

Project Checklist Template

Project Name: Movie Popularity Prediction

Team Members: Aditya Chinnamalla, Shwetha Krishna , Saurabh Sankhe

Document Date: 02/28/2019

Version: Version 1

Activities Checklist			
#	Item	Yes/No	Week
	Understanding the DataSet		
1	Selection of Dataset	Yes	WEEK 1
2	Finalizing the Dataset	Yes	
3	Understanding the Data Source and its reliability	Yes	
4	Loading the Data Dictionary	Yes	
5	Importing the Data in R	Yes	
6	Git Setup	Yes	
	Defining Questions		
7	Decide what to measure(Output Variables)	Yes	WEEK 2
8	Deciding how to measure it	Yes	
9	Data Summary for better understanding	Yes	
10	Identifying dependent and independent variables	Yes	
	Data Exploration		
11	Finding Missing Data	Yes	WEEK 3
12	Identifying Data Spread	Yes	
13	Look for ways to fix missing Data	Yes	
14	Data Visualization	Yes	
	Anomaly Detection		
15	Identifying and transforming the columns with anomalies	Yes	WEEK 4
16	Correlation Plot,Variance Covariance	Yes	
17	Visualizing the output variable	Yes	
	Model Selection		
18	Performing PCA	No	WEEK 5
19	Logistic Regression	Yes	
20	SVM	Yes	
21	Decision Tree	Yes	

22	Random Forest	Yes	
23	Performing various Tests	Yes	

Models Selection done
Chi-square test remaining