

# Trabajo Práctico 2: Machine Learning

[75.06 / 95.58] Organización de Datos  
Segundo cuatrimestre de 2018

## Grupo Datatouille

Alumno	Padrón	Mail
Bojman, Camila	101055	camiboj@gmail.com
del Mazo, Federico	100029	delmazofederico@gmail.com
Hortas, Cecilia	100687	ceci.hortas@gmail.com
Souto, Rodrigo	97649	rnsoutob@gmail.com

<https://github.com/FdelMazo/7506-Datos/>

<https://kaggle.com/datatouille2018/>

## Curso 01

- Argerich, Luis Argerich
- Golmar, Natalia
- Martinelli, Damina Ariel
- Ramos Mejia, Martín Gabriel

## Contents

<b>1</b>	<b>Introducción</b>	<b>1</b>
<b>2</b>	<b>Investigación</b>	<b>1</b>
<b>3</b>	<b>Features buscadas</b>	<b>1</b>
<b>4</b>	<b>Algoritmos utilizados</b>	<b>1</b>
<b>5</b>	<b>Desarrollo</b>	<b>1</b>
<b>6</b>	<b>Resultados obtenidos</b>	<b>1</b>
<b>7</b>	<b>Conclusiones</b>	<b>1</b>

## 1 Introducción

Se propone exponer en el siguiente informe el desarrollo del Trabajo Práctico para predecir la probabilidad de que un usuario de Trocafone realice una conversión en un período determinado de tiempo. Con dicho propósito se utilizan como base dos archivos csv brindados por la empresa. En un primer lugar el archivo `events_up_to_01062018.csv` que contiene la información de los eventos realizados por un conjunto de usuarios hasta el 31 de mayo de 2018. En un segundo lugar el archivo `labels_training_set.csv` que determina si un conjunto de usuarios realizó o no una conversión desde el 01/06/2018 hasta el 15/06/2018.

Se propone como objetivo desarrollar una métrica a fin de evaluar los distintos resultados obtenidos. Se presentan en el informe las distintas decisiones adoptadas para elegir un resultado por sobre otro. Así mismo, se propuso utilizar distintos algoritmos de la librería `sklearn` para elegir el que modelice mejor.

Por otro lado se crearon distintos csv a fin de poder desarrollar de mejor manera el proceso de feature engineering y obtener un pre-procesamiento de los datos que permita encontrar los resultados más altos.

## 2 Investigación

## 3 Features buscadas

## 4 Algoritmos utilizados

## 5 Desarrollo

## 6 Resultados obtenidos

## 7 Conclusiones