**Modeling People’s Voting Behavior with Poll Information**

במאמר מתואר מודל קבלת החלטות חדש להצבעות בוחרים אשר מסוגל לחזות את התנהגות הבוחר כתלות במידע אותו הוא מקבל על המועמדים.

מאמר זה הוא הראשון לספק הערכה אמפירית של תאוריות קבלת החלטות, כאשר החלטת הבוחר מתבססת על מידע קודם. בנוסף, הוא הראשון לבחון את יכולות החיזוי של כל מודל קבלת החלטות שהיו מקובלות עד היום.

המודל נקרא AU (Attainability Utility) והוא מתבסס על מידע קודם על המועמדים, ההיתכנות והתועלת שבחירת מועמד כלשהו תפיק לכל מצביע.

מטרת המאמר היא ללמוד על אסטרטגיות המצביעים ולבדוק איך התנהגותם תואמת מודלים של קבלת החלטות שמצויים בספרות המקובלת אל מול מודל הAU.

חשוב לציין כי המודלים מתבססים על 3 עקרונות:

* מיקסום תוחלת התועלת: מצביע רציונלי יבחר לנסות למקסם את הרווח שהוא יפיק מהצבעה למועמד כלשהוא במידה ויזכה, כתלות בהתפלגות הסיכויים לזכיה של המועמד ולמצביעים אחרים.
* היוריסטיקה בקבלת החלטות: לכל מצב המוצג למצביע (משתנה בתכלות במידע אותו הוא מקבל) ישנה פונק' אשר תמפה את ההחלטה שלו. המצביע לא נתפס כרציונלי ולא בהכרח יש לו תוחלת קרדינלית לכל אירוע אשר יכול להתרחש.
* רציונליות מוגבלת: המצביע יפעל אסטרטגית לפי המציאות אותה הוא תופס ולאו-דווקא לפי המציאות שבפועל.

על מנת להבין את המודלים השונים חשוב להגדיר מושגי בסיס:

הגדרות:

* + תועלת הבוחר: לכל מצביע פונק' תועלת קרדינלית כך ש היא קבוצת המועמדים.

ו- מתאר את תועלת הבוחר כאשר מועמד מנצח.

במקרה של שיוויון תועלת הבוחר תוגדר כך-

כך ש W- יוגדר כקבוצת המנצחים .

* + המידע המקדים שמקבל הבוחר: כאשר מוגדר להיות תוחלת מספר המצביעים שיצביעו למועמד . נגדיר את מספר ההצבעות בסה"כ להיות .
  + אינדקס ההעדפות: אינדקס ההעדפות של כל בוחר יסומן ב כאשר מסמן את מי שהבוחר מעדיף יותר.
  + מודל החלטה: מודל ההחלטה היא פונקציה כאשר היא קבוצת כל פונקציות התועלת הקיימות בקבוצה C.

נגדיר להיות הצבעת הבוחר עם פונקציית תועלת בהנתן מודל החלטה וסקר .

* + הצבעה כנה: נגיד כי כנות היא ב"ת במידע הניתן למצביע ותוגדר בצורה הבאה

המודל AU הושווה עם המודלים הבאים:

* + **K-pragmatist (KP)** – כל בוחר בוחן את K המועמדים המובילים (תוגדר כקבוצה ) ובוחר מבניהן את המועדף עליו ביותר. המודל יוגדר בצורה הבאה:

.

* + **Calculus of Voting (CV)** –כל בוחר רציונלי יצביע למי שימקסמם את תוחלת התועלת שלו. (סיבוכים במודל זה גדלים כאשר אנו מניחים שהבוחר יודע את העדפות האחרים). נסמן את ההתפלגות של תוצאות ההצבעה האמיתיות, כתלות בתוצאות ההצבעה , להיות . נגדיר בוחר להיות פיבוטלי למועמד y על מועמד x – כאשר הצבעה למועמד y תגרום לו לנצח, או להיות בתיקו עם המנצח, כאשר התוצאה של הצבעה אחרת היא שמועמד x ינצח. פונקציית המודל תהיה:

כאשר היא ההסתברות שהבוחר הוא פיבוטלי למועמד y על מועמד x בהתפלגות D.

* **Local Dominance (LD)** – כל בוחר פועל תחת רציונלית מוגבלת כך שיש פרמטר אי וודאות שנסמנו ב- r. הבוחר יצביע רק למומעדים מקבוצת המועמדים הלא נשלטים שכלולה בקבוצת המנצחים האפשריים PW הכוללת את המועמדים שהציון שלהם בהצבעה s היא לפחות . המודל הוא:

ע"ב מודל זה קיים מודל ה- **(LDLB) Local Dominance with leader bias** כאשר והבוחר יצביע למועמד המוביל במקום להיות כנה.

* + **Attainability (AT)** – למועמד כלשהו j נמדוד את היתכנות כמות ההצבעות אשר יובילו לכך שהוא ינצח- כך ש היא תוחלת מספר ההצבעות למועמד j ו הוא פרמטר לכל בוחר. פונקציית המודל היא:

חשוב לציין כי כל המודלים מתבססים על כך שהבוחר פועל עפ"י אסטרטגיה דטרמיניסטית ועקבית והם לא יכולים להסביר לגמרי או לחזות את ההצבעות שמצויות בהן אספקט של אקראיות.

ולבסוף מודל **Attainability Utility (AU)**:

כאשר הוא קבוע אשר נוסף בכדי להתמודד עם תועלת ששווה לאפס.

פרמטר הוא החלופה בין החשיבות היחסית של השגיות לתועלת.

לדוגמא כאשר הבוחר תמיד יצביע למועמד עם מספר הקולות המירבי (leader bias)

וכאשר הבוחר יצביע באופן כנה.

AU מתבסס על מודל אחר המתואר במאמר של Bowman et al.. המודל של Bowman התבסס על הצבעות בינאריות. AU מרחיב אותו ע"י הוספת אפשרות למס' גדול יותר של מועמדים והוספת פרמטר שמבצע התחשבות בסיכוי המועמד להיבחר לבין התועלת אותו מצביע יכול להרוויח מזכיית המועמד.

מתודולוגיית ההשוואה התבצעה על 4 Datasets כאשר 3 מתוכם נלקחו ממאמר קודם של Tal et al.. ב Datasets אלו הבוחרים שיחקו מספר משחקים בהם בחרו בכל סבב את המועמד המועדף עליהם. כל סבב כלל מצביע אנושי אחד אשר קיבל מראש את סדר ההעדפות שלו ומודל רעש של תוחלת ההצבעות של שאר האוכלוסיה כך שפרט זה הוא מלאכותי.

מטרת כל בוחר כאמור היא להצביע באופן שימקסם את התועלת שלו.

הDataset הנוסף נלקח ממאמר של Tyzsler and Schram ובו התמודדו בכל סיבוב 12 מצביעים, כל שלכל מצביע ניתן מידע שלם אודות הבחירות וסדר העדפות ל3 מועמדים שונים. התוצאות של מאמר זה היו מבוססות אמת לעומת המאמר הקודם.

על ארבעת Datasets הללו בוצעו חיזויים על בסיס כל המודלים הנ"ל בכדי לבדוק מיהו המודל עם השגיאה הקטנה ביותר.

**נמצא כי בכל ארבעת ה Datasetsשגיאת מודל הAU הייתה הקטנה ביותר**.

מלבד השוואת מודל הAU מול המודלים הנ"ל בוצעה השוואה אל מול אלגוריתמי למידת מכונה.

מתוך רשימה של מודלים בלמידת מכונה נבחר מודל הRandom Forest (RF) להיות מושווה כנגד מודל הAU מכך שהיה המדוייק ביותר.

יש לציין כי מודל זה (RF) משתמש במספר גדול בהרבה של פיצ'רים (וכמו כל אלגוריתם למידת מכונה יכול ליצור עוד הרבה פיצ'רים) מאשר מודל הAU, ומתבסס על אימון על מידע רחב יותר.

נמצא כי מודל הAU מנצח את מודל הRF ב100 מתוך 335 בוחרים. יתרון הRF נזקף לכך שקיימת קבוצה מסויימת של בוחרים שמודל הAU סיפק ביצועיים פחותים עבורם. בנוסף לעומת מודלי למידת מכונה, מודל הAU ושאר מודלי קבלת ההחלטות מתבססים רק על התצפיות שנבדקות ולא על תצפיות קודמות. כאשר בחרנו את אלגוריתם הRF עם מידע זהה למידע שקיבלו מודלי קבלת ההחלטות, שגיאת המודל RF קפצה משמעותית.

לאחר ניתוח תוצאות הניסויים התקבלו הממצאים הבאים:

1. ישנם בוחרים שקשה יותר לנבא את בחירותיהם - המודלים הנ"ל התבססו על כך שכל מצביע יצביע רק עבור מועמדים שיכולים לנצח (לא נשלטים), אבל במציאות לעומת זאת, מצביעים אכן הצביעו גם למועמד שלא היה כל סיכוי שיבחר. קשה להבין את ההצדקה לבחירה שכזה ולכן אנו מניחים שיש אלמנט אקראיות כלשהו. ולכן, קשה לנבא את בחירותיהם של המצביעים הללו. במאמר סיווגו בוחרים לפי מספר הפעמים שהם ביצעו פעולה נשלטת (הצבעה למועמד שאין סיכוי שיבחר) ומצאו כי בקבוצות בהן מספר הפעולות הנשלטות היה גבוה גם ביצועי מודל הAU נפגמו וגם במודל RF. לעומת זאת בקבוצות אשר מספר הפעולות הנשלטות היה נמוך, ביצועיי המודלים היו טובים משמעותית.

לכן ההשערה שביצוע פעולה נשלטת כנראה מקורה מ"רעש" כלשהו או מתוך אלמנט של אקראיות.

1. הממצא המרכזי הוא שמודל הAU מסביר היטב (בפרט טוב יותר משאר מודלי קבלת ההחלטות אחרים) את ההתנהגות של רוב הבוחרים (למעט אלו אשר בולטים בחוסר עקביותם).

רוב הבוחרים יפעלו בהיוריסטיקה הדומה למודל הAU אשר לאו דווקא מנסה למקסם את תוחלת התועלת, בדומה לקבלת החלטות במישורים אחרים. גם אם במצטבר ניתן להסביר את התפלגות הבוחרים לפי תאוריות רציונליות יותר (כמו CV) עדיין לא ניתן לחזות בוודאות את הבחירה של מצביע יחיד מכיוון שקשה יותר להגדיר מהי פעולה נשלטת עבורו – משמע, כל בוחר מעריך באופן בלתי תלוי כל מועמד.

אנו נקבל תמונה מלאה הרבה יותר של התנהגות הבוחרים ע"י שילוב של פעולות היחיד ופעולות הכלל.

ממצא זה נמצא בקו תואם למחקרים אחרים על קבלת החלטות בחוסר ודאות.

* Add a figure/table that are not included in the paper, that can help clarify some of the definitions/results/proofs

כפי שכבר הזכרנו מאמרים אלו בסיכום, נבחר לציין את המאמר של Tal et al. [28]:

Maor Tal, Reshef Meir, and Ya’akov (Kobi) Gal. 2015.

A Study of Human Behavior in Online Voting. In Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, Istanbul, Turkey, May

מאמר זה היווה בסיס למאמר הנוכחי בכך שחקר התנהגות בוחרים תחת מידע שקיבלו אודות הקלפי, אך מאמר זה לא ביצע השוואות בין חיזויי המודלים. כמו כן, גם הנתונים אשר נאספו במאמר זה שימשו במאמר אותו סיכמנו לצורך ההשוואות ומדידת ביצויי המודלים.

כמו כן גם נבחר לציין את המאמר של Bowman et al. [6]:

Clark Bowman, Jonathan K Hodge, and Ada Yu. 2014. The potential of iterative voting to solve the separability problem in referendum elections. Theory and decision 77, 1 (2014), 111–124.

מאמר זה סיפק מודל אשר היווה בסיס למודל הAU, המודל התחשב בסיכויי ההיתכנות של משתנה בינארי בתלות בתוחלת התועלת של ההצבעה, (מודל זה היה גם שונה בכך שהניח מודל זהה לכל מצביע).

* Find (use Google Scholar) 1 paper citing your paper, and explain why it is cited.
* Include in a separate file the plan for your project