

# Análisis de redes causales en deportes de equipo

Elena Merelo Molina

7 de septiembre de 2022

- 1 Introducción
  - Descripción del problema
  - Definiciones
  
- 2 Contribuciones
  - Resultados
  - Conclusiones y trabajos futuros
  
- 3 Extra

# Motivación

El fútbol es un deporte de equipo *complejo*

Como parte del cuerpo técnico de un equipo habrá que decidir:

- Quién será convocado y quién jugará
- Cuándo y qué cambios habrá
- Alineación por la que se abogará
- A quién fichar o vender, qué hacer en los entrenamientos,...

# Descripción

- ¿Analizar la entropía sobre el grafo de pases nos ayuda a entender mejor el desempeño de un equipo?
- ¿Hasta qué punto es determinante la entropía a nivel de jugador o equipo?
- ¿Es posible ver la entropía reflejada en un tipo de visualización de las redes de pases?

# Metodología

- **Desarrollo ágil**
- *Design thinking*

# Grafos

## Definition

$\mathcal{G} = (V, E)$ , con  $V$  conjunto finito de vértices distinguibles y  $E \subseteq V \times V$  conjunto de aristas. Un par ordenado  $(u, v) \in E$  denota un borde dirigido del vértice  $u$  al vértice  $v$ . Los bordes dirigidos se representarán con flechas.

En nuestro caso, representaremos las redes de pases mediante grafos no dirigidos, donde los nodos serán las jugadoras, y los arcos los pases.

# Redes bayesianas

## Definition (Redes bayesianas)

Una red bayesiana  $\mathfrak{B} = \{\mathcal{G}, \mathbb{P}\}$  está definida por:

- $\mathcal{G} = (V, E)$
- $(\Omega, \mathbb{P})$ .
- $\mathbf{V} = V[i], i = 1, \dots, N, (\Omega, \mathbb{P})$  de tal manera que
$$\mathbb{P}(V[1], \dots, V[N]) = \prod_{i=1}^n \mathbb{P}(V[i] | pa(V[i]))$$

El cálculo de la probabilidad conjunta que se usa en la entropía coincide con el modelo probabilístico de la red bayesiana.

# Entropía

Podemos medir la cantidad de información en una distribución de probabilidad usando la entropía de Shannon.

## Definition (Entropía)

Para  $X$  v.a. discreta  $H(X) := - \sum_x P(x) \log_2[P(x)]$ , donde  $X$  toma valores en  $\mathcal{X}$ ,  $P : \mathcal{X} \rightarrow [0, 1]$ . La entropía conjunta de las variables  $X_1 \dots X_N$  se define por

$$H(X_1, \dots, X_N) := - \sum_{x_1} \dots \sum_{x_N} P(x_1, \dots, x_N) \log_2[P(x_1, \dots, x_N)]$$

Calificaremos a los equipos según la entropía de las redes de pases, que caracteriza su grado de uniformidad o variabilidad.

## Sobre los datos usados

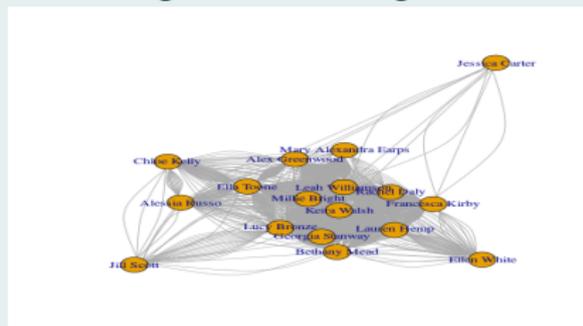
- StatsBomb 360 captura la posición de todos los jugadores por cada evento que ocurre.



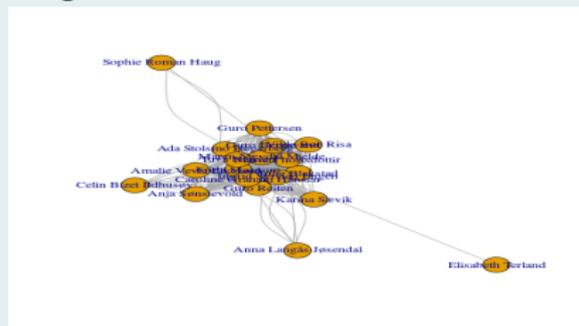
(Nunca antes se habían liberado datos 360 de StatsBomb para fútbol femenino)

# Preprocesamiento de datos

- FreeCompetitions
- FreeMatches
- Guardamos los de Inglaterra y Noruega y les extraemos los eventos
- 2246 pases completos de Inglaterra a lo largo de la EURO 2022 vs 967 de Noruega.
- 444 pases completos de Inglaterra en el partido Inglaterra-Noruega vs 177 de Noruega.

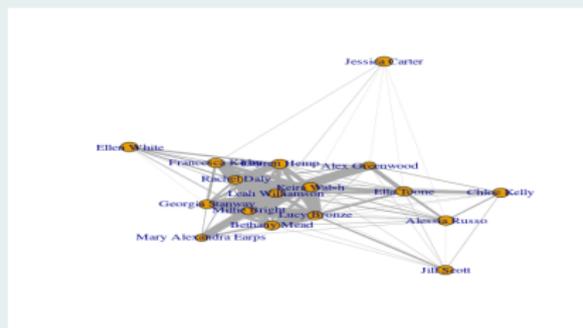


(Red de pases total de Inglaterra)

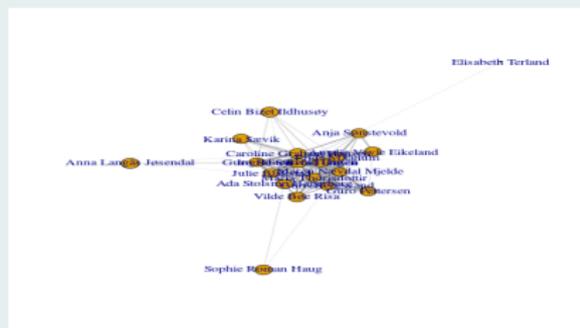


(Red de pases total de Noruega)

# Redes de pases a lo largo de la competición



(Red de pases total simplificada de Inglaterra)



(Red de pases total simplificada de Noruega)

## Nota

La entropía por jugadora es el tamaño de cada nodo. El grosor de los enlaces es el número de pases que ha habido.

## Partido Inglaterra 8 - Noruega 0

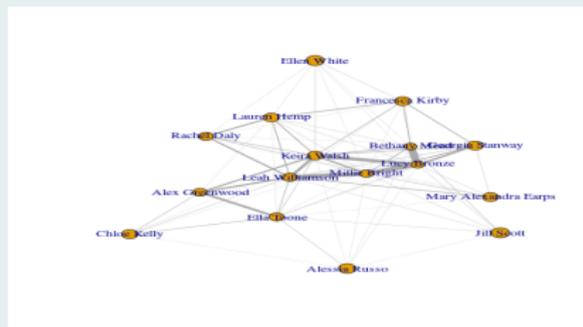


Figure: Red de pases simplificada de Inglaterra

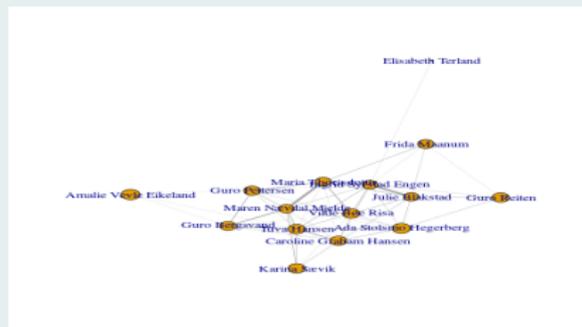


Figure: Red de pases simplificada de Noruega

# Comparativa por jugadoras a lo largo de la competición

Portera			Delantera	Elisabeth Terland	0
Defensa	Mary Alexandra Earps	0.6649	Defensa	Maria Thorisdottir	0.7752
Defensa	Alex Greenwood	0.7233	Defensa	Guro Bergsvand	0.779
Centrocampista	Millie Bright	0.7238	Defensa	Maren Nævdal Mjelde	0.8021
Defensa	Rachel Daly	0.7756	Centrocampista	Guro Retten	0.8261
Centrocampista	Lucy Bronze	0.8	Defensa	Julie Blakstad	0.8482
Defensa	Georgia Stanway	0.813	Centrocampista	Frida Maanum	0.8637
Delantera	Leah Williamson	0.8205	Portera	Guro Pettersen	0.8711
Centrocampista	Bethany Mead	0.8304	Defensa	Anja Sønstevold	0.8837
Delantera	Jill Scott	0.8581	Delantera	Karina Sævik	0.8917
Delantera	Chloe Kelly	0.8641	Delantera	Ada Hegerberg	0.8937
Centrocampista	Lauren Hemp	0.8735	Delantera	Amalie Evjel Eikeland	0.8969
Centrocampista	Francesca Kirby	0.8765	Centrocampista	Ingrid Syrstad Engen	0.897
Centrocampista	Keira Walsh	0.8783	Centrocampista	Tuva Hansen	0.9038
Centrocampista	Alessia Russo	0.8826	Delantera	Caroline Graham Hansen	0.9117
Centrocampista	Ella Toone	0.8881	Delantera	Sophie Roman Haug	0.9183
Delantera	Ellen White	0.9138	Centrocampista	Vilde Bøe Risa	0.9185
Defensa	Jessica Carter	0.9172	Centrocampista	Celin Bizet Ildhusøy	0.9474
			Delantera	Anna Langås Jøsendal	0.9591

(a) Entropía total de Inglaterra

(b) Entropía total de Noruega

## En el partido Inglaterra 8-Noruega 0

Centrocampista	Bethany Mead	0.6506			
Centrocampista	Lucy Bronze	0.6816	Delantera	Elisabeth Terland	0
Defensa	Millie Bright	0.7147	Defensa	Julie Blakstad	0.8037
Centrocampista	Alex Greenwood	0.7609	Centrocampista	Ingrid Syrstad Engen	0.8172
Defensa	Rachel Daly	0.7847	Defensa	Maren Nævdal Mjelde	0.8272
Delantera	Ella Toone	0.7899	Defensa	Maria Thorisdottir	0.8509
Defensa	Leah Williamson	0.8069	Defensa	Guro Bergsvand	0.8618
Centrocampista	Keira Walsh	0.8406	Centrocampista	Frida Maanum	0.875
Delantera	Lauren Hemp	0.8479	Delantera	Caroline Graham Hansen	0.8881
Portera	Mary Alexandra Earps	0.859	Portera	Guro Pettersen	0.8979
Delantera	Francesca Kirby	0.8599	Delantera	Karina Sævik	0.9122
Defensa	Georgia Stanway	0.8655	Centrocampista	Tuva Hansen	0.9189
Centrocampista	Chloe Kelly	0.883	Centrocampista	Vilde Bee Risa	0.9225
Centrocampista	Alessia Russo	0.941	Centrocampista	Guro Reiten	0.9232
Delantera	Ellen White	0.9708	Delantera	Ada Hegerberg	0.964
Centrocampista	Jill Scott	0.9757	Delantera	Amalie Evjelund	1

(c) Entropía total  
de Inglaterra(d) Entropía total  
de Noruega

# Interpretaciones



(White- 0.9708, Mjelde (6)- 0.8272)

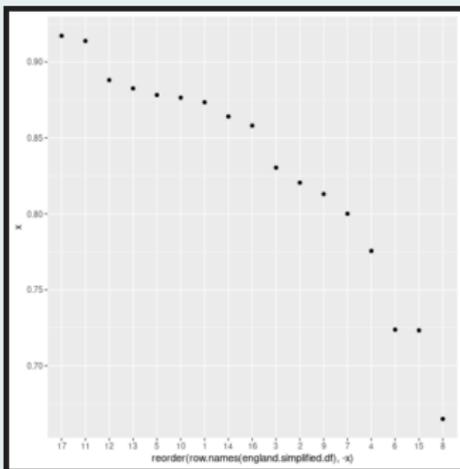


(Mead (7)- 0.6506, White (9), Kirby (14)-  
0.8599, Thorisdottir(3)- 0.8509, Tuva  
Hansen(4)- 0.9189)

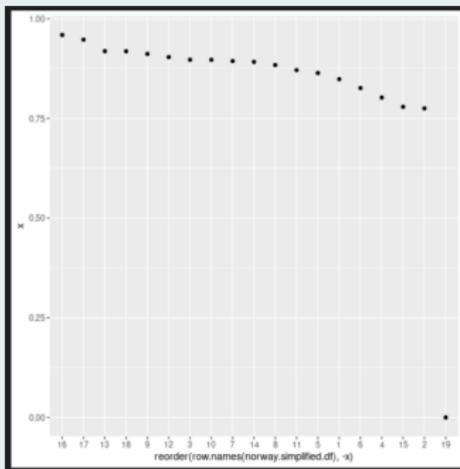
# Resultados

## Principio de máxima entropía de Jayne

Entre todas las distribuciones, debemos elegir aquella con la entropía máxima, esto es, la que tiene mayor variabilidad.



(e) Entropía total de Inglaterra



(f) Entropía total de Noruega

# Conclusiones

- Con la entropía sobre el grafo de pases podemos explicar el desempeño de un equipo
- Entropía a nivel de jugador y de equipo
- Se puede ver la “mejor” o “peor” entropía reflejada en un tipo de visualización de las redes de pases

## Trabajos futuros

Profundización en redes bayesianas para:

- Influencia del rival
- Posesión del balón
- Minutos jugados de cada jugadora
- Entropía espacio-temporal
- Alineación y cambios

# Preguntas

Det er alt, folkens!



## Asignaturas usadas

- Estadística descriptiva e introducción a la probabilidad
- Inferencia estadística
- Probabilidad
- Estadística computacional
- Infraestructura Virtual

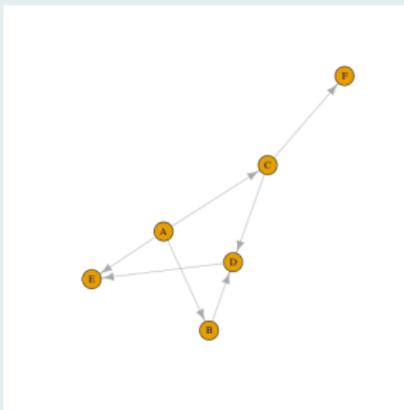
# Costes

Concepto	Coste unitario	Unidades	Total
Amortización portátil	150€	1	150€
Ventilador	14€	1	14€
Costes laborales	2000€	3.5	7000€
Total			7164€

Table: Costes el proyecto en el escenario "analista de datos junior"

# Redes bayesianas

## Ejemplo de grafo dirigido acíclico



(Una red bayesiana consiste en un modelo estructural y un conjunto de probabilidades)

## Entropía + Redes bayesianas

- Para calcular la entropía de Shannon debe cumplirse la propiedad de aditividad
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ; para que sea aditiva la probabilidad de la intersección debe ser 0, pero  $P(A \cap B) = P(A)P(B/A)$

### Entropía de Tsallis

Para modelar el sistema con redes bayesianas, habría que usar una entropía que sí cumpliera esas propiedades. La entropía de Shannon, implementada en igrph, refleja la propiedad (caos/orden/sorpresa) que estamos buscando, lo que fue suficiente para la viabilidad del producto en esta fase.

# Redes causales

- Red bayesiana en la que los padres de cada nodo son sus causas directas.
- Satisface la *condición de Markov causal*

## Condición de Markov causal

- Dadas las causas directas, el fenómeno asociado a un nodo es independiente de los que no tienen efecto sobre él. Esta asunción permite que la distribución conjunta de las variables en una red causal sea factorizada como :

$$P(X_1, \dots, X_N) = \prod_{i=1}^N P(X_i | pa(X_i))$$

# Redes causales

- Red bayesiana en la que los padres de cada nodo son sus causas directas.
- Satisface la *condición de Markov causal*

## Condición de Markov causal

- Dadas las causas directas, el fenómeno asociado a un nodo es independiente de los que no tienen efecto sobre él. Esta asunción permite que la distribución conjunta de las variables en una red causal sea factorizada como :

$$P(X_1, \dots, X_N) = \prod_{i=1}^N P(X_i | pa(X_i))$$

## Entropía de Tsallis

Con la entropía de Tsallis de los pases entre los jugadores, se puede medir la organización asociada al comportamiento de un equipo de fútbol. Fue diseñada para analizar sistemas donde existen correlaciones entre sus microestados

### Definition (Entropía de Tsallis)

Dado un conjunto discreto de probabilidades  $p_i$  con la condición  $\sum_i p_i = 1$ , y con  $q$  cualquier número real, se define la entropía de Tsallis como

$$S_q(p_i) = \frac{k}{q-1} \left( 1 - \sum_{i=1}^N p_i^q \right)$$

donde  $q$  se denomina índice entrópico,  $N$  son los microestados y  $k$  es una constante positiva.

La fórmula se reduce a la de la entropía de Shannon cuando  $q = 1$ .