milan	4-3-3	4-3-3
4	1	1	Inter	Milan	4-3-3	4-5-1
...	...	...	...	...	...	...
195	2	1	Milan	Inter	3-4-3	4-3-3
196	3	3	Milan	Inter	5-3-2	3-4-3
197	1	1	Milan	Inter	3-5-2	3-4-3
198	2	3	Milan	Inter	3-4-3	4-5-1
199	0	3	Milan	Inter	4-4-2	5-4-1

200 rows × 6 columns

- **Tipo di analisi**

## 1) Predizione

La nostra prima analisi si basa sulla predizione dei possibili risultati fra due squadre.

Per fare ciò abbiamo utilizzato la regressione di Poisson che è una forma di modello lineare generalizzato di analisi di regressione.

$$P(X) = \frac{\lambda^X \cdot e^{-\lambda}}{X!}$$

Dove (nel nostro caso) lambda è il numero medio di goals effettuati da una squadra, mentre X è il numero di goals di cui si vuole la probabilità.

## 2) Reportistica

Abbiamo anche effettuato un'analisi in base ai moduli di gioco per mostrare diversi dati statistici come percentuale utilizzo o numero vittorie. Inoltre abbiamo anche prodotto una tabella che mostra le percentuali di vittoria per ogni combinazione di moduli delle due squadre analizzate.

## **Descrizione del prototipo**

- Linguaggio usato:

Python (jupyter notebook)

- Eventuali librerie adottate:

Pandas: manipolazione e analisi di dati

Matplotlib.pyplot: creazione di grafici

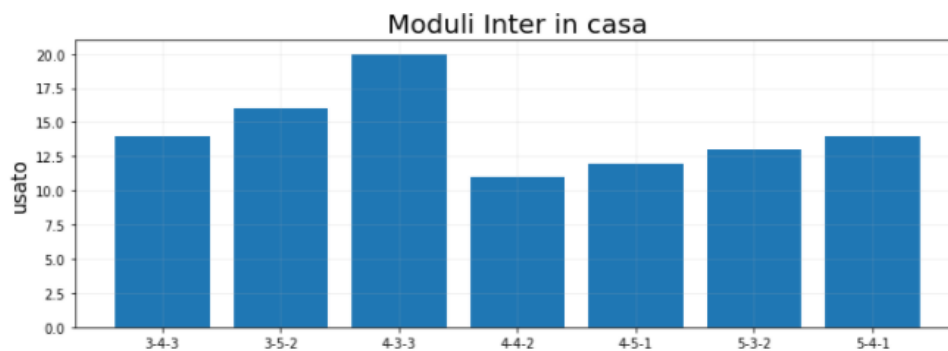
Scipy.stats: per applicare poisson e la pmf (probability mass function)

# Risultati

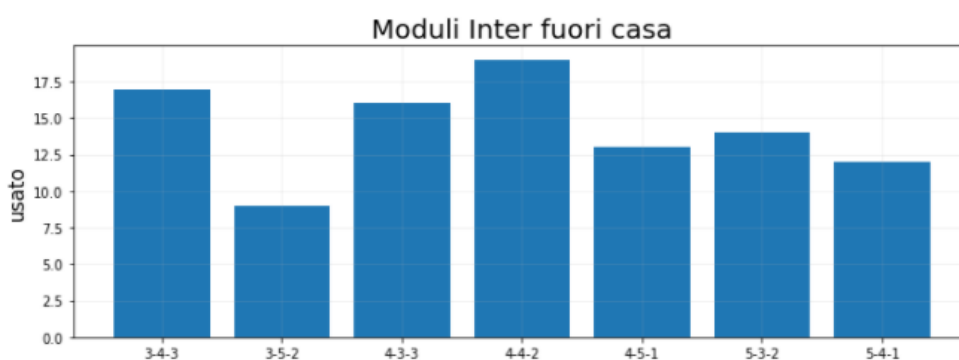
Per come abbiamo strutturato il codice è possibile fare un analisi sia sull'Inter che sul Milan cambiando il parametro che identifica il team nelle funzioni create.

Questi sono i risultati ottenuti effettuando l'analisi sull'Inter:

Grafico che mostra i moduli di gioco più usati e tabella che mostre varie statistiche sui diversi moduli:



	Modulo	PercUse	Wins	Draws	Losses	PercWin
0	3-4-3	14.0	5	2	7	35.71
1	3-5-2	16.0	7	3	6	43.75
2	4-3-3	20.0	11	3	6	55.00
3	4-4-2	11.0	3	4	4	27.27
4	4-5-1	12.0	6	3	3	50.00
5	5-3-2	13.0	4	4	5	30.77
6	5-4-1	14.0	6	4	4	42.86



	Modulo	PercUse	Wins	Draws	Losses	PercWin
0	3-4-3	17.0	10	3	4	58.82
1	3-5-2	9.0	4	3	2	44.44
2	4-3-3	16.0	5	4	7	31.25
3	4-4-2	19.0	6	5	8	31.58
4	4-5-1	13.0	7	2	4	53.85
5	5-3-2	14.0	4	3	7	28.57
6	5-4-1	12.0	6	2	4	50.00

Grafico della regressione di Poisson con tabella che mostra la probabilità dei risultati:

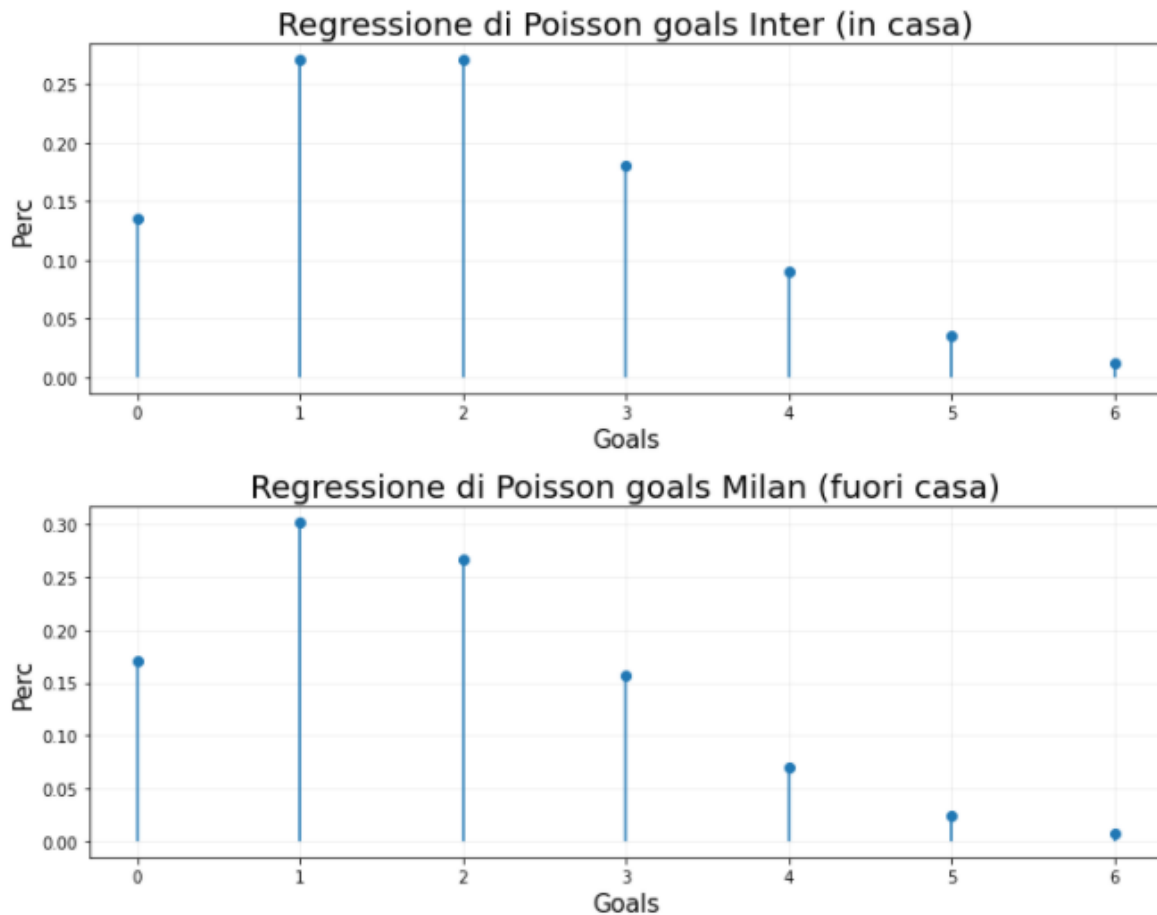


Tabella che mostra la probabilita' dei risultati  
(row indexes = goals Inter, col indexes = goals Milan)

	0	1	2	3	4	5	6
0	2.31	4.08	3.61	2.13	0.94	0.33	0.10
1	4.61	8.16	7.22	4.26	1.89	0.67	0.20
2	4.61	8.16	7.22	4.26	1.89	0.67	0.20
3	3.07	5.44	4.81	2.84	1.26	0.44	0.13
4	1.54	2.72	2.41	1.42	0.63	0.22	0.07
5	0.61	1.09	0.96	0.57	0.25	0.09	0.03
6	0.20	0.36	0.32	0.19	0.08	0.03	0.01

## Tabella percentuale di vincita in base al modulo:

INTER (as home\_t)

Tabella che mostra la percentuale di vincita in base al modulo  
(row indexes = moduli Inter, col indexes = moduli Milan)

	3-4-3	3-5-2	4-3-3	4-4-2	4-5-1	5-3-2	5-4-1
3-4-3	0	100.00	33.33	100.00	33.33	33.33	None
3-5-2	66.67	66.67	50.00	50.00	0	0	0
4-3-3	20.00	50.00	100.00	100.00	50.00	0	66.67
4-4-2	0	0	50.00	50.00	50.00	0	None
4-5-1	0	60.00	0	0	100.00	100.00	None
5-3-2	0	25.00	0	100.00	0	50.00	None
5-4-1	0	66.67	33.33	33.33	50.00	None	None

INTER (as away\_t)

Tabella che mostra la percentuale di vincita in base al modulo  
(row indexes = moduli Milan, col indexes = moduli Inter)

	3-4-3	3-5-2	4-3-3	4-4-2	4-5-1	5-3-2	5-4-1
3-4-3	50.00	100.00	50.00	25.00	66.67	50.00	50.00
3-5-2	33.33	0	0	50.00	100.00	None	None
4-3-3	75.00	33.33	50.00	0	0	50.00	None
4-4-2	100.00	0	33.33	0	0	100.00	None
4-5-1	0	75.00	0	50.00	0	None	None
5-3-2	50.00	0	50.00	50.00	0	50.00	40.00
5-4-1	100.00	0	33.33	100.00	50.00	50.00	None

## Suddivisione del lavoro:

Collaborato insieme alla stesura del codice e al report.

## Bibliografia/Sitografia

Dataset: generato casualmente da noi

Regressione di Poisson: [https://it.wikipedia.org/wiki/Regressione\\_di\\_Poisson](https://it.wikipedia.org/wiki/Regressione_di_Poisson)  
[https://it.wikipedia.org/wiki/Distribuzione\\_di\\_Poisson](https://it.wikipedia.org/wiki/Distribuzione_di_Poisson)