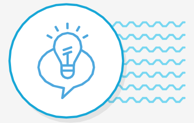
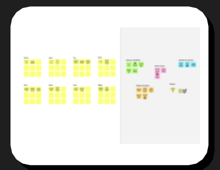


Brainstorm
& idea prioritization

In this Template share ideas and further ideas can be written here to modify accordingly , leader will modify these chart based on mentor feedback.

- 2 months to prepare
- 1 month to collaborate
- 4 Members

Share template feedback



Need some inspiration?

See a finished version of this template to kickstart your work.

Open example

Before we collaborate

We have to make sure wether the IBM management provide us good data , we have to make proper planning , analyzing the problem and learn additional skills like storytelling , stakeholder analysis , etc.

A Team gathering
Devi Priya (team leader) will gather group and instruct , ask
idea and lead the group further.

B Set the goal
- Higher Accuracy.
- Clean Visuals.
- Clean Code.
- More Insights

C Learn how to use the facilitation tools
1. Youtube and IBM sessions to learn concepts.
2. Use documentation to code new concepts.
3. use discord , stackoverflow to clear doubts.

1 Applicant Credibility Prediction for Loan Approval

This data science project will help finance and banking people who give 100's of loan to their applicant and this group project will help stakeholder will come to the number if applicant who are eligible and not eligible by using data visualization , machine learning algorithms and stakeholder will make data driven decisions from this project.

PROBLEM

We are gonna solve this problem by using machine learning algorithms using sci-kit learn and other conventional libraries like spark to handle big data, numpy and pandas for reshaping ,cleaning data,etc.

2 Brainstorm

ideas that come to mind that address your problem statement.

Devi priya
- 1. Get Big data
- 2. Clean values by outlier detection , removing null value by mean/ median
- 3. Remove abnormal data from csv/txt file
- 4. Use Xgboost Regression
- 5.Preprocess data to reduce computation strain
- 6.Evaluate The model.
- 7.Find out which model fits the problem ie. Random forest,svm,logistic regression by scikit learn.

Siva prabhu
- 1.use apache to store big data
- 2.use matplotlib to create clean visuals
- 3. Use Neural Network For this problem.

Maheshwaran
- 1. Use seaborn to visualize data
- 2. Do statistical analysis i.e.Inferential statistics,descriptive statistics,etc
- training model more 3. Use tensorflow for particular use keras for creating layers for neural network

Lokeshwaran
- 1. Try to keep ideas clean and neat
- 2. Do proper Refactoring of code and clean visualization patterns.
- 3. Try to achieve more accuracy by repeated epochs and do parameter tunning.

3 Group ideas

Share ideas and we can make further planning based on mentor feedback.

Devi priya
Use Numpy , pandas , plotly

Devi priya
Use Apache spark to store big data

Devi priya
Use Xgboost for regression

Devi priya
Use aws or azure for model training and deploying model.

Siva prabhu
Use Matplotlib

Siva prabhu
Use Numpy , pandas , Matplotlib

Siva prabhu
Use charts like barchart , piechart , ribbon chart based on data provided

Maheshwaran
Use seaborn for clean visualization , use testing techniques if possible.

Lokeshwaran
Refactor code if possible , use clean visuals and use required libraires to reduce complexity

4 Prioritize

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

20 minutes

Importance
If each of these tasks could get done without any difficulty or cost, which would have the most positive impact?
Devi priya
Use Plotly for interactive graphs or visualization, use xgboost and scikit learn for preprocessing and training model. Use Kaggle to learn code from experienced persons as it is a data science community
Siva prabhu
Use aws , azure to deploy model and training model use seaborn. Use kaggle and github for reference
Lokeshwaran
Clean code ,clean visuals , Higher accuracy.
Maheshwaran
Use seaborn , numpy ,pandas which are commonly used libraries in data science project.
Devi Priya , Sivaprabhu
use Neural network if possible
Feasibility
Regardless of their importance, which tasks are more feasible than others? (Cost, time, effort, complexity, etc.)