

דוגמאות הרצה

ראשית נרשום פסודו קוד בעברית על מנת להבין מה האלגוריתם עושה:

אלגוריתם SHAP-X:

קלט: v ערך של כל קומבינציה של ה"שחקנים", agents רשימה של ה"שחקנים"

פלט: טבלה מחולקת X כך שהיא מקיימת ETX (easy-to-explain)

(1) מאתחל מערך $accum$ של ערכי 0 בגודל $\text{len}(v)$ אשר מכיל את הפרמוטציות כ-KEY

(2) עבור $i=0$ עד $\text{len}(\text{agents})$:

(2.1) עבור j על כל j ששייך לקומבינציות כך ש- $|j|=i+1$:

(2.1.1) מאתחל מערך x של ערכי 0 בגודל $\text{len}(v)$ אשר מכיל את הפרמוטציות כ-KEY

(2.1.2) אם $\text{accum}(j) \neq v(j)$

(2.1.1.1) עבור k על כל ה-" k " ש- j מוכל בהן:

(2.1.1.1.1) $x(k) = v(j) - \text{accum}(j)$

(2.1.1.2) הוסף לא את x

(2.1.1.3) עבור כל ערך ב- $accum$ הוסף את הערכים של x

(3) בצע את shaply על כל המטריצה של X והחזר הודעה מתאימה לפי ההגדרה

דוגמא 1:

קלט:

$v = \{ "": 0, "a": 0, "b": 0, "c": 100, "ab": 300, "ac": 200, "bc": 100, "abc": 500 \}$
 $agents = "abc"$

ראשית $accum = \{ "": 0, "a": 0, "b": 0, "c": 0, "ab": 0, "ac": 0, "bc": 0, "abc": 0 \}$

נעבור על $i=0$ עד 3

ונעבור על כל j ששייך ראשית לקבוצה $\{a,b,c\}$ וראשית $j="a"$

נאתחל את x

(כל פעם שנגיד נאתחל את x נגדיר זאת כך :

$(x = \{ "": 0, "a": 0, "b": 0, "c": 0, "ab": 0, "ac": 0, "bc": 0, "abc": 0 \})$

בשלב זה המצב נראה כך:

			x	accum	v	
			0	0	0	a
			0	0	0	b
			0	0	100	c
			0	0	300	ab
			0	0	200	ac
			0	0	100	bc
			0	0	500	abc

נבדוק האם $v("a")$ שונה מ $accum("a")$ ונראה כי בשניהם 0 נמשיך ל $"b"$ וגם שם

כאשר נגיע ל $"c" j =$ נקבל ש $v("c")$ שונה מ $accum("c")$ ($0 \neq 100$) ולכן נבצעה את פעולה *

(* מוגדרת כך עבור על כל ה- $"k"$ ש- j מוכל בהן בצעה $x(k) = v(j) - accum(j)$ כלומר כל ה- k ש $"c"$

מוכל בהן

ונקבל:

			x	accum	v	
			0	0	0	a
			0	0	0	b
			100	0	100	c
			0	0	300	ab
			100	0	200	ac
			100	0	100	bc
			100	0	500	abc

לאחר מכן נוסיף את x ל X ואת ערכי x לערכי accum ונקבל לאחר אתחול הא מחדש כך:

		X	x	accum	v	
		0	0	0	0	a
		0	0	0	0	b
		100	0	100	100	c
		0	0	0	300	ab
		100	0	100	200	ac
		100	0	100	100	bc
		100	0	100	500	abc

נגיע ל"ab" j=נקבל ש("ab")v שונה מ accum{"ab"} (300 != 0) ולכן נבצעה את פעולה * ונקבל:

		X	x	accum	v	
		0	0	0	0	a
		0	0	0	0	b
		100	0	100	100	c
		0	300	0	300	ab
		100	0	100	200	ac
		100	0	100	100	bc
		100	300	100	500	abc

לאחר מכן נוסיף את x ל X ואת ערכי x לערכי accum ונקבל לאחר אתחול הא מחדש כך:

		X	x	accum	v	
	0	0	0	0	0	a
	0	0	0	0	0	b
	0	100	0	100	100	c
	300	0	0	300	300	ab
	0	100	0	100	200	ac
	0	100	0	100	100	bc
	300	100	0	400	500	abc

נבדוק האם $v("bc")$ שונה מ $accum("bc")$ ונראה כי בשניהם 100 נמשיך ל" ac " j נקבל ש $v("ac")$ שונה מ $accum("ac")$ (200 != 100) ולכן נבצעה את פעולה * ונקבל:

		X			x	accum	v	
		0	0	0	0	0	0	a
		0	0	0	0	0	0	b
		0	100	0	0	100	100	c
		300	0	0	0	300	300	ab
		0	100	100	100	100	200	ac
		0	100	0	0	100	100	bc
		300	100	100	100	400	500	abc

לאחר מכן נוסיף את x ל X ואת ערכי x לערכי $accum$ ונקבל לאחר אתחול h מחדש כך:

		X			x	accum	v	
	0	0	0	0	0	0	0	a
	0	0	0	0	0	0	0	b
	0	0	100	0	0	100	100	c
	0	300	0	0	0	300	300	ab
100	0	100	0	0	0	200	200	ac
0	0	100	0	0	0	100	100	bc
100	300	100	0	0	0	500	500	abc

נראה כי בהמשך $v=accum$ ולכן לא נבצע עוד פעולות.

לבסוף נרי. את המטריצה X על האלגוריתם של שאפלי ונקבל ונחזיר:

	X_3	X_2	X_1	
a	50	150	0	
b	0	150	0	
c	50	0	100	

דוגמאות "טובות" נוספות:

דוגמא 2:

קלט:

$v = \{': 0, 'a': 10, 'b': 10, 'ab': 15\}$

agents="ab"

פלט:

X_3	X_2	X_1	
-2.5	0	10	a
-2.5	10	0	b

דוגמא 3:

קלט:

$v = \{': 0, 'a': 10\}$

agents="a"

פלט:

X_1	
10	a

דוגמא 3:

קלט:

$v = \{': 0, 'a': 100, 'b': 150, 'c': 250, 'ab': 200, 'ac': 250, 'bc': 300, 'abc': 370\}$

agents="abc"

פלט:

X_7	X_6	X_5	X_4	X_3	X_2	X_1	
40	0	-50	-25	0	0	100	a
40	-50	0	-25	0	150	0	b
40	-50	-50	0	250	0	0	c

דוגמא גדולה "גרוע":

דוגמא 4:

קלט:

$v = \{ '': 0, 'a': 50, 'b': 100, 'c': 70, 'd': 80, 'ab': 110, 'ac': 120, 'ad': 150, 'bc': 200, 'bd': 100, 'cd': 140, \{ 'abc': 220, 'abd': 200, 'acd': 300, 'bcd': 280, 'abcd': 300 \}$
agents="abcd"

פלט:

X_14	X_13	X_12	X_11	X_10	X_9	X_8	X_7	X_6	X_5	X_4	X_3	X_2	X_1	
-45	0	30	23.333	3.333	0	0	0	10	-20	0	0	0	50	a
-45	30	0	23.333	3.333	0	-40	15	0	-20	0	0	100	0	b
-45	30	30	0	3.333	-5	0	15	0	0	0	70	0	0	c
-45	30	30	23.333	0	-5	-40	0	10	0	80	0	0	0	d

מדוע דוגמא זו "אינה טובה":

כיוון שכאשר נרצה להביא למשתמש לקרוא יהיה לו המון מקרים ולכן יהיה יכול להיות שהוא יאבד עניין או את החשיבות של סוף התהליך, כמו כן קשה לעקוב אחרי האלגוריתם עצמו אם נרצה לבדוק איך הוא הגיע לשורה מסויימת, בנוסף ניתן לראות שהיה אפשר להפוך את X_{12} - X_{14} לטבלה הנ"ל:

X_13'	X_12'	
0	-15	a
0	-15	b
15	0	c
15	0	ג

שזהו יצוג יותר פשוט כיוון שהצלחנו להחסיר עמודת הסבר.

אפשר לראות שהאלגוריתם לא תמיד מחזיר פתרון קטן ולפעמים ניתן לחסוך עמודות ולמשמעות תשאר זהה מבחינת הסברה למשתמש.