Overview of DFE Run

Automated

December 22, 2022

Abstract

This is an automatically generated report. It is meant ot illustrate the kinds of analysis that can be performed in LightR.

1 Overview

- The report was generated on Jan 1, 2023.
- The number of data sets was 42.
- Summary statistics of number of rows in data set:
  - Minimum 1712
  - Maximum 3252
  - Average 2398.380952381
- The number of formulas per model was 5. Explanation of the formulas is in Table 4.

2 Basic Feature Engineering

- The number of times this was invoked was 42
- The number of errors reported was 0.
- The average time taken (in μseconds) was 51653.857142857

Distribution of occurrence of error codes is shown in Table 1
Explanation of error codes is in Table 6.1

Error Code	Count
0	42
1	0
2	0
3	0

Table 1: Error Counts for basic feature engineering

Formula	Attempts	Successes
0	42	42
1	42	42
2	42	42
3	42	42
4	42	42

Table 2: Error Counts for formula specific feature engineering

3 Formula Specific Feature Engineering

- The number of times this was invoked was 42.
- The number of data sets that had at least one error was 0.
- The average time taken (in μseconds) was 43631.80952381

Distribution of occurrence of error codes (by formula) is shown in Table 2
Explanation of error codes is in Table 6.2
Distribution of errors by formula and type is shown in Table 3

4 Model Building

- The number of data sets for which model building was attempted was X
- The number of models attempted was X
- The number of models that were built was X
- Summary statistics of time (in seconds) to build a model
  - Minimum X
  - Maximum X
  - Average X
- Distribution of build times is in Figure 1

Formula	Error Code	Count
0	0	42
0	1	0
0	2	0
0	3	0
0	4	0
1	0	42
1	1	0
1	2	0
1	3	0
1	4	0
2	0	42
2	1	0
2	2	0
2	3	0
2	4	0
3	0	42
3	1	0
3	2	0
3	3	0
3	4	0
4	0	42
4	1	0
4	2	0
4	3	0
4	4	0

Table 3: Error Counts, broken down by formula and type

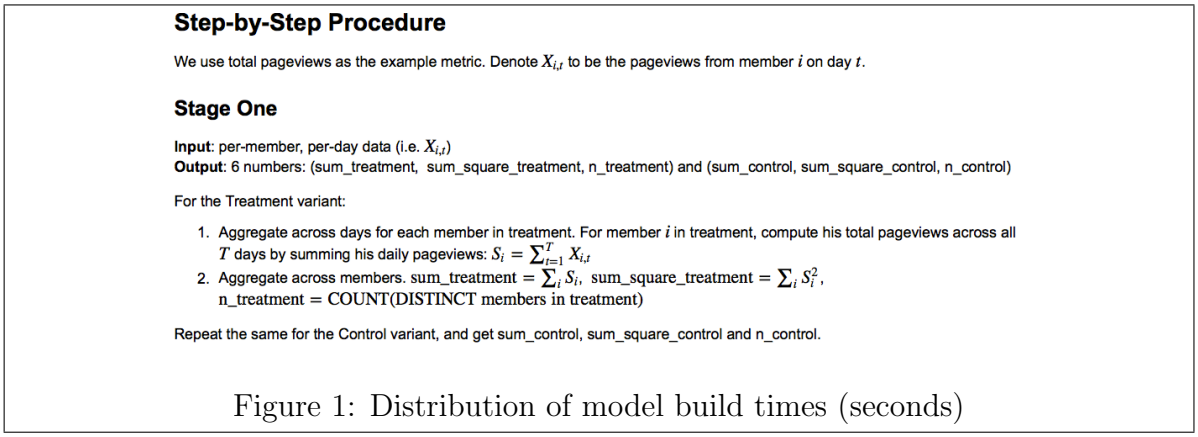

Figure 1: Distribution of model build times (seconds)

Index	Key Explanation	
0	f0	Basic formulas
1	f1	Uses 2 lag components (week 1 and week 2)
2	f2	Uses 2 lag components (week 2 and week 3)
3	f3	Uses 2 lag components (week 3 and week 4)
4	f4	Uses 2 lag components (week 4 and week 5)

Table 4: List of Formulas

Error Code	Explanation
0	No Error
1	insufficient rows in input data frame
2	insufficient rows in input data frame after cleaning
3	insufficient unique values in **toy** component

Table 5: Explanation of Error Codes

5 Explanations of Terms

5.1 Formulas

6 Error Codes

6.1 Basic Feature Engineering

See Table 6.1

6.2 Formula Specific Feature Engineering

See Table 6.2

Error Code	Explanation
0	No Error
1	not enough rows after cleaning
2	Range of **toy** component too small
3	too few uniques in **toy** component
4	too few uniques in lag component

Table 6: Explanation of Error Codes