Using the Belief Landscape Model to Predict Individual Belief Dynamics

Jianghui Li
jli159@syr.edu
Syracuse University
Syracuse, New York, USA

Joshua Introne
jeintron@syr.edu
Syracuse University
Syracuse, New York, USA

ABSTRACT

We demonstrate a classifier that is able to predict how likely it is an individual social media user is to change their professed beliefs. Our classifier uses features based on how people cluster in terms of their expressed beliefs, and we obtain high degrees of accuracy for two distinct datasets. Our results suggest that the frequency with which people change their publicly professed beliefs seems to be a stable property of individuals.

CCS CONCEPTS

• **Applied computing** → **Psychology**; • **General and reference** → **Measurement**.

KEYWORDS

Belief Dynamics, Classification, Computational Social Science

ACM Reference Format:
Jianghui Li and Joshua Introne. 2024. Using the Belief Landscape Model to Predict Individual Belief Dynamics. In *Book of Extended Abstracts of the ACM Collective Intelligence Conference, 2024, June 27–28, 2024, Boston, MA.* ACM, New York, NY, USA, 3 pages.

1 INTRODUCTION

Social media has offered an unprecedented window into the dynamics of collective opinion formation and belief change in online settings. It may also be the case that algorithmic filters [1], misinformation [6], the loss of traditional media gatekeepers [2], and other complex socio-technical factors [10] have led to shifts in population-wide belief dynamics [1]. Of particular concern is that an information environment polluted with misinformation and polarizing content can lead individuals to develop beliefs that diverge from empirically validated reality, possibly leading to real physical harm [7] and political instability [4]. Because of this, there is growing interest in predicting who is likely to fall prey to misinformation [3, 8, 11].

Recent work by Introne [5] offered new methodological tools for characterizing the belief dynamics of online populations. The study introduced the Belief Landscape Framework (BLF), which is a spatial metaphor for representing a population's beliefs and the movements of individuals across the landscape as they express different beliefs. Introne showed that most online users gravitate towards the center of "belief attractors," which tend to draw the bulk of individuals, while a smaller number of users on the periphery of belief attractors update their beliefs more frequently. Here, we build on the Belief Landscape Framework to classify users according to their proclivity for belief change. This approach is unlike prior work, which has tended to use psychometric and other individual features to predict susceptibility to misinformation [e.g., 8]. Predicting levels of individual belief variability is not the same as predicting susceptibility to misinformation, but individuals who are not "anchored'" to cultural belief attractors are more likely to be easily influenced. Our task is thus a first step in developing new techniques for predicting susceptibility to misinformation or other information attacks. We show here that it is possible to classify individuals with high reliability based on analysis of the topological properties of the belief landscape they have traversed.

1.1 Methods

To develop our classifier, we employed two datasets collected from Twitter using the streaming API. One dataset centered on climate change conversations, covering 166,268 tweets from May 24th to October 27th, 2019, and another comprising 264,777 tweets on 'kpop' and 'blm' hashtags from January 1st to December 31st, 2020. These datasets provide a rich source for analyzing belief dynamics in polarized and cross-cultural contexts. Following [5], we processed both datasets to produce belief landscapes and measure the movements of users across them (Figure **??** provides a schematic of our workflow). As discussed above, belief landscapes are organized around attractors with different topologies. Attractors may be deep or shallow, reflecting the size of the population they draw, and attract users from small or large regions of the overall landscape (the basin of attraction).

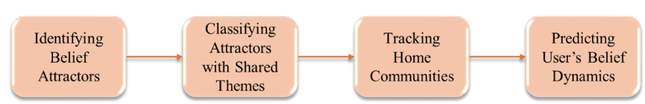

Figure 1: Summary of our data analytic pipeline

Analyzing movements over attractors revealed a higher level of organization, apparent in that users who moved amongst attractors tended to visit distinct subsets and only rarely moved between these subsets. Unlike Introne [5], we interpret these higher-level groups of attractors as mutually exclusive belief systems, and we sought to detect dynamics of movements between these groups.

Because user movements on the landscape were somewhat noisy, we applied a spin-glass clustering algorithm [9] to identify the attractor communities, and then smoothed user transitions between these communities using a Hidden Markov Model, mapping hidden states to the distinct belief communities. We used these transitions to classify users into four distinct classes (Table 1), which became our predictive target.

Table 1: Summary of user count for each class. There are four classes, which are stable user (never changed home location), changer (changed from one home to another), bistable user (changing between two homes), and dynamic user (appeared in more than two homes).

User Class	Climate Change	K-pop & BLM
Class 1 (Stable)	2956	927
Class 2 (Changer)	78	136
Class 3 (Bistable)	80	829
Class 4 (Dynamic)	5	737

In our initial attempts to build a classifier, we found that using the frequency of visits to each underlying attractor yielded a near-perfect classifier. However, such a classifier cannot generalize to different domains. Thus, we sought to build a classifier based on the topology of different attractors. We generated nine features for each landscape by dividing attractors into nine classes based on their depth and width, with three levels in each dimension. Figure 2 shows the attractor class information for several attractors in the belief landscape. We used the amount of time each individual spent in each attractor class as the feature value for each user and compared the performance of simple decision-tree, random forest, and k-nearest neighbors classifiers under 5-fold cross-validation.

1.2 Findings

Class-wise F1 scores from our study are shown in Table 2. Both the random forest and KNN classifiers offered good performance for both datasets and all classes. We also examined the decision-tree classifiers to develop insights into the underlying logic of the classifiers (Figure 2).

Table 2: User classes classification F1 scores using attractor classes (width and depth) as features

Classifier	Climate Change	K-pop & BLM
Decision Tree	0.7022107	0.6812203
	0.8100659	0.7503566
	0.7709280	0.5570560
	0.9204545	0.6941992
Random Forest	**0.9177489**	0.9392151
	0.9726843	0.9594145
	0.9412149	0.9022602
	0.9923037	0.9303215
K-Nearest Neighbors	0.8893754	**0.9480612**
	0.9594382	**0.9675780**
	0.9378589	**0.9388073**
	0.9692860	**0.9498699**

Our findings offer several insights. The topological properties of attractors that people visit on belief landscapes are highly predictive of an individual's proclivity for belief change. The decision tree in Figure 3 illustrates that individuals positioned in shallow and wide belief clusters are more likely to change their beliefs often. Conversely, users situated in regions with densely converging opinions exhibit a lower propensity for belief change, suggesting that consensus has a stabilizing effect on beliefs. This finding corresponds to the prior observation that deeper attractors tend to be 'stickier' than shallow ones [6]. Moreover, we found these results to be similar in both landscapes, suggesting the generalizability of the model across different domains.

Our results also provide evidence that people who change their beliefs frequently tend to skirt around more densely populated cultural belief systems—that is, those that change their beliefs often may also be those that profess beliefs that co-occur relatively infrequently in a population. Additional analysis will be necessary to explore this finding further, but if correct, this may open new lines of inquiry into the belief dynamics of individuals.

REFERENCES

[1] Michael Bailey, Drew M. Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel, and Johannes Stroebel. 2020. Social Networks Shape Beliefs and Behavior: Evidence from Social Distancing During the COVID-19 Pandemic. (12 2020). https://doi.org/10.3386/w28234 DOI: 10.3386/w28234.
[2] Janelle Benham. 2020. Best Practices for Journalistic Balance: Gatekeeping, Imbalance and the Fake News Era. *Journalism Practice* 14, 7 (8 8 2020), 791–811. https://doi.org/10.1080/17512786.2019.1658538
[3] Rex P. Bringula, Annaliza E. Catacutan-Bangit, Manuel B. Garcia, John Paul S. Gonzales, and Arlene Mae C. Valderama. 2022. "Who is gullible to political disinformation?" : predicting susceptibility of university students to fake news. *Journal of Information Technology Politics* 19, 2 (3 4 2022), 165–179. https://doi.org/10.1080/19331681.2021.1945988
[4] Rafael Evangelista and Fernanda Bruno. 2019. WhatsApp and political instability in Brazil: targeted messages and political radicalisation. *Internet Policy Review* 8, 4 (2019), 1–23. https://doi.org/10.14763/2019.4.1434 publisher: Berlin: Alexander von Humboldt Institute for Internet and Society.
[5] Joshua Introne. 2023. Measuring Belief Dynamics on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media* 17, 387–398.
[6] Shan Jiang, Miriam Metzger, Andrew Flanagin, and Christo Wilson. 2020. Modeling and Measuring Expressed (Dis)belief in (Mis)information. *Proceedings of the International AAAI Conference on Web and Social Media* 14 (26 5 2020), 315–326. https://doi.org/10.1609/icwsm.v14i1.7302
[7] Hye Kyung Kim and Edson C. Tandoc. 2022. Consequences of Online Misinformation on COVID-19: Two Potential Pathways and Disparity by eHealth Literacy. *Frontiers in Psychology* 13 (2022). https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.783909 [Online; accessed 2024-03-03].
[8] Kerry Lee. 2004. Age, neuropsychological, and social cognitive measures as predictors of individual differences in susceptibility to the misinformation effect. *Applied Cognitive Psychology* 18, 8 (2004), 997–1019. https://doi.org/10.1002/acp.1075
[9] Jörg Reichardt and Stefan Bornholdt. 2006. Statistical mechanics of community detection. *Physical Review E* 74, 1 (18 7 2006), 016110. https://doi.org/10.1103/PhysRevE.74.016110 publisher: American Physical Society.
[10] Jaime E. Settle. 2018. *Frenemies: How Social Media Polarizes America.* Cambridge University Press.
[11] Tracy Jia Shen, Robert Cowell, Aditi Gupta, Thai Le, Amulya Yadav, and Dongwon Lee. 2019. How Gullible Are You? Predicting Susceptibility to Fake News. *Proceedings of the 10th ACM Conference on Web Science*, 287–288. https://doi.org/10.1145/3292522.3326055 [Online; accessed 2024-03-03].

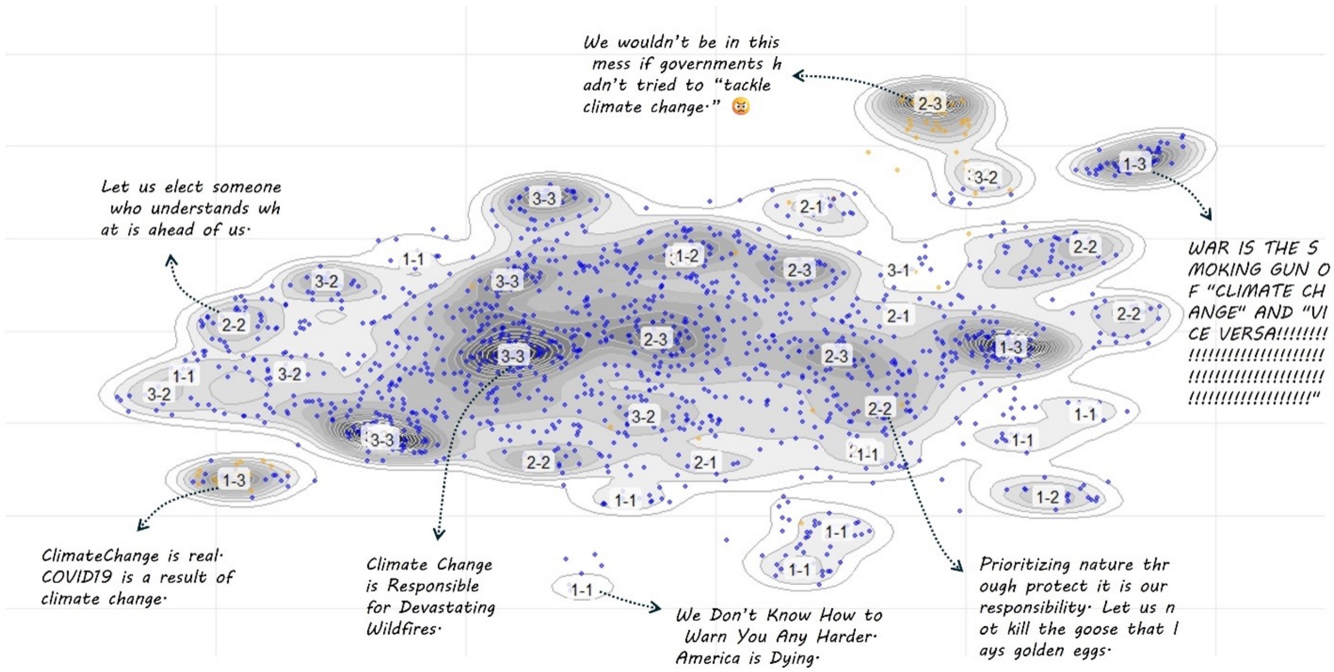

Figure 2: Climate change discussion belief landscape. The dots on the landscape represent a user at a time (only 0.1% are shown). Blue dots are believers, and orange dots are skeptics. The label numbers are attractor classes (9 classes in total) defined by dual quantiles (3 quantiles for each; '1' for low, '2' for medium, and '3' for high) of their width (number to the left) and depth (number to the right).

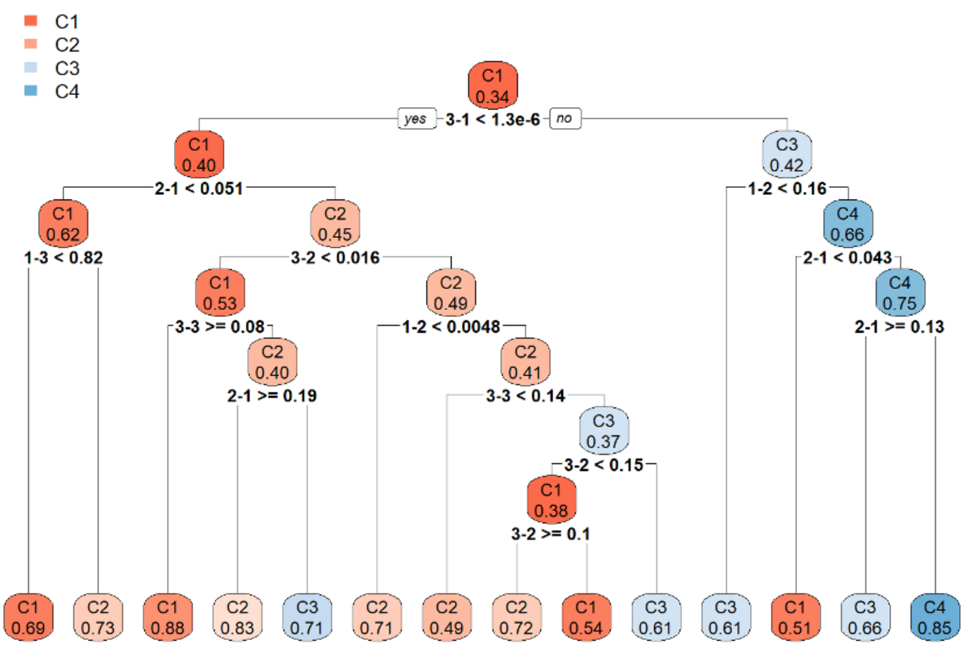

Figure 3: An example of decision tree visualization from climate change data