Predict Student Performance from GAME play

Daphne Ding

slidesgo

CONTENT OUTLINE

001 BACKGROUND

002 PROBLEM STATEMENT

003 DATA PREPARATION & EDA

004 MODELING & EVALUATION

005 CONCLUSION

006 FURTHER WORK

01 BACKGROUND

BACKGROUND

The gamified learning experience offers a powerful approach to education. By incorporating game elements into the learning process, educators can create immersive and interactive environments that foster a love for learning and empower learners to reach their full potential.

This presentation shows part of the results derived from a bigger project that aims to improve students' learning outcomes from game plays. It's a case study of an educational title published by **PBS Wisconsin Education**:

" **Jo Wilder and the Capitol Case** ".

THE GAME

Learn history

FOLLOW A QUEST BOOK

COLLECT CLUES

COMPLETE MISSIONS

NEW LEVEL

THE PROJECT FLOW

01 PLAY

Play the game until level 4

02 TEST

Answer assessment questions

03 CLIMB

Improve learning experience

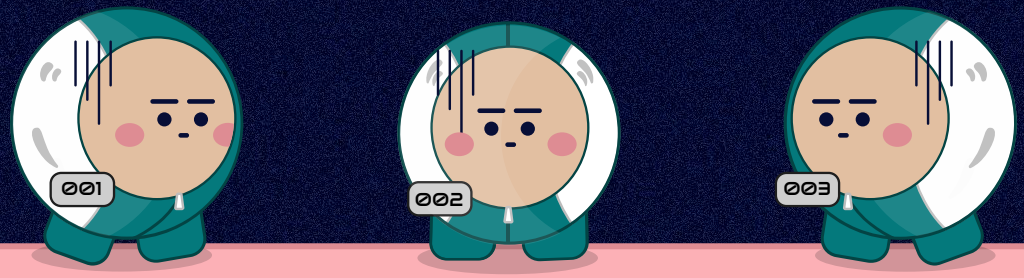
02

PROBLEM STATEMENT

PROBLEM STATEMENT

Develop **a reliable model trained on game logs** for a group of students, and **predict their correctness in answering questions (to assess their learning outcomes)**.

The model will be helpful in **improving the game design** to optimise learning outcomes, as well as **recommending suitable game content** for students' further learning and practice.



03

DATA HANDLING & EDA

DATASET

Kaggle Competition:
Predict Student Performance from Game Play

1. **Game log data:**
   23,000 students
   3.9mil sessions
   20 features

2. **Labels:**
   Correctness of students answering answering 3 questions after completing Level 4 in the game

FEATURE ENGINEERING

FEATURE SELECTION

Dropped features ['fullscreen', 'music_on','hq'']

CATEGORICAL & NUMERICAL COLUMNS

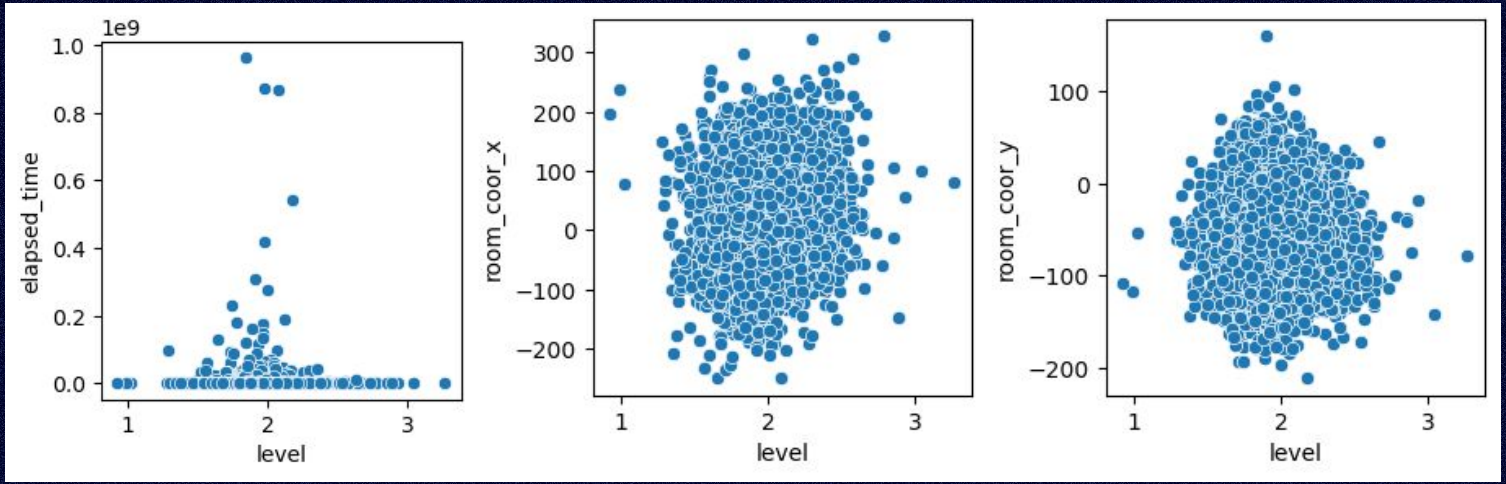
CATEGORICAL = ['event_type', 'event_name','fqid', 'room_fqid', 'text_fqid']
NUMERICAL = ['elapsed_time','level','page','room_coor_x', 'room_coor_y', 'screen_coor_x', 'screen_coor_y', 'hover_duration']

STUDENT SESSION AGGREGATION

Aggregate a student's unique events by sum, activities coordinates by mean and standard deviation.

EDA

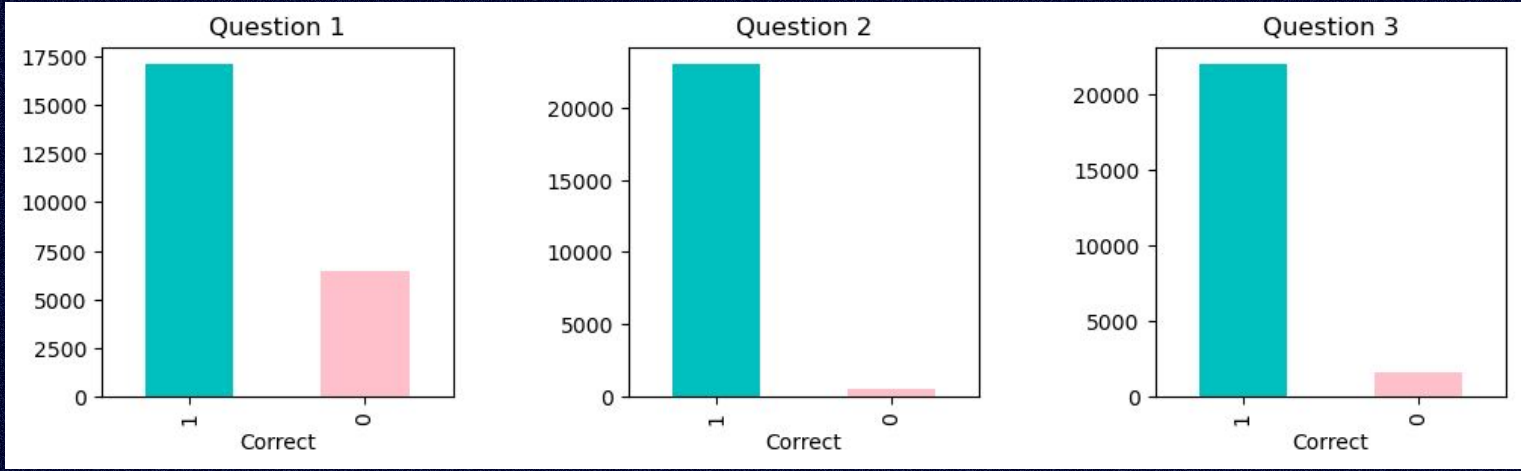
Events vs Game Level



- Students spend more time on more activities at Level 2.

EDA

Answers vs Qns



- Students get more correct answers for Qn 2 and 3.

04

MODELING

MODEL SELECTION

ACCURATE

EFFICIENT

INTERPRETABLE

ROBUST TO NOISE AND OUTLIERS

GRADIENT BOOSTING DECISION TREE

Learns from previous trees

RANDOM FOREST

Builds trees independently

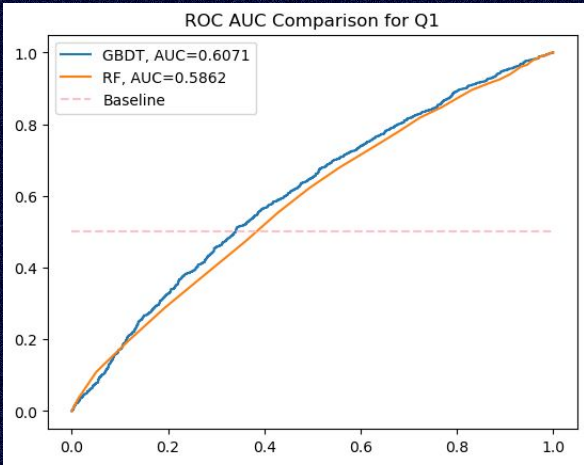
PREPROCESSING

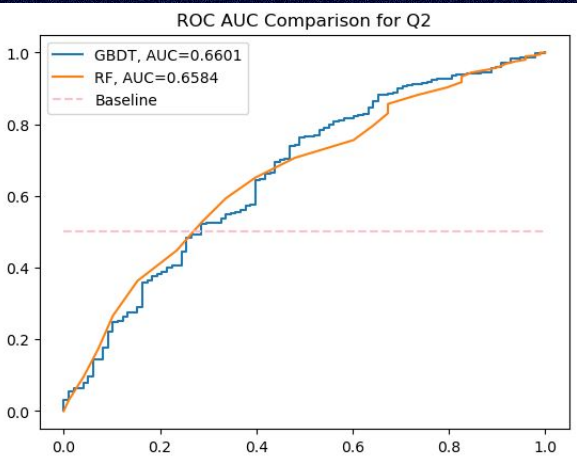
- DATA RESAMPLING
- TRAIN TEST SPLIT
- SEPARATE MODELS FOR QN 1, 2 AND 3

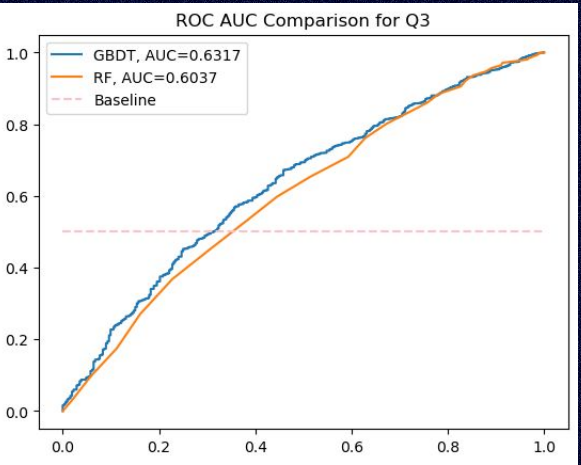
EVALUATION

	BASELINE (50%)	GBDT	RF
Q1 Train Accuracy	68.7%	73.6%	100%
Q1 Test Accuracy		73.2%	72.3%
Q2 Train Accuracy	51.1%	98.2%	99.9%
Q2 Test Accuracy		97.6%	97.9%
Q3 Train Accuracy	53.5%	93.5%	100%
Q3 Test Accuracy		93.4%	93.3%

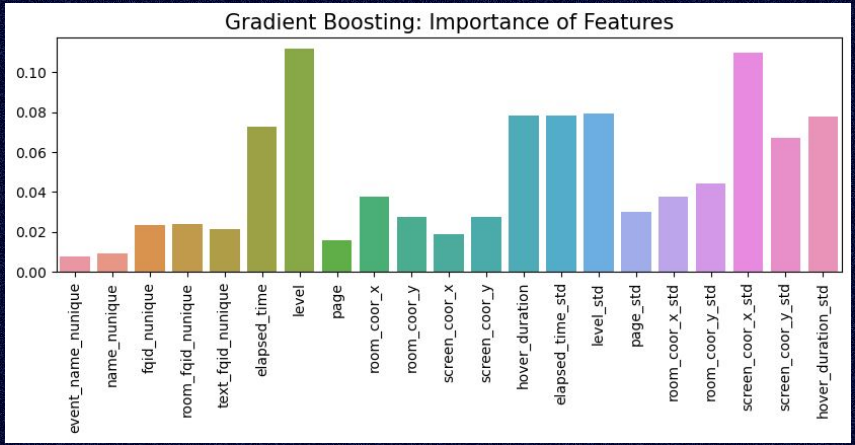
EVALUATION

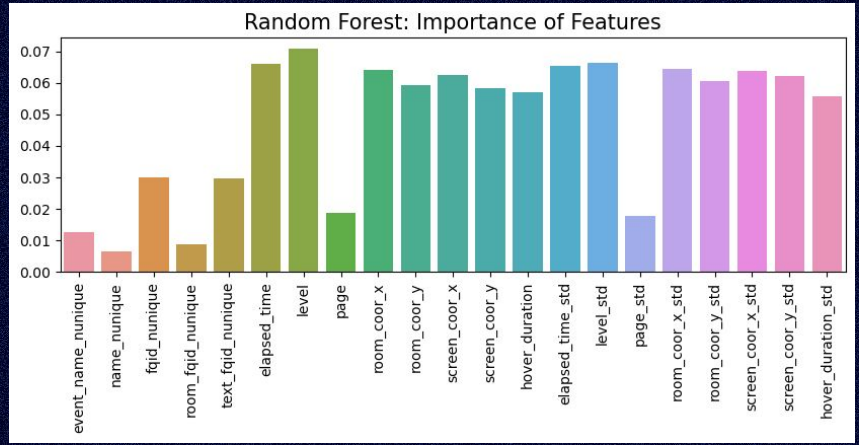

ROC AUC Comparison for Q1
GBDT, AUC=0.6071
RF, AUC=0.5862
Baseline
0.0
0.2
0.4
0.6
0.8
1.0


ROC AUC Comparison for Q2
GBDT, AUC=0.6601
RF, AUC=0.6584
Baseline
0.0
0.2
0.4
0.6
0.8
1.0


ROC AUC Comparison for Q3
GBDT, AUC=0.6317
RF, AUC=0.6037
Baseline
0.0
0.2
0.4
0.6
0.8
1.0

EVALUATION





- **GBDT** takes in the frequency of feature usage during the boosting process.
- **RF** is based on the average contribution of features across the ensemble of trees

05

CONCLUSION

CONCLUSION

GBDT and RF model performance:

- Model accuracies are largely different across 3 questions. Qn 2 and 3 have better prediction outcomes as compared to Qn 1.
- **RF tends to overfit on train set** (although online resources say otherwise).
- **GBDT outperforms RF as shown in ROC/AUC curve.**

CONCLUSION

Possible ways to improve model performance:

- Expand game log data to more levels and include more questions for students to answer at the end of the last level. **More data points** may increase the prediction accuracy. But it will likely introduce more variance to the dataset and the model could be more prone to overfitting.
- **Test other tree models (e.g. CART Model) and adjust parameters of pruning.**

06

FURTHER WORK

FURTHER WORK

In answering to the Problem Statement:

- A model of the best performance performance will be deployed.
- Learning outcomes can be assessed using the game log data. If the failure rate of students is predicted higher than expected, we should look into ways to **improve the game design** in order to achieve optimal learning results.
- Based on different predictions of students learning outcomes, **a recommendation system** can be built to feed weaker students with suitable game content for more practice.

Thanks!

Credit: slidesgo