← BACK TO GRADES

Homework Assignment 3 (week 3)

Faculty Feedback

Score Last published: 8/4/2019 6:36 PM PDT

97 / 100 (97.00%)

Comments

Week 3 Assignment Rubric

Notes: Overall good job. See specific notes below.

Introduction –

Well done! Good job defining the problem and motivating the solution.

Data / Analysis / Models –

Great work with your EDA. You investigated and explored the nature and character of the data and it really brought insight to your subsequent investigation and research!

Results –

Great work with your investigation of the top, most interesting rules! BUT MORE explanations and supporting discussion.

I would encourage you to use visualizations / tables to succinctly describe / summarize results and to support your discussion of the results.

Conclusions –

Conclusions are a bit slim, but hit on the high points.

The following is also a more general review with lots of help and pointers.

Required Elements:

- All Assignments must be written at the graduate academic level.
- R code is required. You have submitted your R – thank you!
- Requirements Noted in the Assignment:
- First perform the necessary **preprocessing steps** required for association rule mining, **specifically the id field needs to be removed and a number of numeric fields need discretization or otherwise converted to nominal.**
- Next perform association rule discovery on the preprocessed data.
- Experiment with different parameters and preprocessing so that you get on the order of 20-30 strong rules, e.g. rules with high lift and confidence which at the same time have relatively good support.
- Don’t forget to report in details what you have tried.
- Finally, **set PEP as the right hand side of the rules**, and see what rules are generated.
- Select the top 5 most “interesting” rules and for each specify the following:
- • Support, Confidence and Lift values
- • An explanation of the pattern and why you believe it is interesting based on the business objectives of the company.
- • Any recommendations based on the discovered rule that might help the company to better understand behavior of its customers or to develop a business opportunity.

- Note that the top 5 most interesting rules are most likely not the top 5 in the strong rules. They are rules, that in addition to having high lift and confidence, also provide some nontrivial, actionable knowledge based on underlying business objectives.
-
- **Headings Required:**
- Introduction.
- The Introduction should properly describe the topic of interest. It should offer good background about the topic, why the topic is important and/or who it is important for, and perhaps what type of research has been performed on this area (briefly).
- No information about the data, variables, types, cleaning, or prep should be in the Intro.
- Analysis
- Subsection 1: Data
- Here you include all the information about the data.
- Include visualizations you used to initially explore and clean the data
- Note all data cleaning and preparation, including discretization (categorization), normalization, etc.
- When you clean, prepare, and explore a dataset, look at each variable, make sure there are no missing or incorrect values, address any outliers, correct data types, etc.
- Subsection 2: Analysis
- Here you will perform association rule mining in a few different ways – such as constraining the RHS and LHS, etc.
- There are many options and tuning which should all appear in this section.
- Results
- Include the top rules for each of the different analyses you performed and talk about what each implies.
- This is a technical area, so you can and should also discuss tuning, errors, different options for conf and sup, etc.
- Conclusions

- This is a non-technical area that explains in clear terms (for anyone) what you found, why it matters, and what it can be used for.

Basis For Grades:

100: This means that your Assignment was amazing. It covered everything – cleaning, prep – analysis that makes sense – visualizations – results (that are true) – etc. There is nothing really left to improve.

95: This means that your Assignment is really good! You covered most of the items noted above and perhaps a few others not noted. You can make some improvement on pre-processing and results analysis, as well as perhaps other visualizations. Overall – you have the idea and you did well.

90: This means that your Assignment is good, but could be a little better. Perhaps add items such as further data cleaning and pre-processing, data normalization, better or more visualizations, and/or more robust conclusions.

85: This means that your Assignment is a good start and largely meets the more general and overall requirements

Below 85 means that the level of 85 above was not quite met and many elements were missing.

Student Submission | Homework Assignment 3 (week 3)

Response

Last submitted: 7/24/2019 5:29 PM PDT

No response

Files | Download all (2)

File Name	Uploaded	Feedback
Ryan_Timbrook_HW3.docx	7/24/2019 5:29 PM PDT	
ist707_week3_hw.R	7/24/2019 5:29 PM PDT	