Understanding Gambling Behaviors influencing the Risk Profile of a Player

Sanjoy Sarkar
Playtech

Playtech

*Playtech is the **largest online games and platform provider** to gambling operators throughout the world.*
Range of products / services offered by Playtech are:

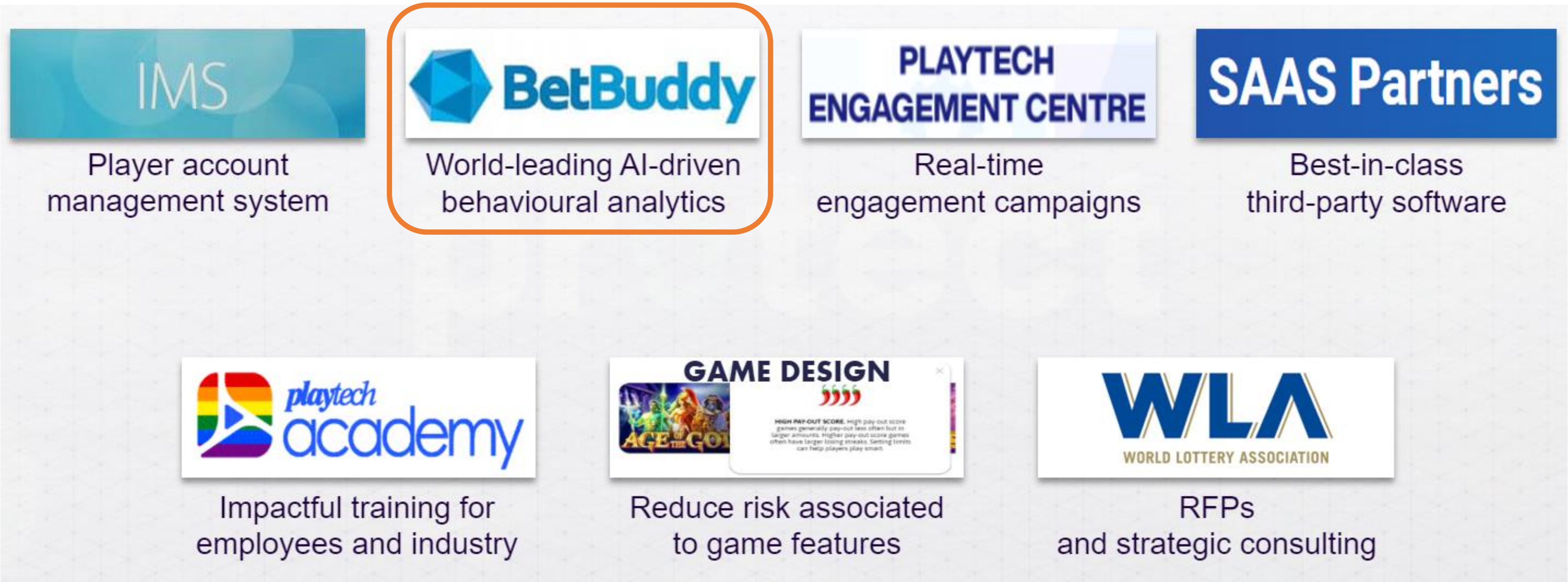
- **Gambling product offerings**:
  - Casino
  - Live Casino
  - Sportsbook
  - Virtual Sportsbook
  - Bingo
  - Poker
  - Lottery
  - Retail
- **Cloud service**
  - GPAS
- **Player account management**
  - IMS
- **Safer gambling**
  - BetBuddy

- **Services**:
  - Hosting
  - Payment advisory
  - Network management
  - Marketing service
  - Consulting service
  - Data driven insight
  - Customer support
  - Enabling sustainable engagements
  - Safer gambling
  - Fraud prevention

https://www.playtech.com

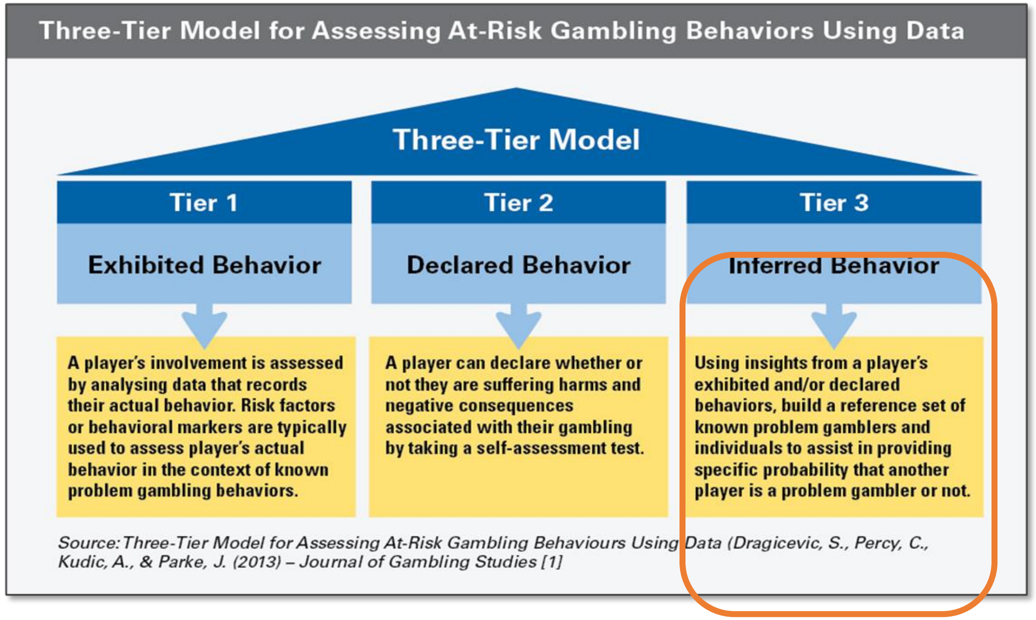
Playtech Protect

Playtech Protect is a ***sophisticated, scalable player protection ecosystem*** *integrated with Playtech's range of products and services. It is designed to* ***provide specialist help and support*** *to gambling operators* ***in the safer gambling space****. It brings together Playtech’s responsible gambling and compliance technology, tools, services, consulting and research capabilities under one roof.*

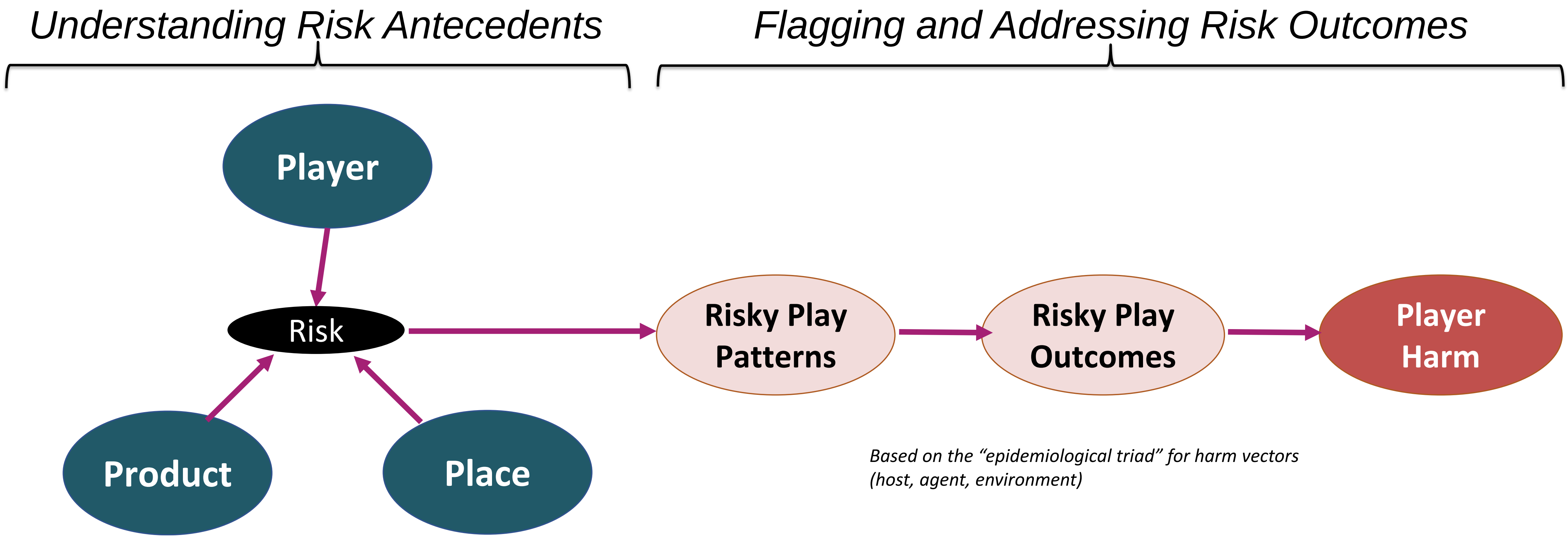


What is BetBuddy?

- BetBuddy is a software solution which identifies players who have the potential of attending gambling related harm.
- BetBuddy uses several behavioral markers to assess gambling related risk for a player
- BetBuddy assigns one of the following four risk categories to players
    - Not Rated
    - Low Risk
    - Moderate Risk
    - High Risk
- BetBuddy takes help of three tier risk assessment models
    - Exhibited behavioral model (applied statistical model)
    - Declared behavioral model (self-test – PGSI)
    - Inferred behavioral model (machine learning model)

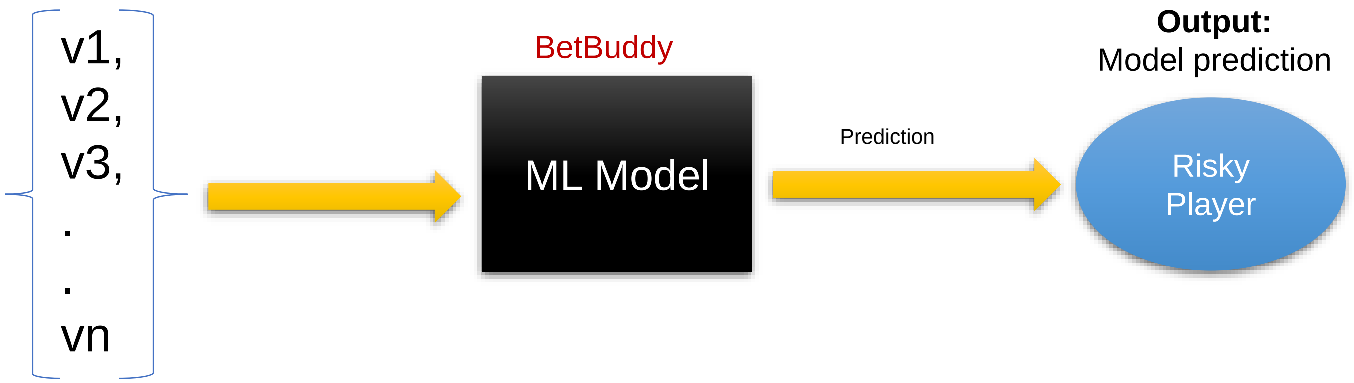


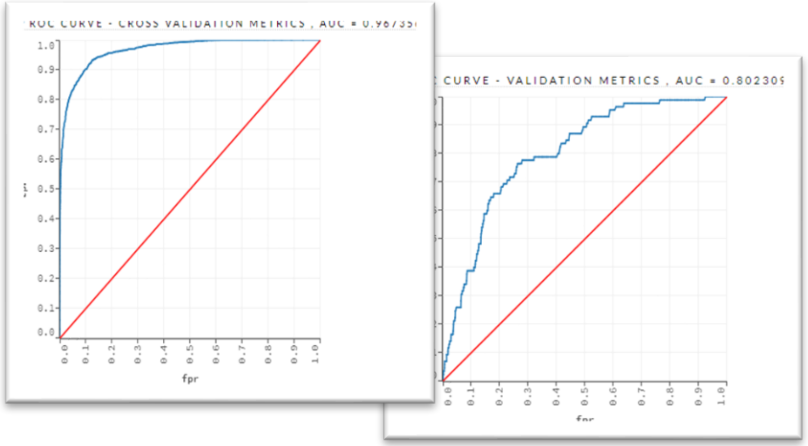
Framework of player risk



Based on the "epidemiological triad" for harm vectors (host, agent, environment)

BetBuddy: Machine Learning Model





Input:
Gambling activities/behavioural data of a player

- The model with 90% cross-validation accuracy and 74% holdout set accuracy, performing well in live environment
- Once model was started predicting risky players, obvious question came up from gambling operators –
  - How do we know model is doing things right?
  - How does the model work?
  - Why does the model predict a player as risky?

BetBuddy: Machine Learning Model

v1,
v2,
v3,
.
.
vn

Input:
Gambling activities/behavioural data of a player

BetBuddy

ML Model

Prediction

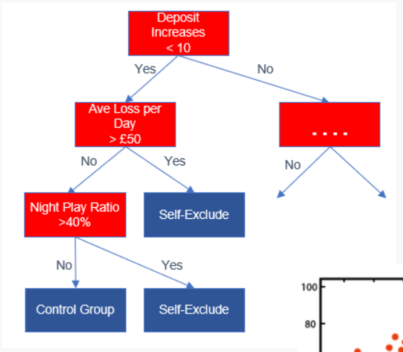
Output:
Model prediction

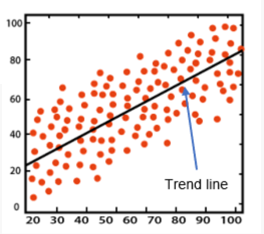
Risky Player

Black Box Model

A black box model typically provides the prediction and associated probability of pattern matching, and nothing else.

Glass Box Model

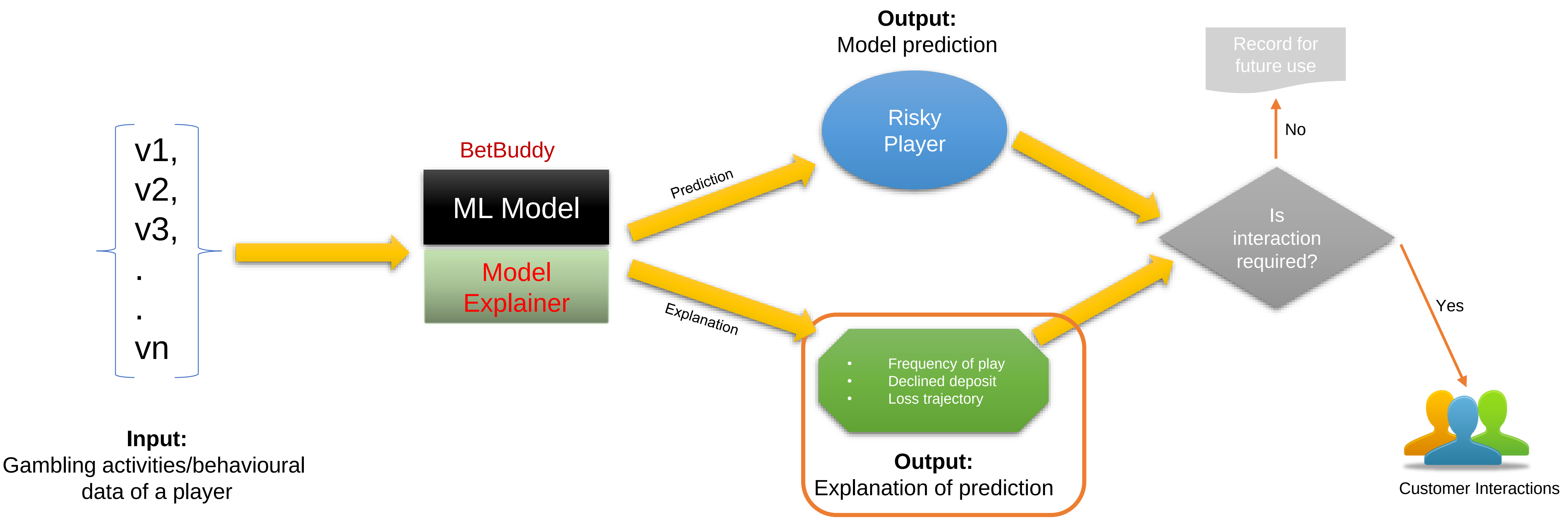




$$y = c_1x_1 + c_2x_2 + c_3x_3 + \ldots. + c_nx_n + c$$

Explainable Boosting Machine

BetBuddy: Machine Learning Model



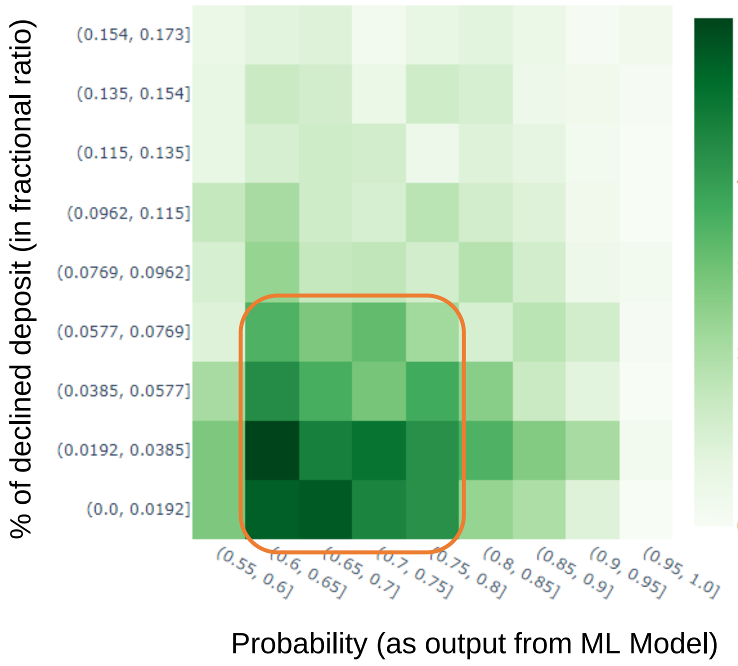
Machine Learning Model Explanation

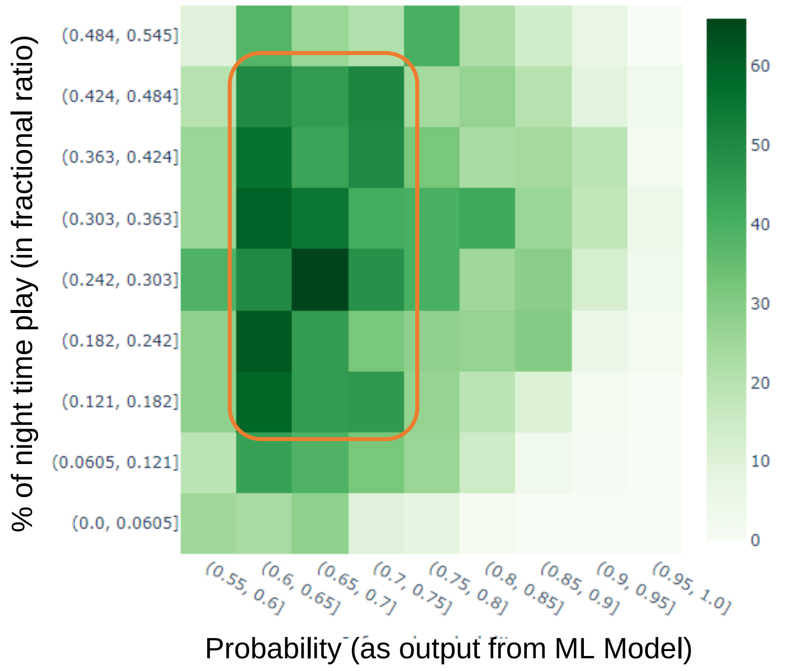
- BetBuddy machine learning models have two types of explanation:
  - **Global level explanation (explainability):** the goal is to provide an ordered list of behavioural features and metrics on their influence on building the model.
  - **Local level explanation (interpretability):** the goal is to provide an ordered list of behavioural features which might have influenced the model decision making (in either positive or negative way) in predicting a player / gambler as either a risky player or a non-risky player from the responsible gambling point of view.
  - For the purpose of model explanation, we are going to focus on two independent gambling behavioural features:
    - **Declined deposit** (% of deposits declined)
    - **Night time play** (% of playing session happened between midnight and 6 a.m. of next morning)

Machine Learning Model

Global Level Explanation: Player frequency-based analysis

Number of risky players from RG perspective



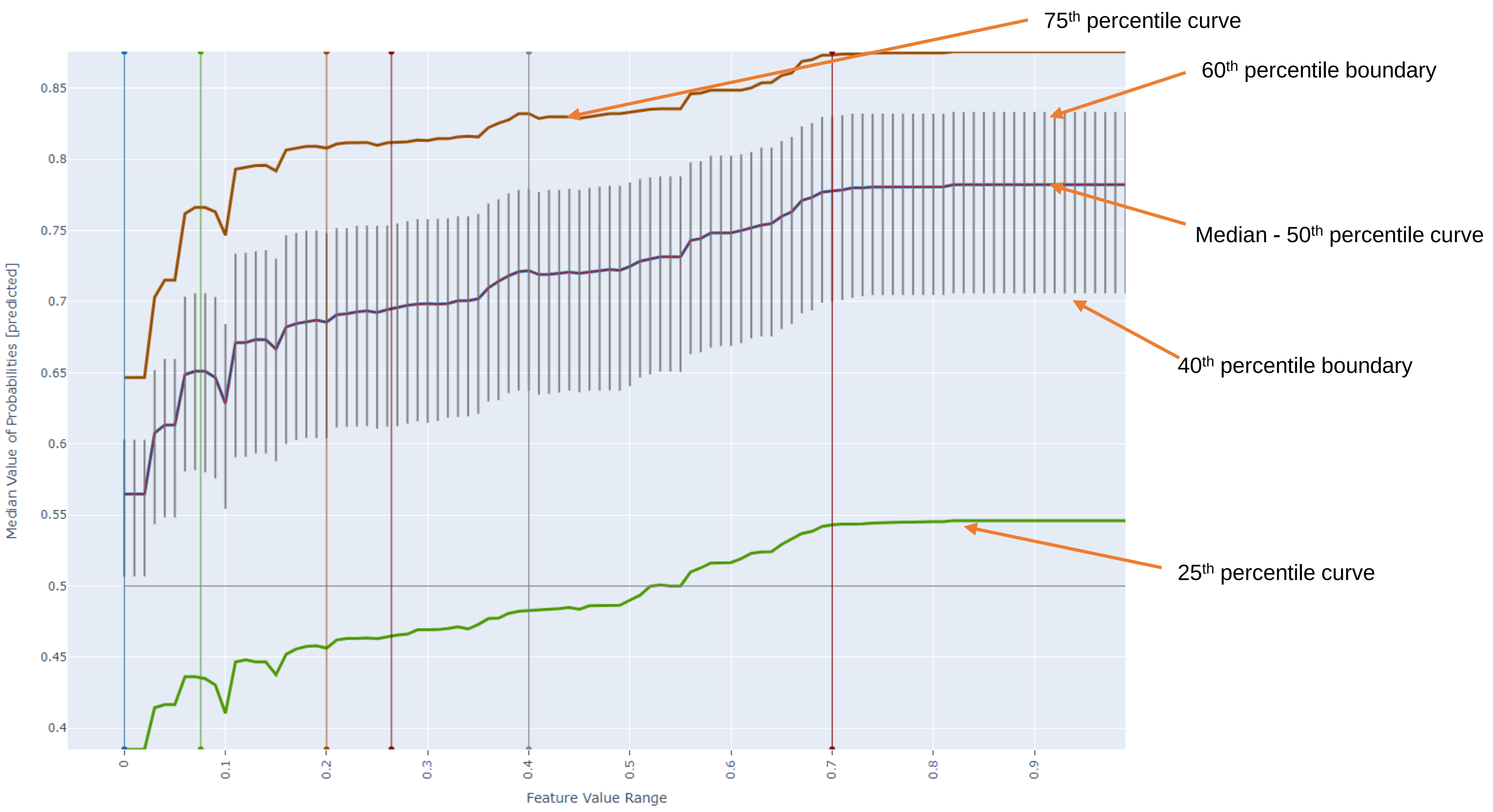


- Declined deposit up to 7.7% has more influence on predicting a player as risky player from responsible gambling perspective.

- Between 6% and 48% of night play has more influence in predicting a player as risky player from responsible gambling perspective.

Machine Learning Model

Global Level Explanation: Feature Risk Curve based analysis

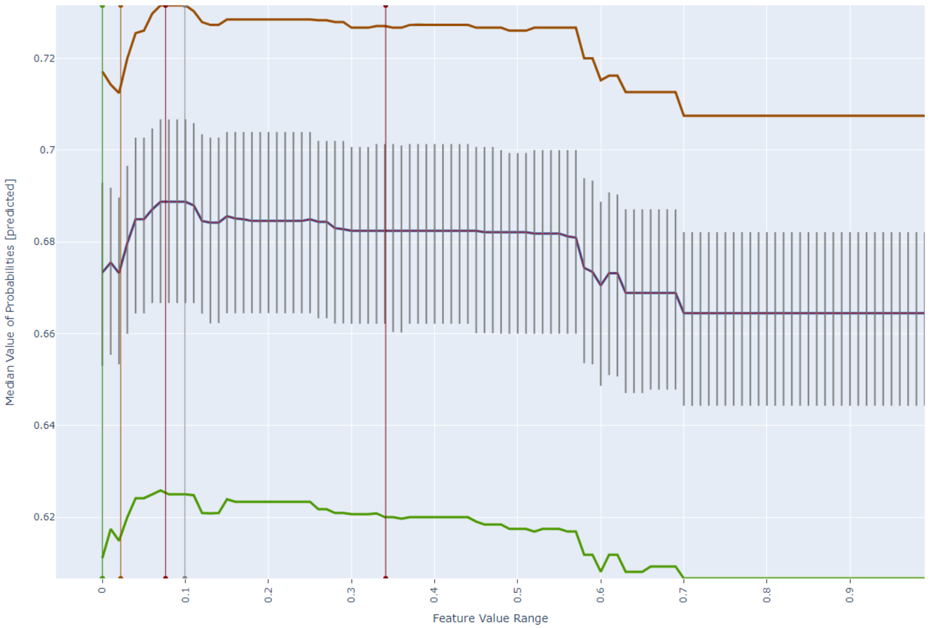


- At the global / model level, the purpose of the feature risk curve is to understand, how the change in values of an uncorrelated behavioural feature, affects the ML model prediction.
- It is somewhat similar to univariate partial dependency plot.
- The grey vertical lines depict the interval of high confidence of prediction
- Coloured vertical lines are 5th ,25th ,50th ,mean (average),75th ,90th percentile values of the feature in the data with which ML model was built

Machine Learning Model

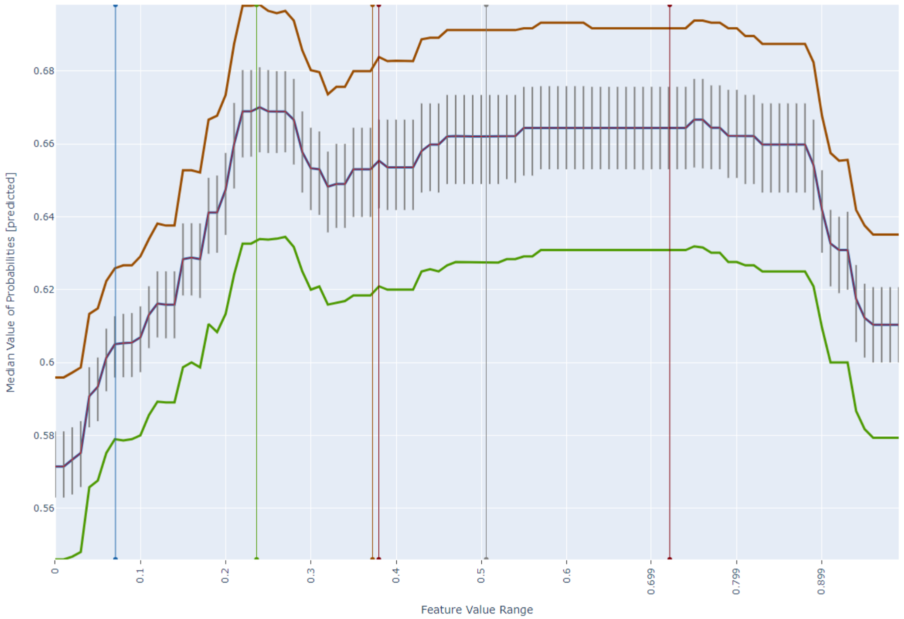
Global Level Explanation: Feature Risk Curve based analysis

% of deposit declined



- As the % of declined deposit increases, slight increase in predictive power is being observed at the model level, after which the trend becomes flat before going down below the predictive power when no deposit was declined

% of night time play



- Increase up to 30% of night play forces a steep increase the predictive power of the model. Then further increase does not bring much change, before reducing heavily at the end.

Machine Learning Model

Local Level Explanation: Industry standard methodologies

- Purpose of the local level explanation is to identify the possible reasons / influencing behavioral features due to which a player is predicted as high risk or moderate risk player from the responsible gambling perspective.

- Gambling operators are more interested in local level explanation.

- Industry standard methods/techniques are available:

    - SHAP (Shapley value, taken from game theory)

    - LIME (Local Interpretable Model-Agnostic Explanation)

- Few other model explanation techniques are slowing becoming popular.

- LIME is more popular with image and sound models whereas SHAP is even popular for models build for tabular data.

- We at Playtech, applied both SHAP and LIME to test local level explanation.

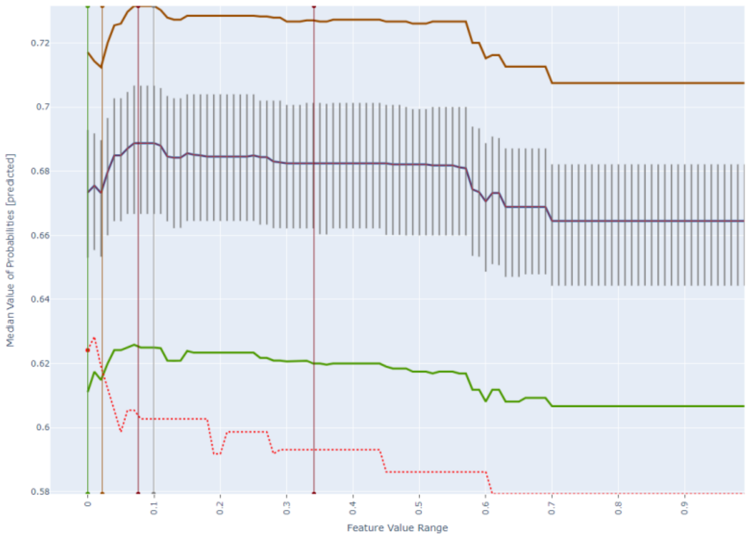
- The result was partly promising. While SHAP explanation shows a feature to be negatively influencing, LIME explanation shows it to be positively influencing.

- Conclusion was drawn that **no single explanation method is fit for all purpose** and use of multiple methods is necessary to determine the set of possible influencers (behavioral features) of model prediction.

Machine Learning Model

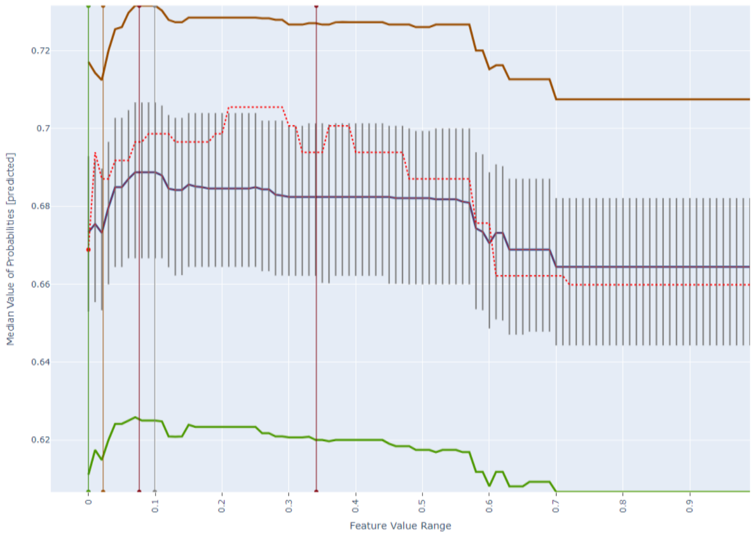
Local Level Explanation: Behaviours of % of deposit declined for three different risky player

Player #1



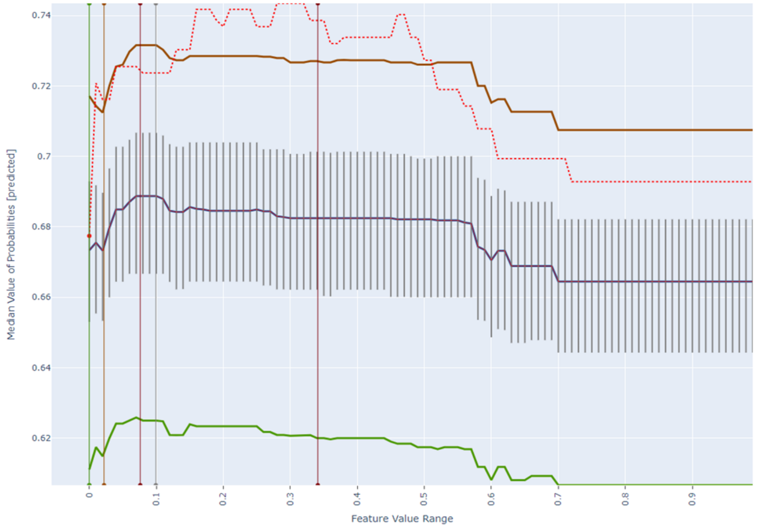
- For this player, % of deposit decline is counter factual and have negative effect on him/her being predicted as a risky player. There are some other behavioral feature(s) which influenced the prediction.
- For this player, behavior of deposit declines is not a concerning reason to initiate customer interaction

Player #2



- Increase up to 30% contributes heavily on increasing the predictability power of the model. The pattern of influence strongly follows the general pattern as seen in the global model level prediction.
- This player's deposits were never declined, and it did not factor into model prediction. Player may be interacted to inform this.

Player #3

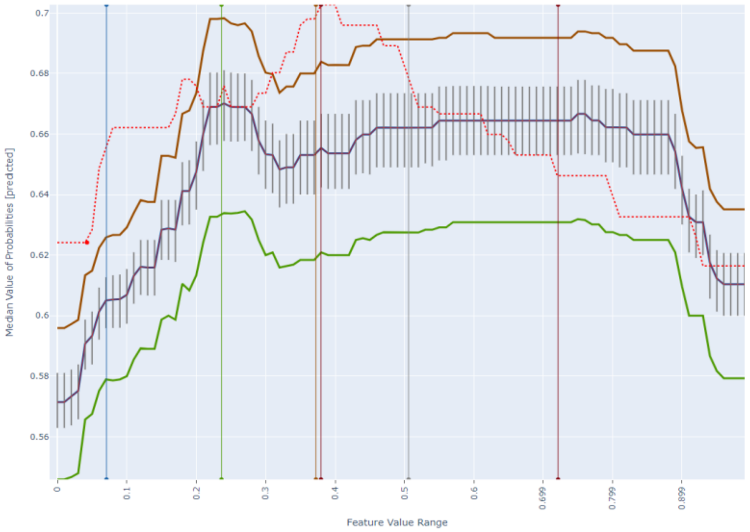


- Increase up to 40% contributes heavily on increasing predictability power of the model. The pattern of influence is on higher side compared to model level pattern of behavior.
- This player's deposits are never declined, and it does not factor into model prediction. Player may be made aware that if the deposit attempts are declined, he/she may have higher chance of being risky player.

Machine Learning Model

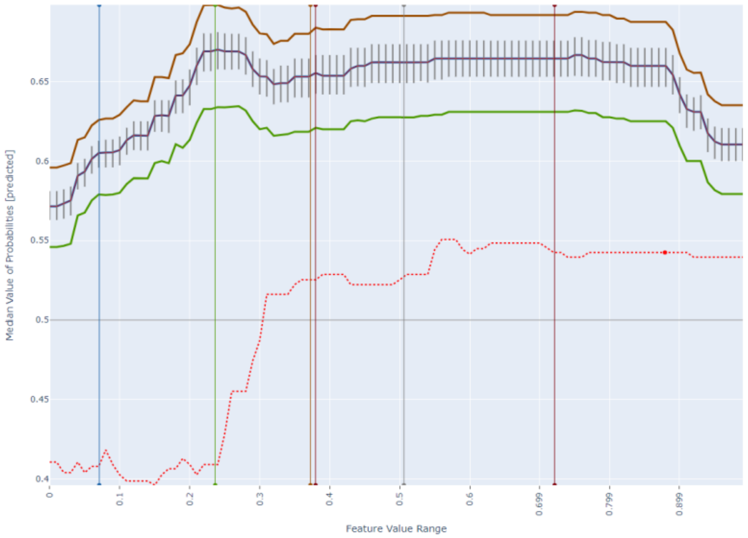
Local Level Explanation: Behaviours of % of night time play for three different risky player

Player #1



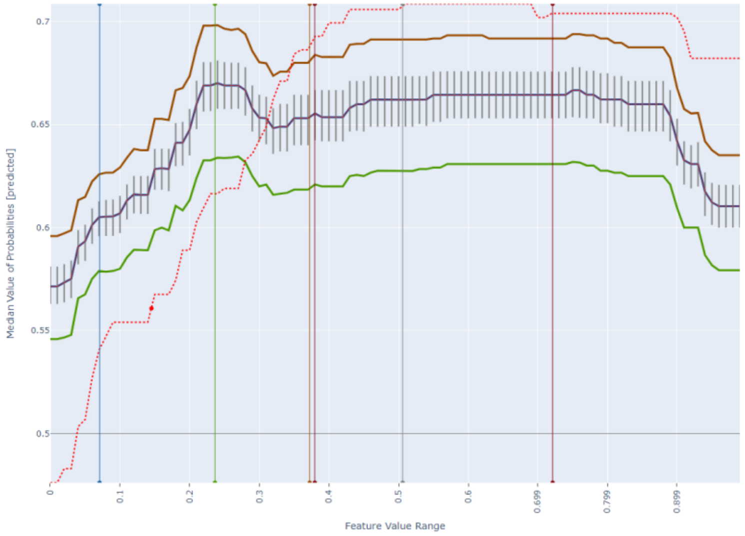
- The player shows the normal trend of increase in the chance of predicted as risky player if % of night play increases from where it is today. However, after 40% it shows a steep decline in the chance of being predicted as risky player.
- Do the player need to be contacted. It may not be, but the player behavior may need be kept on watch

Player #2



- Currently, ~90% of time, the player is playing at night. This may be his/her usual time of play. Hence, this behavior has no effect on model prediction.
- Player is a normal player and may not need to be contacted.

Player #3



- Risk curve of this player does show serious concern. A steady and steep increase in the chance of being risky player is depicted by the curve.
- This player may need to be made aware of the situation, so that he may be able to reduce playing time at night.

Machine Learning Model

A Note on explanation and its usage

- Feature risk curve based analysis adds one more view point to understand the internal working of a black box machine learning model.

- No single explanation method/technique is error free. Multiple methods are to be used to identify possible causal behavioural features.

- Feature risk curve based analysis is an indicative method. It provides additional intelligence on explaining decision of the model.

- It assumes that value of other behavioural features will remain constant . But in practice, all behavioural features in gambling are somewhat weakly or strongly dependant on one another.

- Feature risk curve based analysis will not be effective if there is a strong corelation between a behavioural feature and the subject of interest that machine learning model is trying to predict.

- Applying risk curve analysis on top of SHAP explanation may lead to better identification and understanding on the influence of behavioural features on prediction.

- Business and intuitive intelligence/experience of human are further required to have in-depth understanding of working behaviour of a ML model.

playtech protect

Thank you

Sanjoy Sarkar
sanjoy.sarkar@playtech.com

https://www.playtech.com/playtech-protect