Analysis, Prediction, Recommendation For Churn Rate of a telecom

Raymond Kwok
23 June 2019

Flow

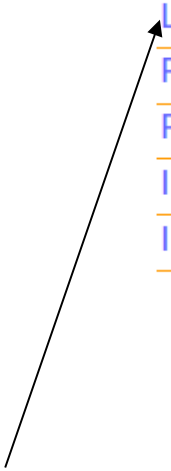
- Get to know the data
- Be creative!
- Make a choice on how to model
- See what features matter more and how
- Find a way to lower churn rate

Hidden column explanation
Let's educatedly guess!

Initially...

uid	Service_05
gender	Service_06
Factor_00	Service_07
Factor_01	Service_08
Factor_02	Service_09
Factor_03	C_01
Service_01	C_02
Service_02	C_03
Service_03	MonthlyCharges
Service_04	Charges

is_churn



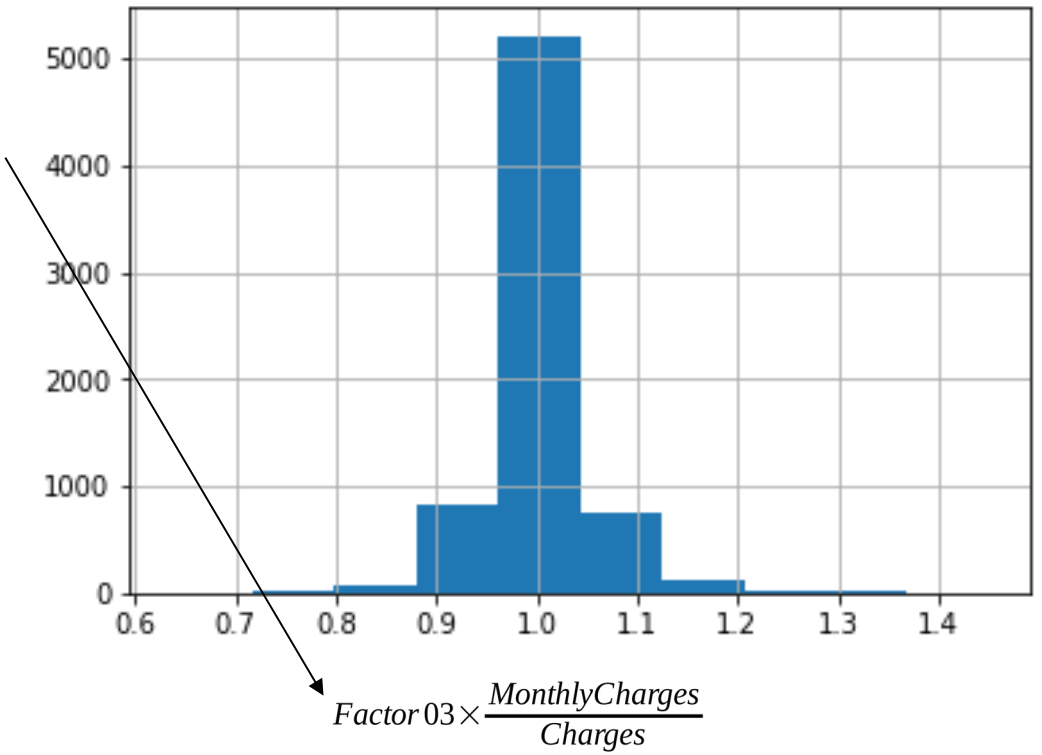
Guessed ...

uid	InternetService_addon2
gender	InternetService_addon3
Factor_FiberGroup	InternetService_addon4
Factor_loyalGroup	InternetService_addon5
Factor_loyalGroupForPhoneService	InternetService_addon6
LifetimeInMonth	SubscriptionPeriod
PhoneService	ExpensivePaymentChoice
PhoneService_addon	PaymentChannel
InternetService	MonthlyCharges
InternetService_addon1	Charges

is_churn

Why Factor_03 → LifetimeInMonths?

- This product is distributed around 1
- Having a unit of month makes sense
- So it's a time!



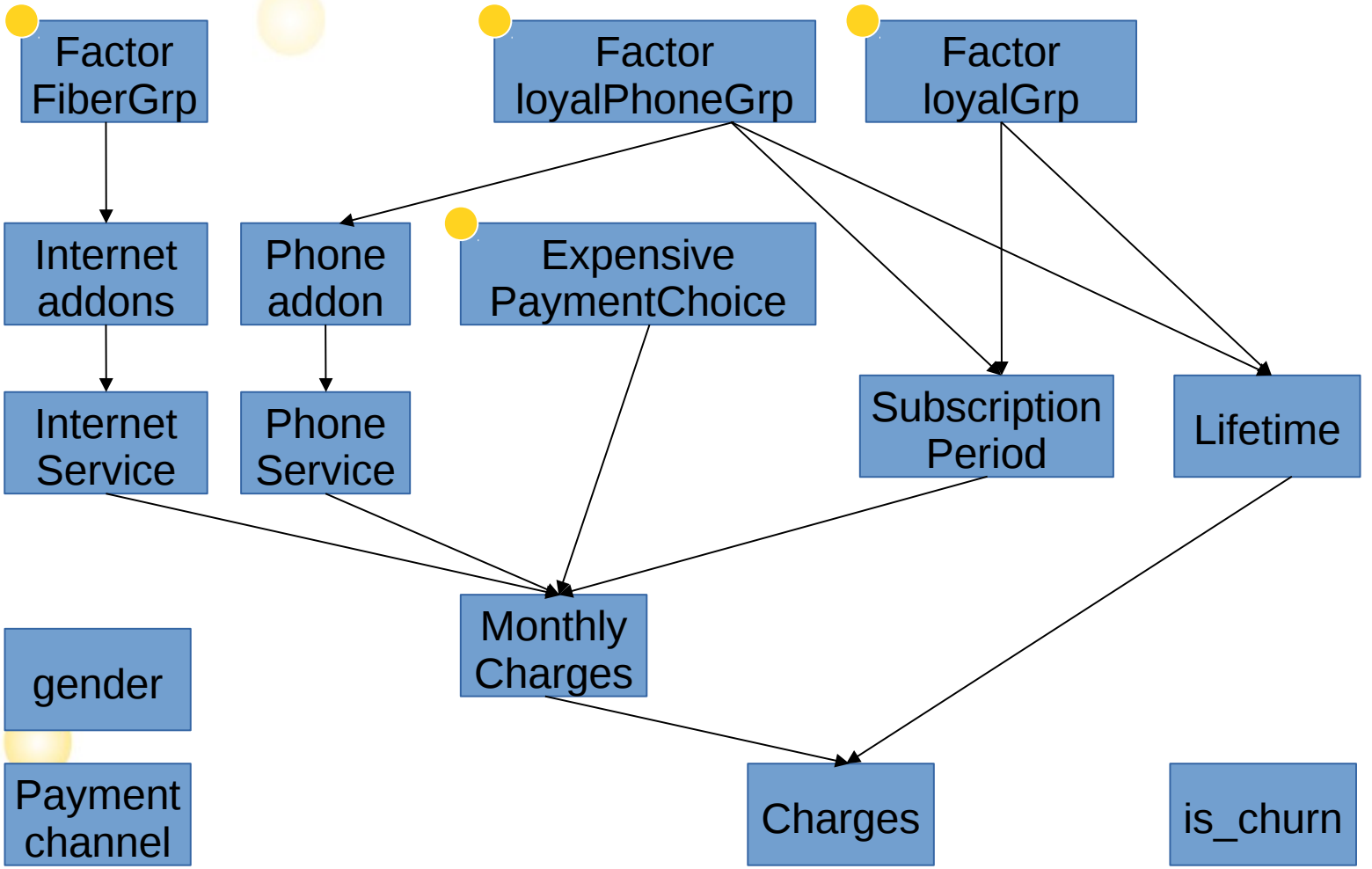
- *Hinted by high correlation between Factor_03 and Charges*

Why Phone / Internet Services?

- Well, it’s a telecom!
- Guess from feature values

Phone		*Internet*							*Billing*		
Service_01	**Service_02**	**Service_03**	**Service_04**	**Service_05**	**service_06**	**Service_07**	**Service_08**	**Service_09**	**C_01**	**C_02**	**C_03**
No	No phone service	DSL	Yes	No	Yes	Yes	Yes	No	Two year	Yes	Bank transfer (automatic)
Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mailed check
Yes	No	DSL	Yes	Yes	Yes	No	Yes	Yes	Two year	No	Mailed check
Yes	Yes	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mailed check
No	No phone service	DSL	Yes	Yes	Yes	Yes	Yes	No	Two year	No	Credit card (automatic)

Seems they are related this way

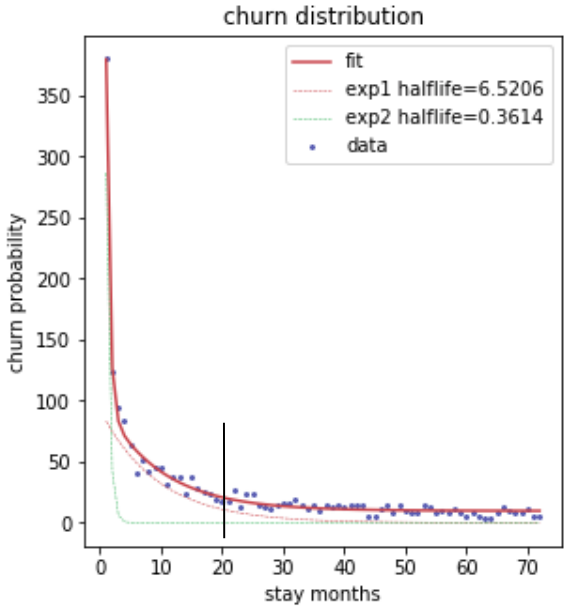


How about being *creative*?

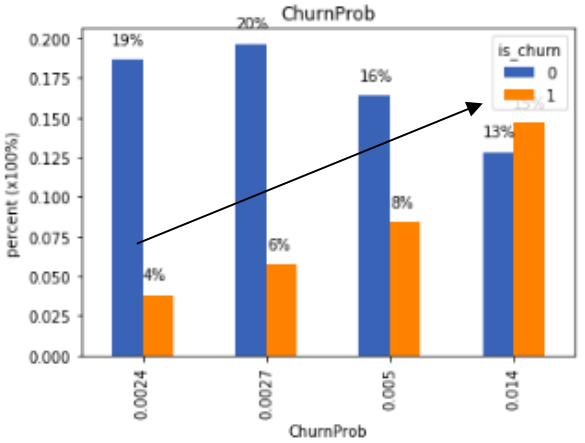
- Lack of client behavior and satisfactory level
- *Squeeze something out from the existing data!*

Let the churn population speaks
1/4: churn probability

- Naive arguments:
  - Clients are very similar
  - Clients leave at similar rates as those who have left
- Assign a churn probability to all clients

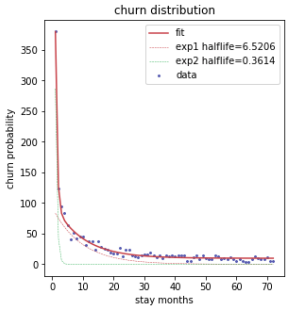


$$ChurnProb(x) = \frac{ChurnDistribution(x)}{\int_0^\infty ChurnDistribution(x)dx}$$



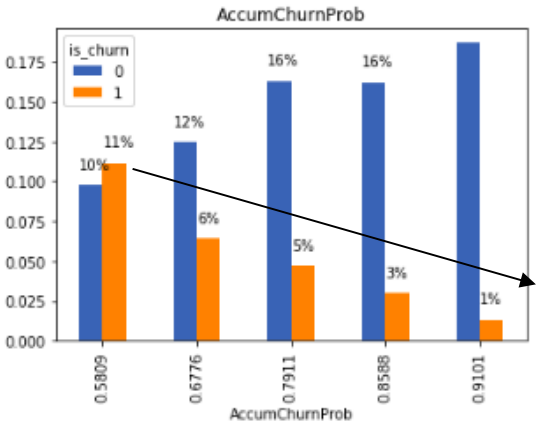
Let the churn population speaks
2/4: accumulative churn probability

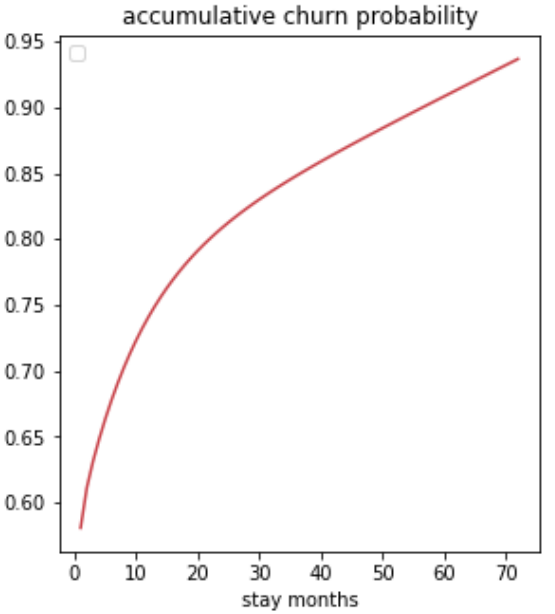
- Accumulate churn probability
- The longer a client stays, the higher the accumulative churn, the more possible it is to leave / stay
- *Non-linearize LifetimeInMonth*



Accumulate / Integrate

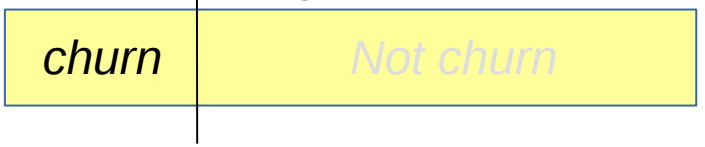
$$AccumChurnProb(x) = \frac{\int_0^x ChurnDistribution(x)dx}{\int_0^\infty ChurnDistribution(x)dx}$$



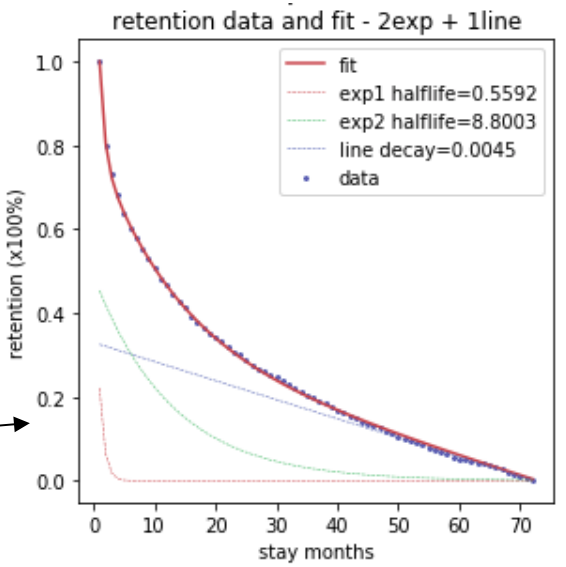


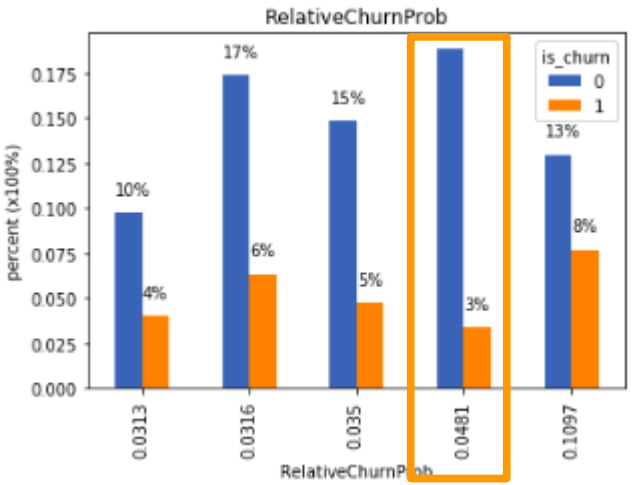
Let the churn population speaks
3/4: relative churn probability

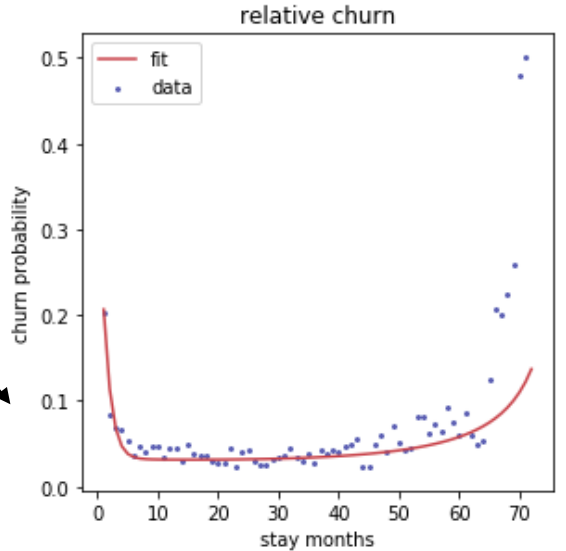
- Retention curve
- 3 processes
  - Exponential (half-life ~1mth)
  - Exponential (half-life ~9mth)
  - Linear (drop 0.45%/mth)



$$y(x) = a\exp(-bx) + c\exp(-dx) - ex + f$$
$$y'(x) = -(ba\exp(-bx) + dc\exp(-dx) + e)$$
$$RelativeChurnProb(x) = \frac{y'(x)}{y(x)}$$





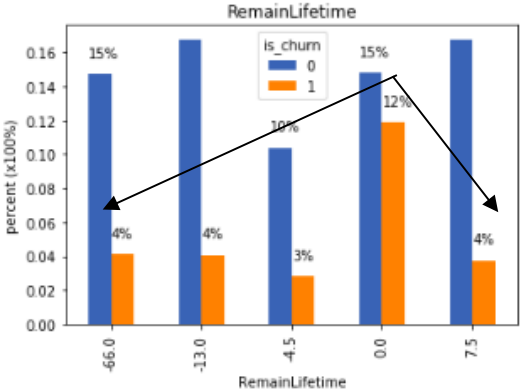


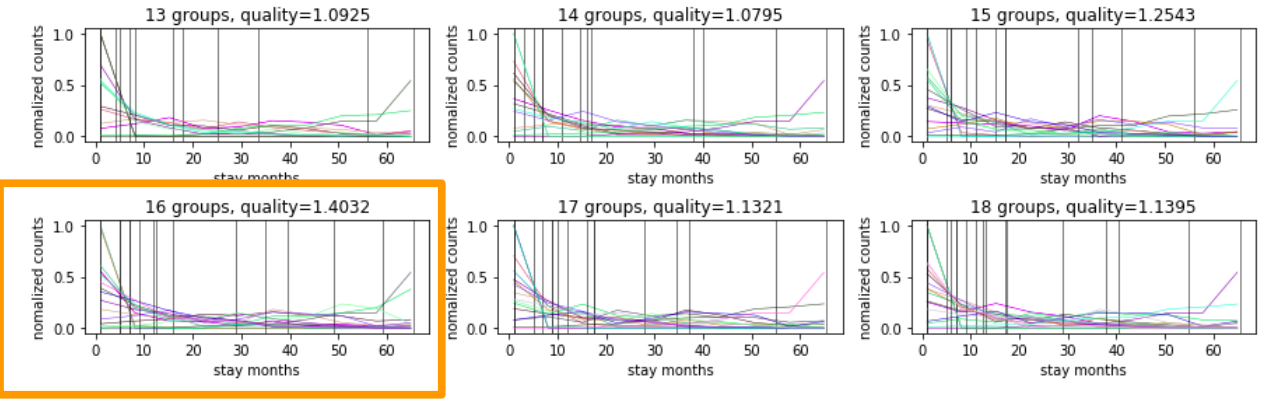
Let the churn population speaks
4/4: expected lifetime

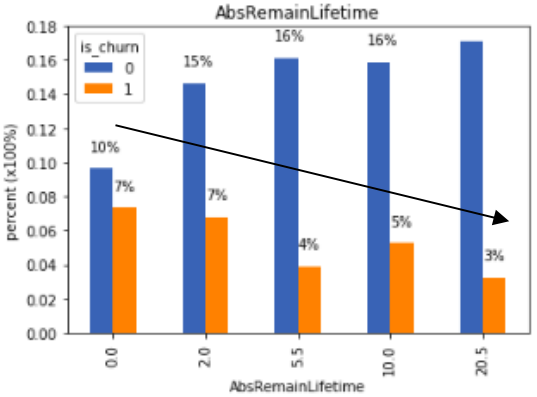
- Un-supervisedly group churn clients and give each group an expected lifetime as median value
- Assign all clients to the groups and get their remaining lifetime

churn | Not churn

$$RemainLifeTime = ExpectedLifeTime - LifeTime$$





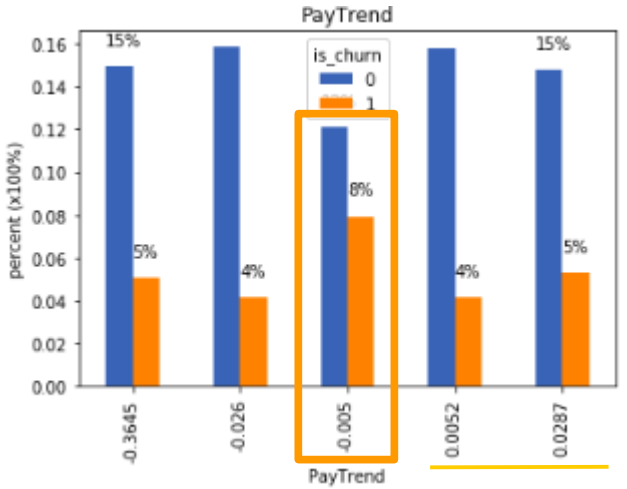


Change of client's pay

$$AvgMonthlyCharge = \frac{Charges}{LifetimeInMonth}$$

$$PayTrend = MonthlyCharge - AvgMonthlyCharge$$

- Positive > Pay more recently than in history > feel satisfied?

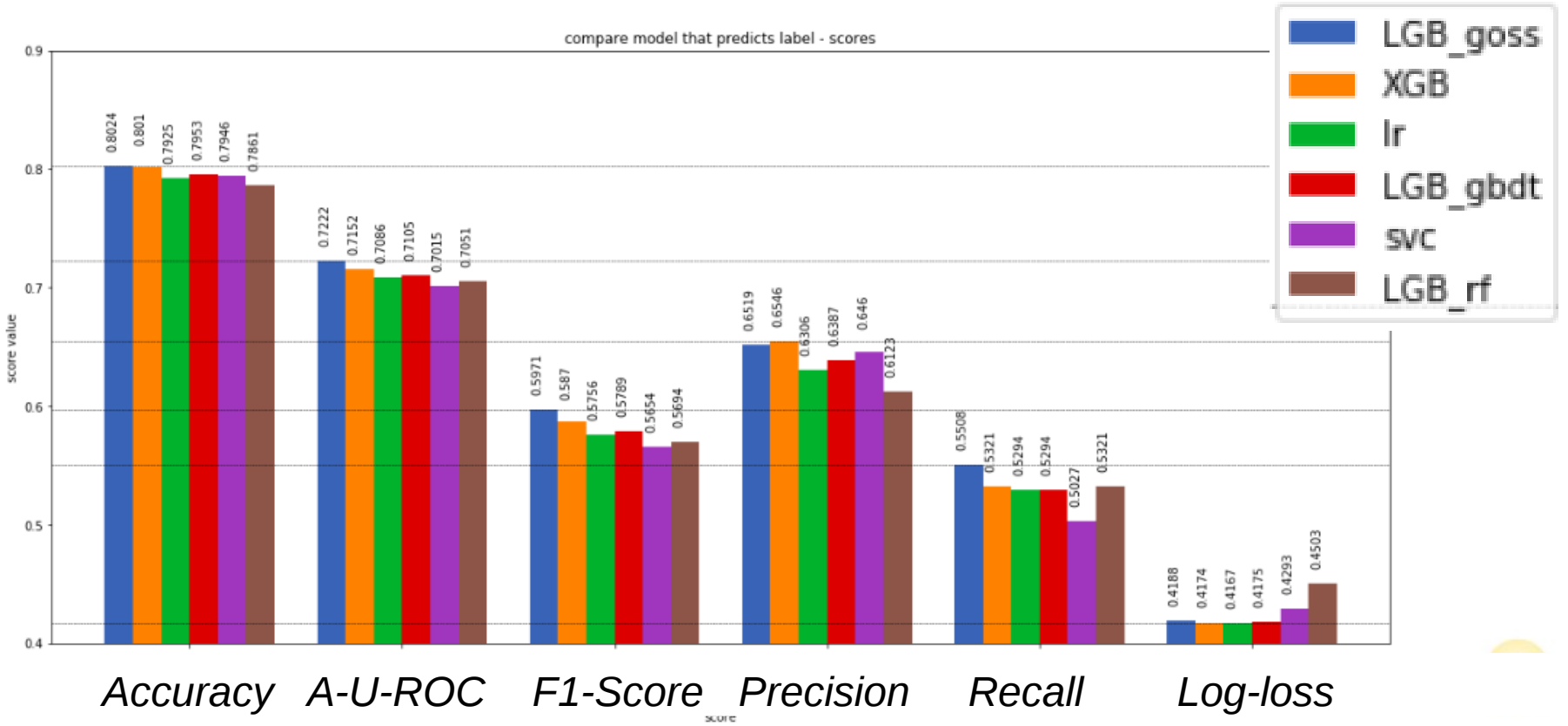


positive

Relatively higher churn at no change
Partly due to clients who stayed for only short period of time 1 month

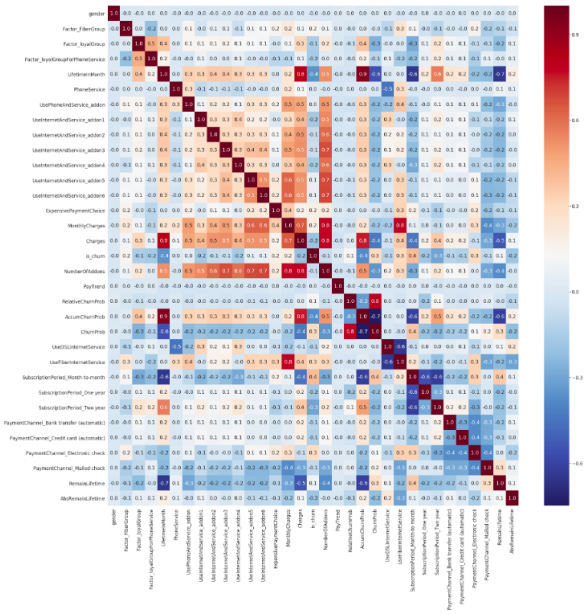
Models performance

- 4:1 train:test shuffled split
- Standard-Scaling learnt from train and apply to both
- *Stratified, 5-Fold CV* Grid-Search for best parameters
- LGB_goss 80.24% accuracy and is selected as the model!



Our data is ready!

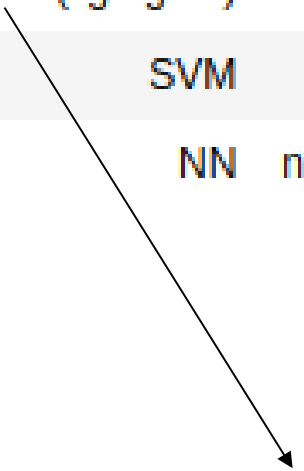
- From 7043 rows to 7032 rows
- From 20 features to 32 features
  - Adding new features
  - Label one-hot encoding
  - Removed uid
- And not miss the label ‘is_churn’



Scary giant correlation matrix

Models

Method	characteristic	feature importance
logistic regression	linear in feature space	by coefficients
GBDT (lightgbm)	non-linear in feature space	by feature frequency
SVM	linear in high dim space	by permutation test
NN	non-linear in high dim space	by permutation test



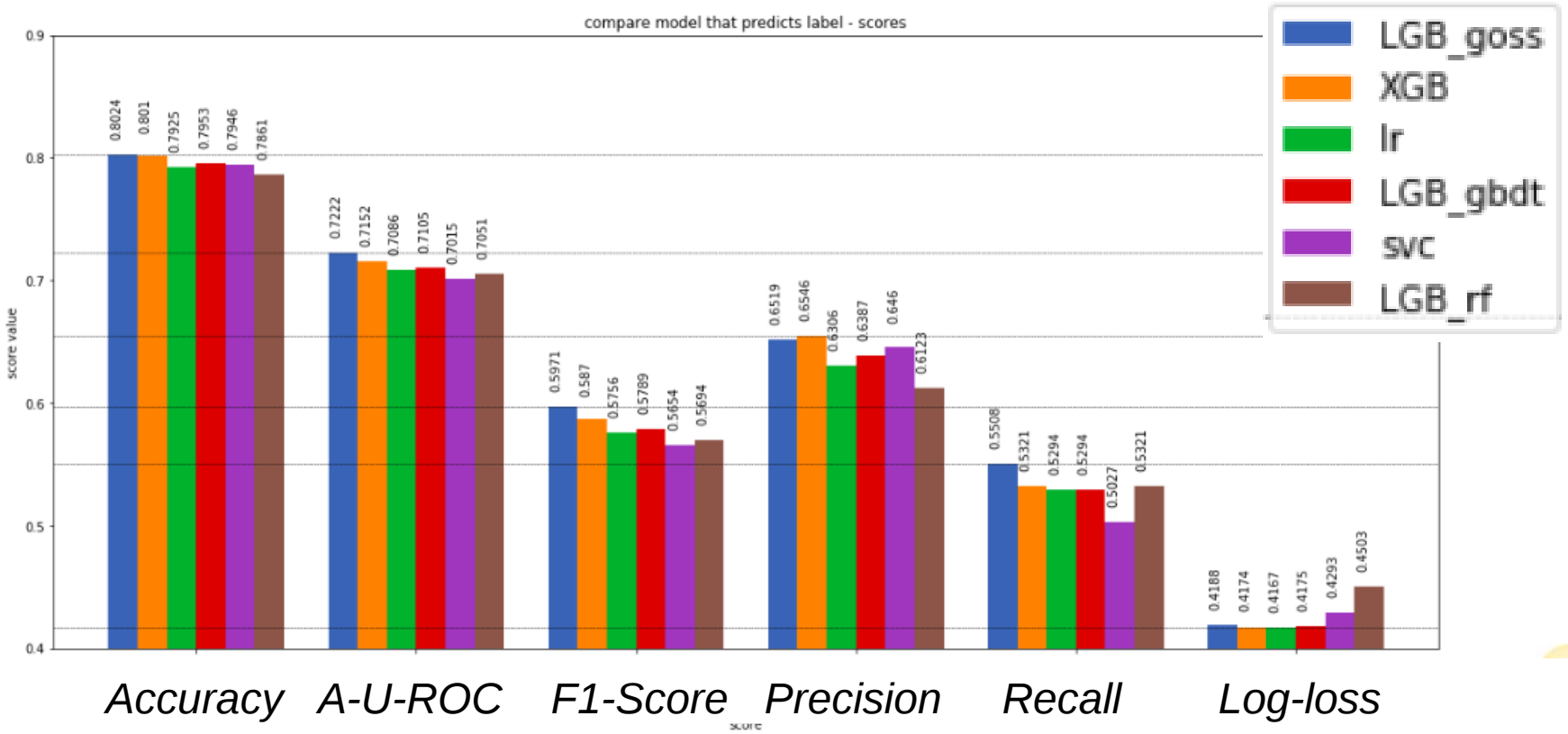
XGBoost
- gbtree (6.1min / 270fits)

LightGBM – faster by grouping up cols and rows
- gbdt (1.9min / 270fits)
- goss – down-sample low loss rows (1.4min / 270fits)
- random forest (7.5sec / 270fits)

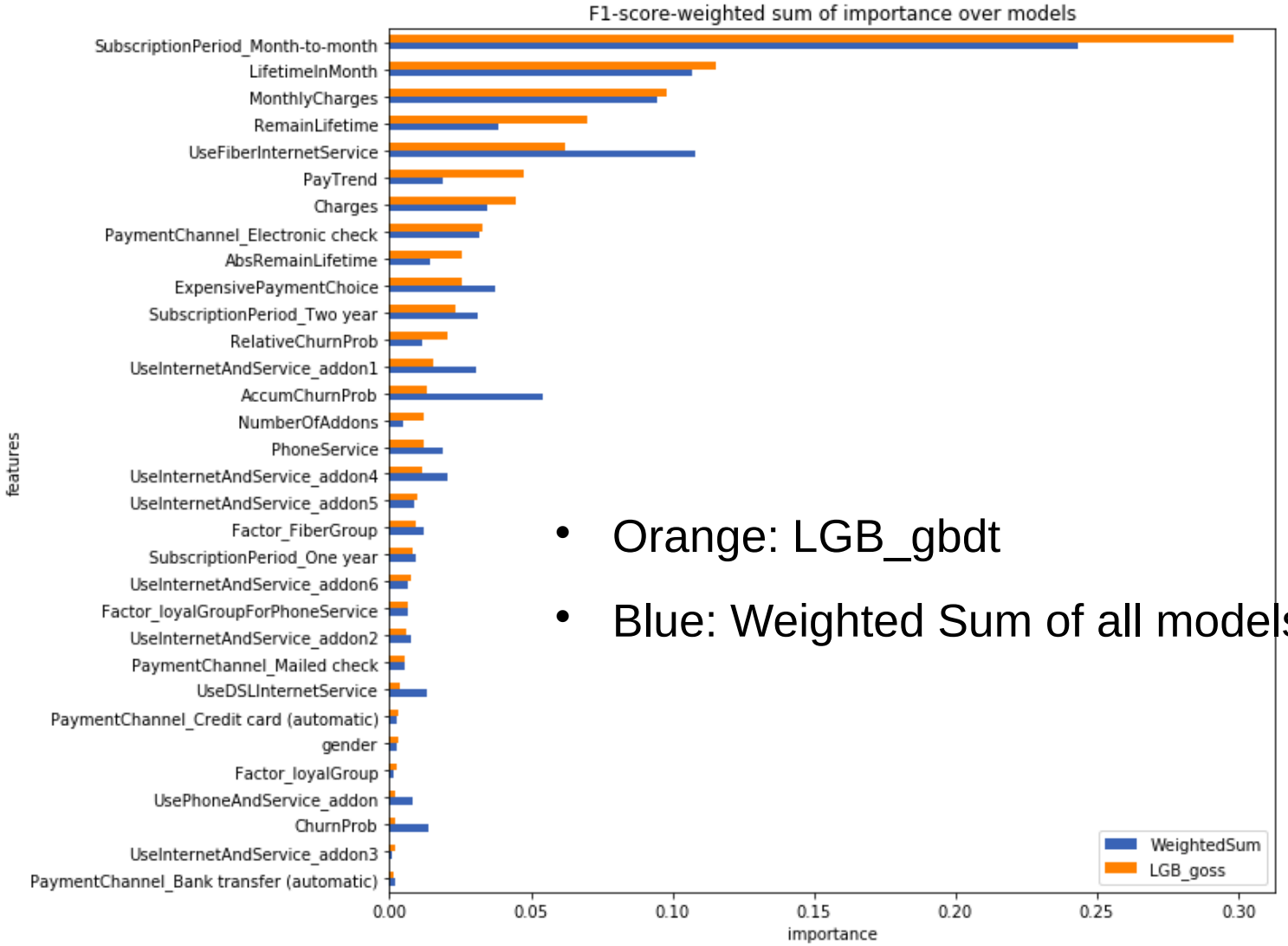
We have 7 models in total for comparison!

Models performance

- 4:1 train:test shuffled split
- Standard-Scaling learnt from train and apply to both
- *Stratified, 5-Fold CV* Grid-Search for best parameters
- LGB_goss 80.24% accuracy and is selected as the model!



Feature importance



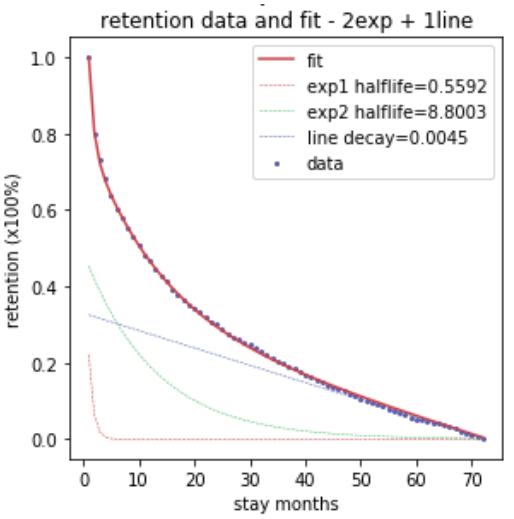
- Orange: LGB_gbdt
- Blue: Weighted Sum of all models

Feature importance X Improve Churn rate

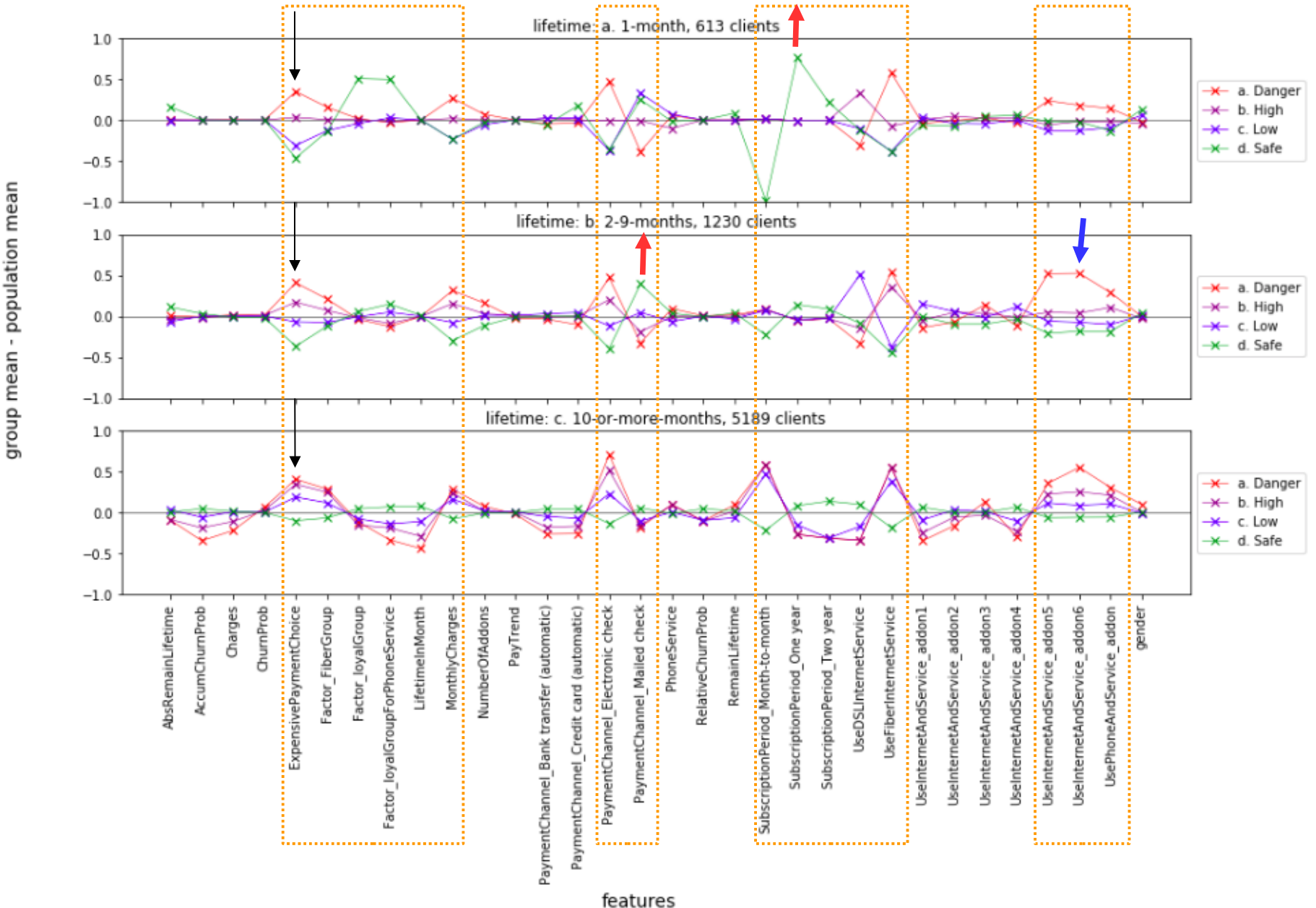
- Model predicts churn probability
- Important features have larger effect to making the probability large or small
- We want to know which feature to tune that will reduce the probability
- We can come up with real-world action plans that can effectively tune the feature values.

Let’s approach the problem in 3 steps 1/3: User segmentation

- Recall: three processes
  - Group 1: 1-month old clients
  - Group 2: 2-to-9-month old
  - Group 3: 10-or-more-month
- Model predicts churn probability
  - Group A: 0.00 >= p > 0.25 : Safe
  - Group B: 0.25 >= p > 0.50 : Low
  - Group C: 0.50 >= p > 0.75 : High
  - Group D: 0.75 >= p >= 1.00 : Danger
- Totaling 3 x 4 = 12 groups

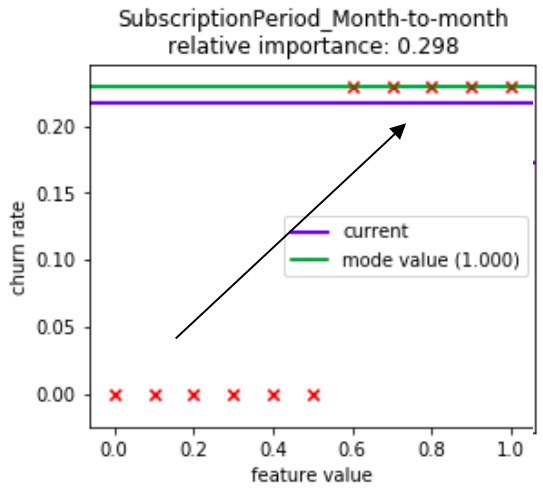


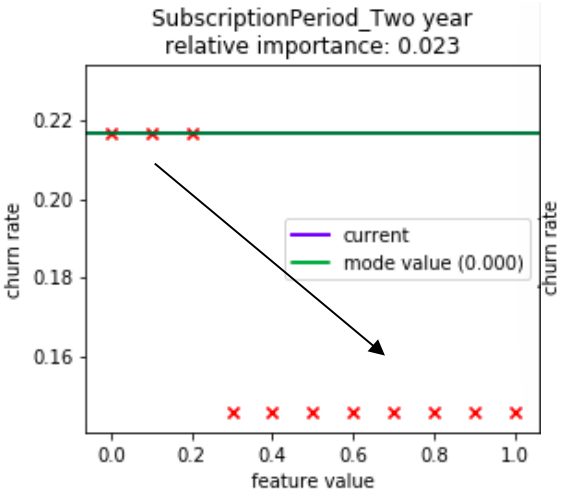
How to reduce Churn Rate?
2/3: Compare averages of each feature

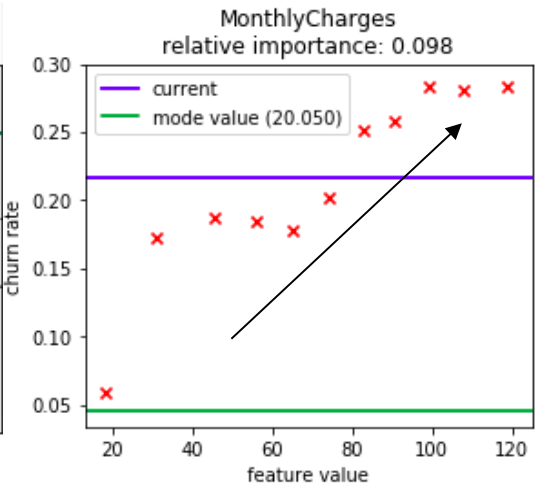


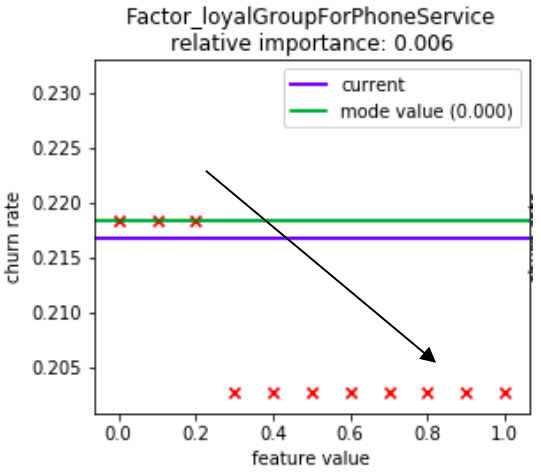
How to reduce Churn Rate?
3/3: Zoom-in each feature

- Hold other features constant, change the value of one feature for all clients



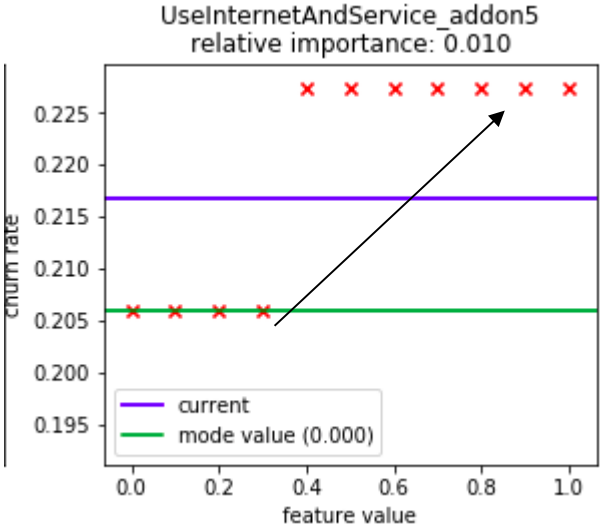


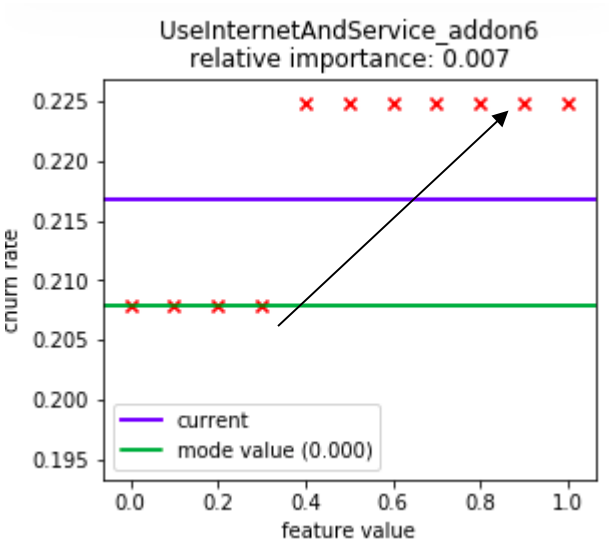




Action: Cheaper long-term plan

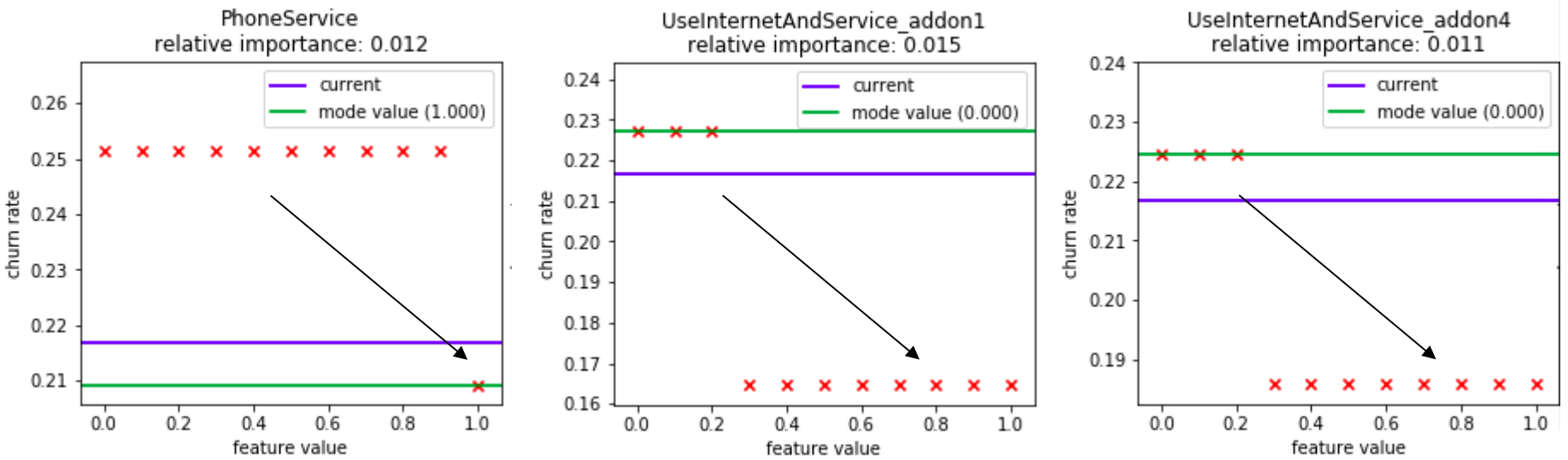
How to reduce Churn Rate?
3/3: Zoom-in each feature





Action: Cheaper long-term plan, particularly avoid Internet service addon 5 and 6 (unless reason for boosting churn rate is found, fixed, verified)

How to reduce Churn Rate?
3/3: Zoom-in each feature



Action: Cheaper long-term plan, particularly avoid Internet service addon 5 and 6 (unless reason for boosting churn rate is found and fixed), however, the company could consider selling more Internet service addon 1,4 and PhoneService to maintain or increase the revenue, as their users stay longer.

=> Can proceed to some AB tests

Summary

- Churn chance of 7032 clients with lifetime from 0-72 months is modeled on their 32 features
- A lightGBM gradient-boosted decision tree model is determined to be the best with accuracy of 80.24%
- Model’s response to the change of features’ values is examined and used as a handle to suggest an action plan to decrease the churn rate
- Model’s dependence on the features is expressed as feature importance