

# Artificial Intelligence and Machine Learning for commercial analysis in the audiovisual sector

## A case study of designing TV schedules

**JUAN DE MEER**

(Chief Information Officer RTVE)

# About RTVE

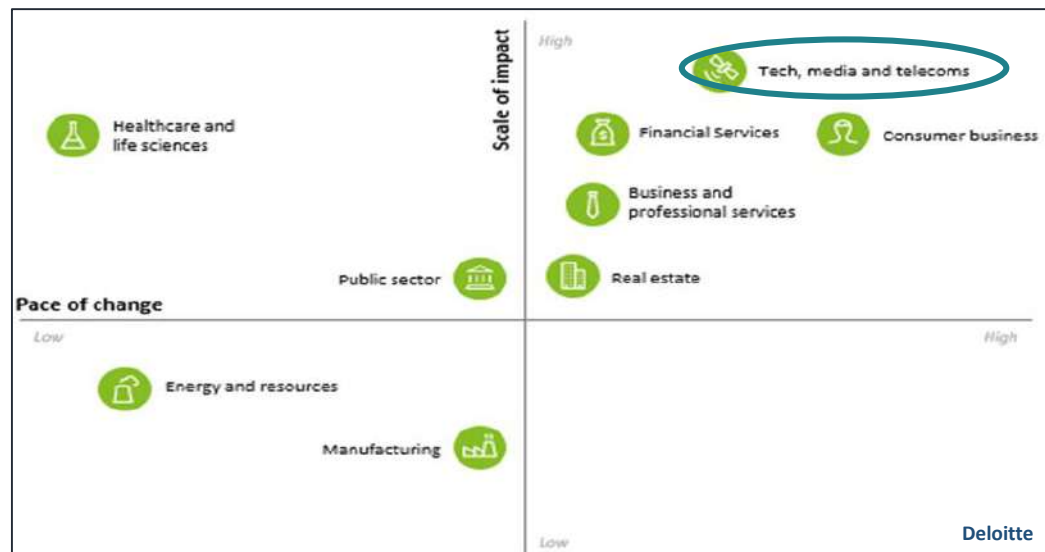


- ❑ 5 TV channels in Spain, and 2 international
- ❑ 6 Radio Channels
- ❑ Digital multimedia platform “rtve.es”
- ❑ Digital TV services like “Red Button” and “PlayZ”

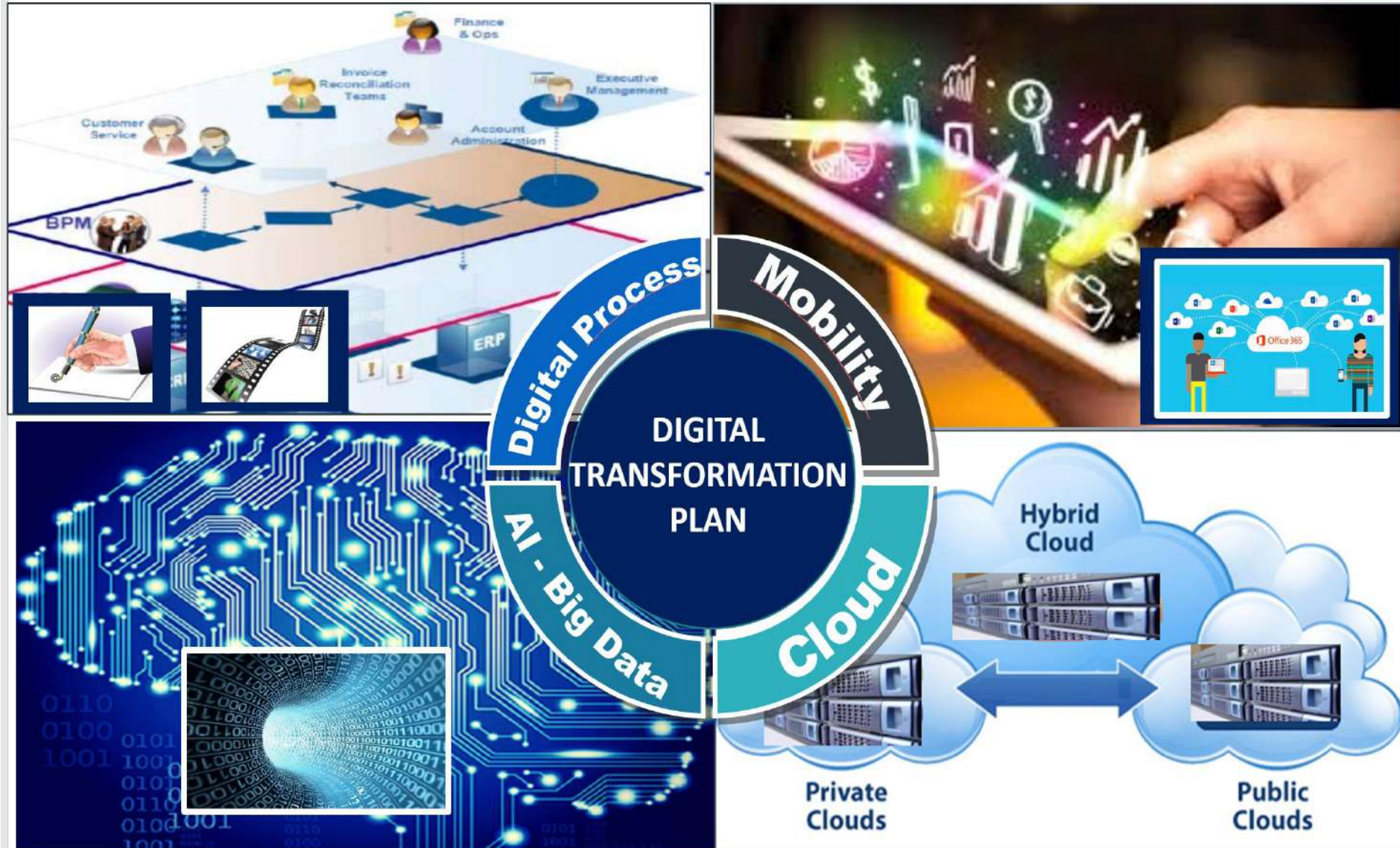


# Digital Transformation in media

- ❑ The media industry has suffered a **deep Digital Transformation of High impact** and **very quickly**.
- ❑ **Competitors** like Netflix, HBO, Amazon, and others are **cannibalizing the audience**.
- ❑ **Competition between TV channels** is very high.



# We must take advantage of Digital transformation technology

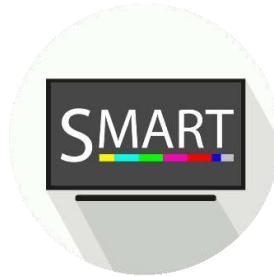


# UC3M and RTVE

1

## Research agreement between UC3M-RTVE

Research in the field of **Big Data**, Linked Data, **Machine Learning** and **AI** applied to the **audiovisual sector**.



2

## Audiovisual sector Digital Transformation

Creation of the **best TV contents** and **services** to **improve the quality of our public service to audience**.



3

## AI-based decision support systems

Make use of **Big Data**, **Machine Learning** and **AI** to **increase audience**.



# AI has many applications in RTVE

1

## Decision Making

Use AI to help **determine the best content to produce, select presenters, make recommendations** to users, etc.

2

## Automatic metadata of content

Create **automatic metadata** for content to improve the **classification, search and reutilization**.

3

## Automatic content generation

Automatic content creation content like **news, TV-reports, sports summaries, etc, without human intervention**.



# Content is very expensive



# We need to optimize the content we have

	21:00	21:30	22:00	22:30	23:00	23:30	24:00
 La 1	21:00 Telediario 2		22:05 TVEmos		22:45 Estoy vivo		23:45 ▶ C... actualida
 La 2	21:00 Los jardines paraíso de Monty Don		22:00 Documaster				
 Antena 3	21:00 Antena 3 Noticias 2	+	+	21:45 El hormiguero 3.0	22:40 New Amsterdam		23:38 ▶ New Amsterdam
 Cuatro	◀ 20:30 Gran Hermano VIP 7: resumen diario	21:30 First Dates			22:45 ▶ Cine Cuatro: La jungla II: Alerta roja		
 Telecinco	◀ + 21:10 Informativos Telecinco	+	+	22:00 ▶ Gran Hermano VIP 7			
 La Sexta	◀ 20:20 Previo Al rojo vivo: La última oportunidad		22:00 ▶ Debate La Sexta: La última oportunidad				

- We have to put the **right Content** in the **right right Time**
- That's not an easy task !
- It depends on **what the other channels do !**



# Share is a 24x7 competition



# The challenge

1

Which TV programs are best for the audience in a given day and hour?

2

Which is the best day of the week and hour for a given TV program?



# Project objectives

Develop an AI-based decision support system  
to design TV schedules  
that maximize the audience

	21:00	21:30	22:00	22:30	23:00	23:30	24:00
<b>1</b> La 1	21:00 Telediario 2		22:05 TVEemos		22:45 Estoy vivo		23:45 ▶ C... actualida
<b>2</b> La 2	21:00 Los jardines paraíso de Monty Don		22:00 Documaster				



 Antena 3	21:00 Antena 3 Noticias 2	+	+	21:45 El hormiguero 3.0	22:40 New Amsterdam	23:38 ▶ New Amsterdam
 Cuatro	◀20:30 Gran Hermano VIP 7: resumen diario	21:30 First Dates			 22:45 ▶ Cine Cuatro: La jungla II: Alerta roja	
 Telecinco	◀+ 21:10 Informativos Telecinco	+	+	22:00 ▶ Gran Hermano VIP 7		
 La Sexta	◀20:20 Previo Al rojo vivo: La última oportunidad	22:00 ▶ Debate La Sexta: La última oportunidad				

# Project steps

- Establish the **initial context**
- Identify existing AI technology** adequate to solve the challenge
- Determine the **Data sources**
- Analyze and design** the system
- Make **prototypes**
- Implement and evaluate** the solution.





# State of Art

- ❑ **Collect data and create model for feature extraction of TV Programs**



- ❑ **There are ways to predict the impact of movies or series**
  - ✓ **Based on the casting of the film.**
  - ✓ **Based on Content genre (comedy, drama...)**
  - ✓ **Based on Social networks opinion.**

**Machine learning techniques are widely used in similar types of classification.**

# State of Art II



## ✓ Genetic Algorithm (Bio-inspired)

- ❑ Solves **optimization problems** that are **not well suited for standard optimization algorithms**:
  - **Inspired by Charles Darwin's theory** of natural evolution
  - Reflects the process of **natural selection** where **fittest individual are selected**
- ❑ Each individual is a **problem solution** -> TV-schedule
- ❑ **Fitness function** that establishes **how individuals fit** -> Audience

# Prediction component

Automatically design TV schedules that optimize audience

## Prediction component

1

Learning module

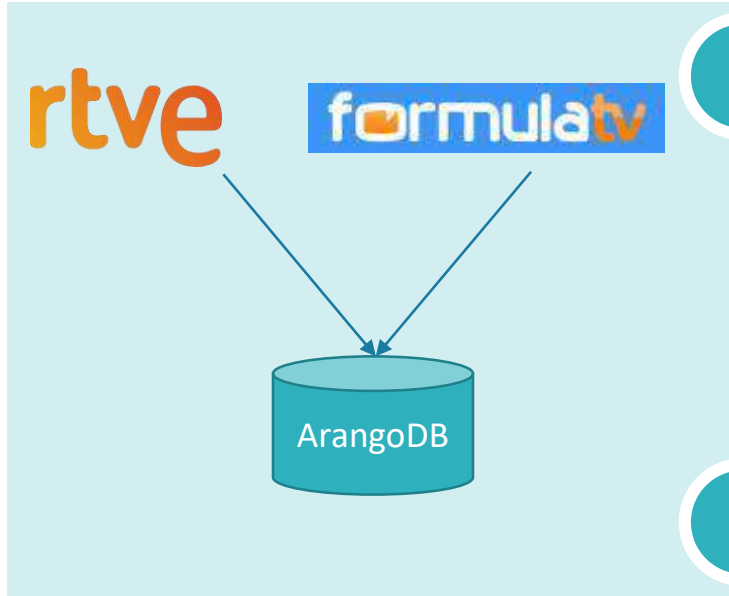
Make use of **classification algorithms** to make the **predictions**

2

Optimization module

**Design TV schedules** using the **genetic algorithm**

# Design – Learning module



1

## Feature extraction

- | RTVE       |                         |
|------------|-------------------------|
| - Name     |                         |
| - Duration | - Genre <sup>1</sup>    |
| - Channel  | - Subgenre <sup>1</sup> |
| - Audience | - Date                  |
| - Share    | - Target <sup>1</sup>   |

- | FórmulaTV  |
|------------|
| - Genre    |
| - Subgenre |
| - Target   |

2

## Create collections and normalize values

3

## Create learning instances. Parwise comparison

Programa 1								Programa 2								Win
name	channel	date	share	audience	genre	subgenre	duration	name	channel	date	share	audience	genre	subgenre	duration	0/1

<sup>1</sup> Only RTVE emissions



# Design – Learning module II

4

## Train models

- Each model stored a **different time slot** information:

- Morning
- Midday
- Afternoon
- **Evening**
- Late Night

- Classification algorithm: **Random Forest**

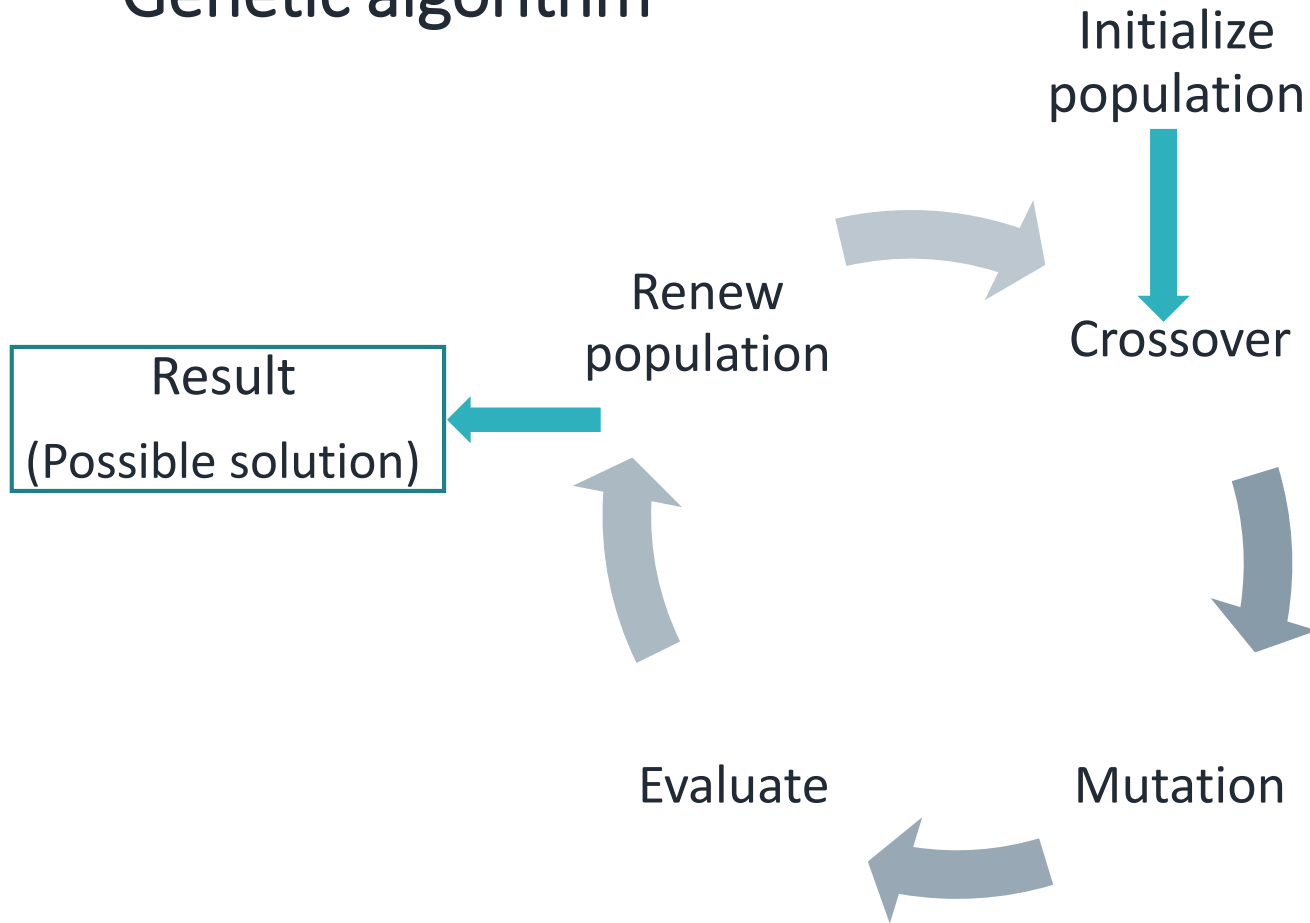
5

## Evaluate models for each time slot

- Morning (98,80%)
- Midday (98,98%)
- Afternoon (99,02%)
- Night (97,02%)
- Late Night (96,75%)

# Design – Optimization module

## Genetic algorithm



# Design – Optimization module II

1

Initialize population: Create random program combinations for a time slot (schedules)

Random TV schedules

Competitors	1	P1	P2	P3	P4	P23	P4	P24	P28	P31	P35	P19	P3	200 ...
		P5	P6	P7	P8	P5	P6	P7	P8	P5	P6	P7	P8	
	5	P9	P10	P11	P12	P9	P10	P11	P12	P9	P10	P11	P12	
		P13	P14	P15	P16	P13	P14	P15	P16	P13	P14	P15	P16	
		P17	P18	P19	P20	P17	P18	P19	P20	P17	P18	P19	P20	

TV Schedule 1

TV Schedule 2

TV Schedule 3

2

Victories->  
(fitness)

10

2

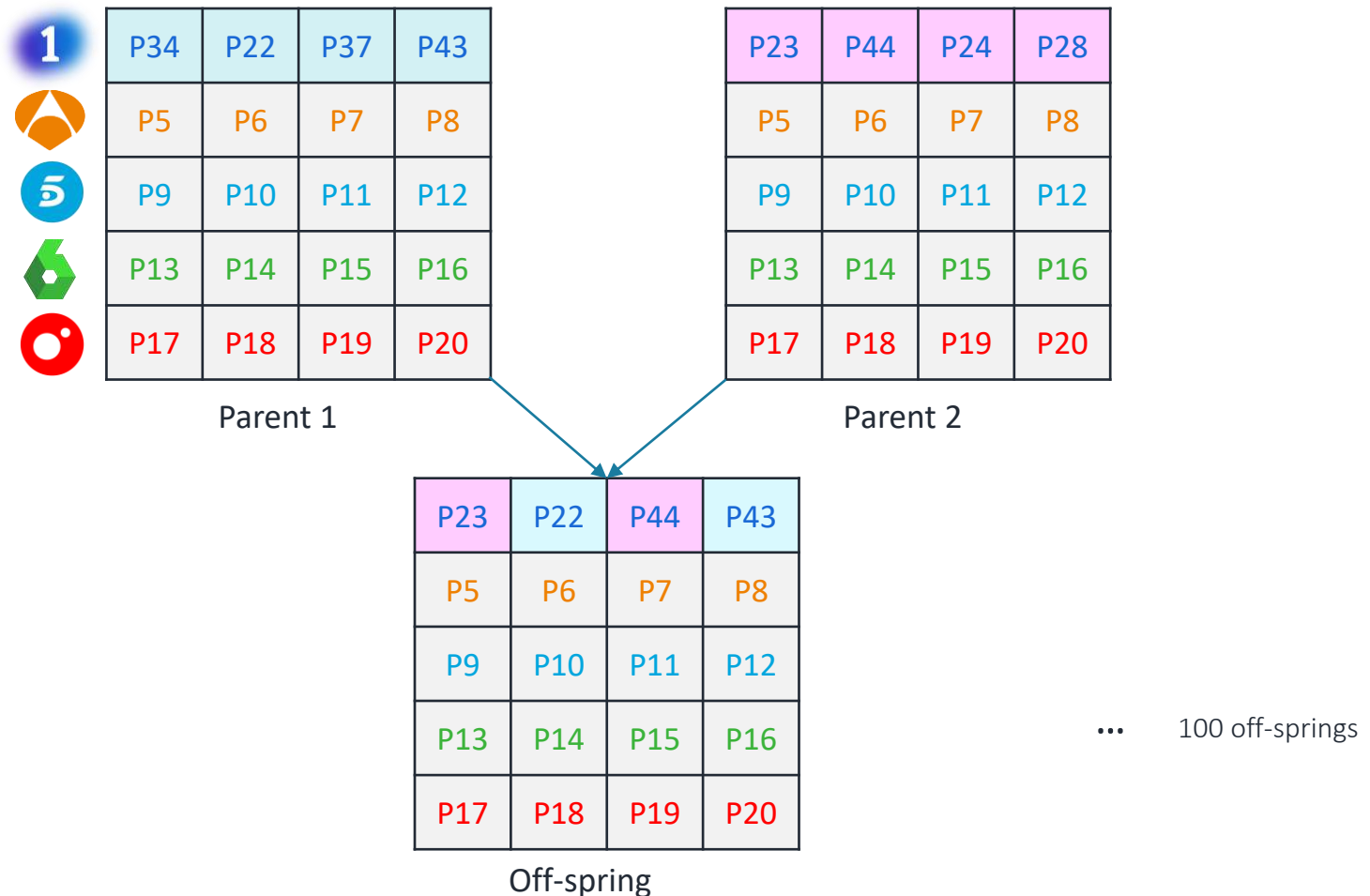
12

- ❑ The Learning module establishes the number of victories achieved in a TV Schedule.
- ❑ Victories are number of programs of the competitors defeated.

# Design – Optimization module III

3

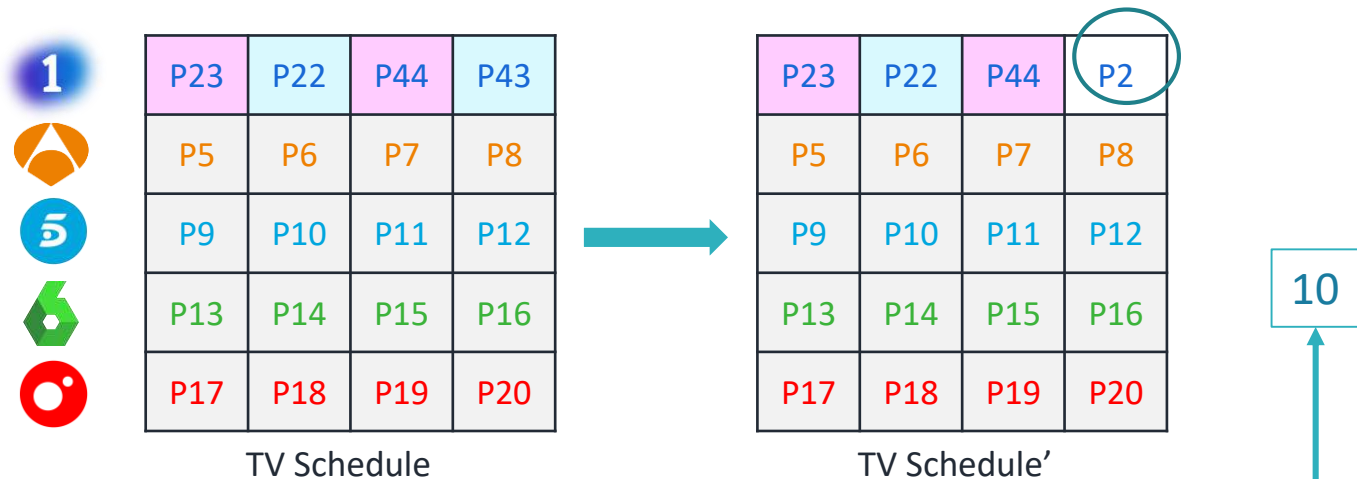
Crossover: Select parents and generate off-springs





# Design – Optimization module IV

## 4 Mutation in some of the individuals (schedules)







## 5 Evaluate

## 6 Renew population

- Replace the entire population of TV schedules by the best 100 at the end of the iteration

# The algorithm proposes the best schedule found

1	P31	P35	P19	P3
	P5	P6	P7	P8
	P9	P10	P11	P12
	P13	P14	P15	P16
	P17	P18	P19	P20

15

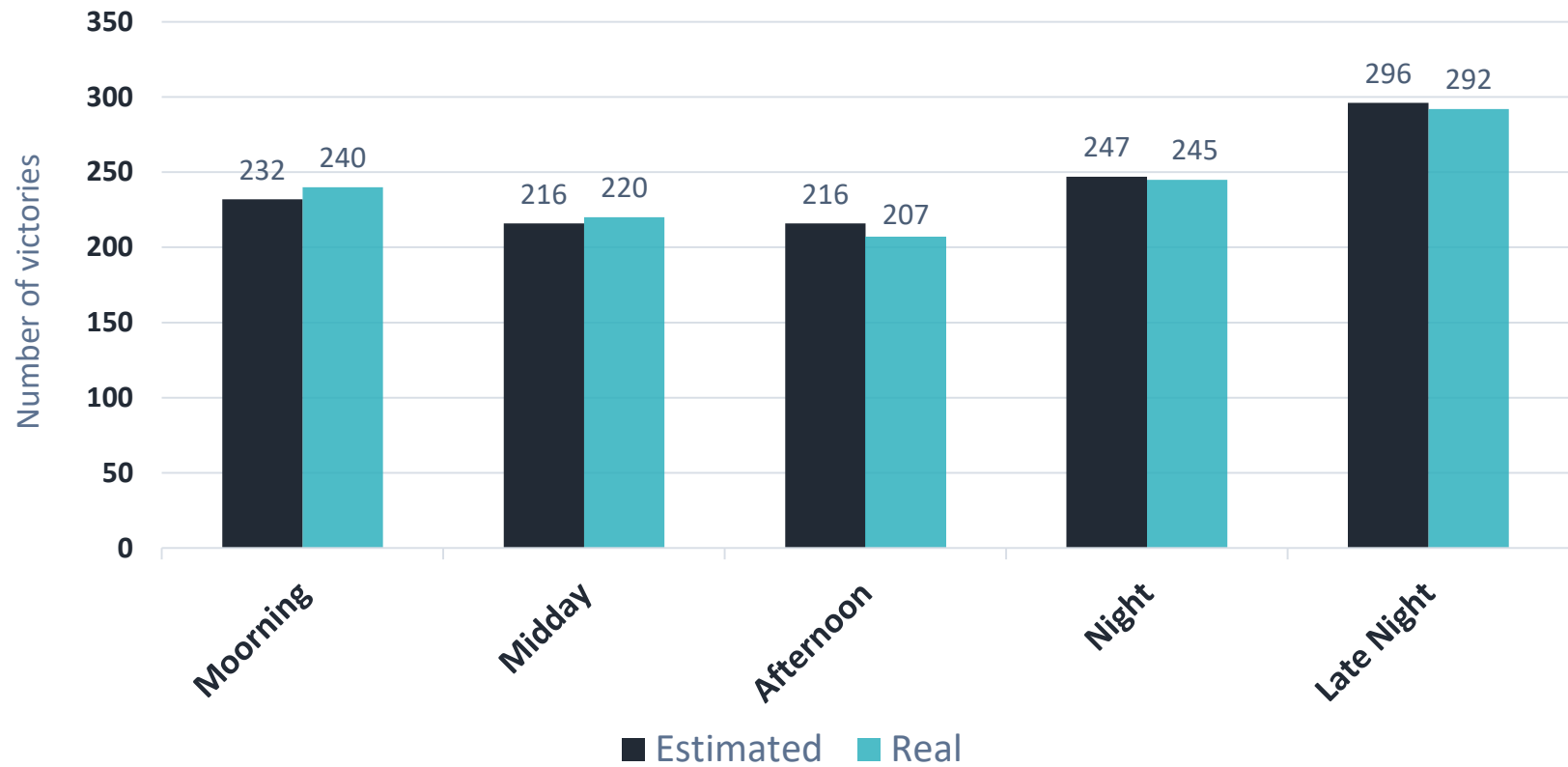
**Victories**

# Demo

The screenshot displays the SMART-RTVE web application interface. On the left is a dark sidebar with a menu containing: Principal, Rankings, Perfiles, Programas, Comparador, Semejanza, Sugeridos, Grafos, **Predicción** (highlighted), and Configuración. The top header includes the SMART-RTVE logo, a search bar with the text "Buscar", and links for "Idioma" and "Admin". The main content area features a five-step progress bar: 1. Tipo de predicción, 2. Datos de entrada, 3. Seleccionar objetivo, 4. Configurar algoritmo, and 5. Resultado. Under step 1, there are two radio button options: "Predecir mejor hora para un programa" (selected) and "Predecir mejor programación". The first option includes the subtext "Obtener una parrilla con un programa determinado para una franja horaria", and the second includes "Obtener la mejor parrilla de programación para una franja horaria". A blue "Siguiete" button is positioned below the first option.

# Experimentation

Nº of estimated victories vs real data



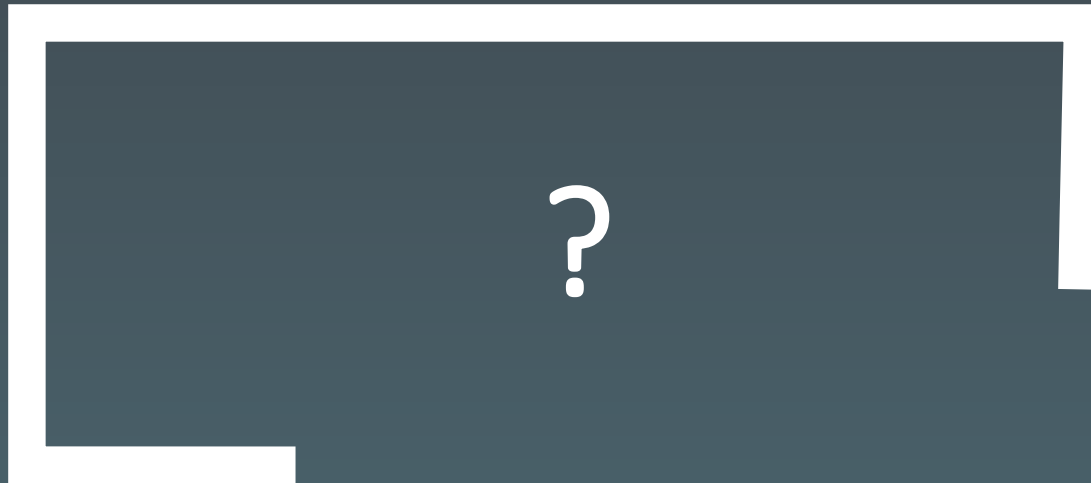
# Conclusions

- ✓ The results are **promising**
- ✓ Prototype **available to assist TV experts** to prepare schedules

# Future Work

- ✓ Take into account certain **restraints and exceptions** that apply to TV schedules.
- ✓ Introduce TV experts knowledge beyond audiences.
- ✓ Take into account **content costs**





## **JUAN DE MEER**

Artificial Intelligence and Machine Learning for commercial analysis in the audiovisual sector: a case study of designing TV schedules.

Eduardo Cibrián, Jose María Álvarez Rodríguez, Valentín Moreno and Juan de Meer