Exploitability computation in “Human Understandable Strategies”

Farzana Beente Yusuf
CAP 5993: Game Theory
Instructor: Dr. Sam Ganzfried
Florida International University

Human Understandable Strategies

Game theory models strategic interaction among different independent rational decision makers with equilibrium computation. Nash equilibrium exhibits stability characteristics where no player has an incentive to deviate from it unilaterally.

But these approaches in general doesn't consider intelligibility aspect, especially in large scale games. With the adaptiveness to huge datasets availability, we aim to focus on incorporating human understandable learning in those approaches to take better decisions and enable information extraction from computing power.

Exploitability

Two player zero sum game is a special type of non-cooperative, competitive game where where one player's loss is exactly equal to the other player's gain i.e Rock-Paper-Scissors.

These games have value in mixed strategies - Minimax theorem by John von Neumann

Strategies are exploitable when any player deviates from the Nash Equilibrium.

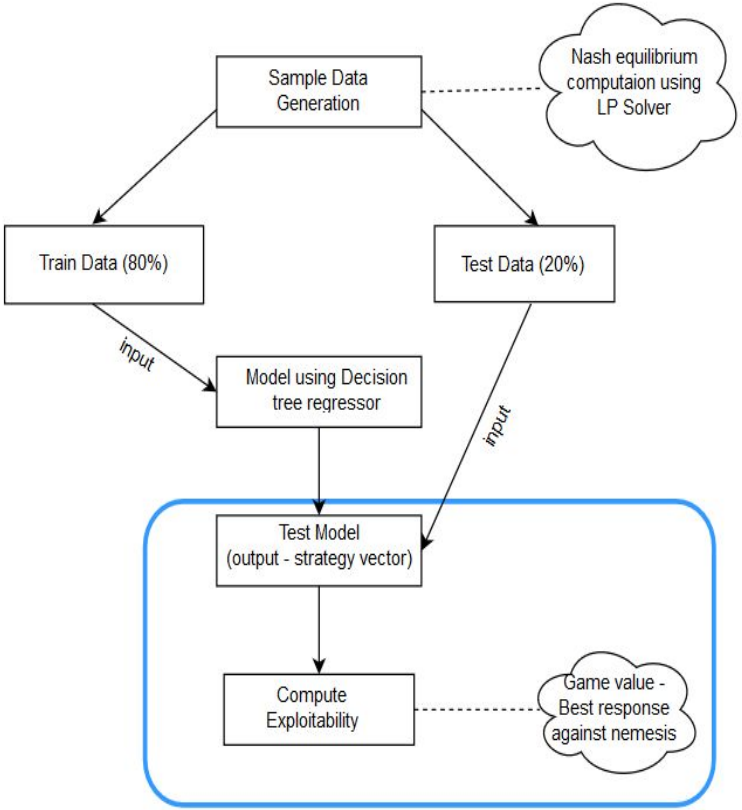
Project Goal

Computation of exploitability in Human interpretable strategies designed by application of learning formulation on Game theoretic solution concept.

General Steps to measure exploitability for the rules adapted a learnt model :

- Extract the strategy from the model using test data
- Find out the best response value against the strategy found at step 1
- Calculate game value from the labeled data and subtract from the best response value of step 2 to figure out the exploitability

Project Overview and Background



Sample Data Generation

- Two player zero sum game - variant of poker
- Deck size 10
- Ante - $1 (0.5 for each player)
- Stack size- 3
- Bet - Any multiple of 0.1- range [0,3]

Nash Equilibrium calculation

- Random probability distribution over 10 card for each player
- Model in sequence form for finding random strategies in extensive form [2]
- Nash equilibrium and game value calculation using Gurobi Optimization solver [3]

Courtesy: Dr. Sam Ganzfried [1]

Background & Modeling

Learning Formulation

Feed train data in decision tree regressor [4] from scikit learn with custom distance function and generate the Model. Earth Movers distance was used to calculate difference between strategies(i.e. probability distribution).

Input (For each player)	**Output (Mixed Strategy vector)**
PDF values over 10 card	31 bet size for 10 card
CDF values over 10 card	31 bet size for 10 card
PDF values over 10 card , specific card no for player 1	31 bet size for single card
CDF values over 10 card , specific card no for player 1	31 bets for single card
PDF values over 10 card , specific card cdf for player 1	31 bet size for single card cdf
CDF values over 10 card , specific card cdf for player 1	31 bet size for single card cdf

Table 1: Input features and output vector for different data representation

Exploitability calculation

- Feed the test data in the model and extract output strategy vector. Ex: Fred gets both winning and losing hand with equal probability and Alice always gest middle hand. If learned strategy ouput for Fred from Model depicts
  - With Jack always Bet 0
  - With King always Bet 1

- Calculate Best Response Payoff of Fred against nemesis
  - Best Response Payoff = 0

- Measure exploitability (Game value - Best Response Payoff)
  - Exploitability = 0.25 - 0 = 0.25

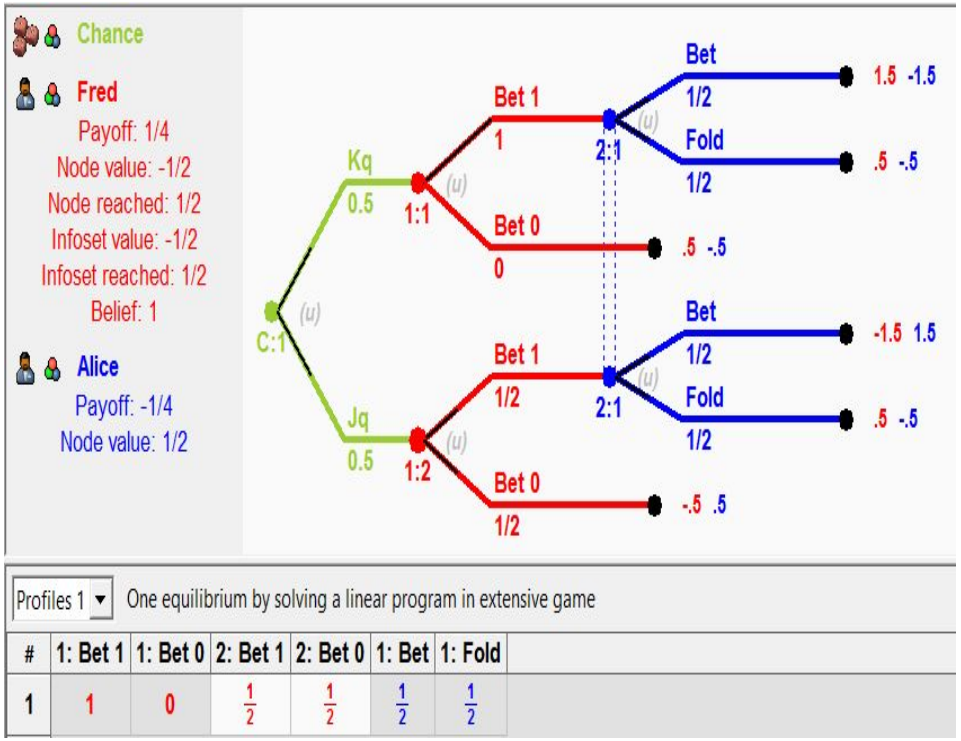


Fig 1 : Nash equilibrium and game value in 3 card poker

Implementation results

Performance measure using exploitability as cost function of learnt model

Features	Exploitability (depth 6)
20 pdf values	0.08739
20 cdf values	0.08970
20 pdf values, 1 card no	0.20036
20 cdf values, 1 card no	0.19871
20 pdf values, card cdf value	0.20788
20 cdf values, card cdf value	0.20231

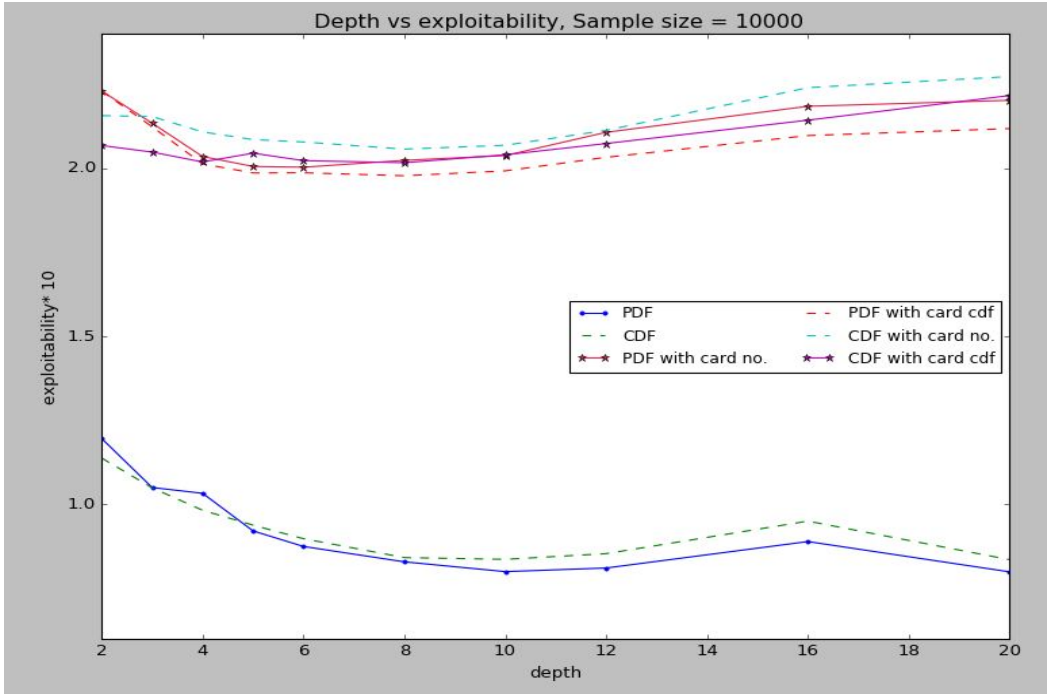


Fig 2- Tree depth vs exploitability plot on different data representation

Conclusion

- A new performance measure metrics was presented and was implemented to calculate the exploitability of different datasets which seems to meet our criteria of human interpretability with low possible cost of exploitability.
- Generic strategies computed using decision tree learning algorithm exhibits seemingly low exploitability for 2 of our representation.

References

1. https://www.cs.cmu.edu/~sganzfri/HumanUnderstandable_17.pdf
2. http://ai.stanford.edu/~koller/Papers/Koller+al:STOC94.pdf
3. http://www.gurobi.com/index
4. http://scikit-learn.org/stable/index.html
5. http://gambit.sourceforge.net/gambit13/gui.html