**raw action series**

# base on single enroll id

## group by course

## group by user

## static resample

### 

* weekdays
* chapter
* event

# attention ratio

# gap

## inside series

### frequency in head and tail

## outside series

### gap between course start/close time and 1st/last action

## 概要

# e : enrollmentc : courseu : user

# raw file

## info

### course info

* Index : course id
* course type
* course categories
* duration

### user info

* Index : user id
* gender
* age
* degree

## label

### test label

### trian label

## log

### test log

### train log

### columns in raw log file

* Index

• enroll id

• user id

• course id

* Feature

• time

• action

• object

• session

# action dict

## enroll 0

### time 0

* Id

• user id

• course id

* Feature

• action

• object

• session

### time 1

* Id

• user id

• course id

* Feature

• action

• object

• session

### time n

## enroll 1

### time 0

* Id

• user id

• course id

* Feature

• action

• object

• session

### time 1

* Id

• user id

• course id

* Feature

• action

• object

• session

### time n

## enroll n

### time 0

* Id

• user id

• course id

* Feature

• action

• object

• session

### time 1

* Id

• user id

• course id

* Feature

• action

• object

• session

### time n

# User behavior statics dataset

## time interval

### static value for

* long intervel <1800 sec
* short interval >1800 sec
* mean,var,rs,ku

### between 1st/last action and course start/end time

* droped

## attention

### user

### course

## state transition

### video

### problem

### courseware

# id correspond table

## enroll id dict

### user id

### course id

## user id dict

### enroll id

### course id

## course id dict

### enroll id

### course id

# query

# useful info dataset

# assemble

# group by course id

## course i

### Index

* user id
* enroll id

### Time

### Feature

* action
* object
* session

# group by user id

## user i

### Index

* course id
* enroll id

### Time

### Feature

* action
* object
* session

# final feature

## gender

## age

## degree

## course type

## course categories

## duration

## time interval

### static value for

* long intervel <1800 sec
* short interval >1800 sec
* mean,var,rs,ku

### between 1st/last action and course start/end time

* droped

## attention

### user

### course

## state transition

### video

### problem

### courseware

## user info

### finish

## course info

### finish

# info

## course info : 6411 items

### Index : course id

### course type

### course categories

### duration

## user info : 9627149 items

### Index : user id

### gender

### age

### degree

### 子主题 5

# split test and train

## test : 67699 items

## train: 157943 items

# User behavior statics dataset

## time interval

### static value for

* long intervel <1800 sec
* short interval >1800 sec
* mean,var,rs,ku

### between 1st/last action and course start/end time

* droped

## attention

### user

### course

## state transition

### video

### problem

### courseware