



## Bank Marketing Project: Presentation 2

Evan Bradley

Oakland University

*edbradley@oakland.edu*

November 2016

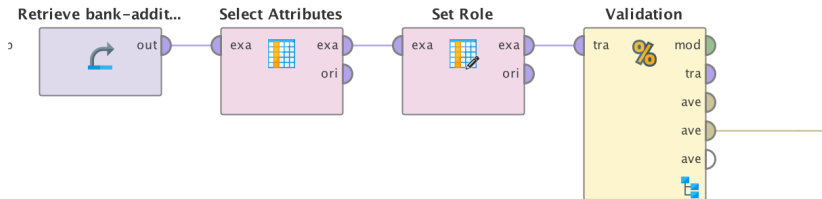


1 Overview of Process

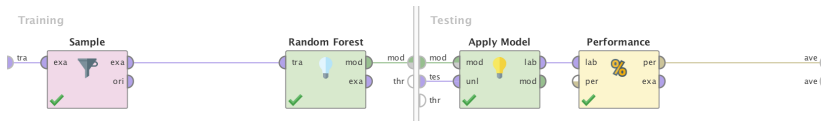
2 Preprocessing

3 Results

# Overall Process



# Cross-validation Process

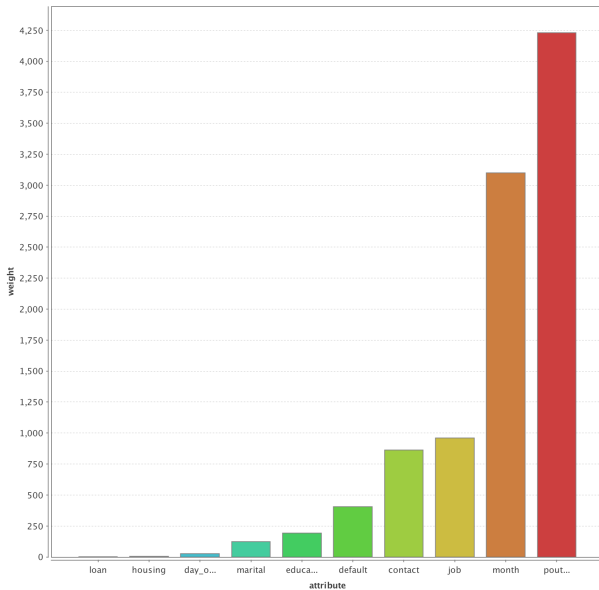


# Categorical variables



attribute	weight
loan	1.094
housing	5.684
day_of_week	26.145
marital	122.655
education	193.106
default	406.578
contact	863.269
job	961.242
month	3101.149
poutcome	4230.524

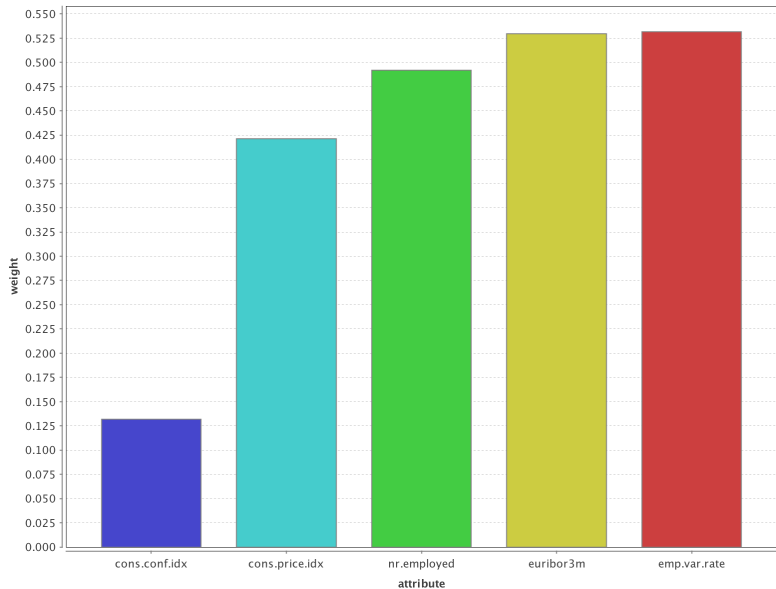
# Categorical variables





attribute	weight
cons.conf.idx	0.132
cons.price.idx	0.421
nr.employed	0.492
euribor3m	0.529
emp.var.rate	0.532

# Economic Indicators








# Attribute removal





Attributes



cons.conf.idx  
cons.price.idx  
day\_of\_week  
duration  
education  
housing  
job  
loan  
marital

Selected Attributes

age  
campaign  
contact  
default  
emp.var.rate  
euribor3m  
month  
nr.employed  
pdays  
poutcome  
previous  
y



class	size
yes	4640
no	4640

# Decision Tree



**accuracy: 89.82% + / - 0.30% (mikro: 89.82%)**

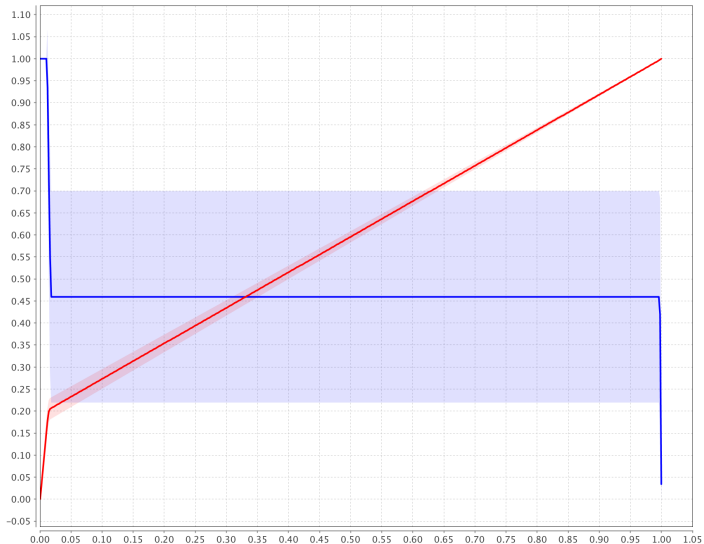
	true no	true yes	class precision
pred. no	36051	3695	90.70%
pred. yes	497	945	65.53%
class recall	98.64%	20.37%	

# Decision Tree



AUC: 0.594 +/- 0.012 (mikro: 0.594) (positive class: yes)

— ROC — ROC (Thresholds)



# Decision Tree



**accuracy: 87.32% + / - 0.58% (mikro: 87.32%)**

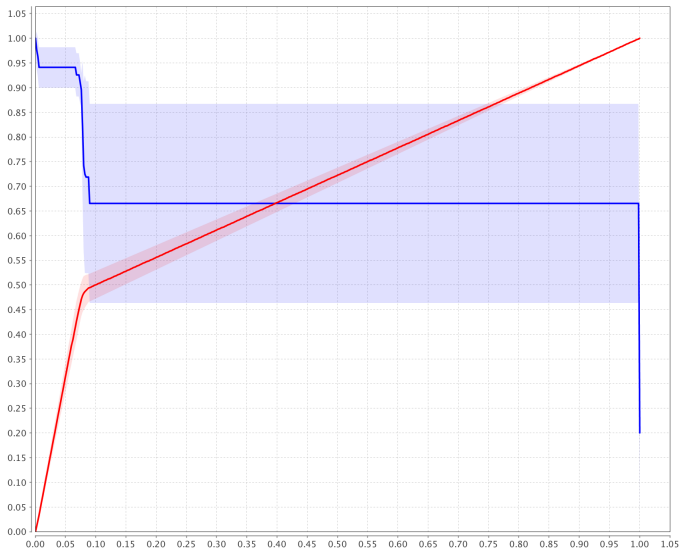
	true no	true yes	class precision
pred. no	33703	2378	93.41%
pred. yes	2845	2262	44.29%
class recall	92.22%	48.75%	

# Decision Tree

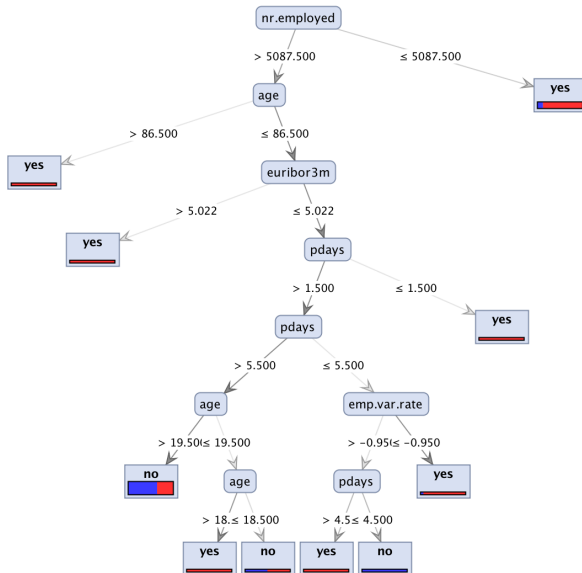


AUC: 0.705  $\pm$  0.014 (mikro: 0.705) (positive class: yes)

— ROC — ROC (Thresholds)















# Decision Tree



# Support Vector Machine



 **SVM (Support Vector Machine)**

kernel type 	<input type="text" value="polynomial"/>	
kernel degree	<input type="text" value="2.0"/>	
C 	<input type="text" value="0.0"/>	
convergence epsilon	<input type="text" value="0.001"/>	
L pos	<input type="text" value="1.0"/>	
L neg	<input type="text" value="1.0"/>	
epsilon	<input type="text" value="0.0"/>	
epsilon plus	<input type="text" value="0.0"/>	
epsilon minus	<input type="text" value="0.0"/>	



# Support Vector Machine



**accuracy: 83.93% + / - 0.99% (mikro: 83.93%)**

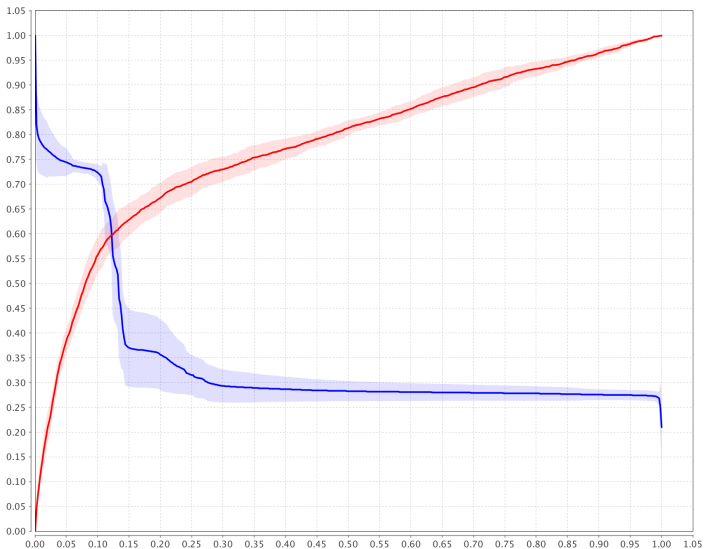
	true no	true yes	class precision
pred. no	31759	1830	94.55%
pred. yes	4789	2810	36.98%
class recall	86.90%	60.56%	

# Support Vector Machine




AUC: 0.778  $\pm$  0.017 (mikro: 0.778) (positive class: yes)








— ROC — ROC (Thresholds)



# Neural Net



 **Neural Net**

hidden layers	 Edit List (1)...	
training cycles	<input type="text" value="500"/>	
learning rate 	<input type="text" value="0.2"/>	
momentum	<input type="text" value="0.2"/>	
error epsilon	<input type="text" value="1.0E-5"/>	



**accuracy: 84.10% +/- 2.11% (mikro: 84.10%)**

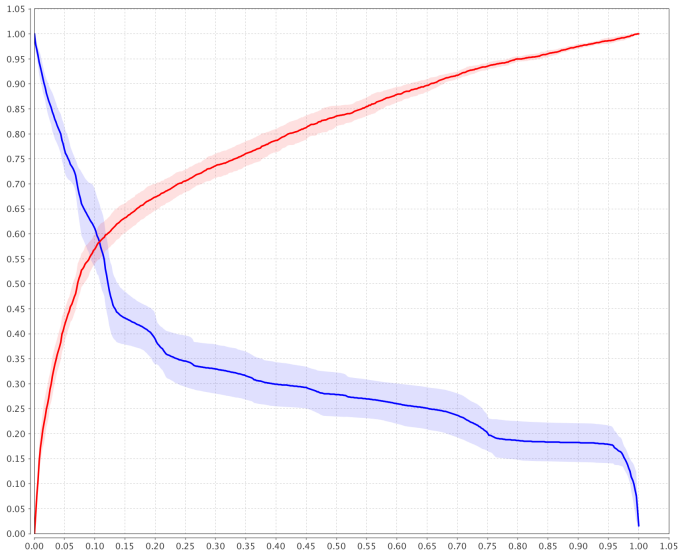
	true no	true yes	class precision
pred. no	31805	1807	94.62%
pred. yes	4743	2833	37.39%
class recall	87.02%	61.06%	

# Neural Net




AUC: 0.793  $\pm$  0.016 (mikro: 0.793) (positive class: yes)


— ROC — ROC (Thresholds)





# Random Forest





 **Random Forest**


number of trees 




criterion 




maximal depth 




☐ apply pruning




☐ apply prepruning



☒ guess subset ratio



voting strategy



# Random Forest



**accuracy: 83.86% +/- 1.05% (mikro: 83.86%)**

	true no	true yes	class precision
pred. no	31627	1725	94.83%
pred. yes	4921	2915	37.20%
class recall	86.54%	62.82%	

# Random Forest



AUC: 0.796 +/- 0.014 (mikro: 0.796) (positive class: yes)

— ROC — ROC (Thresholds)

