

Class 8a: Using Neural Networks in Customer Analytics

Professor Blake McShane

MKTG 482: Customer Analytics
Kellogg School of Management

Customer Analytics Course Structure

Customer Centric Marketing

- Customer Analytics and AI Overview (Class 1)
AI and Analytics,
Why Customer Analytics and AI Needs Customer Centricity

Getting Ready for Analytics

- Using R for Customer Analytics and AI (Class 2)
- Statistics Review (Class 3)

Targeting Customers for Acquisition and Development

- Predicting Response with RFM analysis (Class 4)
- Case Analysis: "Tuango: RFM Analysis for Mobile App Push Messaging" (Class 5)
Lift and Gains
- Predicting Response with Logistic Regression (Class 6)
- Predicting Response with Neural Networks (Class 7)
- Using Neural Networks for Customer Analytics and AI (Class 8)
Training Machine Learning Models
- Case Analysis: Intuit QuickBooks Upgrade: Moving to the Cloud (Class 9)
- Predicting Response with Tree Methods (Class 10)

Targeting based on Incrementality

- From Propensity to Uplift (Class 11)
- The Causality Checklist (Class 12)
- Case Analysis: Creative Gaming Uplift Modeling (Class 13)
- Hyper-Personalization: Next-Product-to-Buy Models (Class 14)

Retaining Customers

- Predicting Attrition (Class 15)
- Linking Analytics with a Business Outcomes Model (Class 16)
- Case Analysis: "S-Mobile: Churn Management" (Class 17)
From Prediction to Action

Selecting the Right Offers

- Design of Experiments / Multivariate Testing (Class 18)
- Case Analysis: "Capital One: Information-Based Credit Card Design" (Class 19)

Scaling Analytics

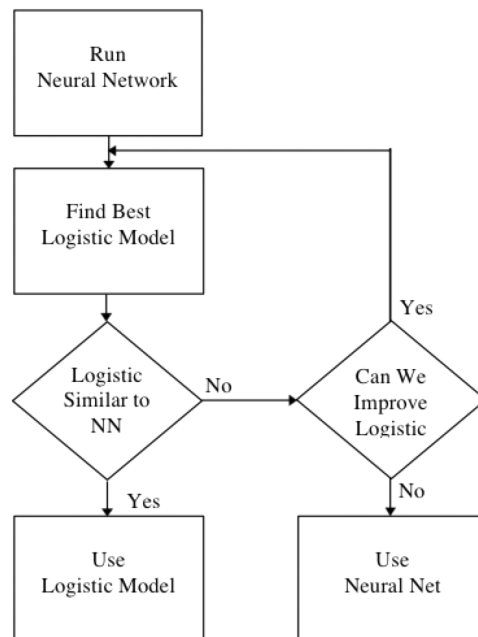
- Scaling Analytics in Practice (Class 20)
Course Wrap-up

From last set of slides ...

But how can we use this for customer analytics in practice?

NNs can be used well in combination with logistic regression

COMBINING ANN AND LOGISTIC REGRESSION



Core idea:

- Use NN as performance benchmark
- Use logistic regression for variable selection and interpretation

We begin by analyzing the prediction of an NN for whether customers buy “The Art History of Florence”

EXAMPLE 1: NN APPLICATION FOR BOOKBINDERS

- Stan Lawton (marketing director) pulls a random sample of 50,000 customers from the Bookbinders database
- Stan mails "The Art History of Florence" to the entire sample
- 4522 customers buy the book
- Plans to use the NN model to determine which customers to target from the entire database (500,000 remaining customers, excluding test group)
- Use R nnet package

For this example we use as input nodes the same variables we used in the logistic regression

EXAMPLE 1: NN APPLICATION FOR BOOKBINDERS

Output Node:

```
$ buyer      : Factor w/ 2 levels "0","1": 1 1 1 1 1 2 1 1 2 1 ...
```

Input Nodes:

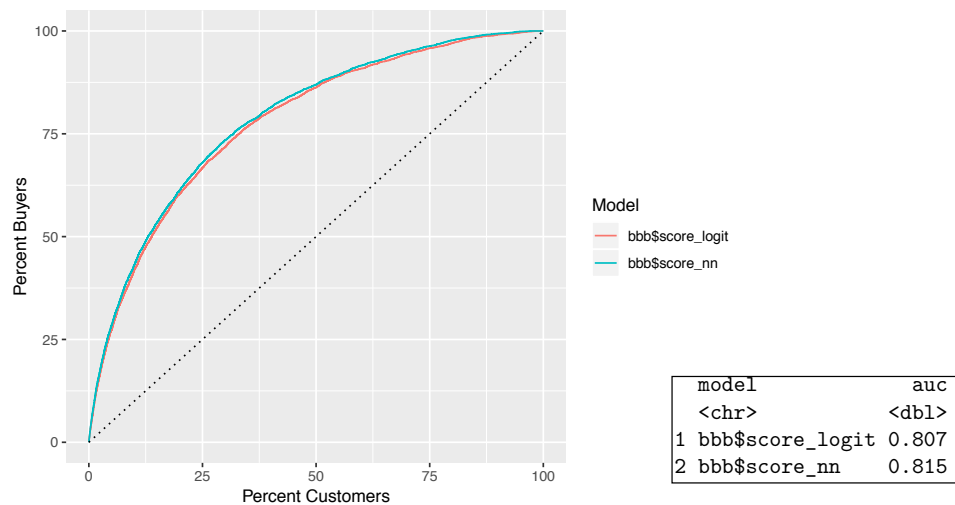
```
$ gender     : Factor w/ 2 levels "F","M": 2 2 1 1 1 1 2 2 1 1 ...
$ last       : int   29 27 15 7 15 7 25 1 5 11 ...
$ book       : int   109 35 25 15 15 15 15 124 130 15 ...
$ nonbook    : int   248 103 147 257 134 98 0 114 288 108 ...
$ total      : int   357 138 172 272 149 113 15 238 418 123 ...
$ child      : int    3 0 0 0 0 0 0 2 0 0 ...
$ youth      : int    2 1 0 0 0 1 0 1 2 1 ...
$ cook       : int    2 0 2 0 1 0 0 2 3 0 ...
$ do_it      : int    0 1 0 0 0 0 1 3 2 0 ...
$ reference  : int    1 0 0 1 0 0 0 0 0 0 ...
$ art        : int    0 0 0 0 0 0 0 0 3 0 ...
$ geog       : int    2 1 0 0 0 0 0 3 1 0 ...
```

Sample: 50,000

R Demo

The NN does about as well as the logistic regression model

GAIN CALCULATION FOR NN MODEL



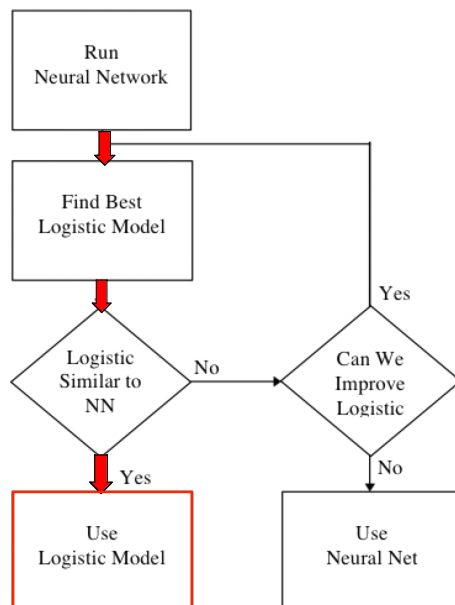
The NN does about as well as the logistic regression model

PROFIT COMPARISON OF NN MODEL

	Mass Mailing	RFM (Independent)	RFM (Sequential)	Logistic	NN
Number of	500,000	500,000	500,000	500,000	500,000
% Mailed	100.000%	46.582%	45.600%	31.120%	30.424%
Response Rate	9.044%	14.001%	14.197%	21.356%	22.048%
Number Mailed	500,000	232,910	228,000	155,600	152,120
Number of Orders	45,220	32,610	32,370	33,230	33,540
Gross Profit (@\$6)	\$271,320	\$195,660	\$194,220	\$199,380	\$201,240
Mailing Cost (@\$.5)	\$250,000	\$116,455	\$114,000	\$77,800	\$76,060
Profit	\$21,320	\$79,205	\$80,220	\$121,580	\$125,180
Return on Mktg.	8.528%	68.013%	70.368%	156.272%	164.581%

Using the NN gives us confidence that our logistic model is capturing behavior well

COMBINING NN AND LOGISTIC REGRESSION



End Result:

- Confidence in predictive performance
- Easy variable interpretation
e.g. "art" purchases
and "male" matter a lot

Consider a second application of NNs

EXAMPLE: FIREWALL TOOL

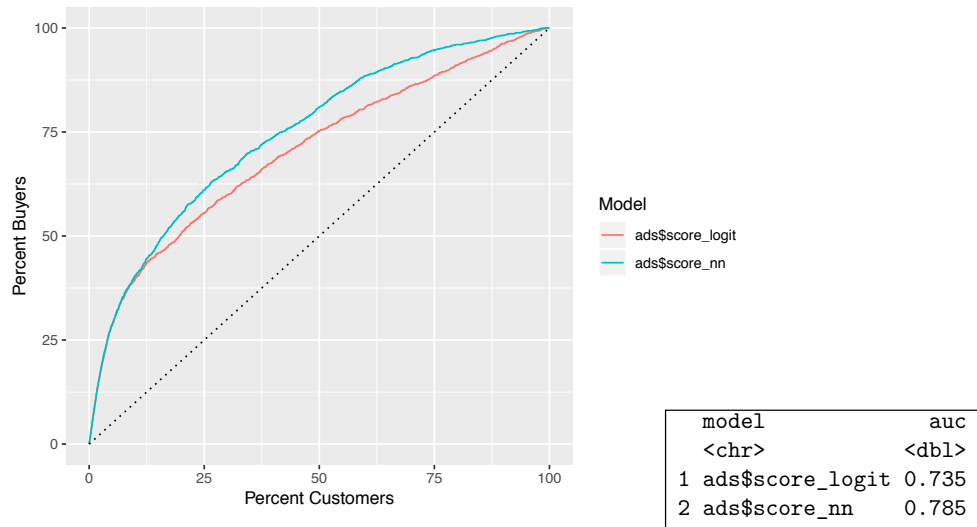
- Firewalls on PC are notoriously hard to manage (require knowledge of IP ports and networking)
- New tool for configuring Windows firewall
- Profit on firewall tool is \$10, cost to target customer is \$1
- We have data on 10,000 customers who have been targeted in test campaign
 - Ad-copy A emphasizes “ease of use” (4,607 customers)
 - Ad-copy B emphasizes “control/options” (5,393 customers)
- Available data:

res	Is 1 if responded to offer, 0 if not
age	age of customer
num purch	total number of purchases
totdol	total dollars spent
adB	Is 1 if Ad-copy B 'control/options', 0 if Ad-copy 'A' ease-of-use'
female	Is 1 if female, 0 if male
- Goal: build predictive model

R Demo

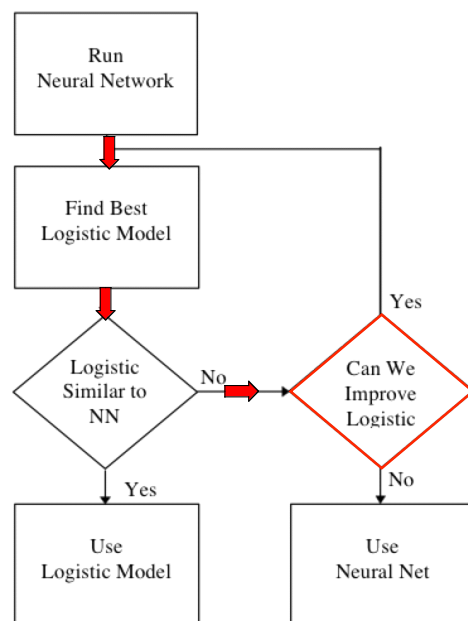
The NN does much better than the logistic regression model

GAIN CALCULATION FOR NN MODEL



The comparison suggests that the logistic model is missing something important

COMBINING NN AND LOGISTIC REGRESSION



Ideas?

-

```
res
-> Is 1 if responded to offer, 0 if not

age
-> age of customer

numpurch
-> total number of purchases

totdol
-> total dollars spent

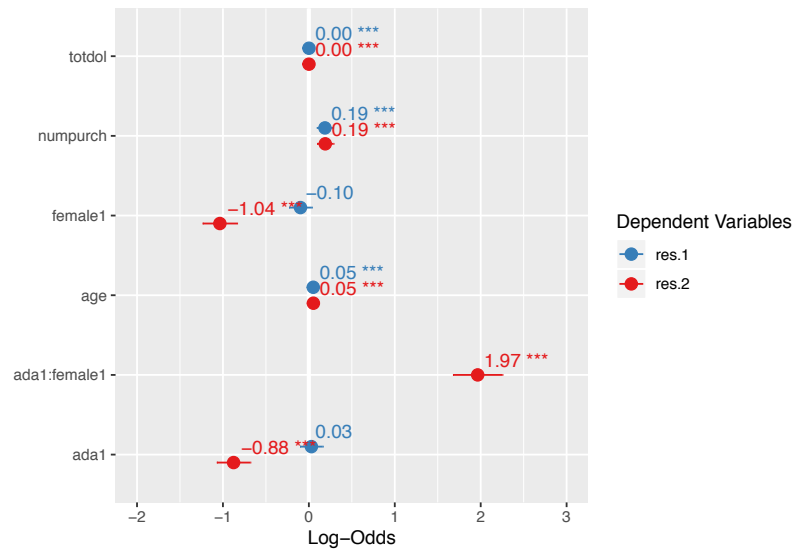
adB
-> Is 0 if Ad-copy A 'ease-of-use',
    1 if Ad-copy B 'control/options'

female
-> Is 1 if female, 0 if male
```

Could interaction effects matter (adA*female)?

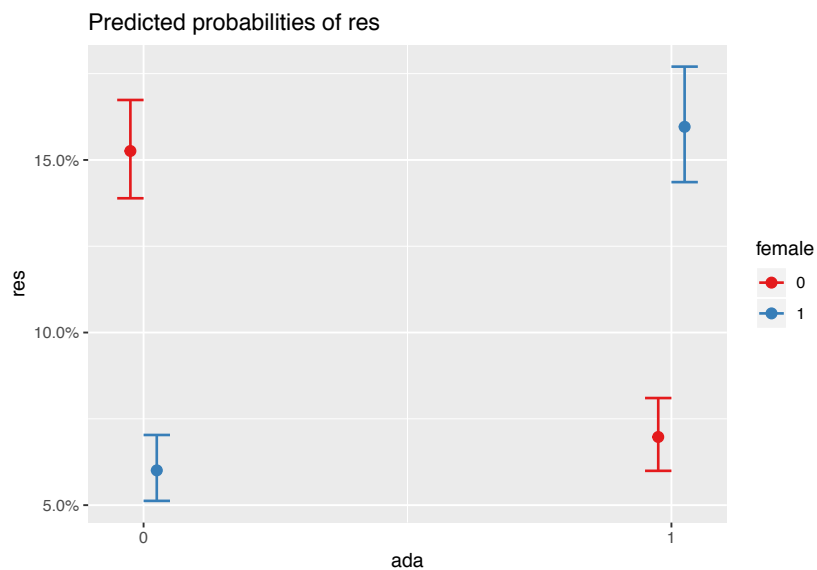
FIREWALL EXAMPLE: RESULTS FROM LOGISTIC REGRESSION

```
\> plot_models(logit_ads, logit_int_ads, show.values = TRUE, transform = NULL)
```



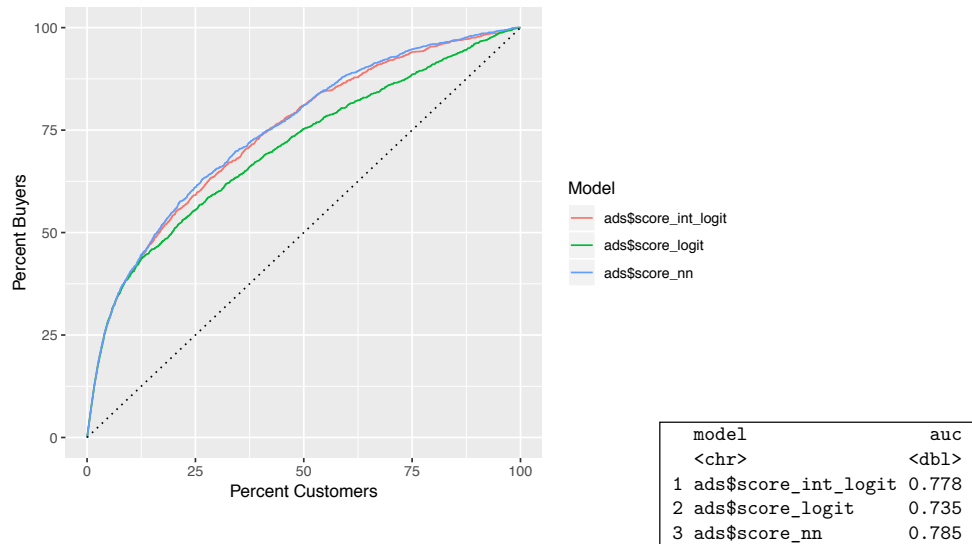
Could interaction effects matter (adA*female)?

FIREWALL EXAMPLE: RESULTS FROM LOGISTIC REGRESSION



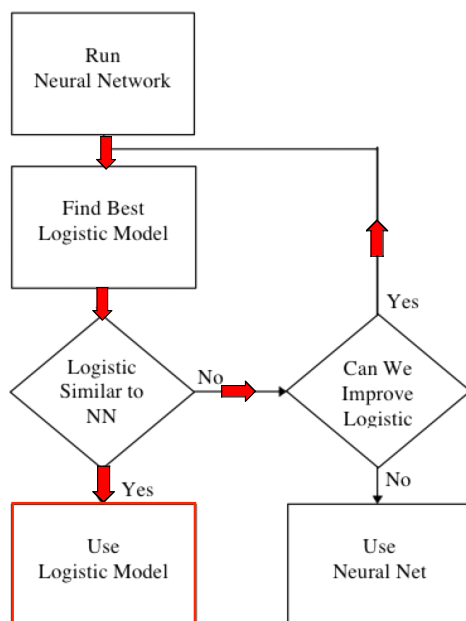
Now the logistic regression performs similar to the NN

GAIN CALCULATION FOR NN MODEL



The iterative process has revealed something important about consumer behavior

COMBINING NN AND LOGISTIC REGRESSION



End result:

- Confidence in predictive performance
- Easy variable interpretation
- We have learned that the ad-copy appealing to "ease-of-use" appeals more to women and the ad-copy appealing to "control/options" appeals more to men ==> use for marketing from now on!

Neural networks work well when predictions are key and there are relatively few inputs

WHEN TO USE NEURAL NETWORKS

- Good choice for classification and prediction tasks when the results of the model are more important than understanding how results come about
- Do not use to inform model building (where decision trees/regressions are very good)
- Do not work well when there are hundreds or thousands of input variables
 - Network has trouble finding pattern
 - Long training phase which never converge to a good solution
- Neural networks work well in combination with regressions (decision trees):
 - Use regression (or decision trees) to find important variables
 - Use variables as inputs for Neural Network
 - Use Neural network for predictions
- For more information, see: James/Witten/Hastie/Tibshirani, "An Introduction to Statistical Learning" and Hastie/Tibshirani/Friedman: "The Elements of Statistical Learning"
<https://web.stanford.edu/~hastie/ElemStatLearn/>