

אפרים גולדין

אפריל 2020

1 מערכת GstatCOVID

מערכת Gstatcovid פותחה על ידי חברת GSTAT, בתאום עם המשרד למודיעין של ממשלת ישראל. מטרת המערכת היא לספק כלי תומך החלטות למשרדי הממשלה האמורים להתמודד עם הלחימה במגפה ובתכנון צעדי היציאה ממנה. כמו כן גם לאפשר לציבור הרחב חשיפה למתודולוגיות המידול המתקדמות שבאמצעותן ניתן לחזות את התפתחות המגפה.

מערכת GstatCOVID שונה ממערכות המידע הרבות המציגות דשבורדים של נתוני התפתחות המגפה בכך שהיא כוללת מעבר לדשבורדים גם מודלים סטטיסטיים מתקדמים המאפשרים לחזות את התפתחות המגפה (כמות חולים, מושגים ועוד), מתחילתה ועד לבלימתה. בכך המערכת נותנת כלים רבי עוצמה למקבלי ההחלטות המתמודדים עם האי הוודאות הרבה לגבי התפתחות המגפה.

התחזיות שבלב המערכת מתעדכנות באופן יומי על בסיס המידע העדכני המצטבר מהעולם לגבי תוכניות השחרור ועל השפעתן על מקדמי ההדבקה בעולם. בכך היא מספקת כלי זמן אמת עדכניים למקבלי ההחלטות.

פתוח המערכת נעשה על ידי צוות של מדעני נתונים (Data Scientists) מצטיינים מ-Gstat שבצעו את העבודה בהתנדבות, כשחלקם היו בחל"ת בתקופת הפתוח. המערכת פותחה כתירוץ של צוות המתנדבים ו-Gstat, למאמץ הלאומי הכולל למלחמה במגפה ולתהליכי היציאה ממנה.

פרופסור מיכאל בינסטוק מהאוניברסיטה העברית חבר לצוות ופתח מודל חדשני וייחודי לכימות ההשפעה של צעדי המדיניות על היקף התפתחות המגפה. עבודתו שחלקים ממנה מיושמים במערכת ושבמסגרתה הוא השתמש להדגמת המתודולוגיה בנתוני ישראל, התפרסמה בימים אלו בעיתונות המקצועית (ראה (Xieer Dai & Michael Beenstock)). [דן להוסיף לינק ל-pdf](#)

המערכת זמינה באינטרנט ופתוחה לכל משתמש שרוצה להשתמש בה ולהבין בצורה מעמיקה את תהליכי ההתפרצות והבלימה של מגפת הקורונה בישראל ובעולם, על בסיס מידול מתקדם.

בהמשך אנו מדגימים את השימוש במודל GstatCOVID, לצורך מתן הערכה מבוססת נתונים עדכניים מהעולם לגבי הסבירות להתפרצות גל שני, ולגבי היקפו במידה ויתפרץ.

2 בעיית חיזוי הגל השני

מדינות רבות בעולם נמצאות בתהליכי יציאה מאמצעי המדיניות הנוקשים שננקטו לבלימת התפשטות מגפת הקורונה ולמניעת הצפת מערכות הבריאות בעולם בביקוש למכשירי הנשמה מעבר להיצע הקיים. אמצעי המדיניות העיקריים שננקטו במינונים שונים ובשלבים שונים של ההתפרצות היו בין השאר איסור התקהלויות, סגירת חנויות ומרכזי מסחר, סגירת שדות תעופה ומעברים בין מדינות, סגירת מוסדות לימוד, הורדת היקף העבודה ברמות שונות, הסגרים, עוצר מלא ועוד.

מדד להתפתחות חומרת הצעדים, המוצג במערכת, הינו Oxford stringency Index (Oxford stringency Index). המדד משקלל את צעדי המדיניות בכל מדינה ובכל יום במספר שבין 0 ל-100, כאשר 0 הוא מצב של אי נקיטת צעדים כלל ו-100 הינו מצב של סגר מלא. המדד מאפשר לחקור את הקשרים בין מועדי הצעדים ועוצמתם במדינות שונות בעולם לבין ההצלחה בבלימת התפשטות המגפה.

כיום נעשים מחקרים רבים למדידת הקשר בין רמת הצעדים שננקטו לבין ההשפעה על קצב התפשטות המגפה. הניתוח נעשה בדרך כלל על ידי בדיקת ההשפעה הדינמית של צעדי המדיניות על הורדת מדד R , או R_0 או R הינו מספר האנשים שמדביק כל חולה, בשלב המחלה שבו הוא מדבק. ידיעה וחיזוי של R , ותחת מספר הנחות נוספות, מאפשרת לחזות את רוב המשתנים הקשורים להתפתחות המגפה, כמו מספר החולים החדשים, מספר החולים הקשים, מספר המתים ועוד.

לו היינו צופים בכל החולים שנדבקו, אמידת R וההשפעה של אמצעי המדיניות על שיעורי הירידה בו בשלבי הבלימה, הייתה יחסית פשוטה, שכן מדובר בתהליך גאומטרי פשוט. אולם בפועל אמידת R הינה מאתגרת שכן אנו יודעים כיום שרבים מבין המדבקים הינם חסרי סימפטומים ולא מדווחים כחולים. בנוסף ישנן בעיות של דיווחים לא מדויקים, ועוד. רוב הגישות המקובלות לאמידה מבוססות לכן על הנחות לגבי הגדלים הבלתי נצפים ומתבססות על המודלים האפידמיולוגיים ממשפח SIR (Monte Carlo Markov Chains models), הנאמדים בדרך כלל באמצעות מודל MCMC שכולל בתוכו הנחות אפריוריות על הגדלים הלא ידועים.

במאמר שהתפרסם על ידי (Xieer Dai & Michael Beenstock) **דן מציע שתוסיף על השמות לינק** **pdf** מוצגת הדרך שבה אנו אמדנו את R במודל שלנו, המבוססת על מודל OLG (Over Lapping Generations). במאמר (<https://cepr.org/content/covid-economics-vetted-and-real-time-papers-0>) דיון ביתרונות והחסרונות של גישת OLG לעומת גישת MCMC מבוססת SIR.

בהינתן אומדנים של R במדינות שונות ניתן לבחון באמצעות ניתוחי חתכים לאורך זמן (Panel data analysis, Longitudinal Data analysis...) את האפקטיביות של אמצעי המדיניות השונים על בסיס נתונים השוואתיים בין מדינות שנקטו צעדי מדיניות שונים בנקודות זמן שונות. דוגמא לעבודה כזו היא:

Xiaohui Chen and Ziyi Qiu "Scenario analysis of non-pharmaceutical interventions on global Covid-19 transmissions " COVID ECONOMICS, ISSUE 7 20 APRIL 2020

בעת כתיבת דברים אלו (סוף אפריל 2020) שיאי המגפה ברוב מדינות העולם וישראל כבר מאחורינו ולכן השאלה העיקרית שעל הפרק הינה חיזוי ההשפעה של אמצעי השחרור על העלייה הצפויה ב- R ועל הסבירות להתפרצות גל שני. אם ההשפעה סימטרית, כלומר ההשפעה של אמצעי המדיניות על הירידה במקדם ההדבקה בעת ההתפרצות היא באותה עוצמה, אולם בכיוון ההפוך בעת תהליך השחרור, ניתן להשתמש במקדמים שנאמדו לתהליך ההתפרצות.

אולם לא היא. הנתונים מסין, שנמצאת בשלבי שחרור מתקדמים ביותר, יחסית לרוב מדינות העולם, מלמדים שלמרות שחלק גדול מצעדי המדיניות שננקטו בזמן ההתפרצות שוחררו מזמן, אין כמעט תופעות של התפרצות מחודשת, או מה שנקרא גל שני, אם כי יש במספר פרובינציות עלייה בשיעורי ההדבקה. גם במדינות כמו סינגפור והונג קונג, שבהן נרשמה התפרצות שנייה, היא נבלמה מהר מאוד ולא הובילה לעלייה משמעותית במקדם ההדבקה.

ההסבר לכך הוא כפי הנראה שכאשר משחררים את אמצעי המדיניות המגבילים, איננו חוזרים לאותו עולם. הציבור ממשיך להגן על עצמו ומתרחק מתוך חשש מהידבקות בעת שהייתו במרחב הציבורי,

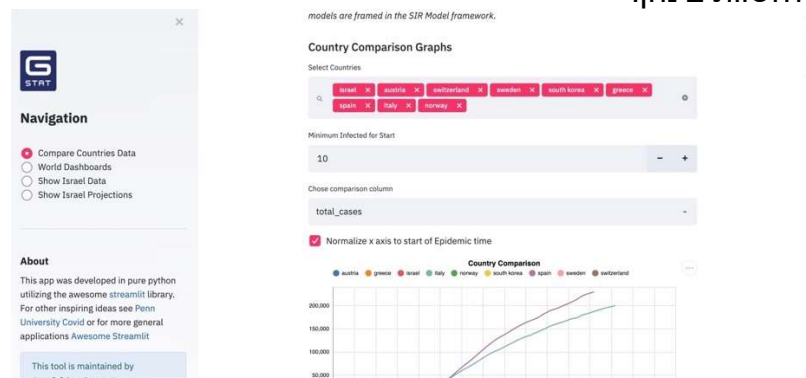
הציבור הולך עם מסכות, סביבת העבודה ואמצעי הבקרה בכניסה למקומות עבודה ולחנויות משתנים ועוד.

מאחר ואין עדיין מספיק דוגמאות ממדינות שונות בעולם על ההשפעה על צעדי שחרור שונים על העלייה ב-R, נקטנו לצורך חיזוי הסבירות של הגל השני בישראל וחיזוי היקפו, בגישה השוואתית המנסה למצות את המירב מהמידע הידוע ממדינות שבהם תהליך השחרור מתקדם יותר במידה זו או אחרת מאשר בישראל.

3 שימוש במערכת GstatCOVID להערכת הסבירות וההיקף של התפרצות גל שני בישראל

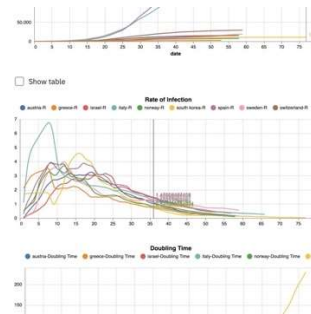
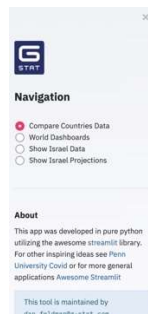
3.1 השוואת התנהגות של מקדם ההדבקה R לאורך תקופת המגפה, במדינות שונות בעולם

מערכת GstatCOVID כוללת יכולות מובנות לחשב את מקדמי ההדבקה במדינות שונות בעולם ולהשוות ביניהן.

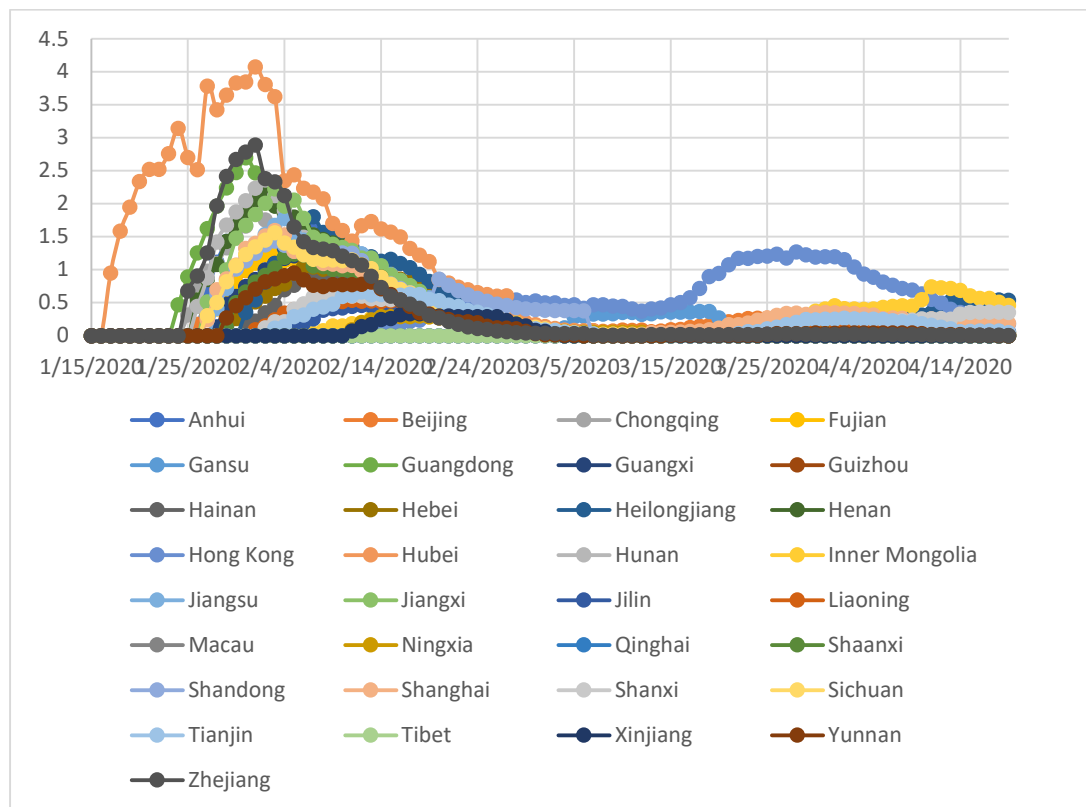


בחירה ב- Compare Countries Data מאפשרת לבחור מדינות להשוואה. ציר הזמן נמדד או בזמן קלנדרי או בזמן קורונה. ברירת המחדל הינה זמן קורונה המוגדר כמספר הימים מאז שנרשמו לראשונה כמות מקרים מאומתים הגדולה או שווה לפרמטר שנקרא Minimum Infected to Start, שבמקרה הנוכחי ערכו נקבע ל-10, אולם ניתן כמובן לשינוי. Choose Comparison Column מאפשר לבחור את המדד להשוואה בין המדינות מתוך מגוון רחב של משתני המגפה (מספר מקרים מאומתים, חולים קשים, מקרי מוות, מספר בדיקות, שיעורים של נתונים אלו למילון תושבים ורבים אחרים). אם נגלול למטה נוכל לראות תמיד את השוואת מקדמי ההדבקה ולאחר מכן את השוואת מקדמי ההכפלה, שהינם טרנספורמציה של מקדמי ההדבקה (ראה דיון במאמר של **Michael Beenstock & Xieer Dai**). גם כאן להוסיף לינק

במקרה הנוכחי בחרנו לצורך השוואת מקדמי ההדבקה אוסף של מדינות שונות מהעולם עם כמויות מקרים שונות. בולט ביותר הוא הדימיון הרב בין תהליכי ההתפתחות של מקדם ההדבקה - התפרצות מהירה עד יום כ-10 לכ -15, ולאחר מכן תהליך דעיכה איטי יותר המתקרב לערך הקרוב לחצי עד קרוב ל-0 בין הימים 55 ל-65.



הגרף הבא לקוח מתוך המאמר של () ומציג את השוואת מקדמי ההדבקה בפרובינציות של סין בין בין תחילת ינואר לאמצע אפריל 2020.



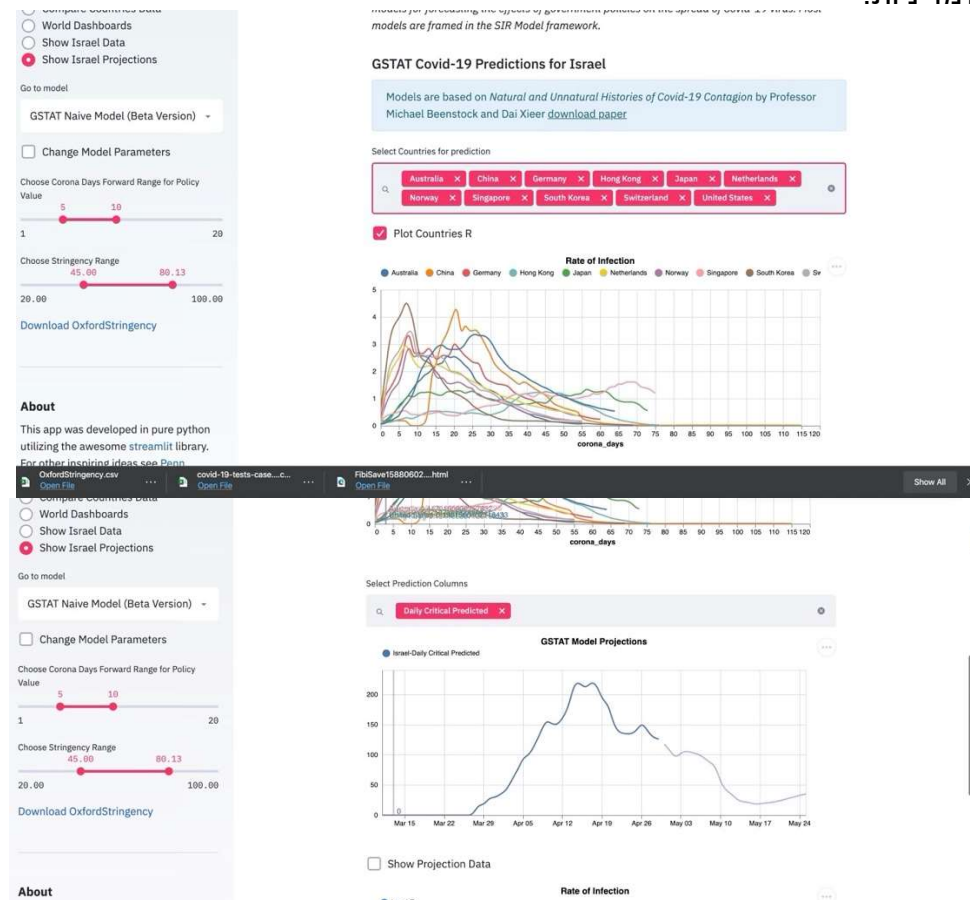
גם כאן אנו רואים תמונה דומה ביותר של התפתחות המגפה, כאשר מבחינים בגל שני בהונג קונג.

3.2 תהליך חיזוי R ומדדי המגפה לישראל

הבסיס לחיזוי של מדדי מחלת הקורונה לישראל מבוסס על לימוד ערכי R במדינות שבהן המגפה בשלבים מתקדמים יותר במידה זו או אחרת מאשר בישראל. תהליך החיזוי מתבצע על ידי בחירת המדינות שיהיו הבסיס לחיזוי ערכי R העתידיים בישראל.

על מנת לבחור ולאחר מדינות דומות לישראל מבחינת מדיניות השחרור ניתן לבחור מדינות שבהן מדד אוקספורד למצב הנוכחי של חומרת הצעדים נמצא בטווח נתון בטווח ימים נתון בעבר. בחירה בטווח מסויים של מדד אוקספורד נעשית באמצעות הסליידר והמערכת מציגה באופן אוטומטי בגרף את ערכי R למדינות העונות לערכים אלו. **דן לא ברור לי מה עושה הסליידר השני אנא הוסף** במקביל נעשה חישוב נוסף. לכל יום קורונה עתידי מחושב הממוצע של מקדם ההדבקה במדינות שנבחרו ומונח שזה יהיה ערכו בעתיד בכל יום קורונה עתידי בישראל. בדרך זו אנו חוזים את R ומתחזית זו נגזרות גם התחזיות של כל שאר המדדים. בררת המחדל לתחזית המוצגת היא תחזית מספר החולים הקשים

בכל יום עתידי בישראל, המהווה חסם עליון לכמות המונשמים, שהינו משתנה הבקרה העיקרי למדיניות.



ניתן לבחור קבוצת מדינות הכוללת מדינות שהיה גל שני של התפרצות בשילוב עם מדינות ללא גל שני, ולקבל תחזית של תרחיש ממוצע. ניתן לבחור רק את המדינות שבהן היה גל שני ולקבל אינדיקטורים למה יקרה בישראל במצב של גל שני, בהנחה שהמהלך יהיה דומה למדינות ההתייחסות שבהן היה גל שני. המשתמש יכול לבחור להשוואה מדינות ספציפיות עם מדינות שחרור ספציפית (חזרה לבית ספר מלאה, רק גני ילדים וכד.), על בסיס מידע חיצוני, ולהתבסס עליהן בחיזוי משתני הקורונה בישראל.

נציג להדגמה שלושה תרחישים שערכנו לחיזוי משתני הקורונה בישראל, במהלך חודש מאי 2020, באמצעות המערכת - אופטימי, ממוצע ופסימי.

תרחיש אופטימי: תחזיות מדדי קורונה בישראל על בסיס מהלך המגפה בסין, דרום קוריאה ושווצריה.

☐ Show Israel Data
☒ Show Israel Projections

Go to model

GSTAT Naive Model (Beta Version) -

☒ Change Model Parameters

GSTAT Model parameters

Tau rate (number of days infectious)

8 - +

Minimum cases for calculation

10 - +

Proportion of asymptomatic

0.25 - +

Daily diagnosis rate

0.0771 - +

Critical Condition Rate

0.05 - +

Recovery Rate

0.5 - +



בהנחה שהמשך ההתפתחות בישראל יהיה כמו בשלושת המדינות הללו, שבהן לא הייתה התפרצות של גל שני, אנו חוזים כמות של 30 חולים קשים בלבד בסוף מאי 2020.

תרחיש ממוצע: התפתחות המגפה בישראל תהיה דומה לממוצע במדינות ללא גל שני (סין, דרום קוריאה ושווצריה) ומדינות שבהן היה גל שני (הונג קונג וסינגפור).



בתרחיש זה ישנה עלייה בכמות החולים הקשים החזויה לסוף מאי 2020 לישראל, אולם לרמה של קצת מעל ל-50, מספר שהינו עדיין נמוך ביותר.

תרחיש פסימי: תחזית מבוססת רק על מדינות שחוו גל שני – הונג קונג וסינגפור



בתרחיש זה כמות החולים הקשים בסוף מאי תתקרב ל-100, בערך כמו בסוף אפריל 2020. גם מספר זה נמוך ובר הכלה. יש לשים לב שגם בהונג קונג ובסינגפור הגל השני לא התפרץ לרמות גבוהות הדומות לשיאי מקדמי ההידבקות בהתפרצויות הגל הראשון בישראל, סין ובמדינות אירופה וארה"ב. כמו כן בהונג קונג הגל השני נבלם מזמן ובסינגפור הגל השני גם במגמת ירידה. בשתי המדינות הגל השני נבע מטעויות ביישום המדיניות. בהונג קונג, לאחר סיום הגל הראשון, לא נעשתה בקרה הדוקה על טיסות שנחתו מאירופה, וההתפרצות מיוחסת ברובה ליבוא מאירופה של חולי קורונה. הערכות מהירה בלמה את הגל. סינגפור התחילה במדיניות סגרים וניטור דומה לזו של דרום קוריאא אולם היישום היה לקוי. ההתפרצות הנוכחית בסינגפור נבעה מהתפשטות מהירה של הנגיף בקרב פועלים זרים שמתגוררים ביחד בצפיפות רבה. מקרים ראשונים שאירעו במעונות אלו הוזנחו, ובדומה לבתי אבות בכל העולם, נרשמה במעונות אלו התפרצות מהירה ביותר.

אם מדינת ישראל תנהיג מדיניות כניסה קפדנית לישראל ותאתר מהר מוקדי התפרצות חדשים הסיכוי לגל שני בישראל בדומה לסינגפור והונג קונג הינו קטן ביותר.

3.3 סכום

הדגמנו את דרך השימוש במערכת GstatCOVID להערכת הסיכוי להתפרצות של גל שני של מגפת הקורונה בישראל ולהיקף הצפוי של החולים הקשים בהתפרצות, בעקבות תהליכי שחרור המשק. הניתוח התבסס על הניסיון הידוע כיום בעולם ממדינות שהחלו תהליכי שחרור לפני ישראל. מהניתוח עולה שגם אם המצב בישראל יהיה דומה למדינות שבהם נרשם גל שני, התוצאות יהיו ברמה ברת הכלה והקפדה על צעדי מדיניות קפדניים בנמל תעופה בן גוריון וניהול סגרים מהיר במוקדי התפרצות חדשים יקטינו לרמה נמוכה ביותר את הסיכוי להתפרצות מחודשת.

במהלך פתוח המערכת בחנו מודלים שונים וגישות שונות. בחרנו לבסוף בגישת הניבוי שנראתה לנו כטובה ביותר מבחינת היכולת לחזות את העתיד ומבחינת מיצוי נכון של המידע החלקי הקיים כיום בעולם על תהליכי השחרור מאמצעי המדיניות.

אנו מצפים ומקווים לקבל ביקורת על המודל ודרך יישומו. אנו משוכנעים שיש דרכים לא פחות טובות משלנו לבצע את החיזוי. GstatCOVID הינה פלטפורמה המאפשרת יישום במקביל של מספר מודלים. אנו מזמינים את כל מי שפתח מודלים בתחום ורוצה לחשוף אותם לציבור, להעלות אותם למערכת

GstatCOVID. אנו יודעים כיום מתחרויות KAGGLE שהמודלים בעלי כושר הניבוי הטוב ביותר היו ברוב המקרים אלו ששילבו מספר גישות חיזוי שונות. שילוב מודלים לקבלת הניבוי הטוב ביותר, דיון ציבורי על המודלים, ודיונים מקצועיים על ההנחות והמתודות יכולים לשפר את טיב הניבויים, וזאת בניגוד למצב שבו נעשה שימוש במודלים שלא נחשפים לביקורת ציבורית ומקצועית.

אנו מודים לאלישע ושאר עובדי המודיעין על התמיכה והליווי שקיבלנו מהם במהלך הפתוח. להסברים ולהבהרות בנושא המערכת והמודלים ניתן להפנות למפתח המערכת דן פלדמן (והרבה תודות לו ולשאר צוות המתנדבים מ-Gstat שעבדו עימו: אלישר חודרוב, רועי עסיס, עוז מזרחי וצוותו ואניה סורקין)

danmsf@gmail.com