

PRATILIPi DATA SCIENTIST ASSIGNMENT

My understanding from assignment details:

1. The aim of the assignment is to predict the prathilipi's that user gonna read later from his/her subscriptions.

My assumptions:

1. Which prathilipi user will read later?

Ans. The story with higher read percent has highest probability to finish it.

- Considering that 75% is the threshold value
- The stories with above 75 read percent have highest probability to read first
- this threshold value can be variable
- By creating target variable turned it as a classification problem

2. Why I choosen Random Forest model ?

Ans. This is one of the best classifier model . We can choose other classifier models also.

- And I got the best results cause dataset has only 3 features and it is not too complex.

3. Alternatives for this problem?

Ans. I tried hard to find out alternate for this problem ,but I couldn't.

Improvements to this model:

1. We can add more detail to the target column so that it is easy to identify the tuple (Not major improvement).
2. We can use data to create recommendations for other users and also notifications for original users to finish the prathilipi.

NOTE : If you expect something beyond this, please provide some more hints and more details of it. I am ready to do it.