

# ביטוח ופיננסים

גיליון מס' 761 ← 17 באוקטובר 2019 // www.inf.org.il

## "ביטוח אלמנטרי יוצר המון הזדמנויות לסוכן"

סו"ב ארז קראוס, חבר ועדת אלמנטר בלשכה, על האתגרים בפניהם עומדים הסוכנים הצעירים, כניסתם לתחום האלמנטרי ועל חברות הביטוח הדיגיטליות • ראיון לקראת כנס אלמנטר ה־33 שיערך באילת בי"א 2019 בנובמבר < עמ' 2

### מקסום עמלות

כך תגדילו את הרווחים ותייעלו את העבודה מול הלקוחות • סו"ב מאיר רוטברג < עמ' 8

### סוכות 2019

פעילי הלשכה על משמעות החג, ארבעת המינים וההווי בסוכה • פרויקט מיוחד < עמ' 5

### משבר בשיאו

נשיא הלשכה סו"ב ליאור רוזנפלד מדבר על משבר הביטוח הסיעודי בישראל ומה ניתן לעשות כדי למנוע אותו < עמ' 4

# למידת מכונה בעולמות המימון וניהול הסיכונים



רועי פולניצר מסביר כיצד יכולים מדעני הנתונים להרחיב את תוצריהם כך ולשפר את עבודתם של ארגונים בהפקת תובנות עסקיות, כלכליות ומחקריות



השלכות מהפכניות על כל תחומי החיים

הצורך להתמודד עם כמויות גדולות של מידע, הנצבר במשך שנים בערוצים רבים, הוליד את הצורך ברכישת יכולות, כלים ושיטות אנליזה, תחקור מידע ובניית מודלים מתמטיים וסטטיסטיים על מנת למקסם את ניצול המידע, לזהות דפוסי התנהגות חוזרים ולחזות התנהגויות עתידיות לצורך הפקת תובנות עסקיות.

מדעני נתונים יכולים להרחיב את תוצרי עבודת המידענות שלהם על ידי ניצול מיטבי של מקורות המידע בהם הם משתמשים ולהוסיף ממד נוסף לארגונים על ידי שימוש ב"מדע הנתונים".

"מדע הנתונים" הינו תחום מבוקש מאוד ובעל השלכות מהפכניות על כל אספקט בחיינו: מקצועית, עסקית, כלכלית, מחקרית ובאופן כללי על אורח החיים שלנו. היכולת לאסוף, לאחסן, לעבד את המידע כדי להפיק ממנו תובנות, להבין טוב יותר את המורכבות של החיים ובעיקר לצפות את העתיד על בסיס הדאטה הקיים, הופך לחיוני לארגונים. בעזרת טכנולוגיה זו של ניתוח מתוחכם

מאמר זה מתמקד בלמידת מכונה (Machine Learning) התחום החם היום במדעי המחשב, באקטואריה ובכינה מלאכותית

של נתונים מעולמות תוכן שונים, שלכאורה לא נראו קשורים קודם, אנחנו יכולים לגלות עולם שלם של משמעויות ותובנות לארגונים.

## פנסיוני פרימיום

השילוב שנותן לך יותר

אפגרייד לפנסיה

פנסיה

מנהלים פרימיום



## טרמינולוגיה משלה

ללמידת מכונה יש את הטרמינולוגיה שלה אשר שונה מהטרמינולוגיה המסורתית המשמשת בסטטיסטיקה. אנו חותמים את המאמר שלנו בסיכום הטרמינולוגיה החדשה. מאפיין הוא משתנה אשר לגביו יש לנו תצפיות. יעד הוא המשתנה אשר עליו אנו רוצים לבצע תחזיות. תוויות הן תצפיות על היעד. למידה בהשגחה היא תחום של למידת מכונה שבמסגרתה אנו משתמשים בנתונים על המאפיינים והיעדים לצורך ניבוי היעד מתוך נתונים חדשים. למידה ללא השגחה היא תחום של למידת מכונה שבמסגרתה אנו מנסים למצוא דפוסים בנתונים על מנת לסייע לנו בהבנת מבנה הנתונים. (בלמידה ללא השגחה אין יעד ועל כן אין גם תוויות).

למידה בהשגחה למחצה היא תחום של למידת מכונה שבמסגרתה אנו מבצעים תחזיות על היעד מתוך נתונים אשר לחלקם יש תוויות (קרי, יש להם ערכים של היעד) וליתר אין תוויות (קרי, אין להם ערכים של היעד). לבסוף, למידה בחיזוקים היא תחום של למידת מכונה שבמסגרתה אנו יוצרים אלגוריתמים לקבלת סדרת החלטות רצופות כאשר מקבל החלטה פועל בתוואי של סביבה משתנה.

הכותב הוא מדען נתונים העושה שימוש במאגרי נתונים גדולים Big Data, תוך פיתוח ושימוש באלגוריתמים של Machine Learning, על מנת לזהות דפוסים ולבצע תחזיות בעולמות תוכן שונים

**פט האימון** - משמש לגיבוש/פיתוח מודלים אלטרנטיביים; **פט האימות** - משמש לבדיקה עד כמה המודלים מכלילים טוב את הנתונים החדשים; **פט הבדיקה** - נשמר בצד לאורך כל התהליך שתואר עד כה ומשמש כמבחן סבירות סופי לרמת הדיוק של המודל הנבחר.

טרם השימוש באלגוריתם של למידת מכונה חשוב מאוד לנקות תחילה את הנתונים. המאפיינים (המשתנים המסבירים) המהווים את הנתונים יכולים להיות נומריים או קטגוריאליים. בכל מקרה עשויים להיות מצבים של חוסר עקביות באופן שבו הנתונים הוכנסו למאגר הנתונים. לפיכך, יש לזהות ולתקן מצבים של חוסר עקביות. חלק מהתצפיות עשויות להיות לא רלוונטיות למשימה הנוכחית ועל כן יש להשמיטן. בנוסף, יש לבדוק שאין תצפיות כפולות או כפילויות בנתונים, דבר שעלול ליצור הטיות. יש להשמיט חריגים אשר נוצרו בוודאות כתוצאה מטעויות הקלדה או מטעויות בהכנסת הנתונים למאגר. לבסוף, יש לטפל בנתונים הסרים באופן שלא יטה את התוצאות.

משפט ביס (נוסחת ביס) הוא תוצאה המשמשת לעיתים כאשר נדרש לכמת את אי הוודאות. משפט ביס הוא דרך להפוך משהו להתניה. נניח שאנו רוצים לדעת מה ההסתברות שמאורע Y יתרחש ונניח שאנו גם יכולים לדעת האם מאורע X אחר שקשור למאורע Y, נקרא לו מאורע X, התרחש או לא. עוד נניח שעל סמך ניסיון אנו יודעים את ההסתברות המותנה שמאורע X יתרחש מותנה בכך שמאורע Y התרחש.

המשיק לתחומי הסטטיסטיקה, האקונומטריקה והאופטימיזציה. למידת מכונה משמשת לפתרון בעיות עסקיות נפוצות, תוך כדי תיאור וקטלוג הבעיה, בניית מודלי חיזוי, עבודה עם מידע לא מאוזן, עבודה עם רגרסיות ומערכות החלטה.

למידת מכונה היא ענף של בינה מלאכותית העוסקת בלמידה מתוך מאגרי מידע גדולים. היא כרוכה בפיתוח וגיבוש אלגוריתמים לצורך ניבוי, סיווג נתונים, או לקבלת סדרת החלטות רצופות תוך אינטראקציה אופטימלית עם הסביבה.

ניתוח סטטיסטי עוסק באופן מסורתי ביצירת השערות (מבלי להסתכל על הנתונים) ולאחר מכן בבדיקת ההשערות באמצעות נתונים. למידת מכונה שונה מניתוח סטטיסטי בכך שהיא לא יוצרת השערות אלא גוזרת את המודל לחלוטין מתוך הנתונים.

## סטים של נתונים

היבט חשוב של למידת מכונה הינו אימות ובדיקה. כלומר, יש לבדוק ולתקן מודלים שנוצרו באמצעו אלגוריתמים של למידת מכונה בעזרת סט נתונים אחר, כזה שלא שימש ליצירת המודל. מודל מורכב מדי עשוי להתאים את עצמו יתר על המידה (ללמוד יותר מדי טוב) לנתונים ששימשו לאימון המודל ובכך הוא עלול שלא להצליח להכליל באופן מספק דיו את הנתונים החדשים. מאידך, מודל פשוט מדי עלול שלא להצליח לתפוס היבטים חשובים של הנתונים. למידת מכונה גורסת שיש לחלק את הנתונים הזמינים לשלושה סטים של נתונים:

blanco

## ביטוח הדירה המקיף של הכשרה חברה לביטוח כל הכיסויים בפוליסה אחת!

**סוכנים וסוכנות יקרים,**  
הכירו את ביטוח הדירה המקיף של הכשרה חברה לביטוח, המציע ראש שקט עם כיסויים מורחבים וייחודיים לנזקי מים, לנזקי דוודים ולנזק תאונתי לבעלי אופניים וקורקינטים חשמליים, במסגרת הפוליסה הקיימת.

**הכשרה**  
חברה לביטוח בע"מ  
יותר דיגיטלי, יותר משתלם

לפרטים נוספים פנו למחוזות \*3453

\*בכפוף לתנאי הפוליסה ולהנחיות החיתום של החברה

