

&lt;edenede2@gmail.com&gt; Eden Eldar

**המשך דיון בנושא הנחיה לפרויקט התזה - עדן אלדר**

18 ביוני 2025 בשעה 9:30

<judith\_somekh@is.haifa.ac.il> **Judith Somekh**  
<edenede2@gmail.com> E E אל:כן  
תודה

On Wed, Jun 18, 2025 at 9:29AM E E &lt;edenede2@gmail.com&gt; wrote:

הי יהודית,

כן, אני יכול להפגש לפגישה קצרה ב14:00.

תרצי שאשלח זימון לפגישת זום ?

<judith\_somekh@is.haifa.ac.il> Judith Somekh מאת 9:24 ב-18 ביוני 2025  
הי עדן,  
אוכל להיפגש לפגישה קצרה ב 14, אתה יכול?  
יהודית

On Tue, Jun 17, 2025 at 8:39 PM E E &lt;edenede2@gmail.com&gt; wrote:

סגור יהודית,  
רק רציתי לוודא

תודה על המענה

עדן

<judith\_somekh@is.haifa.ac.il> Judith Somekh מאת 16:59 ב-17 ביוני 2025  
הי עדן,  
תודה שכתבת,  
לא היית אמור להשתתף בפגישה היום, בנוסף לאור המצב דחיתי את הפגישה לכרגע.  
לגבי מחר אם יתאפשר לי אכתוב לך מייל מחר בבוקר.  
יהודית

On Mon, Jun 16, 2025 at 9:30 PM E E &lt;edenede2@gmail.com&gt; wrote:

היי יהודית

מקווה שאני טופס אותך בזמן טוב וששלומוך טוב ככל הניתן עם התקופה והימים האחרונים.

רק רציתי קצת להתעדכן איתך לגבי המשך פגישות ועבודה בהתאם למצב.  
האם הפגישה שהבנתי שאמורה לקרות מחר זו פגישה שחשבת שאולי אצטרף בשביל להקשיב או שזו פגישה שאת וצוות  
המחקר מקיימים (הייתם אמורים לקיים) ?

בנוסף, בודק לגבי רביעי מאחר וזה עבר לזום, היית רוצה שננסה לקבוע פגישת זום באותה השעה שקבענו להפגש במשרד?

מקווה שיהיה המשך ערב רגוע  
תודה רבה,  
עדן

----- הודעה שהועברה -----

&lt;edenede2@gmail.com&gt; E E מאת:

תאריך: יום ה', 12 ביוני 2025 ב-18:17

נושא: Re: המשך דיון בנושא הנחיה לפרויקט התזה - עדן אלדר  
אל: <Judith Somekh <judith\_somekh@is.haifa.ac.il>

In addition, I am attaching the article about GraphSAGE that inspired my suggestion. If it sounds like a model we can integrate into Shaked's workflow, we can also discuss an interesting aggregation architecture suitable for our goal.

Eden

בתאריך יום ה', 12 ביוני 2025 ב-17:44 מאת E E <edenede2@gmail.com>:

Thank you for forwarding the US group's feedback.

I would happily join Tuesday's Zoom meeting with the American group and schedule a Zoom meeting before Tuesday to discuss the points.

We can do it on Sunday until 10:00 AM, after 12:00 PM, or on Monday.

In the meantime, I would categorize the intriguing remark from the US group into three somewhat related points:

1. The gene permutations in Shaked's paper answer the question "Is the observed module unusually age-sensitive relative to random gene sets of the same size?"

They do not provide a proper null for "Does age itself alter network connectivity?" because the original Young/Old labels stay intact.

To test that directly, we can add a sample-shuffle step: randomly reassign age labels, recompute the correlation matrices and MDC, repeat >1000 times, and compare each observed MDC to this new null distribution. If the four CT-GOC modules (IDs 1, 13, 47, 66) stay significant, our evidence for an age effect is stronger. If some modules lose significance, we will know which modules are unstable.

2. In Shaked's workflow, markers were pre-selected by differential co-expression rather than differential expression. Although MDC itself relies on the young/old split, every downstream step, LASSO feature selection, and all hyperparameter tuning, was executed only inside the inner folds of a nested, stratified cross-validation. Thus, the held-out test fold never saw either the LASSO chosen gene list or parameters estimated from its own labels, greatly reducing information leakage. For complete isolation, I think we could even recompute the MDC and module selection inside each inner loop.

3. My understanding is that the two beta schemes were adopted for the same reason highlighted in Koplev's paper: cross-tissue correlations are much weaker than within-tissue ones. When Koplev tried a single beta, almost all cross-tissue links disappeared, and the resulting modules were overwhelmingly tissue-specific, tuning beta intra to 6 for within tissue edges and a lower beta inter 3 for between tissue edges restored scale-free topology in both layers and recovered biologically meaningful cross-tissue modules. Shaked's workflow follows that logic, each beta chosen separately to reach the same R-squared threshold, so weaker inter-tissue signals are not lost while the stronger intra-tissue structure is left intact. Of course, we can still grid-search beta pairs.

Regarding GraphSAGE, I recently started to become curious about it and the possible applications of GNNs in general. For this specific topic, in my view, GraphSAGE fits our use case because it is inductive and local, as it learns a neighbourhood aggregation rule once, then applies the same rule to any new sub-graph, even if that donor is missing one or more tissues. Practically, we could build a heterogeneous graph where each node is gene x tissue (or module x tissue) and each donor brings only the nodes that correspond to their tissues actually sampled for that person. During training, GraphSAGE samples mini-batches of those per donor graphs, passes messages across intra and inter tissue edges, and produces a donor-level embedding that we map to age or AD risk. This achieves the goal without rewriting pairwise correlation code and, because we can pass the intra/inter flag (or the  $\beta$ -weighted edge weight) as an edge feature, GraphSAGE can learn the optimal weighting itself, addressing the US group's third concern about parameter sensitivity.

Hope that my short opinion was based, and I would be glad to continue developing Shaked's article ideas and algorithm, or to discuss those points together in the meeting. I would point out again that I am new to the field of network analysis, so I would try to learn more about it and be more accurate in my terminology and understanding of the methods for our next week's meeting.

Thank you

Eden

בתאריך יום ד', 11 ביוני 2025 ב-15:24 מאת Judith Somekh <judith\_somekh@is.haifa.ac.il>:

I want to say that I'm impressed by how quickly you grasped this topic.

On Wed, Jun 11, 2025 at 3:22 PM Judith Somekh <judith\_somekh@is.haifa.ac.il> wrote:

Hi Eden,

There is a research group that has all the Alzheimer's data and we plan on collaborating with in the US for improving the algorithm,

We have a meeting next Tuesday 17/6,

These are their comments for the work we published:

"Gene-label permutation is fine if your primary question is whether the chosen module is special relative to random gene sets. However, if the question is whether aging alters connectivity, you need a permutation scheme that breaks the age-network link—i.e. shuffling sample labels or resampling individuals—so that the null distribution reflects what would happen in the absence of an age effect. Therefore, to investigate the effect of aging, we should perform sample shuffling and recompute the correlation and MDC for each shuffled iteration.

Predicting age from omics data is a relatively easy task. I'm curious how differential coexpression analysis contributed to improving prediction accuracy compared to models without pre-marker selection based on differential coexpression. I also wonder how they addressed the potential issue of information leakage from marker selection into the machine learning modeling process as the labels (e.g., old vs. young) were already considered during marker selection. Although the selection was based on differential correlation rather than gene expression levels, the two metrics may still be associated.

In addition, the authors noted that inter-tissue expression correlation tends to be lower hence used a different parameter for those correlations. The resulting modules might thus be highly dependent on the choice of parameters for weighting intra vs. inter-tissue correlation."

Let me know what do you think on that,

GraphSAGE sounds interesting idea though I'm not sure how it solves our problem,

I thought to improve the algorithm by creating the correlation for pairs of tissues each time (for each two genes and two tissues) which eliminate the need to have all the tissues sampled for each individual,

Let me know your thoughts about their comments and generally - We can meet before next Tuesday via Zoom to discuss,

Yudit

On Wed, Jun 11, 2025 at 1:18 PM E E <edenede2@gmail.com> wrote:

אפשר אולי להשתמש במנגנון אטנשן בשביל בסופו של דבר לקבל את הרקמות המעניינות וכך גם להתגבר על החסרון של אינטרפרטציה, אבל אני צריך יותר לקרוא על הנושא ואחפש דוגמאות של השימוש במודל בקונטקסט הזה.

בתאריך יום ד', 11 ביוני 2025 ב-13:16 מאת E E <edenede2@gmail.com>:  
הי יהודית,

תודה, יום רביעי בשעה 11:00 בחדר 111 נשמע מעולה מבחינתי.

אם אני מבין נכון את מדברת על כך שעבדתם במאמר של שקד עם רשת "גלובלית" קבועה שנבנתה על סמך נבדקים שיש להם את כל שלושת הרקמות וכך נוצר מדגם קטן.  
חשבתי האמת שאולי ניתן לפתור את מגבלת החיתוך (הצורך בכל הרקמות לכל תורם): חשבתי ש-GraphSAGE יכול לעקוף אותה. המודל לומד פונקציית אגרגציה מקומית, ולכן אפשר לבנות עבור כל תורם גרף-הטרונגני של "גן-ברקמה" גם כשחסרות רקמות – בלי לוותר על דגימות. זאת אומרת, לכל נבדק יש גרף שמכיל את הצמתים (gene, tissue) והקשרים ביניהם ואז נבדקים שאין להם את הרקמה פשוט לא יכללו את הצמתים הללו.

אני משתדל להיות צנוע ויכול להיות שזה לא ממש הכיוון, אבל נשמע לי שזה כן יכול להיות כיוון מעניין.  
מבחינת הוורקפלו, מחליפים את שלב יצירת הרשת המלאה של שקד בבנייה אוטומטית של גרף פר-תורם, ואז מריצים את GraphSAGE על האוסף כולו. כך נוכל לשלב כמה רקמות או שכבות דאטה נוספות לאותה רשת לחקר אלצהיימר.

אשמח לשמוע מה דעתך ולפרט בפגישה.

תודה רבה,  
עדן

Judith Somekh <judith\_somekh@is.haifa.ac.il> עדן בתאריך יום ד', 11 ביוני 2025 ב-5:57 מאת

הי עדן,

אני נמצאת שבוע הבא ביום ד לסמינר של דנה, אוכל להיפגש איתך לפני הסמינר בשעה 11 יתאים? אני בחדר 111 קומה 1 המטרה היא לשפר את האלגוריתם של שקד (למשל אחת הבעיות שם שהיא חייבת חיתוך שיהיה את כל הרקמות ברשת לאדם ואז יוצא מדגם קטן - תחשוב אם יש לך רעיון לשיפור ונדבר), וכך שנוכל לחבר הרבה רקמות או הרבה סוגים של דאטה לרשת אחת לחקר אלצהיימר יהודית

On Tue, Jun 10, 2025 at 9:03 PM E E <edenede2@gmail.com> wrote:

אהה מעולה,

אני חושב שהמשפט האחרון יצא לי קצת מבולבל..

הכוונה שלי הייתה ששבוע הבא אני יכול כל יום חוץ מחמישי

Judith Somekh <judith\_somekh@is.haifa.ac.il> בתאריך יום ג', 10 ביוני 2025 ב-20:59 מאת

הי עדן

אשלח לך עוד חומר במייל נפרד

בשבוע הבא בחמישי לא אוכל ואח"כ טסה לכנס

ביום ד אגיע לסמינר של דנה

יהודית

On Tue, 10 Jun 2025 at 20:58 E E <edenede2@gmail.com> wrote:

היי יהודית

אכן מעניין אותי ותודה על העדכון

באותה הזדמנות, אשתף שאני עדיין חוקר ומתעמק בעולם התוכן של המאמר של שקד. כרגע זה הביא אותי קצת לחקור ולהתעמק בנושא של מערכות מורכבות וניתוח רשתות בביולוגיה של מערכות, הוא מכיל המון שיטות חדשות מבחינתי ואני הנהנה ללמוד ולהתנסות בהן. הייתי רוצה להמשיך להתעמק בנושא אבל גם הייתי רוצה קצת לשמור על קשר ואולי לדון קצת בשבוע הבא אם זה בסדר מבחינתך.

מה דעתך על יום כלשהו בשבוע הבא ? יש יום מסוים שהוא יותר נוח מבחינתך?

אני כרגע לא אוכל להפגש בשבוע הבא רק בחמישי.

Judith Somekh <judith\_somekh@is.haifa.ac.il> בתאריך יום ג', 10 ביוני 2025 ב-6:35 מאת

הי עדן

דנה תציג את סמינר התיזה שלה ב 18.6

אולי יעניין אותך

יהודית

On Fri, 30 May 2025 at 12:23 Judith Somekh <judith\_somekh@is.haifa.ac.il> wrote:

נסיעה טובה!

On Tue, 27 May 2025 at 21:45 E E <edenede2@gmail.com> wrote:

היי יהודית,

קודם כל תודה רבה על המייל וההבהרה

נעים לשמוע שהועבר רושם טוב.

אני לא תכננתי להרפות, לפני שאני מתעמק בנושא ומבין את רמת הנגישות והסקרנות שלי ביחס לנושא.

קראתי את המאמר של שקד וייבאתי את הדאטה סטים של שלושת הרקמות GTEX וקבצי המטא דאטה על מנת לשחק איתם קצת ולחקור מעט את הנתונים שעבדתם איתם. בינתיים המוטיבציה של המחקר נגישה ומוכרת לי גם מאחר והאנטרקציה הבין מערכתית מתקשרת לי מעט לפרויקט שעשיתי בנושא המנגנון של צום לסרוגין מוביל בסופו של דבר להתייעלות במערכת העצבים ועליה בבטוי BDNF אצל גל ומרענן לחזור לקרוא ספרות שקשורה לנושאים האלו: )

היה לי מעניין לראות את האופן שמידלתם את האנטרקציות בין התאים ובין הרקמות, אני מודה שהמודל באמת עדיין לא מובן לי ב100 אחוז כמו גם עולם התוכן של המאמר ואני צריך לקרוא אותו שוב ולהתעמק יותר. כן עניין אותי ללמוד על האופן בו השתמשתם בכלים ושיטות יחסית סטנדרטיות למידול הנתונים ולמטרת המידול של קבוצות הגנים על פי התהליכים והמערכות.

אשמח לקבל את התזה של דנה ולקרוא גם אותו אחרי שאתעמק מעט יותר במאמר של שקד.  
בנוסף, אני חושב שאולי אחרי שאקרא את שני המאמרים אולי נקבע לדבר שוב עוד שבועיים,  
מאחר ושבוע הבא אני לא בארץ, ולדון קצת בנושאים של המאמרים.  
מה דעתך?

שוב תודה והמשך ערב נעים  
עדן

בתאריך יום ג', 27 במאי 2025 ב-8:04 מאת Judith Somekh <judith\_somekh@is.haifa.ac.il>:  
הי עדן,  
חשוב שהשיחה בינינו אתמול לא תרפה את ידיך  
המטרה היתה להבין מה החוזקות שלך והנטייה הטבעית ולהתאים פרויקט מתאים  
בתחום שלי הבנה ביולוגית יכולה להועיל מאוד  
על אף שאני די עמוסה, אני התרשמתי ממך מהישרה ומהרצינות שלך  
כפי שדיברנו אנא קרא את המאמר ויכולה לשלוח לך גם את התיזה של דנה, כדי שתראה אם הנושאים  
האלו מעניינים אותך  
יהודית

On Mon, 26 May 2025 at 8:53 E E <enedede2@gmail.com> wrote:

היי יהודית  
בוקר אור

מקווה שאני תופס אותך בזמן טוב :

שוחחנו לפני כשבועיים על האופציה להנחיה של פרויקט התזה תחתך ותחת מור,  
קבענו לשוחח כשבועיים לאחר הפגישה על מנת להתעדכן ולחשוב שוב על הנושא.

המשך יום טוב,  
עדן