## **פיתוח מודל התאמה אוטומטית של הזמנות לחשבוניות במערכת SAP**

### **רקע עסקי**

## במערכת ה-SAP של הארגון, תהליך רכש מוצרים וקבלת חשבוניות כולל מספר שלבים:

## **יצירת הזמנת רכש (Purchase Order)** - בחירת פריטים מהקטלוג הפנימי של הארגון

## **קבלת תעודת משלוח (Goods Receipt)** - אישור קבלת הסחורה במחסן

## **קבלת חשבונית מהספק (Vendor Invoice)** - החשבונית מגיעה עם תיאורי פריטים של הספק

## הבעיה המרכזית: תיאורי הפריטים בקטלוג הפנימי במערכת לא תמיד תואמים לתיאורי הפריטים בחשבוניות הספקים, מה שמקשה על התאמה אוטומטית ודורש התערבות ידנית של רכזי הרכש.

### **השאלה העסקית**

## האם קיימת התאמה בין פריטים בהזמנות לפריטים בחשבוניות ובתעודת המשלוח ?

### **מטרת המבחן**

## פיתוח מודל המספק מענה ל:

## האם חשבונית ספק תואמת להזמנת הרכש - התייחס לכלל הפרמטרים (סכום, ח״פ פריט וכו, נסה לבצע התאמה בצורה האמינה ביותר)

## ניתוח התפלגות האישורים בארגון לפי פריטים וספקים שונים

## זיהוי מקרים של אי-התאמה

## **משימות המבחן**

### **חלק א': הכנה וניתוח הנתונים (30%)**

## **טעינת וחקירת הנתונים**:

## טען את קבצי הנתונים מטבלאות

## בצע ניקוי נתונים והתמודד עם ערכים חסרים

## נתח את איכות הנתונים וזהה חריגים

## הצג סטטיסטיקה תיאורית של התהליך

## **ניתוח דפוסי האישורים**:

## נתח את שיעורי האישור לפי ספקים שונים

## זהה מגמות באישורים לפי סוגי פריטים

## בחן את הזמנים הממוצעים לאישור חשבוניות

## זהה הבדלים בין מחלקות או רכזי רכש שונים

## **Feature Engineering**:

## צור משתנים המבוססים על דמיון טקסטואלי בין תיאורי פריטים

## חשב מדדי התאמה בין כמויות ומחירים

## צור משתנים המבוססים על היסטוריית הספק

\*\*במידה ואוכלוסיים הנתונים לא מספקת בעיניך - ניתן להרחיב את אוכלוסיית הנתונים באופן עצמאי

### **חלק ב': פיתוח מודל בסיסי (30%)**

## **מודל Classification בסיסי**:

## פתח מודל לחיזוי אישור/דחיית חשבונית

## בצע Cross-validation ו-hyperparameter tuning

## הערך ביצועים באמצעות metrics מתאימים

## **מודל Text Similarity**:

## פתח מודל להתאמת תיאורי פריטים

## בחן שיטות שונות למדידת דמיון טקסטואלי

## הערך את יעילות השיטות השונות

## **הערכת המודל**:

## נתח את Confusion Matrix ופרש את התוצאות

## זהה את המשתנים החשובים ביותר למודל

## בחן מקרים של False Positives ו-False Negatives

## הסבר את החלטות המודל

### **חלק ג': ניתוח התפלגות אישורים ושיפור המודל (25%)**

## **ניתוח התפלגות אישורים**:

## צור דוח מפורט של שיעורי אישור לפי KPIS להחלטתך

## זהה ספקים או פריטים עם שיעורי אישור חריגים

## נתח קשרים נוספים להחלטתך

* + בצע סבב שיפורים למודל

### **חלק ד': המלצות עסקיות ויזואליזציה (15%)**

## **ויזואליזציה ודאשבורד**:

## צור דאשבורד אינטראקטיבי המציג kpis להחלטתך על סמך השאלה העסקית

## **המלצות תהליכיות**:

## הצע שיפורים לתהליך ההתאמה הקיים

## זהה הזדמנויות לאוטומציה נוספת

## המלץ על קריטריונים לאישור אוטומטי

## הצע KPIs למעקב אחר ביצועי התהליך

## **תיעוד ומסקנות**:

## תעד את כל השלבים בצורה ברורה

## הסבר את ההחלטות הטכניות שקיבלת לאורך כל הדרך

## **הנחיות להגשה**

1. קוד מלא ומתועד ב-Jupyter Notebook או קובץ Python
2. מסמך המסכם את התהליך, הממצאים והמסקנות
3. הצגת התוצר בכלי ויזואליזציה

## **נתונים:**

### **קבצי נתונים בפורמט CSV (מבוססי טבלאות SAP):**

## **purchase\_orders.csv** (מבוסס EKKO/EKPO)

## **goods\_receipts.csv** (מבוסס MIGO)

## **vendor\_invoices.csv** (מבוסס MIRO)

## **invoice\_approvals.csv** (היסטוריית אישורים)

## **vendor\_master.csv** (נתוני ספקים)

## **material\_master.csv** (קטלוג פריטים פנימי)

1. **Historical\_vendor\_performance.csv (**נתונים היסטוריים על ביצועי ספקים**)**

## 

## **הערות נוספות**

## המודל צריך להיות Explainable - רכזי הרכש חייבים להבין מדוע המודל הגיע להחלטה מסוימת

## שקול אספקטים של Bias - האם המודל מפלה לרעת ספקים מסוימים?

## חשוב על ביצועי המודל בזמן אמת - המערכת תצטרך לעבד מאות חשבוניות ביום

\*\*ניתן לבצע ג'ינרוט של נתונים נוספים אם נדרש (כתוספת על הקבצים הקיימים), נא לתעד זאת במבחן