

Weight Loss Prediction

Wenying Gu
Yilin Yang



Outline

Introduction

Challenge

Feature Engineering

Model

Next Steps

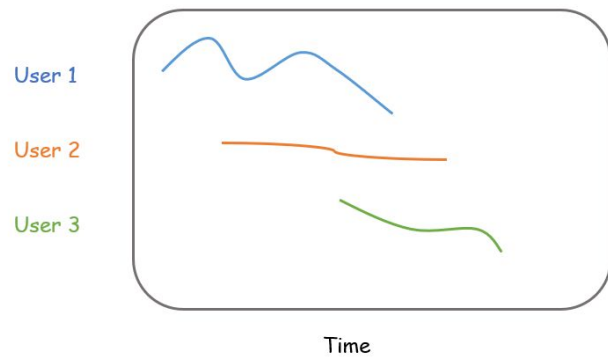


BooHee App

01

- ❑ Report weight related data
- ❑ Track calorie balance
- ❑ Purchase calorie control products
- ❑ Share and receive information from others

Introduction



Challenges



Challenges

02

10 millions user data
causes memory issue

Many null value

How to define loss

Challenges - Memory Issue

Subsample Data By Its Weight Record Duration

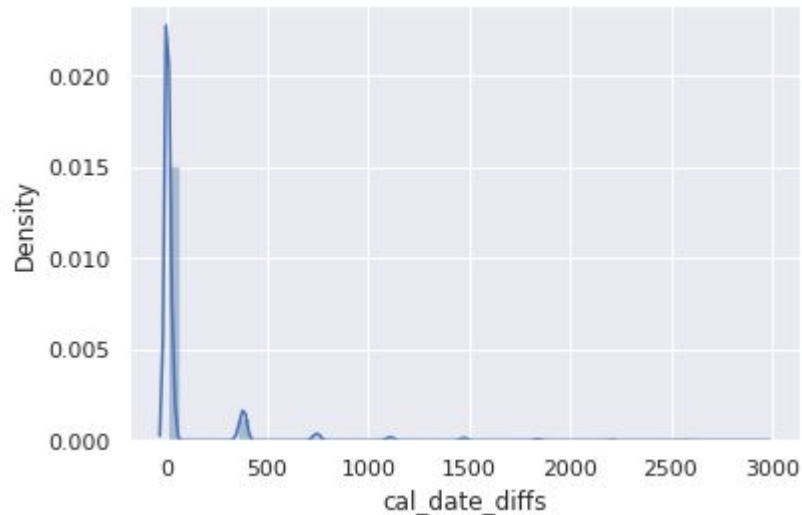
Duration Of A Particular User: last weight record date- first weight record date

User with record: **3,952,879**

User's record duration less than 10 days: **(70.7%)**

User's record duration greater than 400 days: **(3.5%)**

subsample the data by filtering duration days between 300-400 and finally get **298,497(7.6%) users**





Challenges - NA

Fill the NAs with means

- Group By Gender and Age
(Age Range: 0-10, 10-20, 20-30, 30-40, 40-55, 55-70, 70+)
- Group By Age
- No Group



Challenges - Define loss

Training

- Final weight loss = weight record at 250 day - weight record at 1 day
- Users 1 to 200,000 70%

Validation

- Mid weight loss = weight record at last day - weight record at first day
- Users 1 to 200,000 70%

Test

- Final weight loss = weight record at last day - weight record at first day
- Users 200,000 to 300,000 30%

Feature Engineering



Feature Engineering

03

Calorie intake
Daily Comments
Daily Post
Spend per order
Order frequency

General findings after FE



Feature Engineering

For Each User

Calorie intake - $\text{sum}(\text{eating calories} - \text{activity calories}) / \text{number of records}$

Daily Comments - $\text{unique comments} / \text{duration}$

Daily Post - $\text{unique post} / \text{duration}$

Spend per order - $\text{sum}(\text{money}) / \text{order times}$

Order frequency

Feature Engineering - Description Graph

Female: Male: Neutral = 27:2:1

Girls on the BooHee app are more engaged

Overall, people lose weight than gain weight

Among those who clocked in for about 400 days, men lost more weight than women

Age group 31 to 40 is more likely to lose weight regardless gender

It's harder to lose weight as people get older- over 40 since their metabolism is decreasing.



Graph Convolutional Networks

04

Feature Selected:

gender, weight, AgeGroup, bmi
AvgPost, AvgComment, AvgMention,
CaloryIntake, DailyPost, DailyComment,
SpendPerOrder, OrderCount

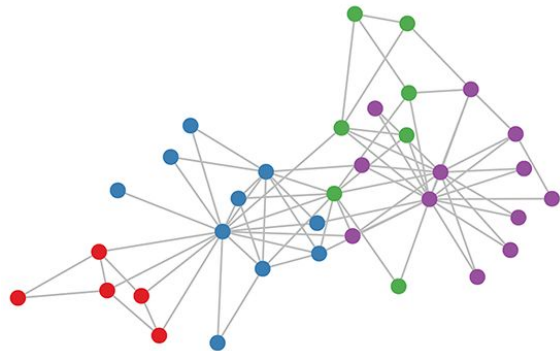
Following Network

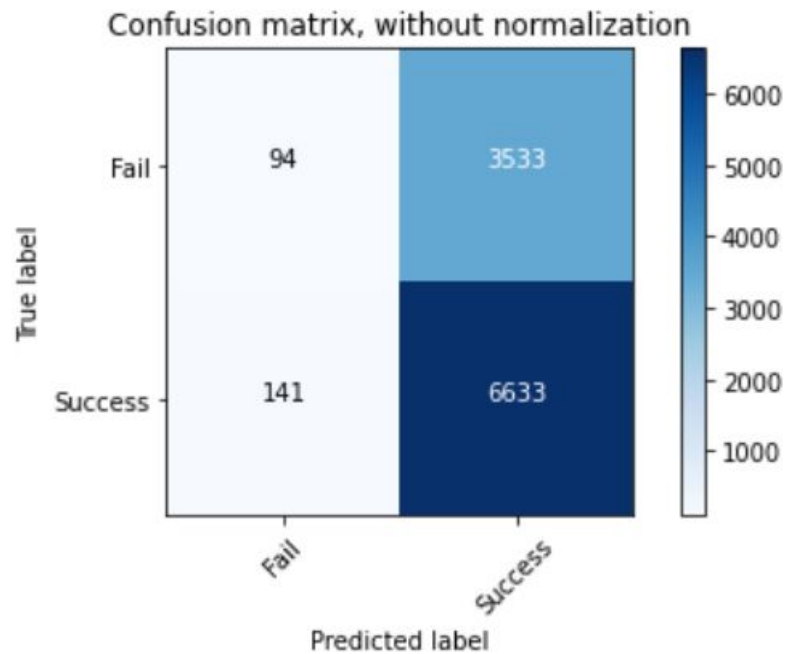
Remove all users with degree < 5

Summarized in a $N \times D$ feature matrix

Hyperparameter tuning

Accuracy: 65%







Next Step

05

Time Series Analysis

Feature Analysis

Thank you for your time !

Q&A