

NEW MOVIELENS SUBMISSION

Status

This assignment is in progress. You still need to complete the peer assessment step.



Your Response due Jan 8, 2021 13:59 HST (in 3 weeks, 3 days)



COMPLETE



Assess Peers due Jan 15, 2021 13:59 HST (in 1 month)

IN PROGRESS (1 OF

5)

Read and assess the following response from one of your peers.

The question for this section

Your submission for this project is three files:

1. Your report in PDF format
2. Your report in Rmd format
3. A script in R format that generates your predicted movie ratings and RMSE score

You may upload the three files directly to the edX platform or submit a GitHub link in the text response box below.

To upload and submit your files press the "Choose Files" button, select three files at once (using the control key on a Windows machine or command key on a Mac) and press "Choose," type a description for each (PDF, Rmd, R), and then press the "Upload files" button. If uploading files, we recommend also providing a link to a GitHub repository containing the three files above in case there is a problem with the upload process.

Note that when downloading files for peer assessments, R and Rmd files will be downloaded as txt files by default.

Your peer's response to the prompt above

The MovieLens project has used the given training set (edx) and validation set and examined various features effects to estimate the best suited model for user rating prediction.

From the detail model analysis it has been observed that regularization brings down the RMSE (root mean squared error) value to 0.8648170.

The optimal RMSE value shows that the Linear regression model with regularized effects with users and movie combined is the most appropriate model to predict the user ratings on the validation set.

Associated Files

Files that were uploaded by you:

[Project Report documentation \(_Movielense_final_report_12122020.pdf \)](#)

[MovieLense RMD file code and output \(MOVIELENS_RMD.Rmd \)](#)

[R script for all data models and graphs plot \(Movielense_RMSE_Projection.R \)](#)

Caution: These files were uploaded by another course learner and have not been verified, screened, approved, reviewed, or endorsed by the site administrator. If you access the files, you do so at your own risk.)

- ▼ **The appropriate files are submitted in the correct formats: a report in both PDF and Rmd format and an R script in either R or Rmd format.**
-

☐ **No files**

No files provided AND/OR the files provided appear to violate the edX Honor Code.

0 POINTS

☐ **Many missing**

Multiple requested files are missing and/or not in the correct formats.

3 POINTS

☐ **One missing**

One file is missing and/or not in the correct format.

5 POINTS

☒ **All correct**

All 3 files were submitted in the requested formats.

10 POINTS

Comments

Preview in LaTeX

CLICK TO PREVIEW YOUR SUBMISSION IN LATEX.

▼ **The report documents the analysis and presents the findings, along with supporting statistics and figures. The report must be written in English and uploaded. The report must include the RMSE generated. The report must include at least the following sections:**

- 1. an introduction/overview/executive summary section that describes the dataset and summarizes the goal of the project and key steps that were performed**
- 2. a methods/analysis section that explains the process and techniques used, including data cleaning, data exploration and visualization, insights gained, and your modeling approach**
- 3. a results section that presents the modeling results and discusses the model performance**
- 4. a conclusion section that gives a brief summary of the report, its limitations, and future work**

☐ **Poor**

The report is either not uploaded or contains very minimal information AND/OR the report appears to violate the edX Honor Code.

0 POINTS

☐ **Many missing sections**

Multiple required sections of the report are missing.

10 POINTS

☐ **Methods/analysis or results missing**

The methods/analysis or the results section of the report is missing or missing significant supporting details. Other sections of the report are present.

15 POINTS

☐ **Intro/overview or conclusion missing**

The introduction/overview or the conclusion section of the report is missing, not well-presented or not consistent with the content.

20 POINTS

☐ **Very hard to follow**

The report includes all required sections, but the report is significantly difficult to follow or missing supporting detail in multiple sections.

20 POINTS

☐ **Hard to follow**

The report includes all required sections, but the report is difficult to follow or missing supporting detail in one section.

25 POINTS

☐ **Minor flaws in multiple sections**

The report includes all required sections and is well-drafted and easy to follow, but with minor flaws in multiple sections.

30 POINTS

☐ **Minor flaws in one section**

The report includes all required sections and is easy to follow, but with minor flaws in one section.

35 POINTS

☐ **Easy to follow**

The report includes all required sections, is easy to follow with good supporting detail throughout, and is insightful and innovative.

40 POINTS

Comments

Preview in LaTeX

CLICK TO PREVIEW YOUR SUBMISSION IN LATEX.

▼ **The code in the R script should be well-commented and easy to follow. You are not required to run the code provided (although you may if you wish), but you should visually inspect it. The code must go beyond simply copying code from previous courses in the series.**

☐ **No code**

No code provided AND/OR the code appears to violate the edX Honor Code.

0 POINTS

☐ **Hard to interpret**

Code appears that it would not run and/or is very difficult to follow or interpret.

10 POINTS

☐ **No comments**

Code appears that it would run without throwing errors, can be followed, is at least mostly consistent with the report, but has no comments or explanation.

15 POINTS

☐ **Insufficient comments**

Code appears that it would run without throwing errors, can be followed, but without sufficient comments or explanations.

20 POINTS

☐ **Easy to follow and well-commented**

Code is easy to follow, is consistent with the report, and is well-commented.

25 POINTS

☐ **Code copies provided course code**

Code simply copies from previous code in the course series without further developing it.

15 POINTS

Comments

Preview in LaTeX

CLICK TO PREVIEW YOUR SUBMISSION IN LATEX.

▼ Provide the appropriate score given the reported RMSE. Please be sure not to use the validation set (the final hold-out test set) for training or regularization - you may wish to create an additional partition of training and test sets from the provided edx dataset to experiment with multiple parameters or use cross-validation.

☐ **No RMSE**

No RMSE reported AND/OR code used to generate the RMSE appears to violate the edX Honor Code.

0 POINTS

☐ **>= 0.90000**

RMSE >= 0.90000 AND/OR the reported RMSE is the result of overtraining (validation set - the final hold-out test set - ratings used for anything except reporting the final RMSE value) AND/OR the reported RMSE is the result of copying provided code from other courses in the series.

5 POINTS

☐ **0.86550 to 0.89999**

The RMSE is greater than or equal to 0.86550 and less than or equal to 0.89999.

10 POINTS

☐ **0.86500 to 0.86549**

The RMSE is greater than or equal to 0.86500 and less than or equal to 0.86549.

15 POINTS

☐ **0.86490 to 0.86499**

The RMSE is greater than or equal to 0.86490 and less than or equal to 0.86499.

20 POINTS

☐ **Less than 0.8649**

The RMSE is less than 0.86490.

25 POINTS

Comments

[Preview in LaTeX](#)

CLICK TO PREVIEW YOUR SUBMISSION IN LATEX.

(Optional) What aspects of this report and algorithm stood out to you? What did it do well? How could it be improved?

I think that this response...

[Preview in LaTeX](#)

CLICK TO PREVIEW YOUR SUBMISSION IN LATEX.

Submit your assessment and move to response #2

▸ Your Grade: Not Completed

MovieLens Grading Rubric

The following is the grading rubric your peers will be using to evaluate your project. There are also opportunities for your peers to provide written feedback as well (required for some categories and optional for others). You are encouraged to give thoughtful, specific written feedback to your peers whenever possible (i.e., more than just "good job" or "not enough detail").

Note that to receive full marks on this project, you **may not simply copy code from other courses in the course series and be done with your analysis**. Your work on this project needs to build on code that is already provided.

After you submit your project, **please check immediately after submitting to make sure that all files were correctly uploaded.** Occasionally, there are file upload failures, and it's easiest to fix if these are caught early.

Files (10 points possible)

The appropriate files are submitted in the correct formats: a report in both PDF and Rmd format and an R script in R format.

- 0 points: No files provided AND/OR the files provided appear to violate the edX Honor Code.
- 3 points: Multiple requested files are missing and/or not in the correct formats.
- 5 points: One file is missing and/or not in the correct format.
- 10 points: All 3 files were submitted in the requested formats.

Report (40 points possible)

The report documents the analysis and presents the findings, along with supporting statistics and figures. The report must be written in English and uploaded. The report must include the RMSE generated. The report must include at least the following sections:

1. an **introduction/overview/executive summary** section that describes the dataset and summarizes the goal of the project and key steps that were performed
 2. a **methods/analysis** section that explains the process and techniques used, including data cleaning, data exploration and visualization, insights gained, and your modeling approach
 3. a **results** section that presents the modeling results and discusses the model performance
 4. a **conclusion** section that gives a brief summary of the report, its limitations and future work
- 0 points: The report is either not uploaded or contains very minimal information AND/OR the report appears to violate the edX Honor Code.
 - 10 points: Multiple required sections of the report are missing.
 - 15 points: The methods/analysis or the results section of the report is missing or missing significant supporting details. Other sections of the report are present.

- 20 points: The introduction/overview or the conclusion section of the report is missing, not well-presented or not consistent with the content.
- 20 points: The report includes all required sections, but the report is significantly difficult to follow or missing supporting detail in multiple sections.
- 25 points: The report includes all required sections, but the report is difficult to follow or missing supporting detail in one section.
- 30 points: The report includes all required sections and is well-drafted and easy to follow, but with minor flaws in multiple sections.
- 35 points: The report includes all required sections and is easy to follow, but with minor flaws in one section.
- 40 points: The report includes all required sections, is easy to follow with good supporting detail throughout, and is insightful and innovative.

Code (25 points)

The code in the R script should be well-commented and easy to follow. You are **not required** to run the code provided (although you may if you wish), but you should visually inspect it.

- 0 points: No code provided AND/OR the code appears to violate the edX Honor Code.
- 10 points: Code appears that it would not run/is very difficult to follow or interpret.
- 15 points: Code appears that it would run without throwing errors, can be followed, is at least mostly consistent with the report, but has no comments or explanation.
- 15 points: Code is simply a copy of code provided in previous courses in the series without expanding on it, but is otherwise well-commented.
- 20 points: Code appears that it would run without throwing errors, can be followed, but without sufficient comments or explanations.
- 25 points: Code is easy to follow, is consistent with the report, and is well-commented.

RMSE (25 points)

Provide the appropriate score given the reported RMSE. *Please be sure not to use the validation set (the final hold-out test set) for training or regularization - you should create an additional partition of training and test sets from the provided edX dataset to experiment with multiple parameters or use cross-validation.*

- 0 points: No RMSE reported AND/OR code used to generate the RMSE appears to violate the edX Honor Code.
- 5 points: RMSE ≥ 0.90000 AND/OR the reported RMSE is the result of overtraining (validation set - the final hold-out test set - ratings used for **anything** except reporting the final RMSE value) AND/OR the reported RMSE is the result of simply copying and running code provided in previous courses in the series.
- 10 points: $0.86550 \leq \text{RMSE} \leq 0.89999$
- 15 points: $0.86500 \leq \text{RMSE} \leq 0.86549$
- 20 points: $0.86490 \leq \text{RMSE} \leq 0.86499$
- 25 points: $\text{RMSE} < 0.86490$

Have a question about the MovieLens project? Need some feedback on the best approach to take or some troubleshooting for a snippet of your code? You can ask your questions here!

You are encouraged to discuss **general approaches** to the MovieLens project. It is okay to post **small snippets** of code if you're having trouble getting a particular piece of code to run. However, you **may not post your entire R script** for the project.

Discussion: MovieLens Project

[Hide Discussion](#)

Topic: MovieLens Project / MovieLens Project

Add a Post

Show all posts
by recent activity

MovieLens project - what is the project???
6

I don't understand, where is the project outline. What are we supposed to do? What am I missin...

Staff I need help - Report in Rmd & Script in R format
4

Hello! I have been working on this project for a while (almost 2 months!). To run each query took...

calculating an effect twice
3

Hello, I initially used regularized movie, user, genre and year effect. Then I attempted to implem...

R Scrip code for submission
2

	<u>Dear All, Regarding this project submission, can someone please suggest what extension is typi...</u>	3
💬	<u>"Your Response" field - what are we supposed to write here?</u> <u>What is it we are supposed to respond to? What is this field?</u>	3
✓	<u>Script and parameters</u> <u>There are two different issues I need some help with. My first question is really simple: is it nece...</u>	4
?	<u>Confirm understanding of bootstrap crossvalidation.</u> <u>To test my understanding of cross validation, I tried both k-fold and bootstrap methods to optim...</u>	1
✓	<u>STAFF! Again! Error message: cannot allocate vector of size 4.9 Gb</u> <u>I have tried to reduce my data to 1/10th of edx to enable me train and move but i keep getting th...</u>	7
?	<u>Linebreak in codechunk</u>	2
?	<u>Algorithms crashing R</u> <u>Hi, I'm working on movielens project and besides those technics presented in course 8 recomm...</u>	1
💬	<u>how to reference data in rmarkdown</u> <u>when generating my report, I keep getting errors about not being able to find an object sum...</u>	3
?	<u>I get a different RMSE when I calculate Movie effects then User Effect compared to User Effect then Movie effects</u>	1
?	<u>Did not add RMD file by mistake.</u> <u>Hello team, I am sorry to take your time for this one. However, I just clicked submit without inclu...</u>	1