



al.rachel@gmail.com

טל: 054-4917344

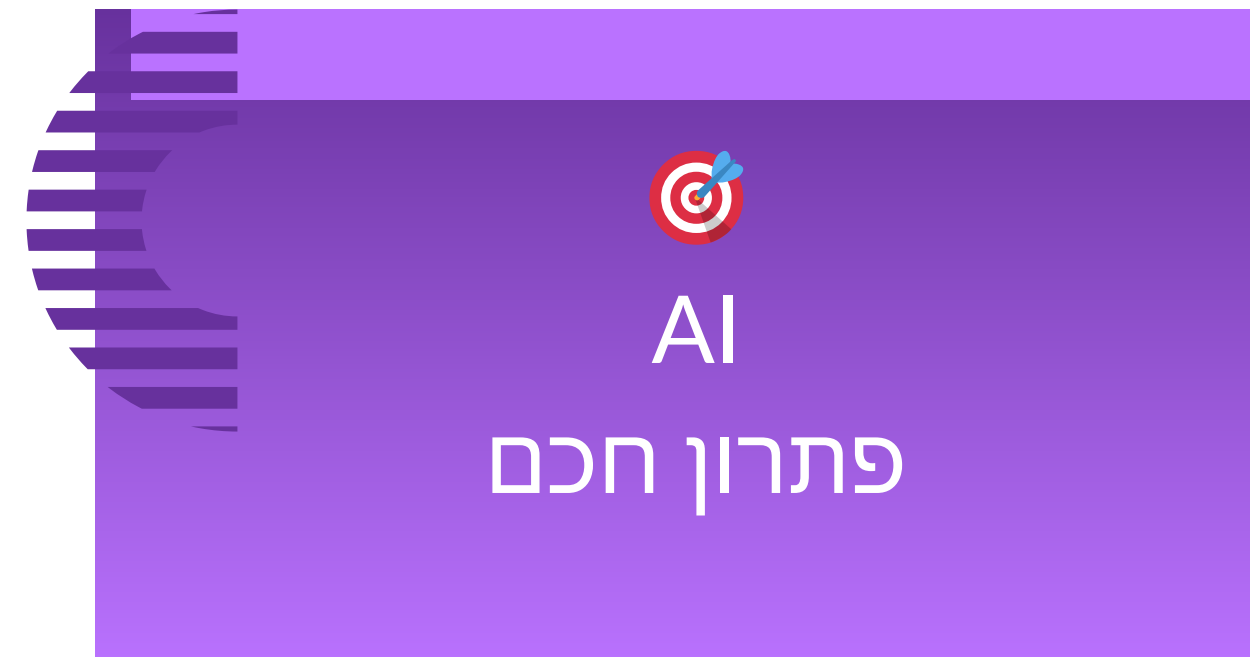
שם: רחל אורנג'

סיווג אוטומטי של משלח יד

מגישה: רחל אורנג'

מעבר מסיווג ידני לטכנולוגיה חכמה

הטמעת בינה מלאכותית בתהליכי סיווג בלמס



הגדרת הבעיה העסקית 🔍

האתגרים

- תהליך ידני איטי וגוזל משאבים 🕒
- אי-עקביות בין סמלנים שונים 🎯
- עלות גבוהה של טיפול ידני 📈
- צורך בעדכונים ותיקונים מתמשכים 🔄
-

המצב הנוכחי

- כל סקר דורש סיווג ידני של משלח היד לפי תקן בינלאומי 🏛️
- עדכונים מרובים של השאלון 👤
- המערכת מאפשרת מילוי חלקי 📝
- כ-10% מהסקרים מקבלים סימול שגוי או חלקי ❌

המטרה: יצירת תהליך AI לביצוע סיווג אוטומטי מדויק ויעיל

ניתוח הנתונים



18

משתנים (features)

62

קודים שונים ל-Target

201

סה"כ רשומות

איכות הנתונים

- הרבה טקסט חופשי
- הזנת מידע לא תמיד במקום המתאים לסוג מידע זה
- ערכים חסרים
- חריגים - קודי משלוח יד עם מקרים בודדים
- מקרים ללא קודי משלוח יד
- שדות קטגוריאליים - עם קבוצות מרובות

הפתרון המוצע 💡

פתרון מבוסס Machine Learning לקביעת סוג משלח יד

- **עיבוד טקסט:** איחוד שדות הטקסטואלים, חילוך מידע וחלוקה לקטגוריות באמצעות NLP, איתור מילות מפתח וכו'
- **Feature Engineering:** יצירת מאפיינים חדשים מחושבים על בסיס המידע הקיים
- קביעת ברירות מחדל
- טיפול בשדות קטגוריה
- טיפול בחריגים
- **אימון ובחירת מודל:** בדיקה והשוואה בין ביצועי מודלים מרובים: Random Forest, XGBoost וכו'
- **הערכת ביצועים:** ובחירת המודל הטוב ביותר
- **Performance Tuning:** אופטימיזציה של פרמטרי המודל

תובנות

- קבוצות קטנות של משלוח יד: קיימים קבוצות קודי משלוח יד עם מקרים בודדים דרוש איסוף מידע נוסף או לחילופין חלוקה לקבוצות על כדי לאפשר חיזוי
- משלוח יד XXXX: הם הסיווגים המורכבים ביותר, נדרש ביצוע סיווג ותחקור יכול לעזור לשיפור המודל
- שיפור איכות הנתונים: קיים מידע שניתן להפוך מטקסט חופשי, לבחירה מתוך רשימת ערכים שדות שמאוד משפיעים על המודל - ניתן להפוך לשדות חובה

הצעות להמשך הטמעת הפתרון

- פיילוט: בשלב ראשון: פיילוט עם מעקב אנושי
- שיפור הסקר: שימוש בתובנות מהמודל כדי לשפר את שאלון הסקר ואיכות הנתונים
- Retraining: אימון תקופתי



al.rachel@gmail.com

טל: 054-4917344

שם: רחל אורנג'

תודה !

מגישה: רחל אורנג'