



שם הפרויקט		מס' פרויקט
שערוך ML לפעפוע דיעות		2021-01-266
מנחה שותף	מנחה אקדמי	
	ד"ר נמרוד טלמון	
חברי הצוות		
		תומר לן
		lan@post.bgu.ac.il

## תקציר

הרשתות החברתיות כיום הם לא פחות מכלי עזר חזק ומשמעותי למקבלי ההחלטות ובעלי עניין שונים. בעידן בו ניתן לדעת מה אדם חושב ומהי עמדתו בנושאים מסוימים רק מתוך סקירה של הרשת החברתית בה הוא נמצא, ניתן לומר שעבור סוגיות מסוימות, שיקוף המצב הקיים והתפלגות הדיעות ברשת עשוי להוביל את מקבלי ההחלטות ובעלי העניין למהלכים שיקדמו את מטרותם. יחד עם זאת, מה שתקף להיום לא בהכרח יהיה רלוונטי למחר כאשר מדובר ברשתות חברתיות, מכיוון שדיעות של משתמשים ברשת בסוגיות שונות עשויות להשתנות מכמה סיבות, וביניהם, חשיפה לדיעה השונה משלהם השייכת למקושרים אליהם ברשת. תהליך כזה בו אדם משנה את עמדתו בעקבות חשיפה לעמדה השונה משלו מתקשר למודלים שונים של פעפוע דיעות ברשתות חברתיות, ומכאן מגיע הצורך לחקור את תהליך הפעפוע, ולפתח מודלים לסיווג התפלגות הדיעות ברשת לאחר פעפוע.

פיתוח המודל החל בשימוש בגרף כסט נתונים מרשת "טוויטר" שכלל את מבנה הקודקודים והקשרים ביניהם. לאחר מכן בוצע פיזור באופן סינטטי של תכונות ("דיעות") בין קודקודי הגרף. אופן פיזור הדיעות בוצע בצורה כזו שתהייה משמעות למרחק בין קודקודים, מתוך נקודת הנחה שקודקודים קרובים יותר יחזיקו ב"דיעות" דומות בהסתברות גבוהה יותר. המודל, המבוסס לימוד מכונה, עשה שימוש במאפיינים שונים מתורת הגרפים ולו הוגדרו שלוש משימות למידה שונות עליהן הוא ינסה לענות. בהינתן גרף דיעות התחלתי, המודל יבצע חיזוי לדיעה הרווחת ברשת בסוף הפעפוע, להתפלגות הדיעות ברשת בסוף הפעפוע, וכן גם לרשימת דיעות לפי סדר העדפה, בסופו של הפעפוע. מודל הפעפוע אשר נעשה בו שימוש הינו Threshold model לשתי משימות הלמידה הראשונות, ובחוק בחירות לפי Borda count למשימת הלמידה השלישית. נעשה שימוש במספר אלגוריתמי לימוד מכונה להשגת אחוז דיוק גבוה ככל שניתן, כגון ANN, KNN ו-SVM.

מודלי הסיווג שפותחו הציגו שבהינתן גרף דיעות התחלתי, ניתן לענות על שלושת משימות הלמידה באחוזי דיוק גבוהים. וכן גם שניתן להצביע על מאפיינים מתוך תורת הגרפים להם הייתה ההשפעה הגבוהה ביותר על תוצאות הפעפוע, כגון Centrality Measures.

**מילות מפתח:** פעפוע דיעות, תורת הגרפים, לימוד מכונה