

LANZAMIENTO

ANALYTICS Y BIG DATA

Ciencia de los Datos aplicada al mundo de los negocios



Sebastián Maldonado

Investigador ISCI y académico de la Facultad de Economía y Negocios de la Universidad de Chile

Carla Vairetti

Investigadora ISCI y académica de la Facultad de Ingeniería y Ciencias Aplicadas de la Universidad de los Andes

ISCI INSTITUTO
SISTEMAS COMPLEJOS
DE INGENIERÍA

FACULTAD DE
ECONOMÍA Y
NEGOCIOS
UNIVERSIDAD DE CHILE

DCS
Departamento de Control
de Calidad y Sistemas de
Información

Universidad de
los Andes

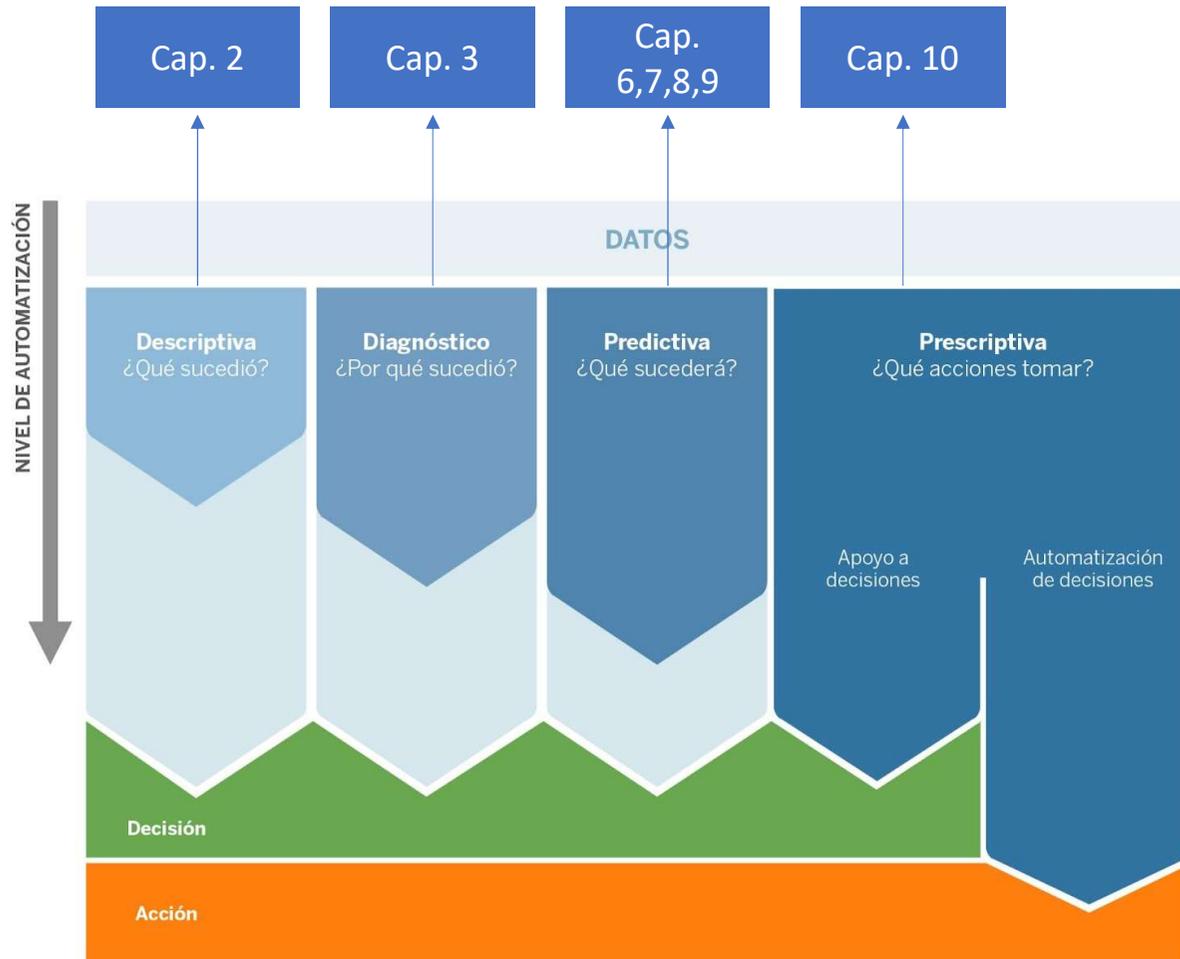
FACULTAD DE
INGENIERÍA Y CIENCIAS
APLICADAS

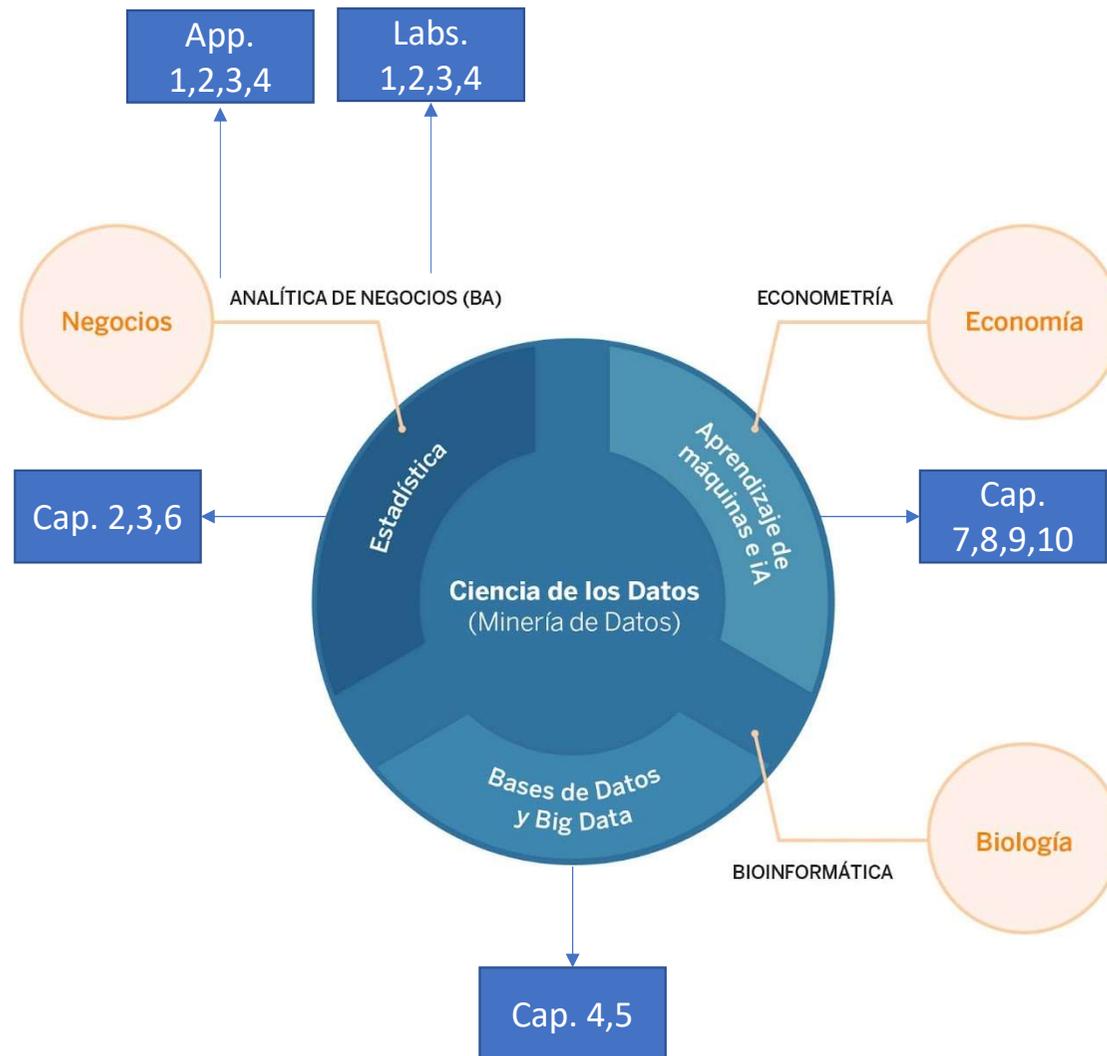

RIL editores

Motivación

- Las empresas enfrentan hoy importantes desafíos para monetizar su información.
- Nuevos datos: textos (reclamos y preferencias) imágenes, *clicks* en sitios web, etc.
- El mundo de los negocios no ha estado ajeno a esta tendencia.
- Empresas que se embarcan en la analítica tienen las siguientes posibilidades:
 - 23 veces más de adquirir clientes,
 - 6 veces más de retener clientes y
 - 19 veces más de ser rentables.







Índice

1. Introducción

2. *Descriptive Analytics* y preprocesamiento

Laboratorio 1: Preprocesamiento

Aplicación 1: Segmentación de clientes y *targeting*

3. *Diagnostic Analytics*

Laboratorio 2: Segmentación de clientes y clustering

4. Fundamentos de las bases de datos

5. *Big Data Analytics*

Aplicación 2: Asignación de créditos

6. *Predictive Analytics* I: Modelos transparentes

Aplicación 3: Fuga de clientes

7. *Predictive Analytics* II: Modelos «caja negra» tradicionales

Aplicación 4: Experiencia del consumidor (CX)

8. *Predictive Analytics* III: Inteligencia artificial

Laboratorio 3: Clasificación

9. *Predictive Analytics* IV: Regresión y series de tiempo

10. *Prescriptive Analytics*

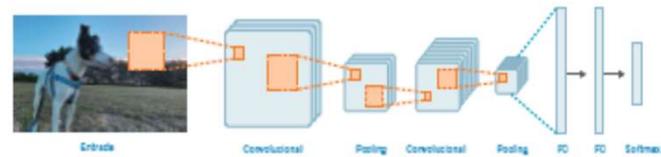
Laboratorio 4: Predicción y decisión de umbral

11. Conclusión



SEBASTIÁN MALDONADO Y CARLA VAIRETTI

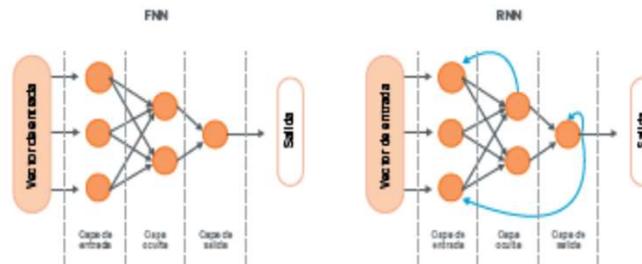
FIGURA 8.14: EJEMPLO DE UNA RED CONVOLUCIONAL TRADICIONAL



8.3.2 Recurrent Neural Networks (RNN)

Esta clase de redes viene a cambiar el paradigma clásico de funcionamiento de las redes neuronales, ya que considera como entrada otras capas de la misma red de manera recurrente. Esto se ilustra en la figura 8.15 con las flechas que apuntan hacia atrás.

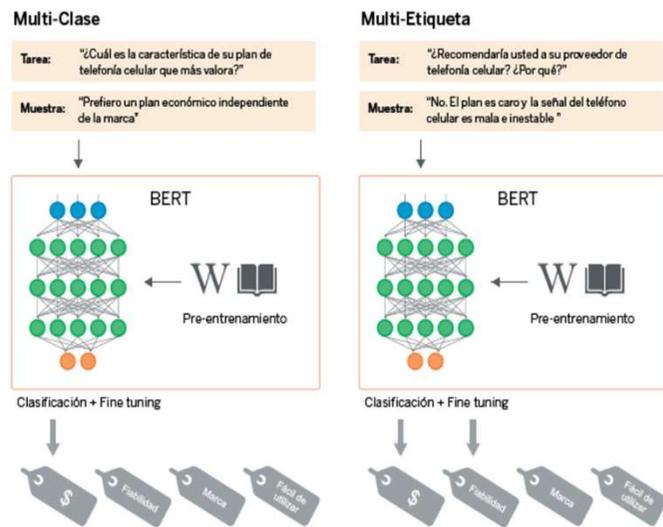
FIGURA 8.15: EJEMPLO DE UNA RED RECURRENTE EN CONTRASTE CON UNA RED TRADICIONAL



Este tipo de redes permite modelar problemas complejos que requieren modelar la secuencia de las observaciones (imágenes, ~~tex-~~

APLICACIÓN 4: EXPERIENCIA DEL CONSUMIDOR (CX)

FIGURA APP4.2: DIFERENCIA ENTRE ENFOQUES MULTICLASE MULTI-LABEL



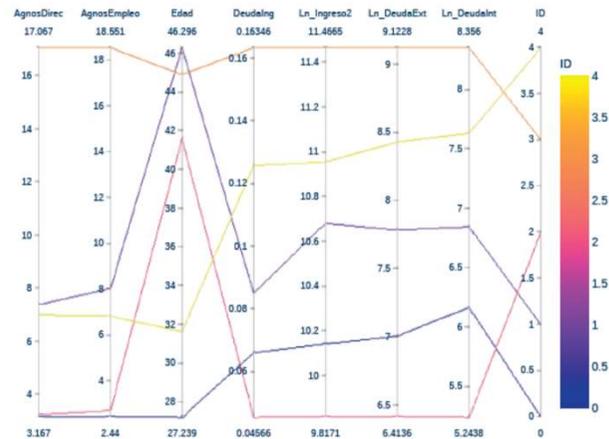
Utilizando un modelo preentrenado como punto de partida, podemos ajustar los pesos de la red con una capa de salida *ad-hoc* para la tarea multietiqueta, mientras alimentamos con los datos nuevos. De esta manera el modelo no solo construye el clasificador que comete menos errores, sino también la mejor representación matemática de nuestros datos, propagando los errores a través del *transformer*. Este proceso también se conoce como *transfer learning*, ya que el conocimiento adquirido por una máquina se transfiere de una tarea a otra.

LABORATORIO 2: SEGMENTACIÓN DE CLIENTES Y CLUSTERING

Vemos que el *cluster* más pequeño tiene 428 clientes (14%), lo que sigue siendo representativo. Vamos ahora al gráfico de coordenadas paralelas:

```
centers = kmeans.cluster_centers_  
centers = pd.DataFrame(centers, columns=data_set.columns)  
centersde = scaler.inverse_transform(centers)  
data_centers = pd.DataFrame(centersde, columns=data_set.columns) data_  
centers['ID'] = np.arange(len(data_centers))  
fig = px.parallel_coordinates(data_centers, labels=data_centers.columns,  
color='ID')  
fig.show()
```

FIGURA 3: GRÁFICO DE COORDENADAS PARALELAS





Libro

Analytics y Big Data: Ciencia de los Datos aplicada al mundo de los negocios

Lanzamiento – Abril 2022

Disponible en: varetti-maldonado.cl/