**General-20231205\_222554-Meeting Recording**

December 5, 2023, 3:25AM

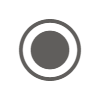
2m 26s

 **Vardhineni, Naveen Rao** 0:04  
Higher we are Group 5 and our project name is home credit default risk team members are Anurag Bharath, Alexis and myself.  
Uh.  
The goal of this project is to predict whether or not to that a customer can be flagged as a high risk applicant.  
We had the following progress, major feature engineering methods such as one hot encoder, adding custom features and we have experimented with various traditional machine learning models.  
Furthermore, in this final phase, we build the neural network model which we have.  
We have validated with all the major classification metrics.  
So the next person can take out.

 **Alexis Perez** 0:34  
Yeah, and feature engineer, we're working to get the raw data and turning into suitable code for machine learning.  
Uh, we have in the first phase year we have the list of the original files.  
We have the light Level 3 files, level 2 files and again Level 3 files split out and then then next section of feature engineering.  
Ohh, we've got the Kaggle data.  
Start within a 12 step process beginning with EDA and ending by training the test.

 **Nampally, Anurag** 1:01  
So previously we implemented logistic regression as baseline model with balance then imbalanced data set.  
Further, we explored various classification models, namely decision trees, bagging classifier, random forest, XG boost, which select K based and smart feature selection algorithms.  
We also performed at least stopping on XG boost to avoid data leakage and overfitting.  
In this phase, we expanded our project by creating single and multilayer deep learning models including linear, sigmoid and Relu as activation function.  
So we used binary cxe plus custom hinged loss with Adam and his duty operate no optimizer.

 **Veldi, Bharath Sri Vardhan** 1:36  
Yeah.  
Coming to the top performing models include a bang classifier over sample random forest and Excel Post, with optimized hyperparameters.  
Oversampling proves effective, particularly for logistic regression.  
Single URL network surprisingly performed well.  
Then multi neural network with Roc as 0.74 and OK ohh exhibit classifier model specifies reveal strong recall 92% and gaggle submission score of 0.735 and for the conclusion except boost with optimized hyperparameters with learning rate 0.1 and estimate is 400.  
Image emerged as the best performing models and for the future scope we want to conduct in-depth EDA, Explore advanced feature engineering, optimized hyperparameters and you know, refile neural networks with optimal loss functions to enhance model performance.  
Thank you.

 **Nampally, Anurag** stopped transcription