Shapley value in a nutshell

Giang Pham

Outline

- Shapley value in theory
  - Game theory
  - Coalition game
  - Shapley value
- Shapley value irl
  - Shapley value in Machine Learning
  - Shapley value in anything else

Modern game theory

- In 1944, 2 serious “gamers” named John von Neumann and Oskar Morgenstern published the book “*Theory of games and economic behavior*”
- In layman’s term, a game is played by several players and their job is to make decision to get the best outcome wrt their concern
- John von Neumann and Oskar Morgenstern help players make decisions easier by:
    - Consider players’ preference
    - Form a function to calculate a single real-value value for each alternatives based on players’ preference

Coalition game in words

- A family of games that the players coaliate together
- For a game of N players we can form any coalition which is a subset of N
- There is a function that returns a real-value value to tell how much a specific coalition can gain
- The value we can gain for the empty set of player is 0

Coalition game (N, v) in formula

- A set of N players $\{1, 2, \dots, n\}$
  - The coalition of all N players is called the grand coalition
- A function $v(S)$:
  - S: a subset of N, e.g $S = \{1, 2, 4\}$
  - v: function that return the real-value payoff value of the coalition S
  - $v(\varnothing) = 0$

Coalition game's problems

Questions:

1. Which coalition we should form?
2. How to distribute the payoff $v(S)$ among all the members of coalition S?

Answers:

1. Regularly we should form the grand coalition with all N players
2. Possibly Shapley value in case of grand coalition

Explain the answers

- Why often the grand coalition? The more the merrier (jk). More often than not:
  - Superadditive games: $v(S \cup T) \geq v(S) + v(T)$ with $S \cap T = \varnothing$
  - Convex games: $v(S \cup T) \geq v(S) + v(T) - v(S \cap T)$
- Why Shapley value? Because of fairness! Is satisfies:
  - Symmetry axiom: if i and j always contribute the same amount to any coalition they form then their payoff should be equal $\psi_i(N, v) = \psi_j(N, v)$
  - Dummy player axiom: if i contribute to any coalition that he forms the same amount he can achieve alone then $\psi_i(N, v) = v(\{i\})$
  - Additivity axiom: with the same set of N player, with 2 payoff functions $v1$ and $v2$, if : $(v1 + v2)(S) = v1(S) + v2(S)$ then $\psi_i(N, v1 + v2) = \psi_i(N, v1) + \psi_i(N, v2)$

Shapley value in formula

Giving a coalition game *(N, v)*, the Shapley value of a player i is:

$$\phi_i(N,v) = \frac{1}{N!} \sum_{S \subseteq N \setminus \{i\}} |S|!(|N| - |S| - 1)!\Big[v(S \cup \{i\}) - v(S)\Big]$$

Terms:

- How much we can gain when add i to the coalition S: $\Big[v(S \cup \{i\}) - v(S)\Big]$
- Number of ways to form coalition S before adding i: $|S|!$
- Number of ways the rest of the players can be added: $(|N| - |S| - 1)!$
- Number of possible sets S: $\sum_{S \subseteq N \setminus \{i\}}$
- Take the average over all the possible ways to order N players: $\frac{1}{N!}$

Shapley value history and meaning (in my own words)

- Proposed by Lloyd S. Shapley (of course) in 1953 and still hot until now
- Shapley value gives us the average contribution of any player i when he forms any coalition with any other players.
    - Consider all possible possibilities
    - Sum all of them and then take the average

Shapley value in ML (this is a link to the original paper)

- Shapley value in feature selection game:
  - Players: Input features of ML model
  - Payoff: Model performance
  - Shapley value: contribution of features to the model performance
- Shapley value in data valuation game:
  - Players: training data points (time?)
  - Payoff: goodness of the fit
  - Shapley value: contribution of data points to the model performance
- Federate learning
  - Players: data owners
  - Payoff: goodness of the fit metric
  - Shapley value: contribution of each data set to the model performance

Shapley value in ML (this is still the same link)

- Shapley value in explainable machine learning
  - Players: input feature
  - Payoff: output result
  - Shapley value: the contribution of each feature of input to the output
- Shapley value in multi-agents in reinforcement learning
  - Players: agents
  - Payoff: global reward
  - Shapley value: contribution of each agent to the global reward
- Shapley value in models assemble
  - Players: models in an assemble
  - Payoff: how correct the prediction of the assemble
  - Shapley value: contribution of each models to the goodness of the prediction

Shapley value in anything else

- Shapley value in computational biology
  - Relevance index for genes: contribution of each gene over some disease
- Shapley value in pollution responsibility
- Shapley value in contribution of database facts to query answers
- … let your imagination fly

As long as you need to measure the contribution of “workers” over the common job you can use coalition game to model and Shapley value to solve it.

Problem of Shapley value

- Computational time
  - Power set
  - Permutation
  - Exponentially grows with number of players

The end!

- Don’t hate the players, hate the game
- Don’t hate the presenter (aka me) if you don’t understand anything, hate the game