

# SiS- Smart Invoice Scan

מגישים: רון בן משה, אסף מוזגו ואוריאל אביטל      מנחה: ד"ר עמי האופטמן

## רקע:

עסקים גדולים וקטנים מבזבזים זמן רב עבור הזנה ידנית של חשבוניות המגיעות לעסק שלהם עקב הוצאות. בתוכנות של היום הזנת החשבוניות הינה כמעט התהליך היחיד שמזן ידנית על ידי מנהלי החשבונות.

מערכת SiS תקלוט חשבוניות אלו ותחזיר את הערכים החשובים להנהלת החשבונות להזנת החשבונית למערכת.

זוהי בעיה קשה מאוד לפתרון כיוון שאין חשבונית שדומה לחברתה, כל חברה מעצבת לעצמה חשבונית כראות עיניה ועל כן קשה למצוא כללים כדי לפענח את כלל החשבוניות.

## תיאור המערכת:

המערכת בנויה ממספר חלקים:

- א. תהליך קליטת החשבונית – המערכת תקלוט את החשבונית או החשבוניות דרך ממשק web הזמין גם לפלאפון ותשלח אותו לשרת.
- ב. תהליך סריקת החשבונית תעבור מיטוב כהכנה לפענוח ע"י שימוש באלגוריתמים (BT.709, Hough Line Transform, ImageMagic) ובכך סריקה שבוצעה בצורה לא מיטבית תעבור מיטוב ותהיה מוכנה לפענוח.
- ג. תהליך תרגום החשבונית – החשבונית תעבור תהליך של פענוח באמצעות טכנולוגית OCR (Tesseract) ולאחר מכן תחפש את הערכים הרלוונטיים על ידי הפעלת מסנן על הטקסט המתקבל מהOCR שיסמן את המילים המרמזות על אותם הנתונים שאנו מחפשים מתוך מאגר שהגדרנו מראש. נבצע התאמה בין הנתונים לבין מילות המפתח על פי הסוג והפורמט (מספר, תאריך, טקסט).

ד. תהליך החזרת הנתונים ללקוח- הלקוח יקבל בקובץ Excel את הפרטים הרלוונטיים וישמור אותם בבסיס הנתונים עבור כל לקוח כמתחייב בחוק.

### **אמצעים טכנולוגיים:**

א. צד השרת נכתב בשפת C# ופותח בסביבת עבודה Visual Studio.

ב. צד הלקוח פותח ב-JavaScript ו-Angular.

ג. ה-Database נבנה ב-MongoDB.

### **תוצאות:**

א. פענוח חשבוניות וקבלות באנגלית עם יכולת דיוק של 85% (ללא תלות במספר העמודים).

ב. תהליך הפענוח השתפר במהלך העבודה ואורך מספר שניות בלבד עד להגעת התוצאה.