# Task 2 submission: Implementation

This is your submission document for [DataCapX Submission 3: Implementation](https://courses.edx.org/courses/course-v1:AdelaideX+DataCapX+3T2018/courseware/a7d3a888e3aa4b2689c5421bc4550619/c8d2611b443049c392c364d5bbbf492b/1?activate_block_id=block-v1%3AAdelaideX%2BDataCapX%2B3T2018%2Btype%40vertical%2Bblock%4022509f00a2814b41882462af3f6ddd3d).   
Save this document on your local machine and include all of your work within the relevant part of the assignment. Once you’ve completed every part of Task 2: Implementation, upload this document via the [Your Response area](https://courses.edx.org/courses/course-v1:AdelaideX+DataCapX+3T2018/courseware/a7d3a888e3aa4b2689c5421bc4550619/c8d2611b443049c392c364d5bbbf492b/1?activate_block_id=block-v1%3AAdelaideX%2BDataCapX%2B3T2018%2Btype%40vertical%2Bblock%4022509f00a2814b41882462af3f6ddd3d).

**Note: Don’t forget to save your scripts related to this task and submit them when you submit this document.**

# Checklist

* Have you answered every question?
* Have you shown all of your working, including evidence of your code?
* Have you included all R output as script (.r) files?
* Have you clearly stated conclusions where required?
* Have you saved your code in a script?

# Task 2 submission: Implementation

Prepare the dataset and build a classifier

1. Read in the reddit dataset correctly and provide evidence of your code.   
   [3 points]

Include the name/s of any associated code (.r) script file/s related to Step 1 that are included in your .zip file:

1. Build a classifier based on the six factors and provide the classifier information. Provide the correct analysis about the factors.   
   [3 points]

Include the name/s of any associated code (.r) script file/s related to Step 2 that are included in your .zip file:

1. Build a classifier based on the six factors and the additional factor. Provide the learned classifier information. Provide the correct comparison analysis between the built model and the model built in the previous step.   
   [2 points]

Include the name/s of any associated code (.r) script file/s related to Step 3 that are included in your .zip file:

**Model/feature selection**

1. Perform model selection and produce the correct selection results. Provide your result and associated code.   
   [4 points]

Include the name/s of any associated code (.r) script file/s related to Step 1 that are included in your .zip file:

1. Perform model selection using another method and produce the correct selection results. Provide your result and associated code  
   [4 points]

Include the name/s of any associated code (.r) script file/s related to Step 2 that are included in your .zip file:

Verifying your chosen method

1. Provide an appropriate definition of the criterion. Provide code to test the performance of the learned classifier. Compare it with the baseline method. Provide evidence of your code.  
   [5 points]

Include the name/s of any associated code (.r) script file/s related to Step 1 that are included in your .zip file:

1. Perform feature/model selection with the criterion developed in Step 1 of *Verifying your chosen method*. Conclude whether the selection results are the same as those in *Model/feature selection*. Provide your selection results and associated code. Provide your selection results and associated code.   
   [5 points]

Include the name/s of any associated code (.r) script file/s related to Step 2 that are included in your .zip file:

Total points possible for Task 2: Implementation 26