

<yan465@gmail.com> Yaniv Lavi

**תרגול מדעי הנתונים - 29/4**

הודעה אחת

26 באפריל 2022 בשעה 17:30

ניב בר <ItCenter@hit.ac.il>

שלום לכולם,

עד ה- 1/5 התבקשתם לשלוח אליי את הצעות הפרויקט, כמה דברים בנוגע להגשות:

- מבקש לנסות ולהקדים לפני התאריך המדובר, כיוון שיש הצעות פרויקט שמוגשות בפעם הראשונה ואינן מאושרות, ונדרשות לעבור תיקון בהתאם להחלטות שלי. קבוצה/סטודנט שלא יקבלו אישור ממני בזמן, יגררו הורדת ציון בהתאם להנחיות באישור הצעת הפרויקט.

- לפני שתעלו את הטופס לדרייב, מבקש לדייק בסעיפים שהתבקשתם למלא. ישנן טעויות רבות הנוגעות לוחמר הנלמד, אז כמה נקודות חשובות: ניסוח שאלת המחקר צריכה להתבסס על עמודת מטרה שיש לכם בדאטה סט, כלומר על מידע קיים. אין לנסח שאלת מחקר על הקשר בין X ל-Y. בהתאם לעמודת המטרה, התבקשתם לפרט את בעיית המידת המכונה. שימו לב - בעיות קלסיפיקציה - סיווג קטגוריות, ערכים בינאריים. בעיות רגרסיה (למעט רגרסיה לוגיסטית המתייחסת לסיווג בינארי) - מיועדות לערכים רציפים. נסו לוודא כי הדאטה אכן ניתן להרכשה מהאתרים שצייתם, ואני מציע אף לנסות לעשות הרכשה קטנה כדי לוודא שאכן אפשרי.

- בתרגול הקרוב, ב-29/4, אקיים בשני התרגולים הפרונטליים שלי הכנה לקראת בניית הפרויקט. נעבור על השלבים הנדרשים בפרויקט, אתן טיפים ואקסטרה פיצ'רים שיוכלו לשפר את רמת הפרויקט שלכם. בנוסף נראה גם פרויקט לדוגמה וקשיים שאתם עלולים להתקל בהם. בחלק השני של התרגול של כל קבוצה, יינתן מענה לשאלות שיש לכם לגבי הפרויקט.

בהצלחה לכולם!

ניב בר.