APPLIED DATA SCIENCE

SMART LENDER – APPLICANT CREDIBILITY PREDICTION FOR LOAN APPROVAL

TEAM MEMBERS - PRIYA HARSHINI R, RIDUVARSHINI A S, NITHARSHAN C V, BINDESH GUGAN S

LITERATURE SURVEY

1. PAPER DETAILS: S. Barua, D. Gavandi, P. Sangle, L. Shinde and J. Ramteke, "Swindle: Predicting the Probability of Loan Defaults using CatBoost Algorithm," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 1710-1715, doi: 10.1109/ICCMC51019.2021.9418277.
   - DESCRIPTION: Swindle executes CatBoost algorithm for prediction of loan defaults. Besides this a verification module for the documents using Tesseract and Camelot and also a personalized loan module is done, by implementing this the main goal of mitigating the risk of the financial institutes in issuing loans to defaulters and unauthorized customers is figured out.

2. PAPER DETAILS: R. E. Turkson, E. Y. Baagyere and G. E. Wenya, "A machine learning approach for predicting bank credit worthiness," 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR), 2016, pp. 1-7, doi: 10.1109/ICAIPR.2016.7585216.
   - DESCRIPTION: They have taken multiple machine learning algorithms to examine a real bank credit data for comparative analysis and have choose algorithm which is a best fit for learning bank credit data. The achieved final prediction accuracy of the model is more than 80%. Furthermore, from a total of 23 features the most important features that determine whether a customer will default or pay is extracted. The result signifies that these features can accurately determine the credit worthiness of the customers.

3. PAPER DETAILS: T. M. Alam et al., "An Investigation of Credit Card Default Prediction in the Imbalanced Datasets," in IEEE Access, vol. 8, pp. 201173-201198, 2020, doi: 10.1109/ACCESS.2020.3033784.
   - DESCRIPTION: In this paper they have used multiple datasets for building the machine learning model using algorithms such as Random Forest, Bagging, K-Nearest Neighbor, Gradient Boosted, etc, They have not only implemented machine learning techniques, but in addition to the they have implemented under sampling and oversampling techniques for better accuracies.

4. PAPER DETAILS: Y. -R. Chen, J. -S. Leu, S. -A. Huang, J. -T. Wang and J. -I. Takada, "Predicting Default Risk on Peer-to-Peer Lending Imbalanced Datasets," in IEEE Access, vol. 9, pp. 73103-73109, 2021, doi: 10.1109/ACCESS.2021.3079701.
   - DESCRIPTION: In this paper, they utilize not only several machine learning schemes for predicting the default risk of P2P lending but also re-sampling and cost-sensitive mechanisms to process imbalanced datasets. The re-sampling method makes the datasets becoming balance by changing the distribution class.

5. PAPER DETAILS: Arun, K. et al. “Loan Approval Prediction based on Machine Learning Approach.” (2016).
   - DESCRIPTION: They have implemented 6 machine learning models – Decision Random Forest, Support Vector Machine Model, Ada boost, Neural network. Using this they have predicted whether loan can be approved/not.

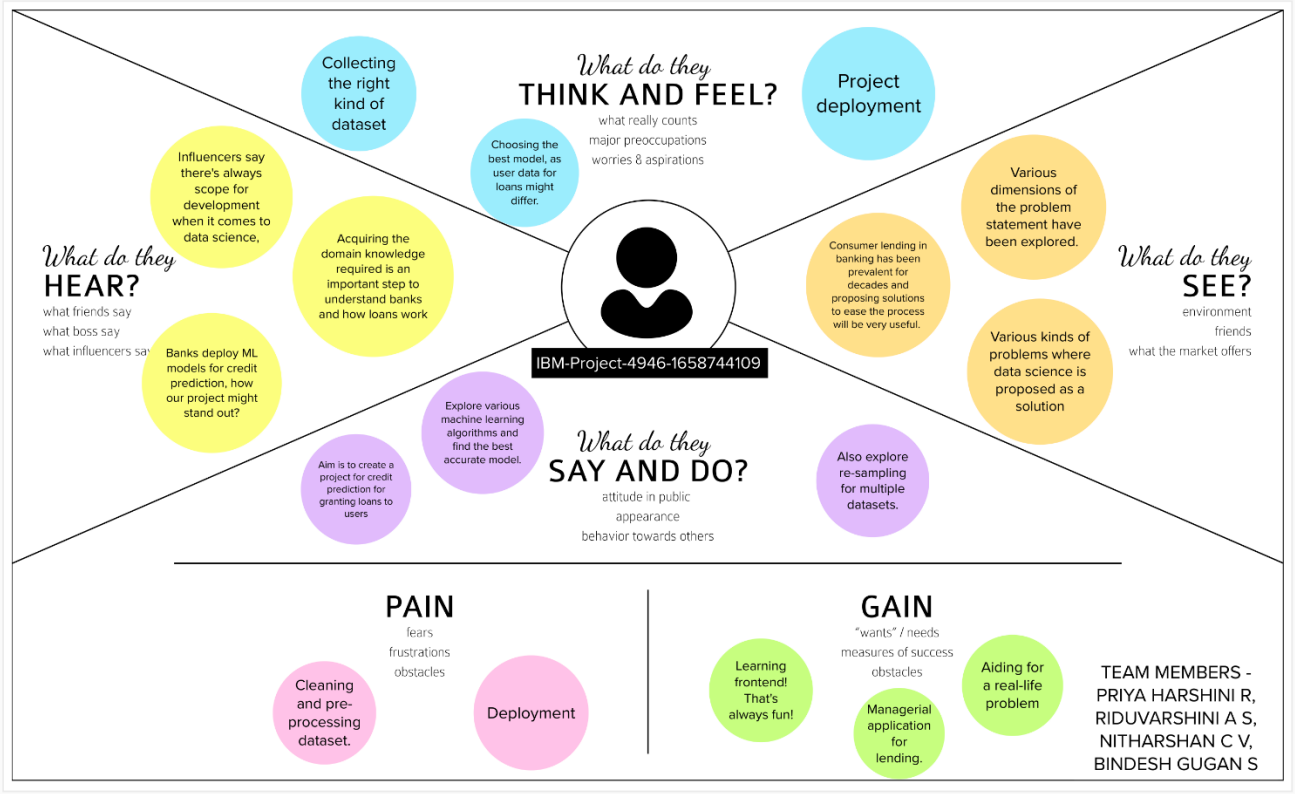
6. PAPER DETAILS: J. G. Ponsam, S. V. J. Bella Gracia, G. Geetha, S. Karpaselvi and K. Nimala, "Credit Risk Analysis using LightGBM and a comparative study of popular algorithms," 2021 4th International Conference on Computing and Communications Technologies (ICCCT), 2021, pp. 634-641, doi: 10.1109/ICCCT53315.2021.9711896.
   - DESCRIPTION: Credit Scoring can be done by using state-of-the-art Machine Learning models. Machine Learning and Data Science are becoming increasingly crucial in the fin-tech world. Popular machine learning algorithms such as Random Forest and Linear Support Vector Machines are being used currently. They’re looking to explore further into credit risk analysis with LightGBM algorithm as a choice.

Empathy Map Canvas

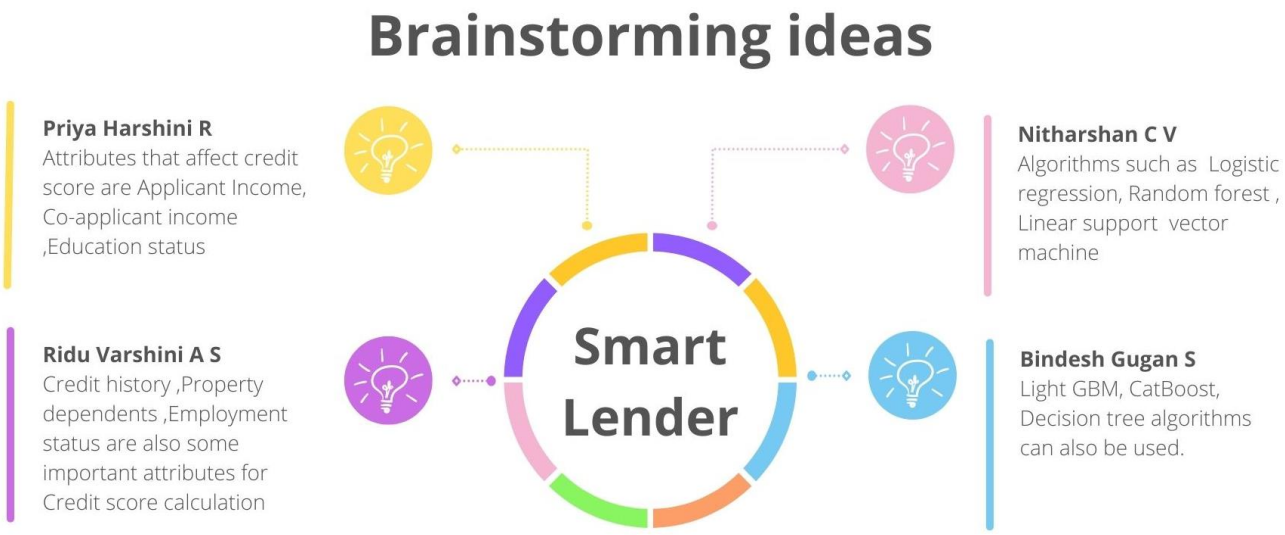
Gain insight and understanding on solving customer problems.

1

Build empathy and keep your focus on the user by putting yourself in their shoes.



IDEATION


Brainstorming ideas
Priya Harshini R
Attributes that affect credit score are Applicant Income, Co-applicant income ,Education status
Nitharshan C V
Algorithms such as Logistic regression, Random forest , Linear support vector machine
Ridu Varshini A S
Credit history ,Property dependents ,Employment status are also some important attributes for Credit score calculation
Bindesh Gugan S
Light GBM, CatBoost, Decision tree algorithms can also be used.
Smart Lender