

מור יפה ודנית קליש

1. מוטיבציה

הפרויקט מתבצע בחברת AVG Mobile, אשר מתמחה בתחום מערך ההגנה על מחשבים בייתים וניידים מפני וירוסים ותוכנות זדוניות. עבודה זו עוסקת בבניית המודל לומד אשר יאמוד את הזמן הדרוש לבניית יכולות חדשות ומוצרים חדשים בפיתוח תוכנה. מדובר בתהליך שמטרתו לחזות את המאמצים הנדרשים לפתח או לתחזק תוכנה. המוטיבציה שלנו לפרויקט נבעה מתוך ההבנה שקיים צורך אמיתי בכלי משוב ובכלים אשר ינטרלו הטיה אנושית ויהוו כלי עזר חכם לחיזוי נכון ומדויק של פיתוח יכולות אשר יכול למנוע עיכובים בהמשכו של הפרויקט. כמו כן להביא להקטנת התלות הקיימת בפיתוחים עתידיים ובמשאבים מוגבלים הן בהיבט התקציבי והן בהיבט של לוח הזמנים.

2. סקירת ספרות

סקירת הספרות של הפרויקט עסקה בתיאור מודלים לומדים והתפתחותם לאורך השנים. ניתן דגש על כיצד הצורך במודלים לומדים נהיה מורכב יותר לאורך השנים בשל התפתחות עולם התוכנה. בנוסף תוארו מספר שיטות לפיתוח תוכנה: הערכת מומחים, מודל הערכה מבוסס תוכנה ותהליך הערכה המבוסס על מודל והערכה. מסקנתנו מסקר הספרות להביא לכך שהמודל יאופיין על בסיס נוסחאות ונתוני עבר ובכך להקטין טעויות הערכה של הגורם האנושי. וכך למנוע בזבוז משאבים בחברות התוכנה.

3. מתודולוגיה

כדי להשיג את מטרת הפרויקט, בניית מודל לומד, ביצענו שליפה של כל נתוני הפרויקט מתוך מערכת JIRA. בבוצע טיוב לרשומות אלו, בחרנו את השדות שלפיהם נוכל להסיק לגבי זמן פיתוח יכולות/מוצרים. התבצעה המרה של נתונים טקסטואליים לנתונים כמותיים וערכים קטגוריים. זאת על מנת שהאלגוריתם J48 ממערכת WEKA יוכל לבצע הרצה ולייצא את מסקנותיו מתוך הנתונים. הנתונים הני"ל נחקרו לפי CROSS VALIDATION - תהליך ניסוי לשערוך ביצועים עתידיים. בתהליך זה מתבצעת חלוקה מדגם של נתונים לשתי קבוצות: קבוצת אימון אשר מהווה 90% מהנתונים וקבוצת בדיקות שמהווה 10% מהנתונים, עליה מאמתים את התוצאות, מבצעים ממוצע של התוצאות. שלב אחרון של מחקר הנתונים ועיבודם היא בדיקת איכות החיזוי ע"י הרצה של רשומות חדשות ע"פ המודל שנבנה.

4. ממצאים

לשם בניית המודל היינו צריכים לבחון משתנה תלוי – האם התבצעה הערכת יתר או הערכת חסר של משאבים נדרשים. הוחלט לאחר הרצות רבות של האלגוריתם J48 שיקוטלג לשני סוגים:

- הערכת יתר: חיזוי חיובי, זמן הפיתוח המתוכנן היה גבוה מזמן הפיתוח בפועל.
- הערכת חסר: חיזוי שלילי, זמן פיתוח המתוכנן היה נמוך מזמן הפיתוח בפועל.

תהליך בניית המודל הצריך קבלת תוצאות מהימנות עם אחוז מסוים על מנת שנוכל להעריך את המשתנה התלוי ע"י המשתנים הבלתי תלויים. לשם כך, היה צורך לבצע שינויים בנתונים, מחיקת רשומות עם שדות ריקים, ביצוע הפרש בין הזמן המתוכנן לזמן בפועל ולפי כך לקטלג כחיזוי חיובי או חיזוי שלילי. הטיוב התבצע ל-523 מתוך 574 הרשומות שהוגדרו כ"הושלם" – Done. מודל חיזוי סופי מאפשר לנו לדעת אם אפשר אנשי הצוות מבצעים הערכת זמנים נכונה או לא. כ-80% מהמקרים שבו אנשי הצוות הגדירו מעבר לזמן שבפועל נדרש, המודל צדק והצביע על הערכת יתר של זמן פיתוח. מצב זה ללא ספק גורם לחריגה ניכרת ממסגרת התכנון כבר בשלבים הראשונים של מחזור החיים של הפרויקט. ב-28% מהמקרים בהם המודל חזה שהתבצעה הערכת חסר במתן זמנים, הוא טעה, כלומר לא כדאי להסתמך על המודל במצב זה.

5. סיכום

כאמור, התקבלו תוצאות חיוביות לגבי המודל ומשמעותיו בהחלט מספקת ובעלת משמעות לגבי המשך שימוש בחברה. ניתן ואף רצוי להשתמש במודל בפרויקט הבא על מנת לבחון את מתן הזמנים של אנשי הצוות. הערך המוסף אשר אנו הסטודנטים מקבלים מפרויקט היא היכרות עם עולם הבינה המלאכותית וחשיבותו לעולם ניהול הפרויקטים. הבנה שבקרה נכונה ומדויקת יותר תביא לאיסוף ביצועי שעות ועלויות טובים יותר כך שפרויקטים יסתיימו במועד המבוקש, ללא עיכובים ובהתאם לתכנון, תוך חיזוי נכון יותר של היכולות העתידיות שאמורות להתווסף למוצר.