

2/1/22

עיצוב ופיתוח מערכות נתוני עתק ומחשוב ענן
פרויקט סיום תשפ"ב

כללי:

בתקופה הנוכחית, בנוסף לנתונים המופקים ממערכות תפעוליות, מאותרים ונאספים נתונים רבים ממגוון מקורות כבסיס לניתוח, גיבוש תובנות, אימוץ אסטרטגיות וקבלת החלטות מבוססת ראיות.

מערכות Big Data Analytics דורשות משאבים ארגוניים רבים ולכן שימוש בספקי שירותי ענן לאחסון ועיבוד הם פרקטיקה נפוצה. בתכנון מערכת מבוצרת מסוג זה אין מניעה לטווח מגוון שירותים ממספר ספקים, בשילוב שירותים שמקורם בארגון עצמו, לכלל פתרון שלם.

בפרויקט הנוכחי ניצור פתרון לעיבוד ותצוגת נתונים בהשראת תבנית למדה.

מהות הפתרון:

מערכת הבנויה משלוש תת מערכות שיחדיו מאפשרות ניטור שיחות נכנסות ל- Call Center בגישת Near Real Time ובאמצעות Dashboard המציג מדדי מפתח, וכן מאפשר סיווג סוג השיחה מוקדם בכדי לנתב את אופי השיחה המומלץ לנותן המענה.

דרישות פונקציונאליות:

1. נתוני לקוחות (שם, ת.ז, תאריך לידה, כתובת, מין, רשימת מוצרים להם הוא מנוי) ישמרו במסד נתונים תפעולי.
2. המערכת תקבל נתוני שיחות Call Center ותדאג לאחסנם, לעבדם ולהציגם (תמונה 1 כדוגמא).
3. הנתונים הנצברים במערכת בעקבות שיחה:
 - סוג תקופה (חגים, חופשות קיץ, ימים רגילים)
 - שעת השיחה.
 - עיר המתקשר.
 - מין המתקשר.
 - גיל המתקשר.
 - מספר שיחות בחודש האחרון.
 - מוצר נדון (אינטרנט ביתי, כבלים, סלולרי, כולם).
 - נושא השיחה (הצטרפות, שירות, תלונה, בקשת ניתוק)
4. המערכת תציג ב- Real-Time Dashboard את מספר השיחות הממתינות כרגע וזמן המתנה ממוצע בעשר הדקות האחרונות באמצעות גרפים \ Widgets.
5. המערכת תאפשר להציג בטבלה ובגראף את מדדי מספר השיחות הממתינות וזמני המתנה מתחילת היום ברמת אגרגציה של 5 דק'.
6. המערכת תאפשר צפייה במדדים סיכומיים לגבי היום (טבלה) : מספר שיחות הצטרפות, מספר שיחות תלונה, מספר בקשות ניתוק.
7. נתוני השיחות יישמרו לצורך תחקירים וניתוח.
8. המערכת תאפשר למידה מהנתונים וחיזוי אופי שיחה נכנסת (הצטרפות, תלונה, בקשת ניתוק) מיד עם קבלתה.

דרישות לא פונקציונליות:

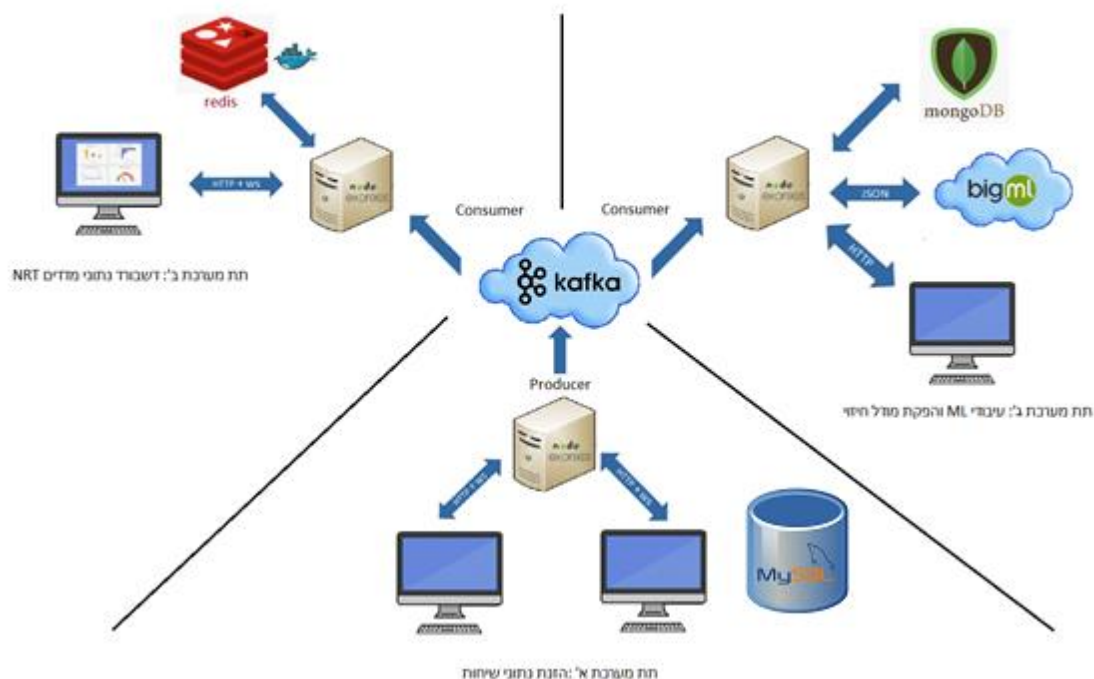
1. ניתן לממש מסד נתונים תפעולי בכל תשתית ובכל מיקום (לדוגמא MySQL לוקאלי/ענן\ Docker)
2. המערכת תפותח בתפיסה מונחית שירותים בגישת [Micro Services](https://www.cloudkarafka.com/).
3. המערכת תתבסס על שלוש תת מערכות, Kafka ישמש כ- Message Broker לאינטגרציה ביניהם.
4. שלוש תת מערכות הן:
 - א': תת מערכת קליטת נתונים.
 - ב': תת מערכת Dashboards ונתוני זמן אמת.
 - ג': תת מערכת אחסון כל נתוני השיחות ויצירת מודל חיזוי באמצעות למידת מכונה.
5. השתמשו בתשתיות Node.js/express, Kafka, Redis, MongoDB וחבילות D3.js או Chart.js לפיתוח הפתרון.
6. Kafka מומלץ להריץ בענן ולהשתמש בספק: <https://www.cloudkarafka.com/>
7. Redis מומלץ להריץ במסגרת Docker עם ה-image https://hub.docker.com/_/redis
8. MongoDB ניתן להריץ לוקאלי, ע"ג Docker Image או משירות ענן כגון: <https://www.mongodb.com/>
9. תת מערכת ב' זוכרת רק נתוני תאריך נוכחי (מחצות עד חצות) על ידי מסד נתונים מסוג Redis ומתאפסת בחצות או על ידי בקשה מהמשתמש.
10. תת מערכת ג' זוכרת את כל נתוני השיחות היסטורית ע"ג MongoDB.
11. מעבר לטעינה ראשונית באמצעות בקשת HTTP, תת מערכת א' ותת מערכת ב' מעדכנים נתונים בינם לבין הדפדפן באמצעות פרוטוקול Socket.io.
12. ארגון השירותים בסביבת Node.js תתבסס על תבנית MVC (השתמשו ב-express כמרכיב View המחולל תצוגה).
13. תת מערכת ג' תאפשר לנתח את הנתונים ההיסטוריים באמצעות אלגוריתם למידת מכונה וליצור מודל חיזוי **נושא שיחה**, על ידי שימוש בשירותי BigML.com וחבילת Node המתאימה. במידה וקיים קושי ליצור מודל מול שירות BigML ניתן להשתמש בחבילת Node לניתוח לוקאלי.

Total Waiting calls:

Answer Call

Start Time	City	Gender	Age	Prev. calls	Product	Topic	End
25-Jun-2020, 9:03	ירושלים	גבר	56	4	אינטרנט	ניתוק	סיום
25-Jun-2020, 9:03	ירושלים	אישה	14	14	כבלים	תלונה	סיום
25-Jun-2020, 9:03	ירושלים	אישה	26	6	אינטרנט	שירות	סיום
25-Jun-2020, 9:04	ירושלים	גבר	67	7	אינטרנט	שירות	סיום
25-Jun-2020, 9:04	ירושלים	גבר	32	2	סלולרי	ניתוק	סיום
25-Jun-2020, 9:04	ירושלים	גבר	33	3	סלולרי	ניתוק	סיום
25-Jun-2020, 9:04	ירושלים	אישה	45	5	כבלים	הצטרפות	סיום

תמונה 1: מסך קליטת נתונים לדוגמא



דיאגרמה 1 : שלושת תת המערכות בפתרון

הערות:

1. ההגשה בזוגות אך ניתן גם בשלישייה.
2. נא להרשם להגשה כבר עכשיו [בלינק הזה](#).
3. מימוש למידת מכונה באמצעות BigML הינה אופציה למגישים בזוגות אך חובה לשלישייה.
4. מוזמנים להשתמש בדוגמאות הקוד במודל אך יש להתאימם לפתרון שלכם.
5. את הקוד ותאור הפרויקט יש לתעד ב- GitHub.
6. תאריך ההגשה : 24.4 באמצעות זום. מועד שני 1.5 מיועד למאחרים.
7. בונוס: פריסת תת מערכת א' ע"ג ספק שירותי אירוח כגון <http://heroku.com> כמתואר כאן: <https://www.geeksforgeeks.org/deploying-node-applications/?ref=rp>

בהצלחה.