Ideation Phase

Brainstorming & Idea Prioritization

Date	28-10-2023
Team ID	Team-591779
Project Name	Walmart Sales Analysis for Retail Industry with Machine Learning
Maximum Marks	4 Marks

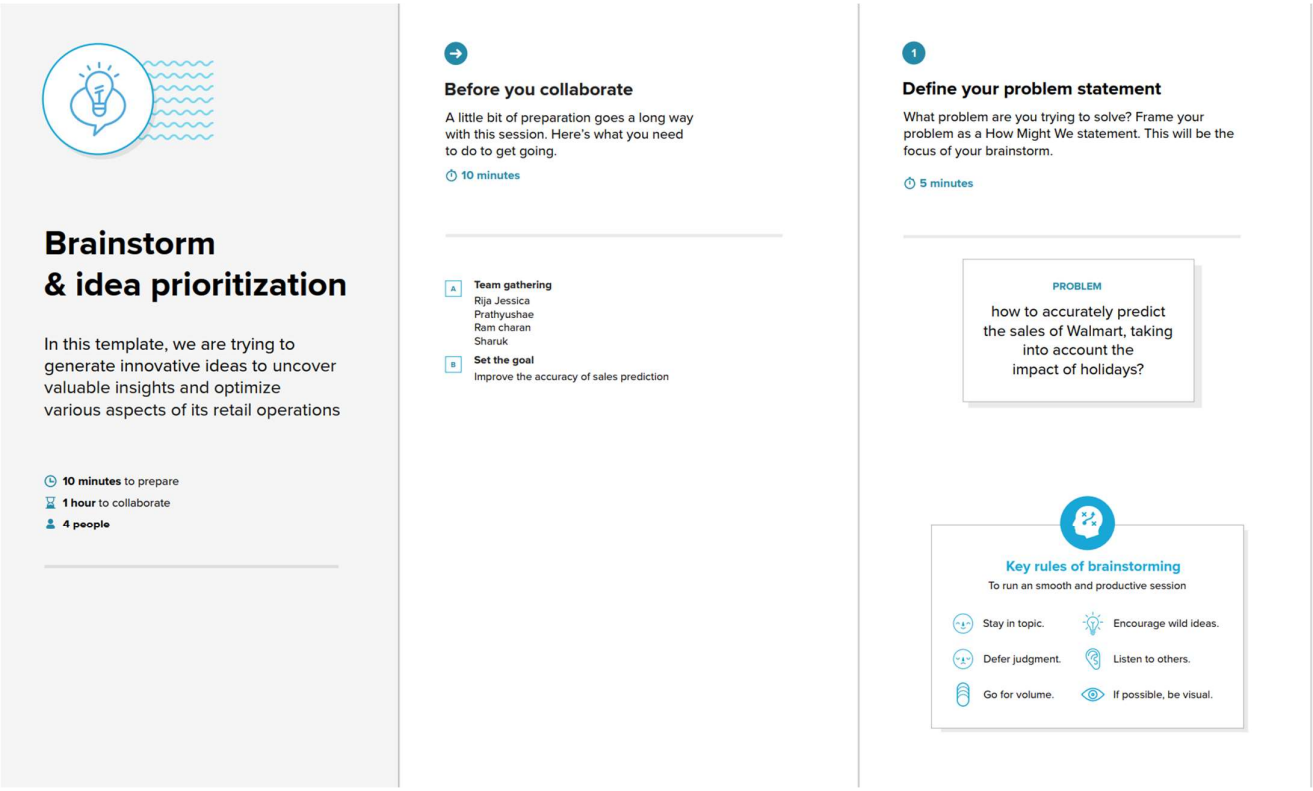
Brainstorming & Idea Prioritization:

The objective of brainstorming for Walmart data could be to generate innovative ideas and insights related to improving sales forecasting, enhancing customer experience, optimizing inventory management, or any other relevant business aspect.

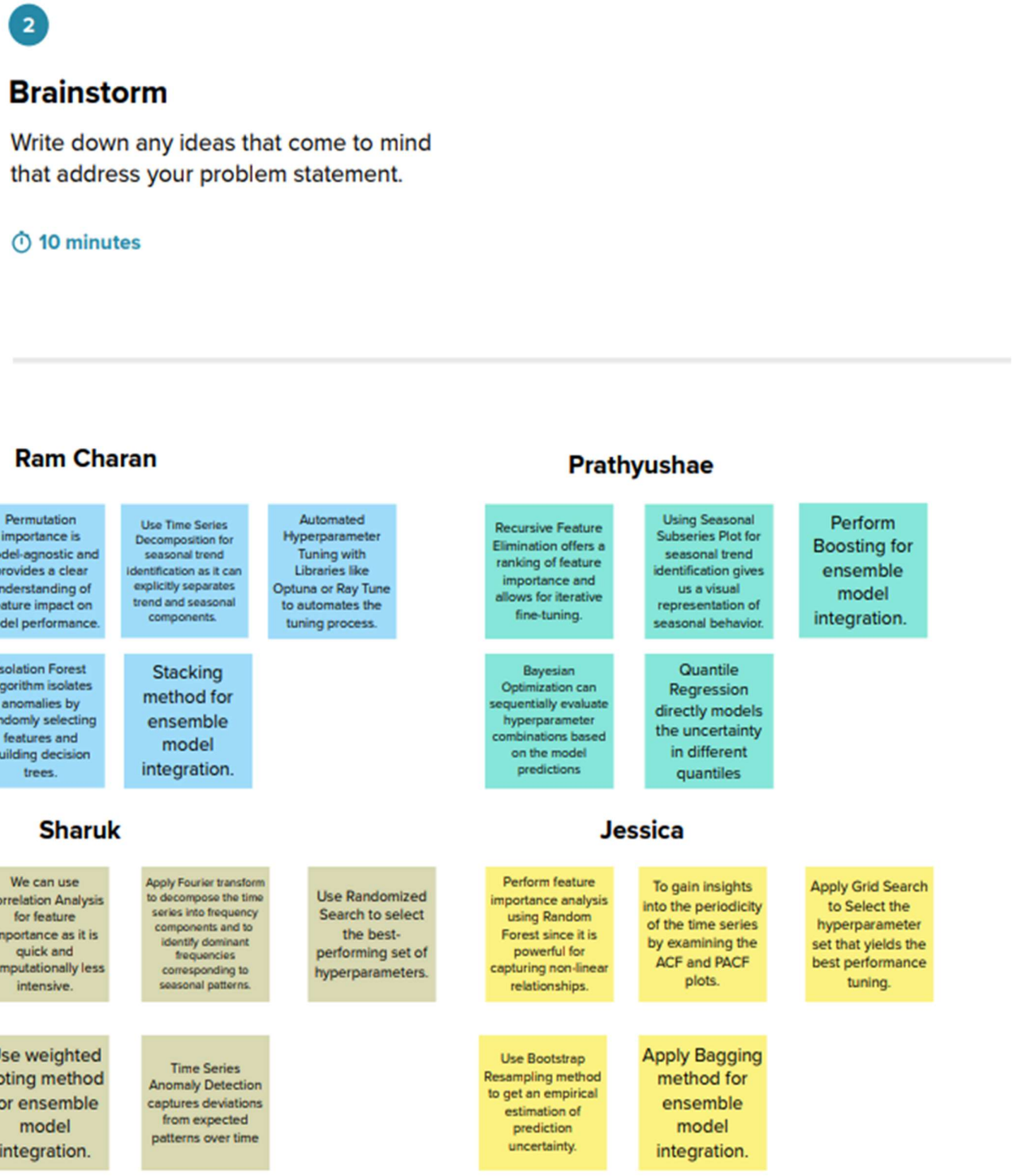
Reference:

https://app.mural.co/t/walmart8615/m/walmart8615/1698904530359/3daad31688cf710edcda862f098cee469daf357c?sender=u948e88338da34e7f8d8d1103

Step-1: Team Gathering, Collaboration and Select the Problem Statement



Step-2: Brainstorm, Idea Listing and Grouping



3

Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.

20 minutes

Feature Importance

- Perform feature importance analysis using Random Forest since it is powerful for capturing non-linear relationships.
- We can use Correlation Analysis for feature importance as it is quick and computationally less intensive.
- Recursive Feature Elimination offers a ranking of feature importance and allows for iterative fine-tuning.
- Permutation importance is model-agnostic and provides a clear understanding of feature impact on model performance.

Seasonal Trend

- Using Seasonal Subseries Plot for seasonal trend identification gives us a visual representation of seasonal behavior.
- Use Time Series Decomposition for seasonal trend identification as it can explicitly separates trend and seasonal components.
- To gain insights into the periodicity of the time series by examining the ACF and PACF plots.
- Apply Fourier transform to decompose the time series into frequency components and to identify dominant frequencies corresponding to seasonal patterns.

Hyperparameter tuning

- Use Randomized Search to select the best-performing set of hyperparameters.
- Bayesian Optimization can sequentially evaluate hyperparameter combinations based on the model predictions
- Apply Grid Search to Select the hyperparameter set that yields the best performance tuning.
- Automated Hyperparameter Tuning with Libraries like Optuna or Ray Tune to automates the tuning process.

Predictive Uncertainty Analysis

- Use Bootstrap Resampling method to get an empirical estimation of prediction uncertainty.
- Quantile Regression directly models the uncertainty in different quantiles

Ensemble model Integration

- Apply Bagging method for ensemble model integration.
- Use weighted voting method for ensemble model integration.
- Stacking method for ensemble model integration.
- Perform Boosting for ensemble model integration.

Anomaly Detection

- Isolation Forest algorithm isolates anomalies by randomly selecting features and building decision trees.
- Time Series Anomaly Detection captures deviations from expected patterns over time

Step-3: Idea Prioritization

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

TIP

Participants can use their cursors to point at where sticky notes should go on the grid. The facilitator can confirm the spot by using the laser pointer holding the **H key** on the keyboard.