# VERI MADENCILIĞİ

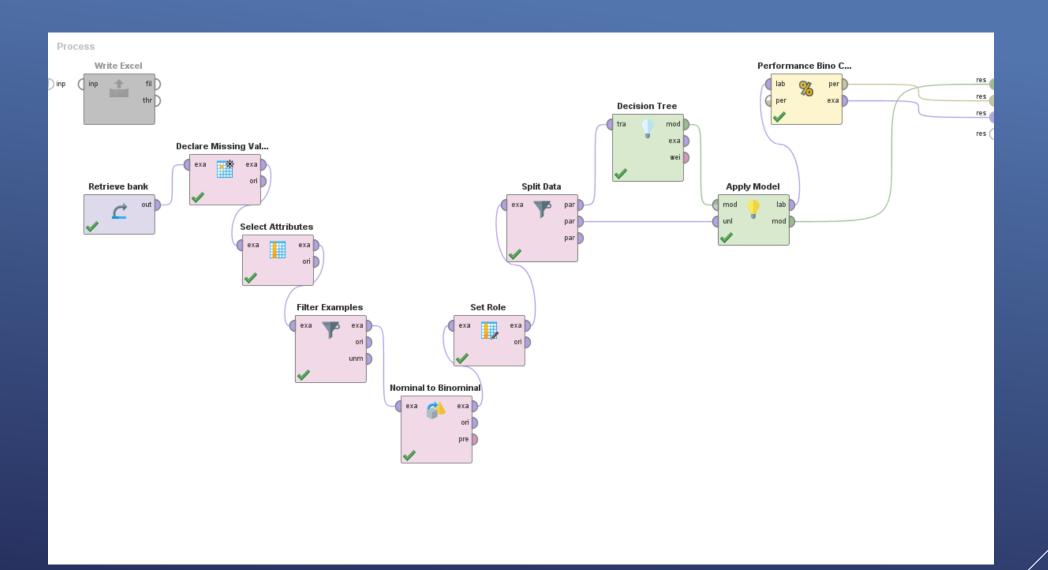
Yasir Erkam Özdemir

# ÇALIŞMADA KULLANILAN VERİLER

- ▶ Bu çalışmada kullanılan veriler, Portekizli bir bankacılık kurumunun doğrudan pazarlama kampanyaları ile ilgilidir. Pazarlama kampanyaları telefon görüşmelerine dayanıyordu. Müşterinin kampanyaya abone olup olmayacağını belirlemek için genellikle aynı müşteriyi birden fazla kez aramak gerekiyordu.
- ► Veriler uci.edu üzerinden alınmıştır : https://archive.ics.uci.edu/ml/datasets/Bank+Marketing#

#### VERISETINDEKI ATTRIBUTE'LER

- ► Input variables:
- # bank client data:
- ▶ 1 age (numeric)
- 2 job : type of job (categorical: "admin.","unknown","unemployed","management","housemaid","entrepreneur","student", "blue-collar","self-employed","retired","technician","services")
- > 3 marital: marital status (categorical: "married", "divorced", "single"; note: "divorced" means divorced or widowed)
- ▶ 4 education (categorical: "unknown", "secondary", "primary", "tertiary")
- ▶ 5 default: has credit in default? (binary: "yes", "no")
- 6 balance: average yearly balance, in euros (numeric)
- ▶ 7 housing: has housing loan? (binary: "yes", "no")
- ▶ 8 loan: has personal loan? (binary: "yes", "no")
- # related with the last contact of the current campaign:
- 9 contact: contact communication type (categorical: "unknown", "telephone", "cellular")
- ▶ 10 day: last contact day of the month (numeric)
- 11 month: last contact month of year (categorical: "jan", "feb", "mar", ..., "nov", "dec")
- 12 duration: last contact duration, in seconds (numeric)
- # other attributes:
- 13 campaign: number of contacts performed during this campaign and for this client (numeric, includes last contact)
- ▶ 14 pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric, -1 means client was not previously contacted)
- ▶ 15 previous: number of contacts performed before this campaign and for this client (numeric)
- ▶ 16 poutcome: outcome of the previous marketing campaign (categorical: "unknown", "other", "failure", "success")
- Output variable (desired target):
- ▶ 17 y has the client subscribed a term deposit? (binary: "yes", "no")



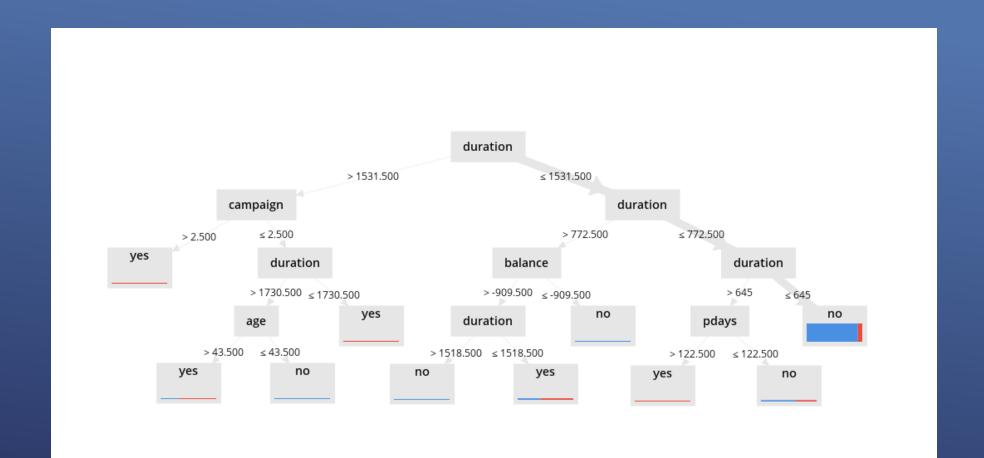
### RAPID MINER DIZAYN

## VERI ÖN HAZIRLIĞI

- Verisetinde bazı özelliklerin bilinmeyen değerlerine unknown değeri girilmiş. Bu nedenle verisetinde hiç bilinmeyen değer yok olarak görünüyor. Bu değerleri temizlemek için Rapid Miner'daki Declare Missing Value operatörüyle unknown değerler bilinmeyen olarak değiştiriliyor.
- ▶ Böylece bilinmeyen değerler ortaya çıkıyor. Bu değerleri handle etmek için **Select Attributes** ve **Filter Examples** operatörleri kullanılmıştır.
- ► Verisetinde çok fazla bilinmeyen içeren **contact** ve **poutcome** attributes'leri Select Attributes operatörüyle verisetinden çıkarılmıştır.
- Ayrıca **job** ve **education** attributes'lerindeki unknown içeren satırlar Filter Examples operatörüyle verisetinden çıkarılmıştır.
- y attribute'denki değerler Binomonal Classification performans ölçümüyle ölçüm yapabilmek için Nominal to Binominal operatörüyle binomale çevirilmiştir.
- Son olarak y attribute'ü Set Role operatörüyle çıktı labeli olarak ayarlanmıştır.

# DIĞER OPERATÖRLER

- ▶ **Split Data** operatörüyle veriseti test ve öğrenme seti olarak ayrılmıştır.
- ▶ **Decision Tree** operatörüyle öğrenme seti üzerinde Decision Tree algoritmasıyla öğrenme gerçekleştirilmiştir.
- ▶ **Apply Model** operatörü kullanılarak öğrenmeyi yapan Decision Tree modeli ve test seti girdi olarak alınıp model uygulanmıştır.
- Performance Binominal Classification operatörü kullanılarak performans ölçümü yapılmıştır.



# KARAR AĞACI

# PERFORMANCE

accuracy: 89.10%			
	true no	true yes	class precision
pred. no	1116	113	90.81%
pred. yes	28	36	56.25%
class recall	97.55%	24.16%	

