Predicting Subscriber Dissatisfaction and Improving Retention in the Wireless Telecommunications Industry

Michael C. Mozer[+*]
Richard Wolniewicz[*]
Robert Dodier[*]
Lian Yan[*]
David B. Grimes[+*]
Eric Johnson[*]
Howard Kaushansky[*]

[+]Department of Computer Science
University of Colorado, Boulder

[*]Athene Software
Boulder, Colorado

The Wireless Industry

Extremely dynamic and competitive market

- Penetration rate 25% in 1998, 50% in 2004 vs. 71% in France (3Q 2003)
- Some local markets have as many as five carriers.
- Carriers announce new rates and promotions almost every month.
- New services and technologies are constantly introduced.

Competition has resulted in high rate of *churn*—customers switching from one carrier to another.

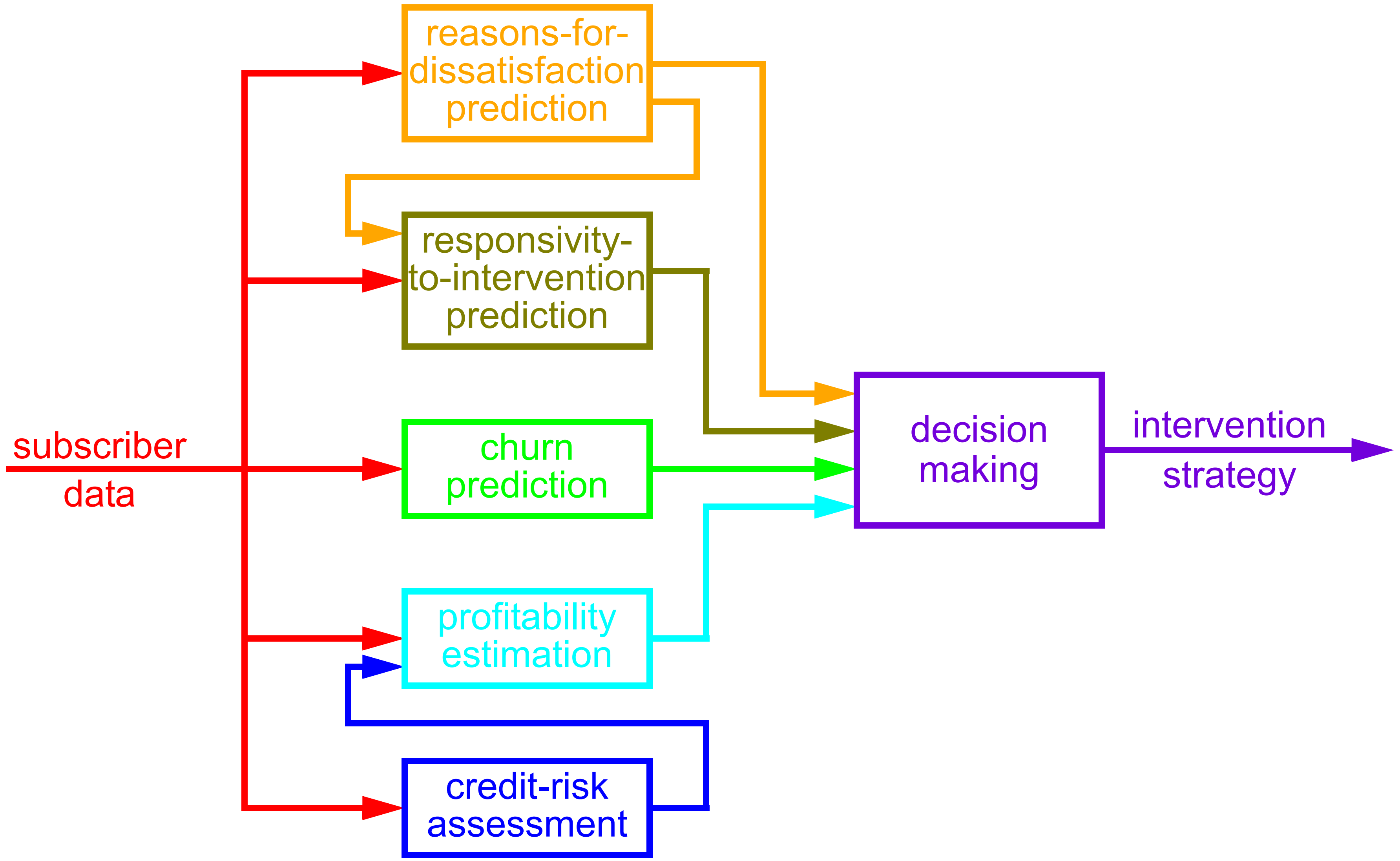
Monthly churn rates in US are ~2% of customer base.

- Feb 2004: 14% of AT&T customers thinking of churning in next 3 months

Churn cost industry nearly $10 billion in 2001.

- Signing new subscriber costs 5 times as much as retaining existing one.
- For carrier with 5M subscribers with an annual churn rate of 30%, that is a lost revenue of $870M. Cutting churn in half will save $435M.

Decision-Making Framework



Factors Influencing Subscriber Satisfaction

Factor	Importance	Nature of data required for prediction
call quality	21%	network
pricing options	18%	market, billing
corporate capability	17%	market, customer service
customer service	17%	customer service
credibility / customer communications	10%	market, customer service
roaming / coverage	7%	network
handset	4%	application
billing	3%	billing
cost of roaming	3%	market, billing

Information Sources Related to Churn

Network

- usage patterns (peak/off peak, number and duration of calls, location of calls)
- dropped calls
- quality of service

Billing

- base fee
- charge for minutes beyond prepaid limit
- roaming charges

Customer Service

- nature of complaints and resolution

Application for Service

- rate plan
- handset type
- credit history
- active services (number, type, avenue of activation, cancellation dates)
- customer classification (corporate vs. retail)

Market

- competitor rate plans

Demographics

- population density
- average income

Data Set

Provided by national wireless carrier

Account profile

- 46,744 subscribers, primarily small businesses
- average revenue per subscriber = $234
- no long-term contracts
- four state region
- 20% in major metropolitan areas

Time period

- Training data from October to December, 1998
- Test involved predicting churn in January or February 1999
- 6.2% churn rate

Data Representation

Naive

134 variables $\rightarrow$ 148 element vector

discrete one-of-*n* variables translated to an *n*-dimensional subvector
e.g., credit classification

Sophisticated

134 variables $\rightarrow$ 73 element vector

collapsed across some variables
e.g., different types of calls to customer service

expanded some variables
e.g., length of time with carrier

transformations
e.g., ratios, regression coefficients

Methodology

Ten-fold cross validation

Model classes

- logit regression
- neural network
- decision tree

Representations

- naive
- sophisticated

Model combination techniques

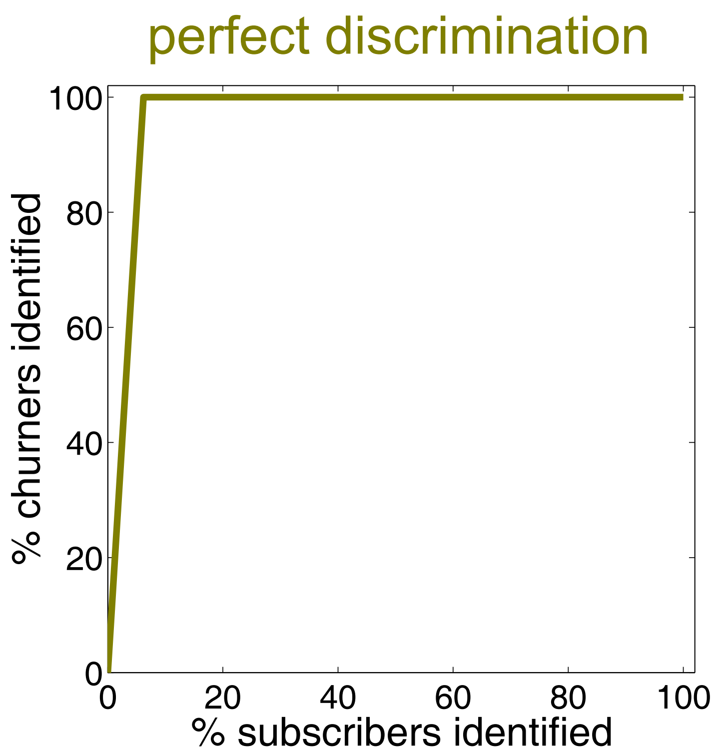
- single model
- majority vote
- Adaboost

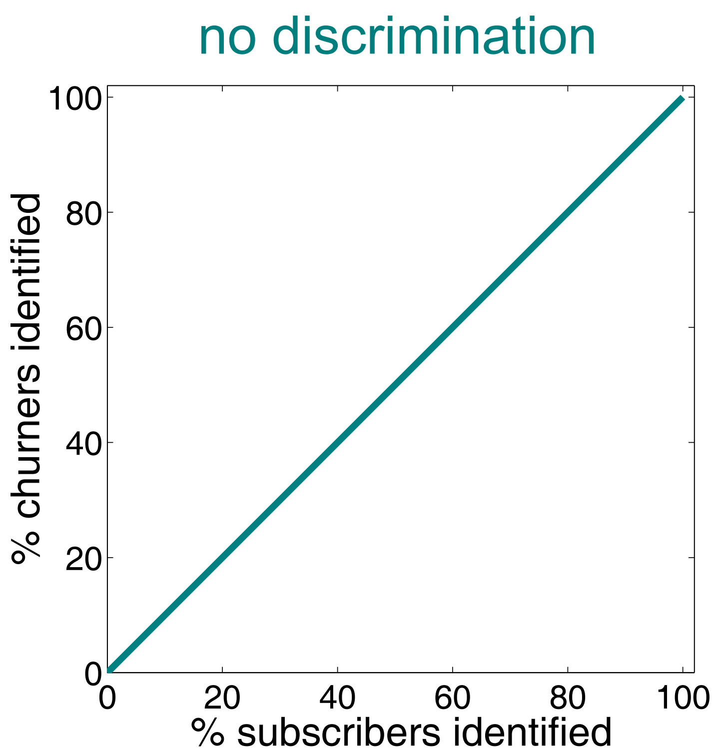
Lift Curve

Interpret predictor output as a probability of churn.

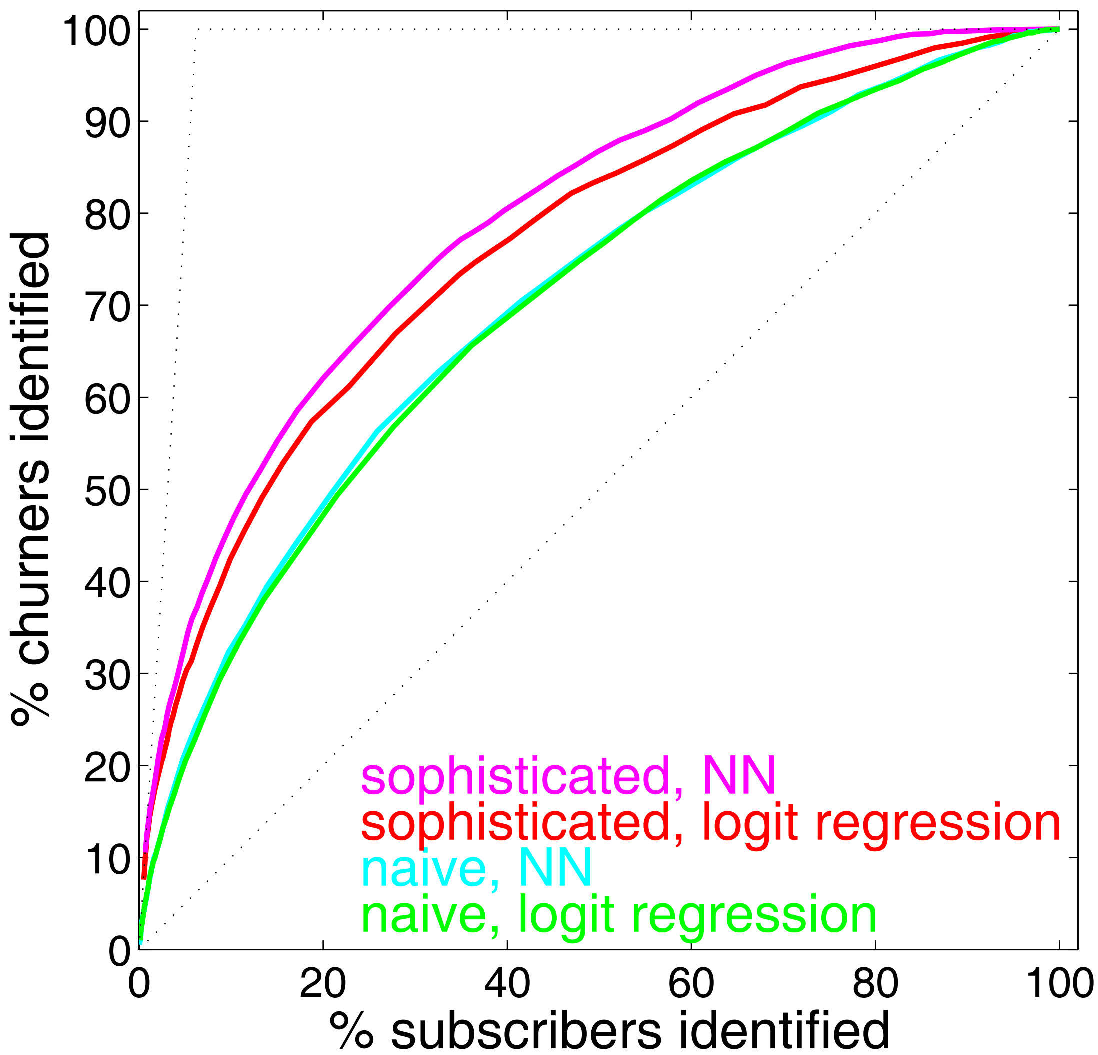
For a given probability threshold, determine
- **fraction of all subscribers above threshold, and**
- **fraction of all churners above threshold.**

Plot one quantity against the other for various thresholds.



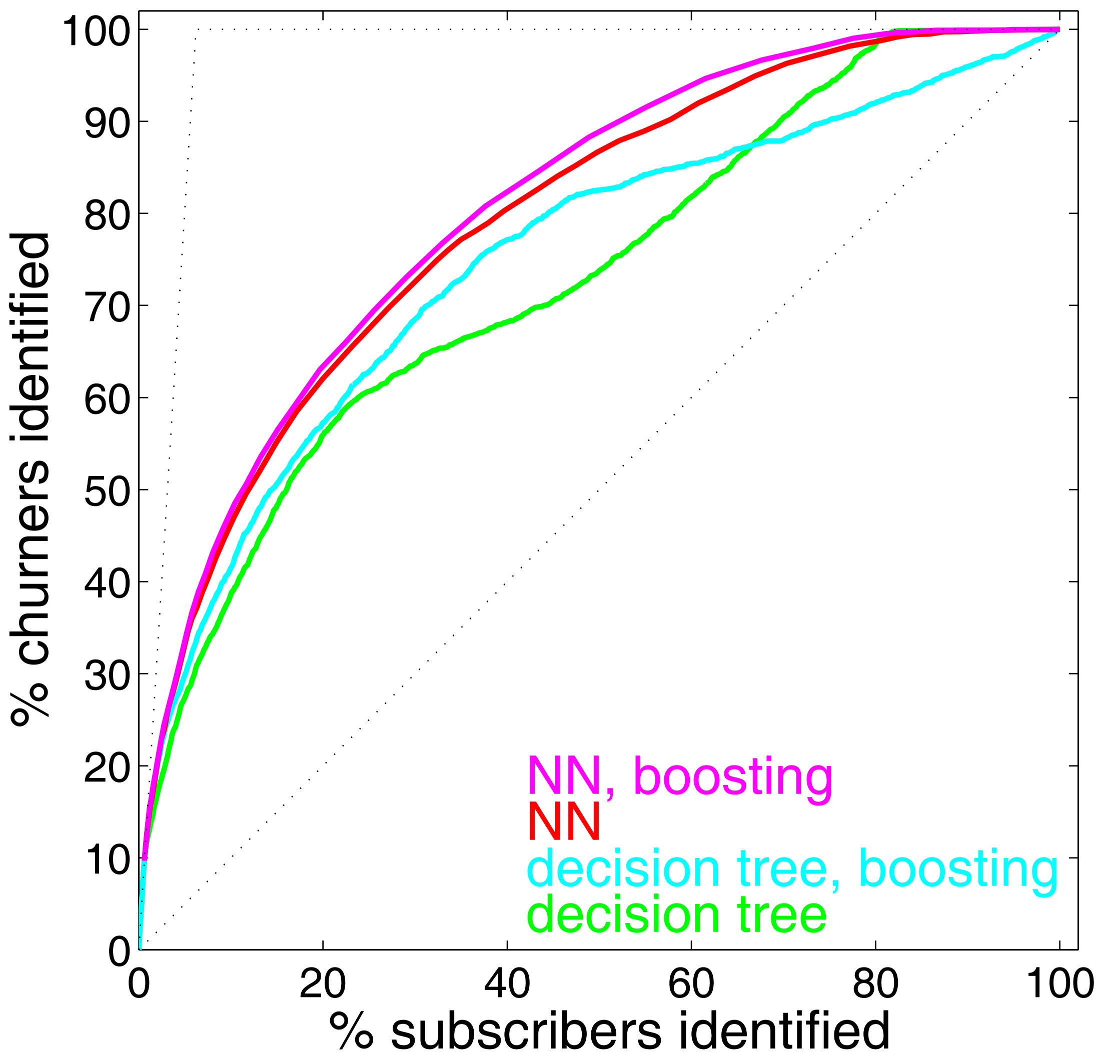


Results



100
90
80
70
60
50
40
30
20
10
0
% churners identified
0
20
40
60
80
100
% subscribers identified
sophisticated, NN
sophisticated, logit regression
naive, NN
naive, logit regression

Results



100
90
80
70
60
50
40
30
20
10
0
% churners identified
0
20
40
60
80
100
% subscribers identified
NN, boosting
NN
decision tree, boosting
decision tree

Decision Making

Should a given subscriber be contacted and offered some incentive to remain with the carrier?

Offer incentive to all subscribers with churn probability > θ

Select θ to maximize expected cost savings to carrier

Expected savings depends on

C_I	cost to carrier of providing incentive
H	time horizon over which incentive affects subscriber's behavior (assume 6 months)
P_I	reduction in probability that subscriber will leave within time horizon as a result of incentive
C_L	lost revenue that results from churn (assume $500 acquisition cost + income over H months)

Expected savings also depends on statistics from predictor

$N_{pc/ac}$ # subscribers predicted to churn who actually churn barring intervention

$N_{ps/ac}$ # subscribers predicted to stay who actually churn barring intervention

$N_{pc/as}$ # subscribers predicted to churn who actually stay

$N_{ps/as}$ # subscribers predicted to stay who actually stay

Net cost to carrier of performing no intervention

$$net_{NI} = (N_{pc/ac} + N_{ps/ac}) C_L$$

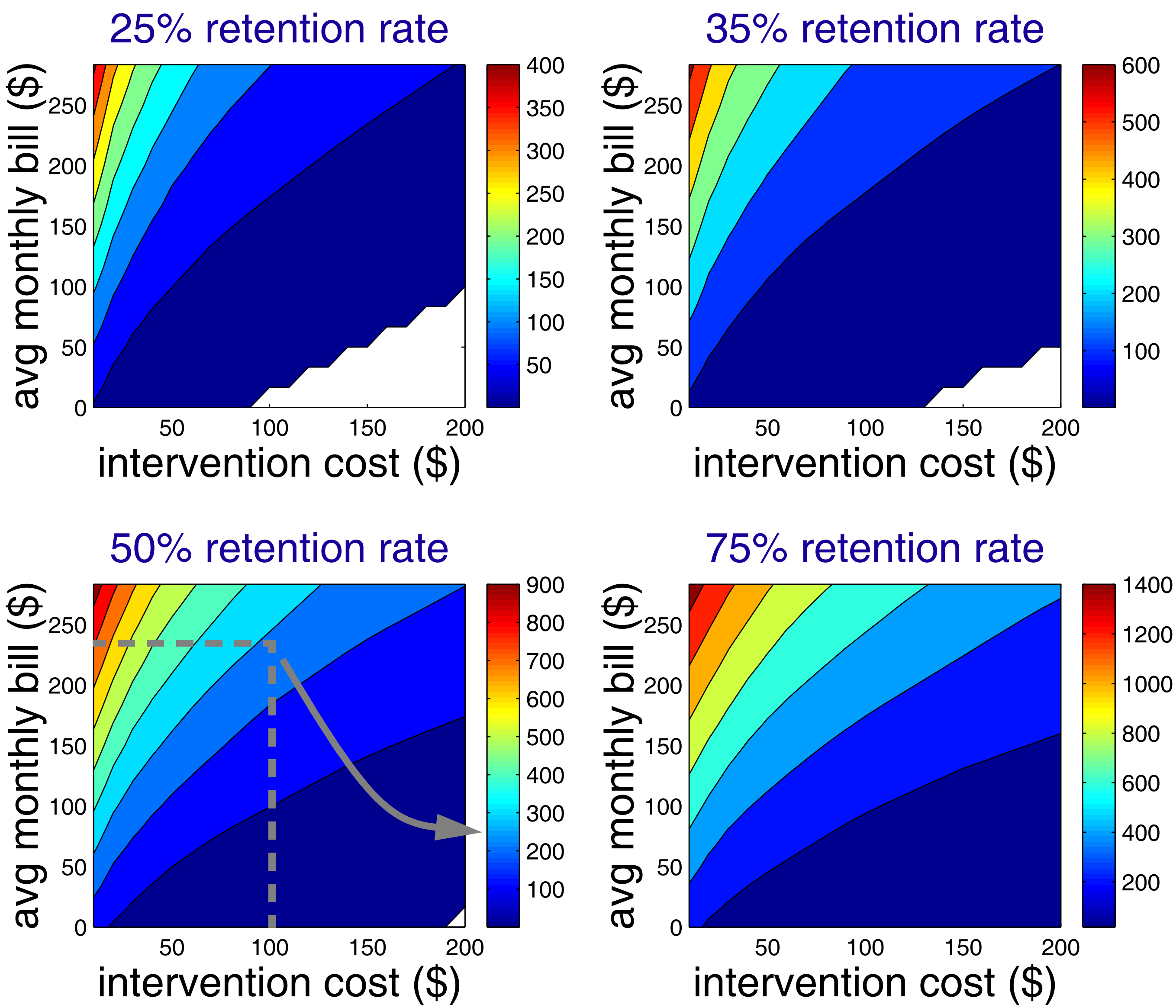
Net cost to carrier of performing intervention

$$net_I = (N_{pc/ac} + N_{pc/as}) C_I + (P_I N_{pc/ac} + N_{ps/ac}) C_L$$

Savings per churnable subscriber

$$savings = (net_{NI} + net_I) / (N_{pc/ac} + N_{ps/ac})$$

Expected Savings Per Churnable Subscriber Under Various Assumptions Concerning Intervention Cost and Resulting Retention Rate



Real World Testing

Six week experiment

- Control and treatment groups
- Treatment group contacted based on our recommendation
- Churn rate 3.7% in control group, 2.2% treatment group → 40% retention
- Intervention cost = \$92 (\$17 for incentive, \$75 for call center)
- By decision-theoretic framework, the savings per churnable customer is \$417.

From subscriber in treatment condition with .81 churn score:

...I am writing this letter in regards to an employee there that I feel deserves special recognition. Your representative, Alicia Holmes, has single handedly encouraged me to stay on with X as my cellular service provider. She is professional, competent, polite, an expert with her skills and knowledge of your services... She turned a bad experience with X into a good one. If not for her I would have left X as soon as possible...

Observations

Positive correlation with churn

- average monthly bill
- total number of calls
- credit class of customer
- ratio of recent to earlier monthly bills

Negative correlation with churn

- number of active dispatch and messaging services
- time with carrier
- average number of billed minutes per month

Mutual information between each variable and churn is extremely low (<< .01 bit).

Domain intuitions almost always supported by mutual information scores and prediction accuracy.

In many data sets, few higher-order regularities (given appropriate representation).

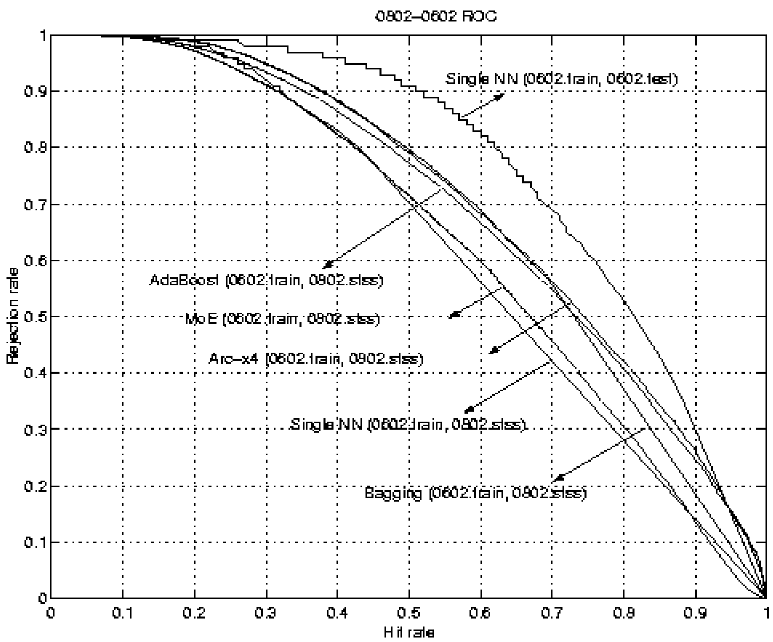
Because of redundancy and low information content, input pruning is feasible with no loss in accuracy.

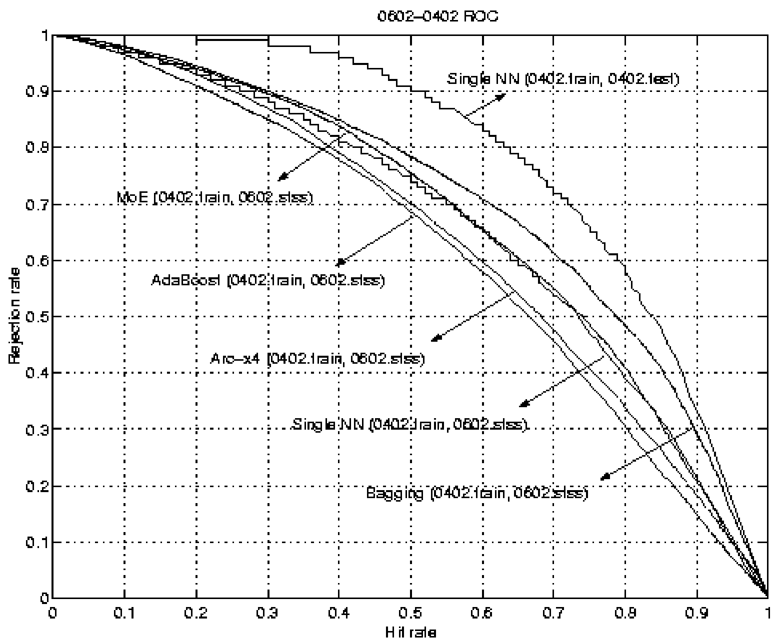
Further Experiments

Test window shifted in time

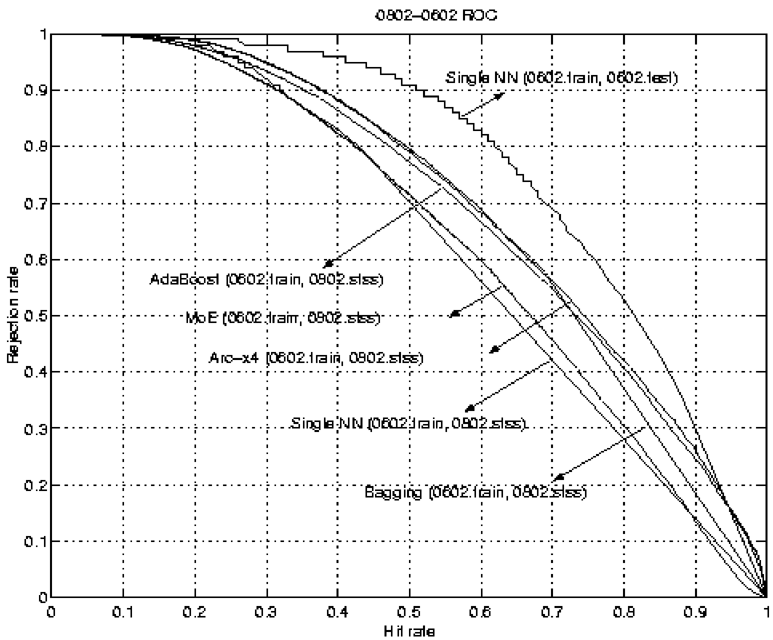
Replicability of results over time

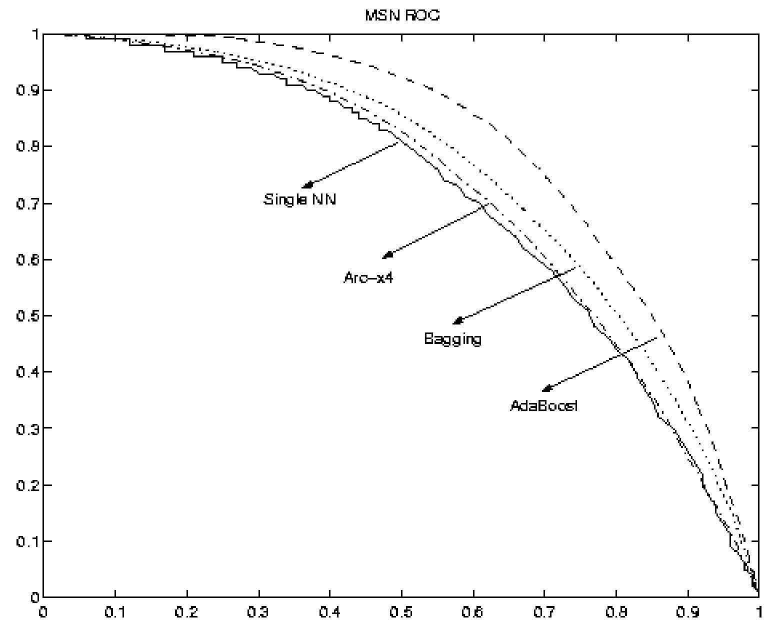
Comparison of various ensemble techniques



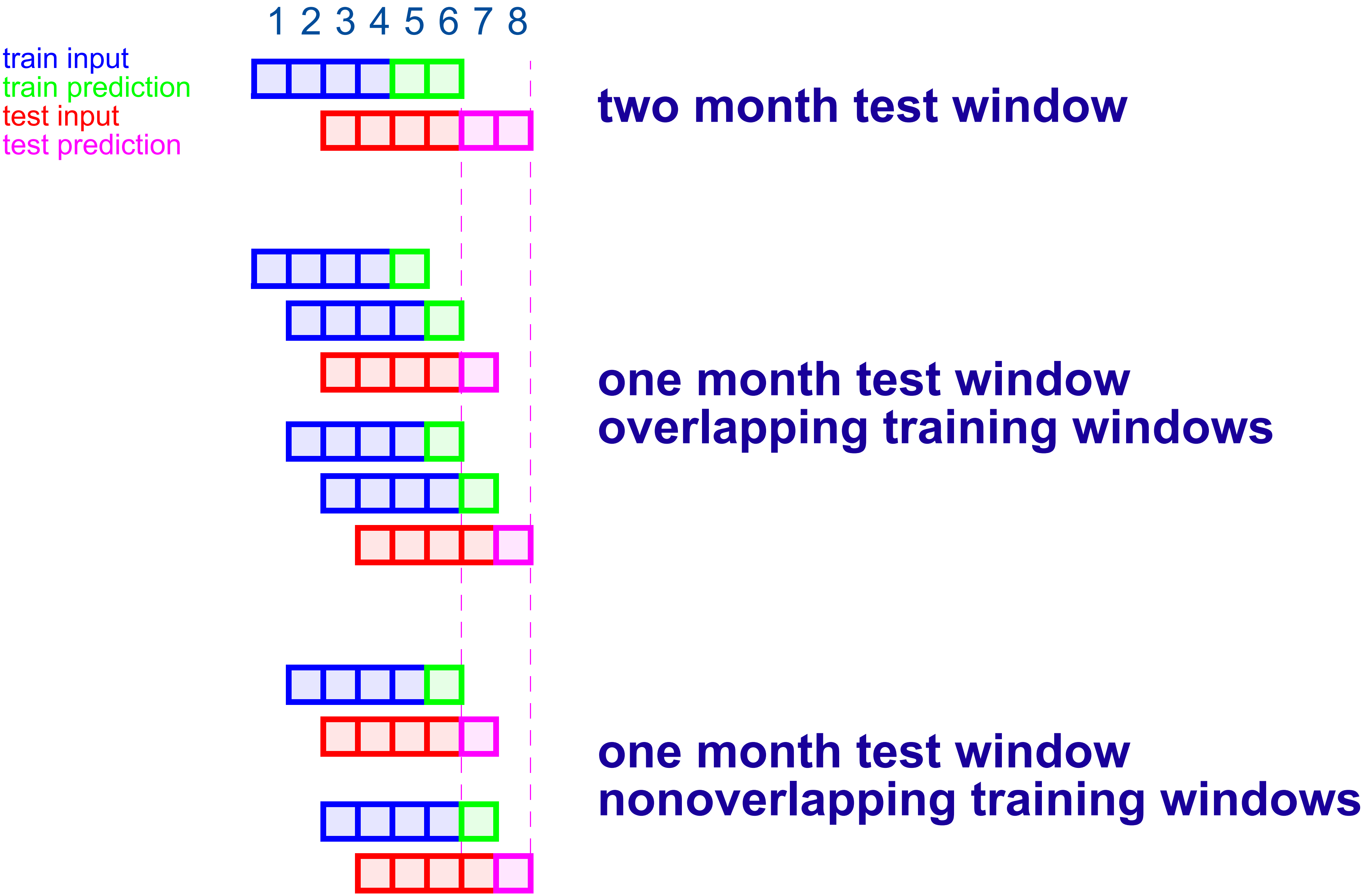


Comparison of Wireless and ISP Data Sets

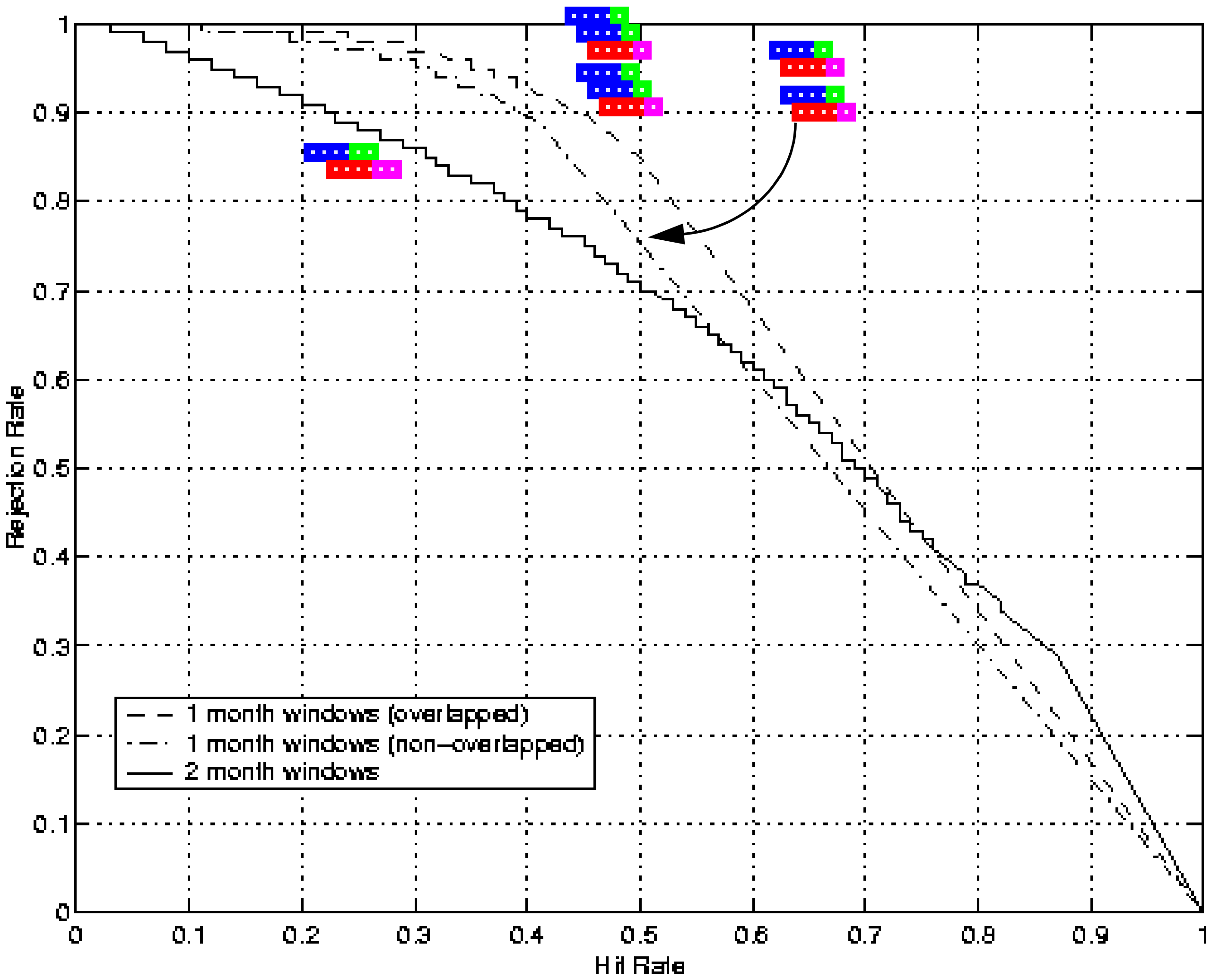




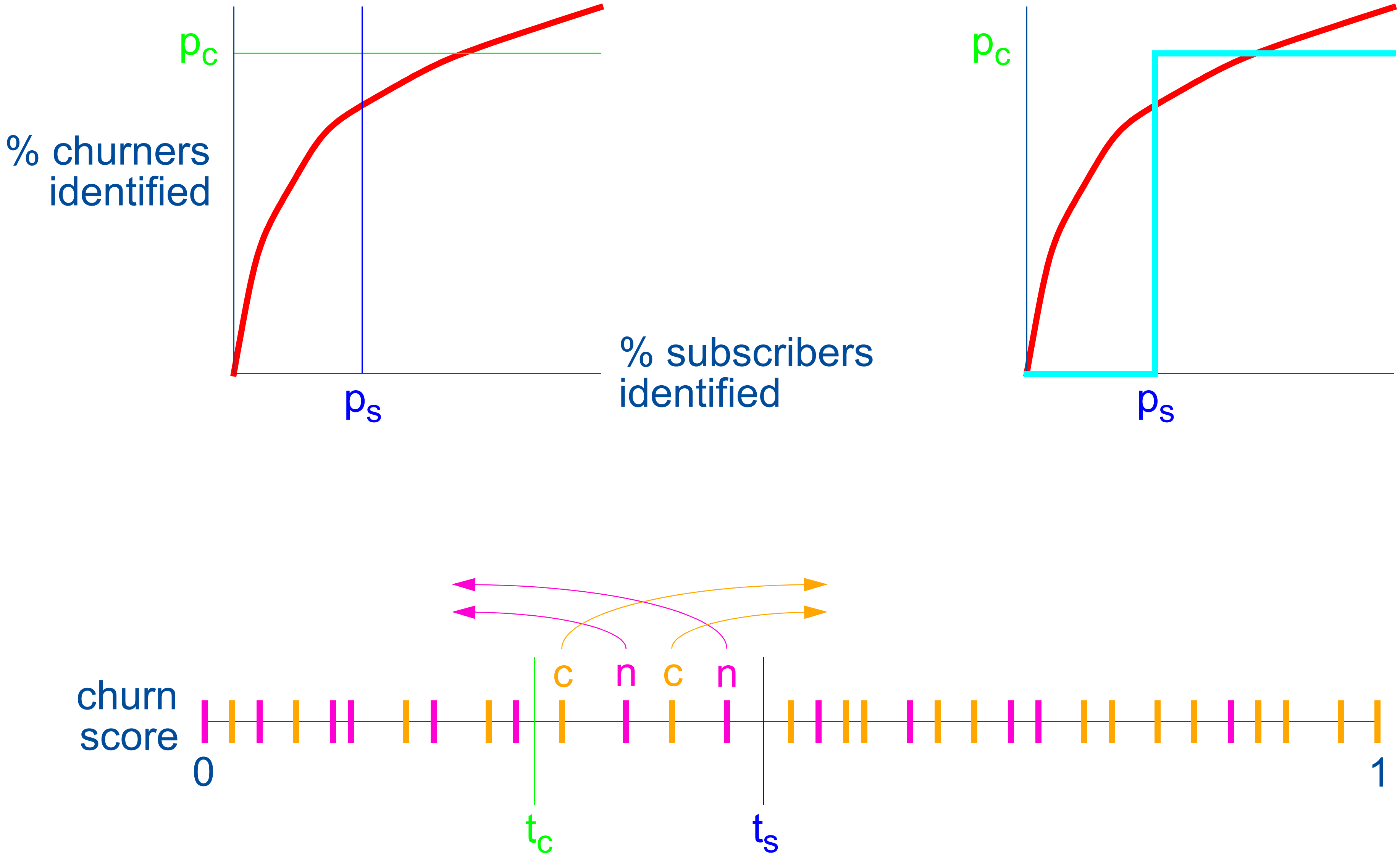
Test Window Experiment



Test Window Experiment



Optimizing Performance on Lift Curve



Decision Network for Subscriber Profitability