תרגיל 3 – Data Collection & Network Analysis

בתרגיל זה נתרגל ניתוח רשתות.

1. ענו על השאלות הבאות עבור הרכיב הגדול שברשת הנתונים של האנטומיה של גריי:
   1. (15%) למי מהשחקנים ישנו מדד מרכזיות גבוה ביותר לפי:
      1. Betweenness
      2. Closeness
      3. Eigenvector
   2. (15%) הפעילו שני אלגוריתמים לזיהוי קהילות. ציינו מהם האלגוריתמים שהפעלתם וענו על השאלות הבאות עבור כל אחד מהאלגוריתמים:
      1. הדפיסו את הרשתות לפי קוד צבעים התואם את הקהילות.
      2. כמה קהילות התקבלו ע"י כל אלגוריתם. ומה גודלן?
      3. מה ערך ה modularity שהתקבל?
2. השתמשו בAPI להורדת נתונים מתוך רשת חברתית (כדוגמת facebook, twitter וכדומה):
   1. (20%) תארו את תהליך איסוף הנתונים ואת הנתונים שאספתם.
   2. (20%) הגדירו מהם הצמתים והקשתות ברשת שלכם (ניתן לבחון קשרים בין משתמשים, מילים, נושאים ועוד).
   3. (15%) יצרו מתוך הנתונים שאספתם גרף שמכיל עד 200 צמתים.
   4. (15%) ענו על אותן שאלות שנשאלו בשאלה 1.

* על מנת לקבל את מלוא הנקודות עליכם לצרף את קטע הקוד ששימש אתכם ל crawling, ליצירת הרשת ולמענה על השאלות.

פורמט הגשה:

* יש להגיש את התרגיל במאגר github
* המאגר יכיל את הנתונים שנאספו לצד קוד רלוונטי ודו"ח מסכם הכתוב בפורמט RMD
  + הדו"ח ייכתב בשפת markdown וישמש כקובץ README למאגר כשילוב של מלל לצד קוד ותוצאה (הסתייעו בקובץ RMARKDOWN והמירו אותו לקובץ MD + תיקיית תמונות)

**יש להגיש את התרגיל בזוגות עד ל- 24/12/2017.**

תרגילים שיוגשו באיחור "יקנסו" באופן הבא: 5 נקודות יורדו מציון התרגיל לכל יום איחור.

ההגשה תתבצע בתיקיה המתאימה באתר הקורס באופן הבא:

1. על כל אחד מבני הזוג להגיש את העבודה.
2. יש להגיש קובץ ובו קישור למאגר ה Githubשל התרגיל וכן שמות השותפים לתרגיל. שם הקובץ יהיה מהסוג ID1\_ID2.zip (תעודות הזהות של המגישים מופרדות באמצעות קו תחתון).

בהצלחה ובהנאה!