



# איכות

ע"י כריית ביג-דטה

אדית אורי - MSc תעשייה וניהול

[edith@dataalert.co.il](mailto:edith@dataalert.co.il)

# אדית אורי

- מהנדסת תעשייה וניהול, MSc – מיזם דטאלרט להתראות מוקדמות
- מפתחת אלגוריתם GT data mining
- רע"נ איכות באגודת תו"נ, לשכת המהנדסים.

[edith@datalert.co.il](mailto:edith@datalert.co.il)



## – big data

### בעיה או הזדמנות אדירה?

- יש היום נתונים בכמות וברמה טובה יותר מאי פעם... אך בקרת האיכות לא נהנית מהשפע הזה ☹️

- הסיבה: בעיה במודל. ב"א מחיבת נתונים מפוקחים, בעוד שביג דטה הם **unsupervised**.

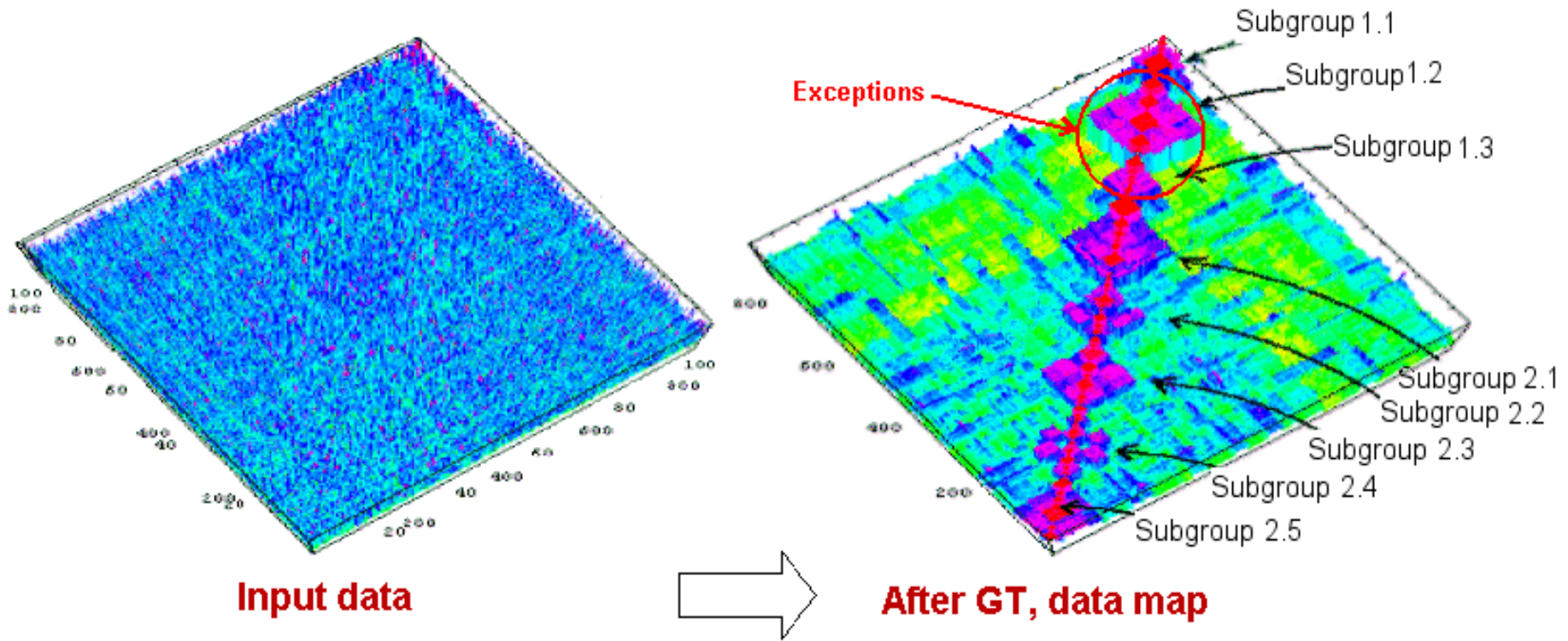
- הפתרון: מודל שמתאים לנתונים לא-מפוקחים...



# אתגר ביג דטה וסביבת IT

- + אינטגרציה מביאה **נתונים לא-מפוקחים**
  - + מידע ממקורות חיצוניים
  - + שינויים מהירים במוצרים בטכנולוגיה ובארגון
  - + משתנים לא-כמותיים
  - + תהליכים שתמיד קשה למדוד – כגון Reworks
- 
- = עיבוד יקר ואיטי

# Self Clustering מודל של GT



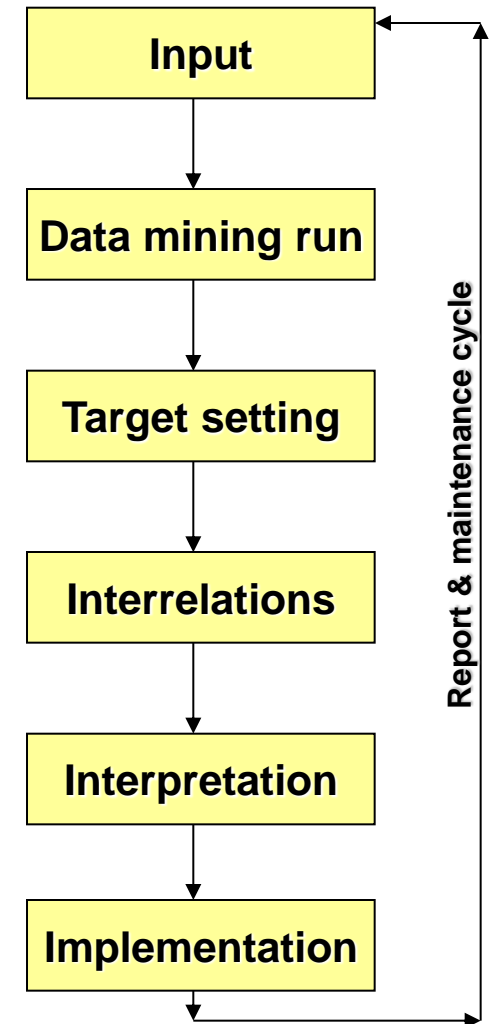
אילוסטרציה של קבוצות ב-GT – זיהוי המבנה הרב מימדי

# פתרון GT data mining

self clustering - מקזז רעשים וטעויות,  
ובכך מגדיל את האיכות של נתונים לא-מפוקחים  
ומאפשר להשתמש בהם **ללא "ניקוי"**!  
ולהקטין בהרבה את העלויות והלו"ז...

# תהליך העבודה עם GT

1. גיבוש הנתונים לקלט, ומסירתם בפורמט טקסט, טבלא שטוחה.
2. כריית נתונים ב-GT, והפקת מפת קבוצות טיפוסיות, אפיון והגדרת גורמי מפתח.
3. דיון משותף על הממצאים דלעיל, והחלטה לגבי נוסח המטרה והתופעות שמעונינים לחקור.
4. מציאת גורמי מפתח, וניסוח הקשרים ביניהם (במרחב הרב מימדי). בדיקת הנוסחאות על סט נתוני מבחן.
5. דיון מפורט על הממצאים דלעיל, עריכת סבב הבהרות במידה ונדרש, וקבלת אישור.
6. הגשת דוח מסכם.
7. יישום, הטמעה, אחזקה.



# התרומה המיוחדת של GT לאיכות

✓ חשיפת **דפוסים חריגים וסמויים**

✓ אבחון סיכונים חדשים

✓ גילוי מוקדם

✓ תגובה מהירה לשינויים

וגם

+ חיזוי פרמטרי שמסיע ל- **scalability**

+ יישום יעיל בלי מומחים יקרים או שינוי ארגוני



# דוגמא - ניתוח תקלות בחב' תקשורת

## הבעיה מבחינת החברה:

- יש יותר מדי תקלות
- התקלות עולות בכסף ובסיכון לנטישת לקוחות
- קשה להפיק מסקנות מנתוני התפעול, בגלל **כמויות המידע, מגוון גדול של שירותים, ושינויים תכופים** בהגדרותיהם ובהסכמים עם לקוחות.

Big-data def.: **V**olume, **V**elocity, **V**ariety

# המשך דוגמא – תקלות בחב' תקשורת

## מקורות מידע

- מח' IT
- שיווק
- ספקי שירות
- מרכזיות
- הנה"ח, Billing
- נציגי שירות ומער' CRM
- כח אדם



# המשך דוגמא – תקלות בחב' תקשורת

## ביג-דטה

- נתוני בקרה עם פירוט תקלות ותיקון חשבונות
- פרטי לקוחות והשירותים שהם מקבלים
- לוג תקשורת
- תהליכי חיוב כולל תחנות שירות אוטומטי/אנושי
- פרטי תחנות השירות ותכנות השירות שלרשותם
- נתונים על עומס במרכזי השירות, מצבת עובדים וכד'
- פרטי ספקים ומחירי השירותים שלהם
- שינויים בטבלאות המחירים, שירותים והסכמים
- דוחות מכירה ו-CRM..



# המשך דוגמא – תקלות בחב' תקשורת לכאורה מצב חסר מוצא

אין מודל אנליטיקס לביג דטה  
והפתרונות הטכנולוגיים לאחסון-אחזור הם כבדים  
ויקרים

...

מה הסיכוי לפתור את זה בעזרת PC?

# המשך דוגמא – תקלות בחב' תקשורת

## איתור הבעיה בעזרת GT

אחת מתת-הקבוצות ש-GT חושף - "ספקי סלולאר קטנים חדשים", מכילה שיעור גבוה של תקלות. מאפיני התת-קבוצה:-

- מחירים גבוהים או היעדר מחיר
- שימוש במסלקה
- תלונות חוזרות
- וראש צוות מסוים בשירות-לקוחות...

מה האפיון הזה אומר לנו?

# המשך דוגמא – תקלות בחב' תקשורת

## שורש הבעיה

קצה החוט הוא הגילוי המוזר, שהמערכת בכלל מאפשרת שיחות שאין להן פרטי חיוב תשלום.

מריבוי התלונות במקרים אלה, ניתן להסיק, שחיובי התשלום שגויים; ומהעובדה שמדובר בספקים חדשים ניתן להבין שפרטי ההסכמים איתם לא ברורים.

על התבנות האלה הולכים לשוחח עם אנשי הארגון...

# המשך דוגמא – תקלות בחב' תקשורת

## ההסבר של אנשי הארגון

- אכן, בתת קבוצה "ספקי סלולר קטנים חדשים" ישנם ספקים שהתעריפים שלהם עוד לא הוגדרו סופית.
- התקלות גורמות לשרשרת פעולות ארוכה במסלקה, ובמח' הנה"ח ושירות לקוחות.
- הערה: ראש הצוות ששמו נזכר, לא קשור לבעיה אלא לפתרון – הוא נחשב למומחה, ומעבירים לטיפולו את כל המקרים הקשים.

# המשך דוגמא – תקלות בחב' תקשורת

**מדוע לא גילו את גורמי הבעיה עד כה!?**

© ב"ניקוי" מדגם הבקרה, נזרקים החוצה השגויים החריגים שחסרים בהם שדות. כך שלא ניתן להבחין בהם אפילו אם עולה החשד.

© כל אחד ממאפייני הבעיה כשלעצמו נראה תקין  
(ספקים קטנים הם עפ"ר מצוינים, תכנות המסלקה הטובות בסוגן, הנה"ח מדויקת...)

© טעות: במערכות IT נוטים להאשים את הגורם האנושי, אך כאן הוא דווקא היה הצד החיובי.



# המשך דוגמא – תקלות בחב' תקשורת

## מסקנות אופרטיביות

1. מניעה  
להגדיר בתת קבוצה "ספקי סלולאר קטנים" תבנית הסכם בסיסי כברירת מחדל.
2. מעקב וגילוי מוקדם  
לעקוב על סימנים ראשוניים לתקלות (כגון עיכובים) בתת קבוצה זו.
3. טיפול מהיר  
למנות מומחה-אחראי שאליו יועברו ישירות התקלות של תת-הקבוצה.

# המשך דוגמא – תקלות בחב' תקשורת

## תבנות

1. איתור גורמי השגויים בתת-קבוצה "ספקי סלולר קטנים חדשים" מאפשר להרחיב את הבדיקה ולאתר שגויי חיוב בספקים חדשים של כל הקבוצות, כנוהל שגרה.
2. מאפייני התלונות החוזרות של תת-קבוצה זו, יכולים לשמש כאינדיקטור לטעויות מערכת בכלל.
3. לבדוק רשומות בעלות מאפייני תת-קבוצה זו אבל "0" תלונות; הן חשודות כטעויות סמויות לטובת הלקוח.

✓ **כאשר הבקרה ממוקדת, היא נהיית פשוטה...**

# הצעה לתקן ISO כריית נתונים ל-QC

1. אבחון קבוצות/דפוסי ההתנהגות היררכיים.
2. חשיפת שורשי תקלות, ע"י ניתוח היררכיה.
3. מניעה, מעקב ותיקון (כלל אצבע: כמחצית התקלות ניתנות למניעה)
4. בחירת גורמים שמופיעים מוקדם בתור אינדיקטורים למעקב ואיתות תקלות.
5. הפניה אוטומטית לטיפול באיתותי תקלות.
6. נוהלי עבודה לקוח-ספק כריית נתונים: אחריות, הגדרות, דיונים, אישורים, יישום, הפעלה, ואחזקה.



# תודה

**Edith Ohri**, MSc Eng., Datalert

הבית של GT data mining

טל": 054-3170161

דואל: [edith@datalert.co.il](mailto:edith@datalert.co.il)