

הוספת מידע: הרגע שבו למידה חישובית הפסיקה לחקות את המציאות והתחילה להפוך אותה / ערן הדס

פורסם:

הדס, ערן. (סתיו, 2018). [הוספת מידע: הרגע שבו למידה חישובית הפסיקה לחקות את המציאות והתחילה להפוך אותה](#). מארב מוסף 24

מאז ומתמיד עסקה אמנות בדמיון, ופעמים רבות יצרה עולם של מציאות חלופית ובדויה. בעידן הבינה המלאכותית המכונות, או ליתר דיוק התוכנות, הן אלה שיוצרות עולמות בדיוניים של מציאות מדומה. לפיכך יכולה אמנות לשמש כגשר בין המציאות הפיזית לבין זו של הבינה המלאכותית. מצד שני, אמנות היא זירה שבה ניתן לבקר את הטכנולוגיה ואת השפעתה על התרבות שלנו מנקודת מבט ייחודית. על-פי רוב, טכנולוגיה נמדדת מהיבטים כלכליים (כמה רווח היא גורפת) או פונקציונליים (כמה מידע, מה המהירות וכו'). אמנות יכולה לדון בדרכים שבה טכנולוגיה משפיעה על המציאות שלנו, בתכונותיה האתיות והאסתטיות, ובשימושים הלא פונקציונליים שלה או אלו שלא לשמם היא נוצרה. בטקסט הזה אדון בהשלכותיה של הלמידה החישובית, שהיא צורת הבינה המלאכותית שנכנסה לחיינו ובמידת מה השתלטה עליהם בשנים האחרונות.

1. בעיות בלמידה חישובית

למידה חישובית היא למידה הבנויה על הכללה מתוך דוגמאות. זה החלק שקל לאנשים להבין. אך רבים מרימים גבה כשהם שומעים שהיא אינה מבוססת על חוקים הקשורים לבעיה. לדוגמה, למודל שבוחר מסלול

אופטימלי לנסיעה לא מזינים אף חוק תנועה. למודל שמחולל טקסט תו-אחר-תו אין מושג אם אוסף התווים שייצר בין שני רווחים הוא מילה. ישנם כמובן חוקים למערכות למידה חישובית (המכונים מודלים), אבל הם חוקים מתמטיים בלבד, ועל-כן הם דורשים לייצג את הדוגמאות כיצורים מספריים שהמודל יודע לעבד. וכשזה המצב, האם אנחנו כפופים לחוקי המציאות, או שחוקי המתמטיקה עשויים לכפות עלינו חוקים אחרים?

רעיון ההכללה, שעליו מתבססת הלמידה החישובית, ידוע לנו כבעייתי, למשל כשאנחנו מתייחסים לבני אדם. הכללה ביחס לקבוצת אנשים על-סמך התנהגות של מספר אנשים בודדים היא דבר פסול חברתית, מפני שאנחנו מייחסים לאנשים תכונות על סמך אנשים אחרים, וזהו שיפוט מוקדם (pre-judgement) או דעה קדומה. לעומת זאת, כשמודל למידה חישובית צריך לזהות עצמים שמופיעים בתמונה, אנחנו מצפים שמתוך התמונות שכבר ראה יכליל לתמונות שלא ראה. לדוגמה, לאחר שהמודל אומן לדעת שאלפי תמונות של שולחן מייצגות שולחנות, נצפה מן המודל שיידע שגם תמונה של שולחן שטרם ראה מייצגת שולחן. אם ניקח למשל שדות תעופה, ידוע לנו שמותקנות בהם מערכות למידה חישובית[1]. האם הן לומדות איך אמור להיראות או להתנהג מחבל מדוגמאות עבר, ולנבא עבור כל אחת ואחד מאתנו האם אנחנו מחבלים או לא?

היות שהחוקים המתמטיים הם רק מסגרת שאינה מודעת לתכנים שאותם היא מייצגת, מהותן של מערכות הלמידה החישובית מצויה בנתונים. במקרים רבים, ככל שהנתונים רבים יותר, כך המערכת עובדת טוב יותר. הגישה לנתונים רבים יותר היא לרוב נוחה יותר לתאגידים ולממשלות מאשר לאדם הפרטי, אך גם כשהמידע חופשי, נדרשות רמות שונות של עיבוד מידע כדי להצליח לאמן את הנתונים. כיום קיימת היררכיה של יחידת עיבוד מרכזית (CPU) במחשבים אישיים, יחידת עיבוד גרפית (GPU) לחברות שעוסקות בלמידה חישובית ויחידת עיבוד טנזורית (TPU) לחברת גוגל עבור מערכות ענק כמו מודל התרגום האוטומטי שלה. מה שגוגל יכולה לעשות, לחברה קטנה יהיה קשה, ולאדם פרטי בלתי-אפשרי.

גם בתסריט שבו כולם אוחזים באותה יכולת חישובית קיימות בעיות. קיימים לא מעט מקרים שבהם יישומים אומנו, בגלל העובדה שפותחו בעמק הסיליקון במפרץ סן-פרנסיסקו, על אנשים שעבדו שם. כך יצא שמערכות מסוימות מכלילות היטב ביחס לגברים לבנים, אבל מצליחות פחות על אוכלוסיות אחרות. מקרים לא נעימים אירעו למצלמת רשת שאמורה היתה לאפשר מצב של מעקב אחרי פני האדם המצולם, ועבדה רק על אנשים לבנים[2], או לתוכנת תיוג שלא זיהתה אנשים שחורים כבני אדם[3].

בעיות מתרחשות גם במקרים שבהם אין החמצת אוכלוסיה או מדגם חלקי. הדוגמה הבולטת היא אלגוריתם Word2Vec, שאומן על אוסף נתונים (dataset) בשם Google-News המכיל טקסטים ממאות אלפי מאמרים. האלגוריתם ממפה מילים למרחב מתמטי, וזכה לפרסום רב על שום העובדה שניתן להשתמש בו לגילוי אנלוגיות: על השאילתה "צרפת לפריז היא כמו איטליה ל-" משיב האלגוריתם רומא. יחסית במהרה התגלה שעל השאילתה "גבר לדוקטור הוא כמו אישה ל-" משיב האלגוריתם אחות[4]. במהרה התגלו הטיות רבות במידע שהוא כביכול אובייקטיבי. בעקבות תוצאות כאלה קיימת מודעות מסוימת בעולם התוכנה שלמערכות למידה חישובית יש אחריות לא רק לייצג את המידע, אלא גם לתקן בו הטיות (biases).

איל גרוס ואני הצגנו בפסטיבל פרינטסקרין 2015 עבודה בשם Word2Dream[5], שבה המשתמשים מזינים למחשב טקסט, והמחשב לוקח כנקודת מוצא את הטקסט, ומפליג מן המילים שלו למילים אחרות באופן אסוציאטיבי בהתאם למקומות אליו מוביל אלגוריתם Word2Vec. אחת התוצאות המעניינות היא שגם כשלא נראה בשלב ראשון שיש הטייה בנתונים, שרשרת אסוציאציות תוביל אותנו לכיוון מוטה. לכן, גם כשמתקנים הטיות נקודתית, התמונה הגדולה עשויה שלא להשתנות. לדוגמה, כשמזינים את נאום "יש לי חלום" של מרטין לותר קינג, ה"חשיבה" האסוציאטיבית של המחשב מובילה לציונות, לפמיניזם ולנאציזם. למעשה, המניע שלנו ליצירת הפרויקט היה להראות שיש משהו מלאכותי וכוללני

ביכולות האסוציאטיביות שלנו כבני-אדם, ושהמידע שבו אנו מלעיטים את הרשת מביא להנצחת הדעות הקדומות שלנו בבינה המלאכותית.

2. למידה חישובית והוספת מידע

עולם הבעיות איתן מנסה להתמודד למידה חישובית הולך וגדל. בתחילת העשור, בעיקר לאור הצלחתן של שיטות למידה חישובית לזהות עצמים בתמונות, החלו להתפתח מודלים שחיקו יכולות אנושיות. למעשה, באותם רגעים הפך המותג "למידה עמוקה", שמייצג למידה חישובית באמצעות רשתות נוירונים, למילה נרדפת לבינה מלאכותית, כלומר בעיות שבתפיסה האנושית נחשבות לכאלה שמחשב לא מסוגל לפתור אבל אדם כן. לכן הושם דגש על ראייה ממוחשבת, זיהוי דיבור, תרגום אוטומטי, ניסיון לנהל שיחה ועוד יכולות אנושיות, כשחוקרים ניסו להבין האם בעקבות ההתפתחויות הטכנולוגיות והאלגוריתמיות בלמידה חישובית מתקבלות תוצאות טובות יותר מאשר קודם לכן. במקרים רבים התשובה היתה חיובית, כמו בראייה ממוחשבת או בזיהוי דיבור, שם במקרים רבים משיגה למידה חישובית תוצאות דומות לאלו של בני אדם. במקרים אחרים התוצאה טובה יותר מאלגוריתמים אחרים, אך עדיין לא מתקרבת ליכולות אנושיות, כמו בתרגום אוטומטי. במקרים רבים הלמידה החישובית לא מניבה תוצאות טובות.

בשנים האחרונות החלה הלמידה החישובית להיכנס לסוג חדש של בעיות, שבהן המטרה היא לקבל קלט עם כמות מידע מסוימת, ולייצר פלט עם כמות מידע גדולה יותר. דוגמה אחת היא מודל שצובע תמונה שמגיעה בשחור לבן, כלומר הוא מקבל תמונה בשחור לבן ואמור להחזיר את אותה תמונה בצבעים. דוגמה אחרת היא מודל שמקבל תמונה שצולמה באמצעות מצלמה רגילה, ומחזיר את אותה תמונה באיכות גבוהה יותר או כאילו צולמה ממצלמה איכותית (DSLR). מחקר חדשני מצא דרך לקחת תמונה שצולמה בתנאי תאורה חשוכים כמעט לגמרי, ולשחזר ממנה כביכול כיצד היתה נראית בתנאי תאורה טובים יותר[6].

התרבות המודלים שאכנה "הוספת מידע", נובעת מהתפתחויות באלגוריתמים ובטכנולוגיות בלמידה חישובית. פופולריים במיוחד הם מודלים שמממשים שינוי סגנון (Style Transfer) של תמונה, כלומר לוקחים תמונה ומייצרים אותה מחדש בסגנון של תמונה אחרת. האמן והתכניתן ג'ין קוגאן מרבה להשתמש בטכניקה הזו[7]. אחת מפריצות הדרך בלמידה עמוקה היא מודל שנקרא (GAN (Generative Adversarial Network שבו (מבחינה רעיונית) שתי ישויות למידה חישובית מתמודדות אחת מול השנייה, וכל אחת משתפרת באמצעות המשוב שהיא מקבלת מרעותה.

איל גרוס ואני הצגנו בפסטיבל פרינטסקרין 2018 פרויקט בשם האוצר האלקטרוני המבוסס על [8]GAN. הפרויקט מייצג התמודדות בין צייר לבין אוצר. הצייר מנסה ליצור דיוקן עשוי ירקות מתוך תמונת המשתמשים, והאוצר מאשר או פוסל כל ניסיון, וכך הם לומדים יחדיו. המחשבה שהניעה אותנו ליצור את הפרויקט היתה העובדה שכמו לבני האדם, גם לבינה המלאכותית יש צורך בדיון, במשוב ובתחרות כדי להגיע להישגים. משהו בהתנהלות "שניים אוחזין" הזו, שמצד אחד דורשת תלות באחר ומצד שני כורכת הישגיות בתחרותיות, הרגיש לנו כמו מאפיין תרבותי מערבי, לעתים בלתי מורגש, שראוי להפנות אליו זרקור.

ישנן טענות[9] שהמוטיבציה לפרויקטי הוספת מידע היא דווקא כלכלית ולא טכנולוגית. יצרנים יעדיפו לייצר מצלמה זולה מבחינת חומרה, שעליה מותקנת תוכנה יקרה שמסירה רעשים ומשפרת איכות, על-פני ייצור של מצלמה יקרה מבחינת חומרה. עלות החומרה משתקפת בכל מכשיר, בעוד עלות התוכנה היא חד-פעמית, ואפס לכל מכשיר נוסף.

מה שמייחד את המודלים החדשים הללו הוא העובדה שהם מייצרים מציאות שלא בהכרח קיימת. כשהופכים תמונה צבעונית לתמונה בשחור לבן, ברור לנו שמדובר בייצוג שהוא אמנם לא אמיתי (כי המציאות אינה בשחור לבן), אבל הוא כן אינדקסיקלי (indexical). כלומר, ברור לנו שהמעבר מצבע לשחור-לבן הוא פרי תהליך חד-ערכי, ושקל לנו לומר לגביו, במידה שנראה את המקור לעומת התוצאה, האם התוצאה היא אמנם ייצוג

נאמן של המקור בשחור לבן. לעומת זאת, כשהופכים תמונה בשחור-לבן לתמונה צבעונית, אובדת האינדקסיקליות. אם בתמונה המקורית יש משאית בגוון אפור כלשהו, אין לנו שום דרך לוודא אם התוצר צריך לייצג משאית ירוקה או כחולה או אדומה. היות שהמידע לא נמצא, ניבוי הוא בגדר המצאה. זו לא עוד טעות במודל או במידע, אלא דבר שונה: יש כאן יצירת מציאות חדשה, בדויה (fabricated), שאולי מנסה לחקות את המציאות המוכרת, אבל היא אינה אותה מציאות. גם אם ניבוי צבע המשאית "יקלע" לצבע הנכון, זה לא יהיה שחזור של המציאות, אלא בנייה של מציאות אחרת, דומה..

3. מציאות "הפוכה"

למידה חישובית דורשת דוגמאות. הדרך הבסיסית ביותר וגם הפופולרית ביותר כרגע לאמן מודל למידה חישובית היא בצורת למידה מונחית (Supervised Learning). בצורה זו מזינים למודל דוגמאות לקלטים ולפלטים המצופים. לדוגמה, בבעיה של זיהוי עצמים בתמונה, מזינים הן את התמונה, והן את העצמים בתמונה, כאשר אנו מאמנים את המערכת שעבור התמונה הזו אלו העצמים הנמצאים בה. קיימות אפשרויות אחרות ללמידה, למשל בפרויקט האוצר האלקטרוני לא מקשרים בין הקלט לפלט, אלא רק נותנים אוסף קלטים ואוסף פלטים. לעתים יש רק אוסף קלטים, למידה בלתי-מונחית, אך עדיין רוב המודלים כיום הם מונחים.

כדי ליצור מודל מונחה, יש צורך לאסוף מידע ולעתים גם לייצר אותו. במקרה של זיהוי עצמים בתמונה, אפשר לאסוף תמונות מהרשת בקלות יחסית. אך כדי להשיג לכל תמונה את העצמים הנמצאים בה אנחנו צריכים לפתור את אותה בעיה שלשמה נבנה המודל. במקרים כאלו מתבצע תיוג ידני, או שמשתמשים באלגוריתמים נקודתיים שפותרים את הבעיה לאוסף מקרים מצומצם. בכל מקרה ישנו מאמץ מסוים לעבור מכיוון הקלטים לכיוון הפלטים. ברגע שקיים פלט לכל קלט ניתן להזין אותם יחד למודל לשם אימון.

במודלים של הוספת המידע מתרחש תהליך דומה אבל שונה. כדי לאמן מודל להפוך תמונה בשחור-לבן לתמונה צבעונית, ניתן אמנם לאסוף בקלות תמונות בשחור-לבן, אבל תהיה זו מלאכה לא פשוטה להפוך אותן באופן ידני לצבעוניות. צביעה ידנית של תמונה היא מלאכה שדורשת מחקר ומחשבה ועל כן צפויה לארוך זמן רב. יחד עם זאת, בקלות ניתן להשיג אוסף תמונות בצבע ולהפוך אותן, באמצעות מחשב, לתמונות בשחור לבן. כך משיגים את שני החלקים של המידע הנדרש למודל. רק שהפעם את הקלט (התמונה הצבעונית) מזינים בתור הפלט המצופה, ואת הפלט (התמונה בשחור-לבן) מזינים כקלט. המהלך הזה הוא לא רק בדיה של מציאות, אלא היפוך שלה: הקלט הופך לפלט והפלט הופך לקלט, הסיבה הופכת לתוצאה והתוצאה לסיבה, הלפני לאחרי והאחרי ללפני.

דוגמה משעשעת לתוצאות של הוספת מידע, או השלמת מציאות, היא פרויקט של האמן קלמנט ואלה בשם גלויות מגוגל ארץ[10]. פרויקט צילומי הלווין של גוגל מצלם מהחלל, ולכן תמונות שכנות מצולמות מזוויות מעט שונות. כדי ליצור תמונה חלקה, נבנה מודל שמנסה להשלים את החסר ולחבר את הצילומים. התוצאה היא גשרים מעופפים, מכוניות מרחפות וכדומה. דוגמה נוספת הקשורה לפרויקט המיפוי של גוגל הוא האי סנדי, שמופיע אמנם בגוגל, ואמיתי לחלוטין מבחינה אלגוריתמית, אך לא קיים במציאות[11].

מעבר לקוריוז שבדוגמאות שלעיל, הבעיה הגדולה היא שהמודלים שעובדים בשיטת המציאות "ההפוכה", מייצרים מציאות אמיתית. אפליקציות שבהן אנחנו משתמשים באופן קבוע לניווט מכילות מציאות לא נכונה, ואנחנו עלולים לנסוע על הגשרים הללו או לשוט לאי הלא קיים. יתר על כן, בעולם שמתקדם לכיוון עזרים חושיים (כמו גוגל גלאס), ושבסביר להניח שבקרוב נרכיב על עצמנו משקפיים שמתמשות בלמידה חישובית כדי "לשפר" את התמונה, אנו מתקרבים למצב שבו איננו יכולים לסמוך על החושים שלנו. וכל זה לא מפני שהאלגוריתם טועה, אלא מפני שאנחנו בוחרים במודלים

של הוספת מידע. אנחנו בוחרים, בידועין, להתנתק מהמציאות שלנו לטובת מציאות חלופית, שלפחות במושגי סיבה ותוצאה היא מציאות הפוכה.

[1]

<https://gcn.com/articles/2018/04/20/ct-scans-luggage-machine-learning.aspx>

<https://www.theatlantic.com/technology/archive/2016/04/the-underlying-bias-of-facial-recognition-systems/476991>

[2]

<https://www.theverge.com/2018/1/12/16882408/google-racist-google-rillas-photo-recognition-algorithm-ai>

[3]

<https://www.technologyreview.com/s/602025/how-vector-space-mathematics-reveals-the-hidden-sexism-in-language>

[/http://eranhadas.com/word2dream](http://eranhadas.com/word2dream) [5]

<https://arxiv.org/pdf/1805.01934.pdf> [6]

[/http://genekogan.com/works/style-transfer](http://genekogan.com/works/style-transfer) [7]

<https://www.youtube.com/watch?v=4sZsx4FpMxg> [8]

[9]

<https://www.e-flux.com/journal/60/61045/proxy-politics-signal-and-noise>

[/http://www.postcards-from-google-earth.com](http://www.postcards-from-google-earth.com) [10]

[11]

<https://www.livescience.com/28822-sandy-island-undiscovered.html>

