

Course Sequence Analysis

Tim Ransom

Randi Sims

2023-01-27

Contents

1	midfield	5
1.1	Exploring the tables	5
1.2	collection of utility functions	5
1.3	Pulling student course sequences	5
1.4	Visualizing a student's course sequence	5
2	modeling course sequences	7
2.1	Random forest model	7
3	model visualization	9

Chapter 1

midfield

1.1 Exploring the tables

1.2 collection of utility functions

```
did_student_graduate <- function(mcid) {  
  return(degree %>% filter(mcid == mcid) %>% nrow() > 0)  
}
```

1.3 Pulling student course sequences

```
# convert to tibble  
course <- tibble(course) %>% select(mcid, abbrev, number, term_course) %>% nest_by(mcid)
```

1.4 Visualizing a student's course sequence

TODO: visualize a single students path to graduation

Chapter 2

modeling course sequences

We've already gotten our course sequences, lets use them to train a model!

2.1 Random forest model

```
library(caret)

## Loading required package: lattice
##
## Attaching package: 'caret'
## The following object is masked from 'package:purrr':
##
##      lift
# https://topepo.github.io/caret/train-models-by-tag.html#random-forest

https://stackoverflow.com/questions/57939453/building-a-randomforest-with-caret
```


Chapter 3

model visualization

good science comm uses visuals