

Temporal Arcs of Mental Health: Patterns Behind Changes in Depression over Time



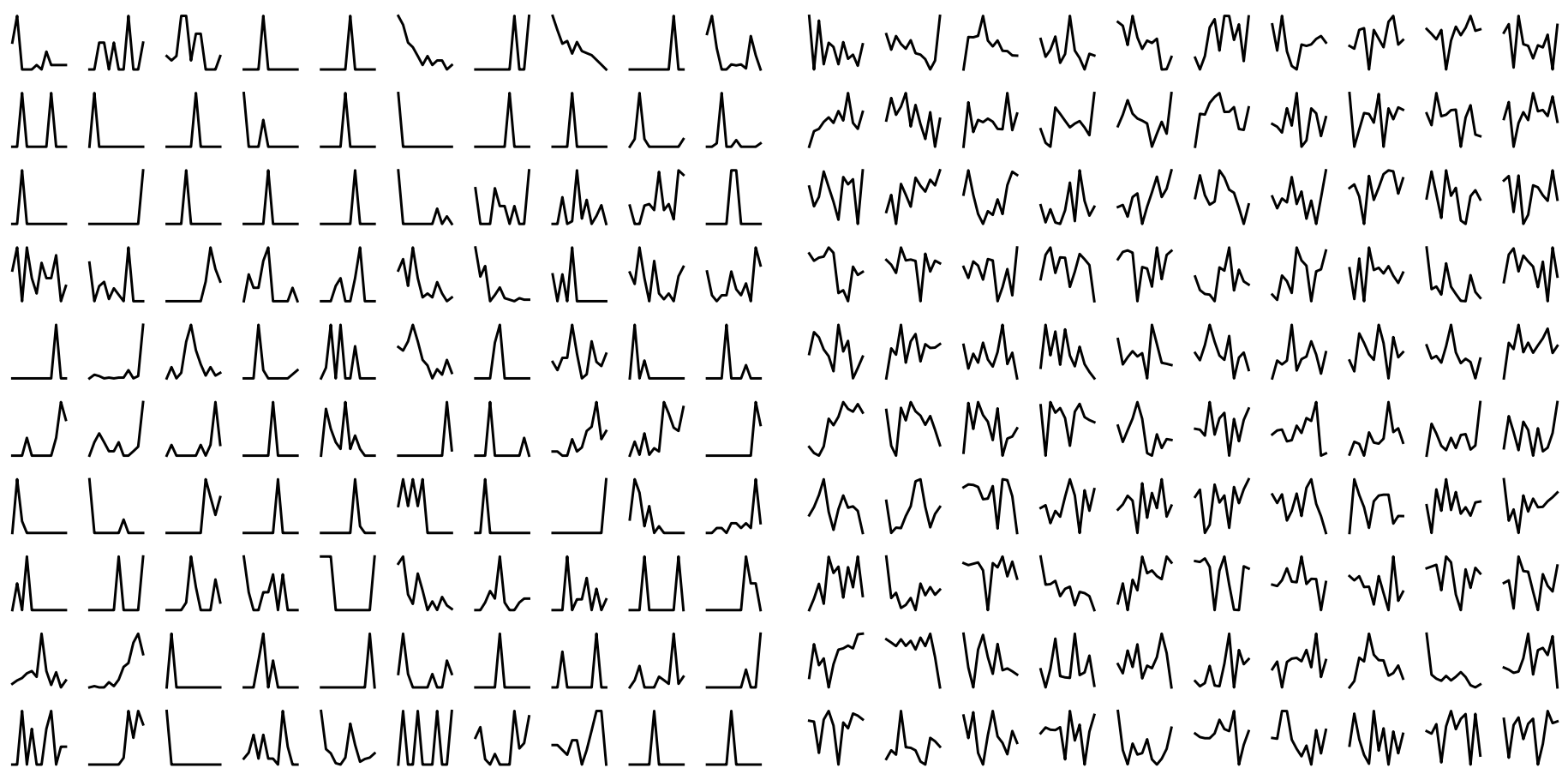
Laura Biester, James W. Pennebaker, Rada Mihalcea
lbiester@middlebury.edu, pennebaker@utexas.edu, mihalcea@umich.edu

Motivation

- Depression is more common in the colder, darker months
- Seasonal affective disorder (SAD) diagnosis:
Regular onset of major depressive episodes at the same time of year, with full remission between episodes [1]
- In computational studies:
 - Search keywords like “depression” peak in the winter [2]
 - Depression classifiers tend to predict higher levels of depression in the winter than summer [3]

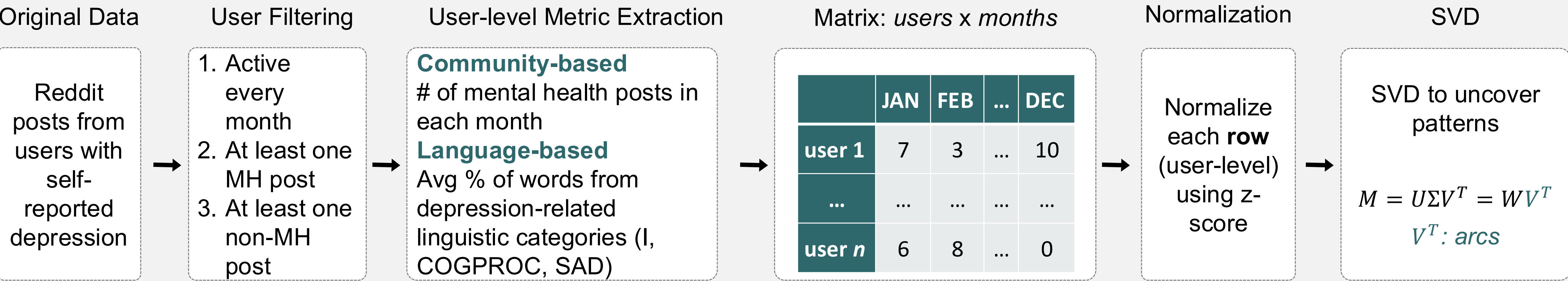
Motivating question: How does mental health change over time and across seasons, and how do those changes vary among users?

[1] American Psychiatric Association. (2022). Diagnostic and statistical manual of mental disorders (5th ed., text rev.).
[2] Yang, A. C., Huang, N. E., Peng, C.-K., & Tsai, S.-J. (2010). Do Seasons Have an Influence on the Incidence of Depression? The Use of an Internet Search Engine Query Data as a Proxy of Human Affect. PLOS ONE, 5(10), 1–7.
[3] Schwartz, H. A., Eichstaedt, J., Kern, M. L., Park, G., Sap, M., Stillwell, D., ... Ungar, L. (2014). Towards Assessing Changes in Degree of Depression through Facebook. Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 118–125.



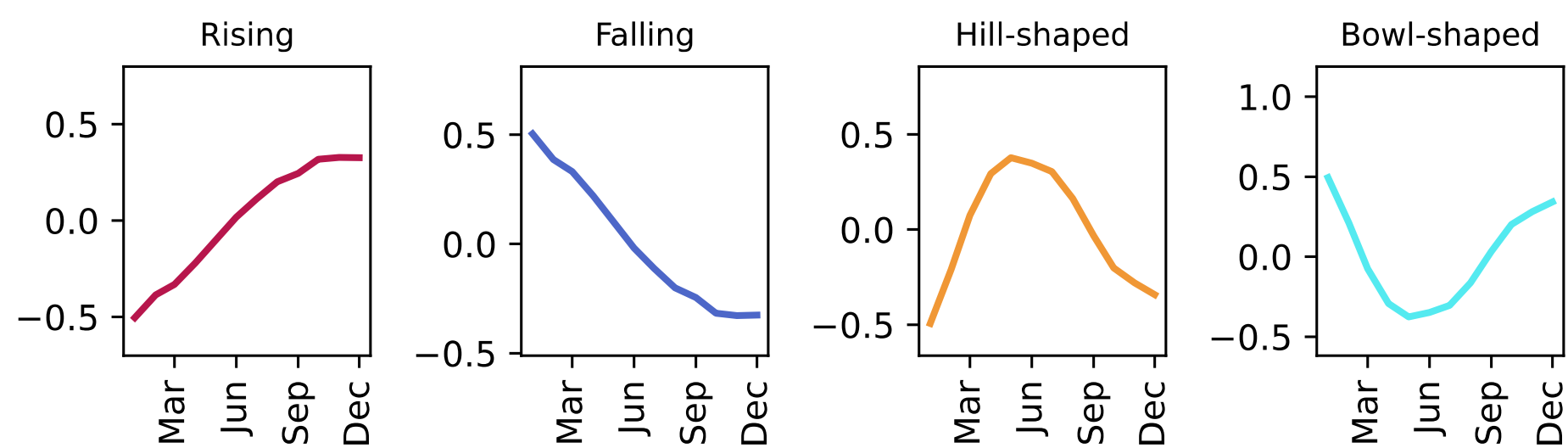
Individual user’s yearly trajectories for a **community-based metric** (left) and a **language-based metric** (right) are noisy

Identifying Temporal Arcs

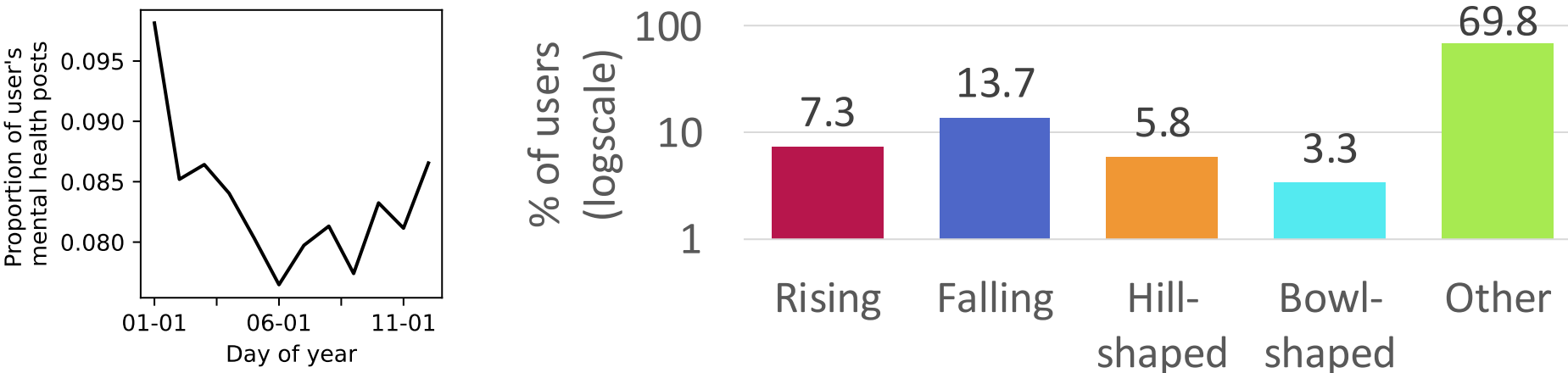


Analysis of Temporal Arcs

Community-Based Arcs



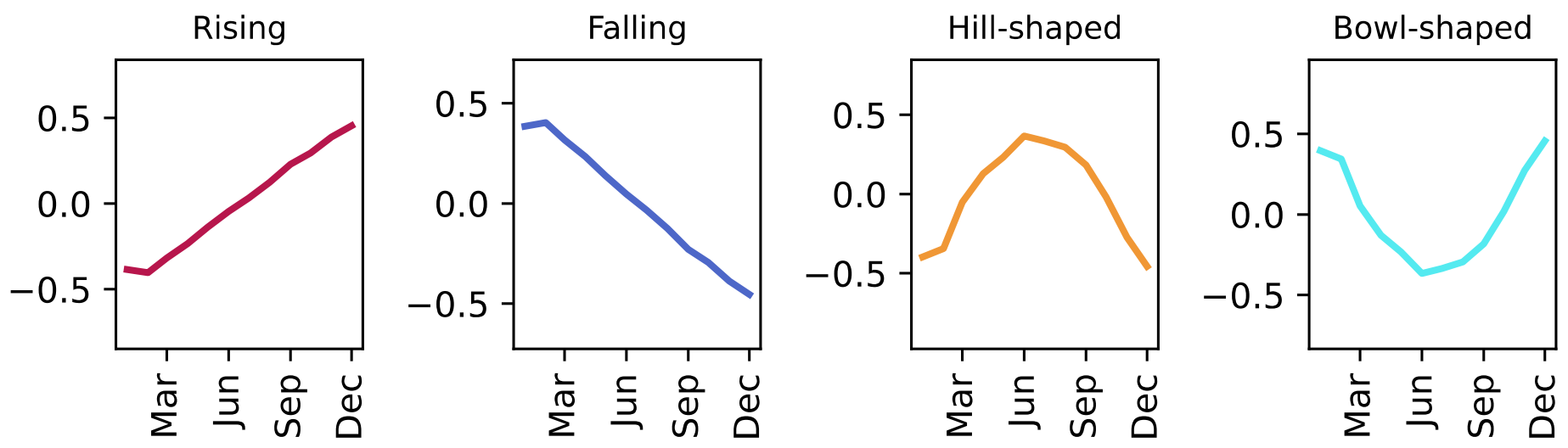
Four temporal arcs of mental health based on the community-based metric. The arc is shown in color and the 40 individual user’s times series with the highest weight for each arc are shown in grey.



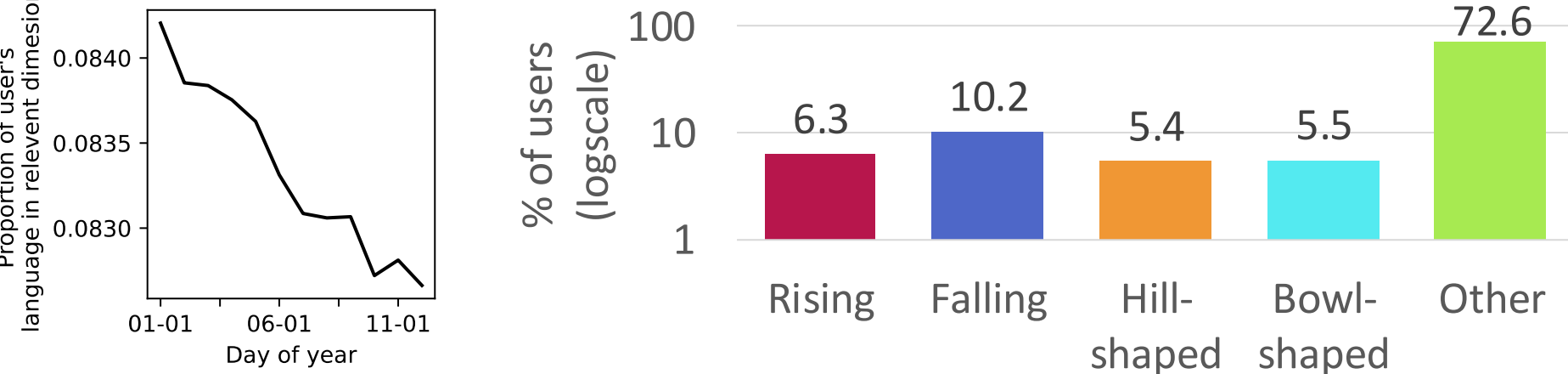
mean of user-level distributions

percentage of users associated with each arc

Language-Based Arcs



Four temporal arcs of mental health based on the language-based metric. The arc is shown in color and the 40 individual user’s times series with the highest weight for each arc are shown in grey.



mean of user-level distributions

percentage of users associated with each arc

Takeaways

- Depression-related metrics on Reddit (both community-based and language-based) follow four core patterns across a year: **rising**, **falling**, **hill-shaped** and **bowl-shaped**
- The dominant pattern for individual users does not necessarily reflect the mean pattern for a metric across all users
- Many users are most strongly associated with **other** singular vectors beyond the core arcs
- The **bowl-shaped** arc (reminiscent of SAD) is relatively uncommon with the community-based metric, but more frequent with the language-based metric

Future Work

- Improvement of arcs and analysis:
 - Consideration of the full mixture of components that describes each user’s behavior
 - Additional methods (beyond SVD) for extracting arcs from data
- Downstream use of temporal arcs:
 - Do **hill-shaped** or **bowl-shaped** arcs tend to fit the definition of SAD in that they repeat themselves year after year?
 - Can we predict which arc a user will be associated with in year *n* using data from only years *n-1* and earlier?
 - Are there factors that are associated with a change from **rising** to **falling** (or vice-versa)?