



| | | |
|------------------------|------------------------|------------------------|
| שם הפרויקט | | מס' פרויקט |
| Sportector | | 2022-01-212 |
| מנחה שותף | מנחה אקדמי | |
| | ד"ר נמרוד טלמון | |
| חברי הצוות | | |
| אושרי פורטל | רוית עמוס | ברק הרן |
| portalo@post.bgu.ac.il | ravitam@post.bgu.ac.il | haranba@post.bgu.ac.il |

תקציר

נושא הפרויקט הינו פיתוח יישום לחיזוי פציעות ספורט של אצנים מקצועיים, היישום שפתחנו נקרא Sportector.

אצנים מקצועיים, לאורך הקריירה חווים פציעת ספורט. פציעת ספורט יכולה לנבוע מחבלה (טראומה), מתיחה חדה, עומס יתר וחוסר הדרגתיות באימונים. זיהוי מוקדם של אפשרות לפציעת ספורט מונע מהספורטאי תקופת התאוששות ארוכה, אובדן כושר והפסדים כספיים.

מטרתו של Sportector הינה להתריע מבעוד מועד על פציעה אפשרית של אצנים. מטרה זו מיושמת באמצעות אלגוריתם Machine Learning, כאשר ממשק המשתמש הינו פלטפורמת WEB, באמצעותה המשתמש מזין את נתוניו האישיים ומקבל בחזרה את סיכוי הפציעה שלו באימון הבא.

מודל לימוד המכונה בו השתמשנו לחיזוי, אומן על בסיס נתונים רחב, שכלל מידע שנאסף במשך 7 שנים, על אימונים של 74 מתאמנים מקצועיים. כל אימון מיוצג בבסיס הנתונים ע"י 10 פיצ'רים אובייקטיבים וסובייקטיביים.

טרם אימון המודל, ביצענו מספר שלבים מקדימים, ניתוח בסיס הנתונים אשר כלל ניתוחים סטטיסטיים על הפיצ'רים ובדיקת השפעתם על משתנה המטרה, בדיקת נכונות הנתונים כמו ערכים חסרים ובדיקת חוסר איזון בבסיס הנתונים.

מניתוח הנתונים הסקנו כי לפיצ'רים סובייקטיביים- נתונים עליהם מעיד המתאמן בעצמו לאחר האימון, יש השפעה גבוהה יותר על משתנה המטרה. כמו כן יש משמעות לתקופה עליה צריך להסתכל בשביל לחזות פציעה- עד שישה ימים לפני האימון.

בבדיקת איכות הנתונים עלה כי בסיס הנתונים אינו מאוזן, על מנת לקבל מודל אמין ולא מוטה, השתמשנו בשיטת under sampling לזיון הנתונים. בסקירת הספרות שביצענו מצאנו כי האלגוריתמים היעילים לחיזוי בעיות בעולם הרפואה הינם אלגוריתמים מבוססי עצי החלטה, ולכן בדקנו 2 סוגים של אלגוריתמים- XGBoost ו-RandomForstClassifier.

הערכת האלגוריתמים התבצעה ע"י מדדים AUC, ו F1_score. נמצא שאלגוריתם Random Forest משיג תוצאות חיזוי טובות יותר עם מדד AUC של 0.75 לעומת 0.7 ב-XGBoost.

תוצר הפרויקט הינו אתר אינטרנטי, פיתוח האתר נעשה בשפות Python, HTML ו-CSS. המשתמש (אצן/מאמן) מתעד את האימונים שלו באזור האישי באתר, נתונים אלו נשמרים בבסיס הנתונים על גבי MySQL. נתוני האימון משמשים אותנו לחיזוי הפציעה באמצעות אלגוריתם לימוד המכונה, וכן לאימון תמידי של האלגוריתם. כמו כן למשתמש יש אפשרות בכל עת לקבל מידע על האימונים הקודמים, ולצפות בתרשימים המציגים את קצב ההתקדמות שלו.

מילות מפתח: Machine Learning, XGBoost, RandomForstClassifier, Python, WEB