Estimating the Political Orientation of Twitter Users in Homophilic Networks

Morteza Shahrezaye, Orestis Papakyriakopoulos
Juan Carlos Medina Serrano, Simon Hegelich
Bavarian School of Public Policy at Technical University of Munich
Richard-Wagner street 1
80333 Munich, Germany
{morteza.shahrezaye, orestis.papakyriakopoulos, juan.medina}@tum.de, simon.hegelich@hfp.tum.de

Abstract

There have been many efforts to estimate the political orientation of citizens and political actors. With the burst of online social media use in the last two decades, this topic has undergone major changes. Many researchers and political campaigns have attempted to measure and estimate the political orientation of online social media users. In this paper, we use a combination of metric learning algorithms and label propagation methods to estimate the political orientation of Twitter users. We argue that the metric learning algorithm dramatically increases the accuracy of our model by accentuating the effect of homophilic networks. Homophilic networks are user clusters formed due to cognitive motivational processes linked with cognitive biases. We apply our method to a sample of Twitter users in Germany's six-party political sphere. Our method obtains a significant accuracy of 62% using only 40 observations of training data for each political party.

Introduction

Measuring and estimating the political orientation of normal citizens and political actors has always been a relevant question. The answer to this question is essential for electoral campaigns (Gayo Avello, Metaxas, and Mustafaraj 2011; Dokoohaki et al. 2015; Papakyriakopoulos et al. 2018), agenda setting, policy making (McCombs 2014), and research purposes (Golbeck and Hansen 2011; Barberá 2014; Hegelich and Shahrezaye 2015). The methodological efforts to answer this crucial question possess three qualities.

The first quality is related to the number and type of inputs in the algorithm: What type of features are considered while estimating the latent political orientation of the users? The second quality is if the method is designed to estimate the political orientation of a specific group of political actors (Wong et al. 2013; Groseclose and Milyo 2005) or a more general group of citizens (Barberá 2014). If a method is designed based on a specific group of political actors or citizens, it cannot be generalized to estimate the political orientation of other groups of political actors or citizens. Cohen and Ruths have presented that methods that have accuracy greater than 90% in estimating if a Twitter user is a Democrat or Republican, would have accuracy level of less than 65% when applied on general Twitter users. The last quality is if the method measures the political orientation on a one dimensional or a multidimensional latent space. Most of the literature has been designed based on the two-party political system of the United States. Thus, they are inherently designed to estimate a one-dimensional latent variable.

In this work, we use a combination of metric learning algorithms and label propagation methods to estimate the political orientation of Twitter users. Our method has three distinguishing features. First, the method requires a minimal number of features as training data because it exploits the homophilic structure of social networks (Geschke, Lorenz, and Holtz 2018; Madsen, Bailey, and Pilditch 2018). Second, the proposed method estimates on a multidimensional latent space; therefore, the proposed method can be used to estimate the political orientation of users in a multiparty political system. The third feature is that our method is extendable to multiple groups or cluster of users. Our method can estimate the political orientation of users even if the target users have zero political activity on the platform.

Methodology

We use a combination of metric learning algorithms with label propagation methods to estimate the political orientation of Twitter users. The goal of label propagation algorithms is to estimate the labels of a large set of unlabeled observations from the small set of labeled observations.

Suppose there are l labeled observations $(x_1, y_1), \dots, (x_l, y_l)$ and u unlabeled observations such that $l < u$, and $n = l + u$. Consider a connected graph $G = (V, E)$ with nodes $L = \{1, \dots, l\}$ and $U = \{l + 1, \dots, l + u\}$ corresponding, respectively, to the labeled or training observations and unlabeled or test observations. A label propagation algorithm propagates the labels for the set U, based on the distances between its observations to the observations in L. Within the label propagation algorithm, the labels of the vertices in set L would be fixed, but the labels of the set U would be estimated based on a function of their distance to set L.

Let n be the total number of Twitter users we have including l users for whom we already know their political orientation and u users for whom we want to estimate their political orientation. We use only the structure of the friends' network to estimate the political orientations. Let F be the set of friends of all n users with size m. Therefore, we can create the binary matrix A with dimension $n \times m$, which would represent the friends of each of the n users. Before constructing graph G from matrix A, we transform matrix A by using a proper metric learning algorithm.

The reason for transforming matrix A is that we believe there are hidden information within the network structure, which we could use to increase the estimation accuracy. By contrast with the rational choice theory, the human judgment is influenced by various cognitive biases, prior judgments, environmental features, and stimulus-feedback loops (Kenrick et al. 2010; Donkin, Heathcote, and Brown 2015). Cognitive biases reproduce human judgments that could be systematically different from rational reasoning (Kahneman and Tversky 1973; Haselton, Nettle, and Murray 2015). The cognitive biases make the human brain process the information in a distorted manner compared with an objective reality (Sharot, Korn, and Dolan 2011). Although there is a list of cognitive biases that affect the online activity of the users, we are specifically interested in cognitive biases related to self-categorization. Self-categorization describes the motivations and circumstances under which communities with shared identities form. The self-categorization theory articulates that the spectrum of human behavior can be analyzed from a pure interpersonal or individualistic and a pure intergroup or collectivist perspective. Humans have the desire for a positive and secure self-concept; therefore, they connect with individuals that confirm their pre-existing attitudes, verify their self-views, and increase their social identity. The aforementioned behaviour is called confirmation bias (Geschke, Lorenz, and Holtz 2018). In addition, “If we are to accept that people are motivated to have a positive self-concept, it flows naturally that people should be motivated to think of their groups as good groups” (Hornsey 2008). Striving for a positive and secure self-concept, humans’ collectivist behaviors contribute to the formation of online and offline communities with shared social identities (Ridings and Gefen 2004). Consequently, users with similar labels, that is, similar political preferences, are expected to be relatively closer to each other. Therefore, if we were to supposedly apply a k-nearest neighbors learning method, it makes sense to use a distance function that interprets similar users closer to each other. Instead of using an off-the-shelf distance function such as Euclidean distance, we use an alternative distance function that guarantees higher accuracy for the labeled or training observations after running the learning method.

A brief description of the steps of our method is as follows. First, we acquire matrix A, which includes the labeled observations and the unlabeled observations as rows. Second, we learn the optimized distance or metric function that guarantees higher accuracy within the labeled observations by exhausting the special structure of homophilic networks. We transform matrix A by using the learned metric to construct graph G. Finally, we apply the learning method or the label propagation algorithm.

Metric Learning for Large Margin Nearest Neighbor Classification (LMNN)

The accuracy of each learning algorithm is a function of the distance function or the metric used to compute the distance between the observations. The metric learning algorithm we use is based on the following: a precise k-nearest neighbors classification will correctly classify a labeled observation if its k-nearest neighbors share the same label. The algorithm then attempts to increase the number of labeled observations with this property by learning a linear transformation of the input space that proceeds the final learning method. The linear transformation of *LMNN* is derived by maximizing a loss function with two terms. The first term minimizes the large distances between observations within class, and the second term maximizes the distances between the observation between the classes (Weinberger and Saul 2009).

In general, metric learning algorithms estimate the positive semidefinite transformation matrix $\mathcal{M}$ such that the distance between two observations, x_i and x_j, is derived by the Mahalanobis distance,

$$d_{\mathcal{M}}(x_i, x_j) = \sqrt{(x_i - x_j)^T \mathcal{M} (x_i - x_j)}$$

which follows certain features. If we replace $\mathcal{M}$ with the identity matrix, the resulting metric would be Euclidean metric. *LMNN* learns a linear transformation matrix $\mathcal{M}$, such that the training or labeled observation satisfies the following items (Weinberger and Saul 2009):

- Each labeled observation should share the same label as its k- nearest neighbors. This is achieved by introducing a loss function that penalizes large distances between observations belonging to the same class,

  $$\epsilon_{pull}(L) = \sum_{j \rightsquigarrow i} ||L(\bar{x}_i - \bar{x}_j)||^2$$

  where $j \rightsquigarrow i$ indicates that j is an observation that we desire to be close to i, and L is the function representing the transformation by matrix $\mathcal{M}$.

- The labeled observations with different labels should be significantly separated. This separation is achieved by introducing a loss function that penalizes small distances between observations belonging to different classes,

  $$\epsilon_{push}(L) = \sum_{i,j \rightsquigarrow i} \sum_{l} [1 + ||L(\bar{x}_i - \bar{x}_j)||^2 - ||L(\bar{x}_i - \bar{x}_l)||^2]$$

  where the inner sum iterates over all the observations with a different class to i, and l invades the perimeter of i and j plus unit margin. In other words, the observation l satisfies

  $$||L(\bar{x}_i - \bar{x}_l)||^2 \leq ||L(\bar{x}_i - \bar{x}_j)||^2 + 1$$

The final loss function is a weighted combination of the two defined components,

$$\epsilon(L) = (1 - \mu)\epsilon_{pull}(L) + \mu \epsilon_{push}(L)$$

Although the general loss function above is not convex, by limiting the solution space to positive semidefinite matrices, the loss function will be a convex function.

The solution to the minimization of the loss function, given the labeled subset of A, is the desirable matrix $\mathcal{M}$. We transform matrix A to obtain matrix $A_{\mathcal{M}}$ by

$$A_{\mathcal{M}} = A \times \mathcal{M}$$

We construct graph G using the $A_{\mathcal{M}}$ of size $n \times m$ by using the nearest neighbor graph method. In other words, using n rows of $A_{\mathcal{M}}$, we define n vertices of G and then define edges between each vertex and its k_G nearest neighbors by using the Euclidean distance function.

Label Propagation Using Gaussian Fields and Harmonic Functions

The goal of applying a label propagation algorithm to a graph is to estimate the labels of unlabeled vertices by using their connections to the few labeled vertices. This problem is usually formulated as an iterative process within which the labels are gradually diffused over the matrix, such that the state of the graph would converge to a stationary state. This iterative process might have an analytical solution that would be more efficient than applying the algorithm iteratively (Barrett et al. 1994; Zhu and Ghahramani 2002). The most crucial implication of a label propagation algorithm for our question regarding estimating political orientation of Twitter users is that the only requirement for estimating the political requirement of a user is that the user should be connected to graph G. Hence, the user should not necessarily have politicians or other political actors as friends.

The algorithm we use for label propagation is based on Zhu, Ghahramani, and Lafferty. Let the simple graph $G = (V, E)$ and the set of the labeled and unlabeled vertices, L and U, be as defined. The goal is to compute the real-valued function $f : V \to \mathbb{R}$ on the simple graph G. f must assign the same given labels for the set L or $f_l(i) \equiv y_i$ for $i \in l$. To estimate the function f they defined the energy function

$$E(f) = \frac{1}{2} \sum_{i,j} w_{i,j} (f(i) - f(j))^2$$

and the Gaussian field

$$p_\beta(f) = \frac{-e^{\beta E(f)}}{Z_\beta}$$

where β is an inverse temperature function and $Z_\beta = \int_f exp(-\beta E(f))$ which normalizes over all functions constrained to the constraint $f_l(i) \equiv y_i$ on the labeled vertices. Then, they demonstrate the result of the minimization

$$f = \arg\min_f E(f)$$

which is a harmonic function that satisfies the constraint $f_l(i) \equiv y_i$ on the labeled vertices. The harmonic property implies that the value of f at each unlabeled vertex is the average of f at neighboring vertices. Therefore, the estimated labels would be a function of the similarity of all neighboring vertices.

The estimated f has an interpretation within the framework of random walks. The estimated $f(i)$ for an unlabeled vertex $i \in U$ would be a vector of size equal to number of possible classes. The jth element of $f(i)$ would be the probability that a particle that started at vertex i would first hit a vertex with class j. Therefore, the resulting algorithm can be used to estimate the political orientation of a user in a multidimensional latent space.

Data and Results

Data Preparation

We require two sets of data for training and testing. We acquire both sets from the public Twitter API. In the first step, we obtained the list of all the members of the main and local German parliaments who are available on Twitter. This list contains 623 Twitter users from one of the six parties CDU/CSU, SPD, Grüne, Linke, FDP and AfD.

From a database of German political Tweets, we obtained a list of 400,000 random Twitter users. We downloaded the list of all their friends and their last 4,000 Tweets by using the public API. We counted how many times each user retweeted the Tweets of members of each of the political parties we acquired in the first step. If a user has retweeted a minimum of five Tweets from members of party j but no retweets from other parties, we tag this user as a user with a political orientation to party j. From the 400,000 initial users, we could label 8,146 based on the mentioned heuristic.

To reduce the complexity of the computations, we reduced the sample size to 50,000 from 400,000. Thus, we created matrix A using 50,000 random users including all of the 8,146 labeled users. Matrix A has at this step 50,000 rows as users, which we want to use for our training and test set, and 7,194,153 columns as the friends. To further reduce the complexity of the computations, we removed the friends who are friends of less than 0.01% of the users. The final matrix A has the dimension 50,000$\times$ 552,136.

We confirm that our test data has a minor bias in the sense that we already know our test data includes users who have engaged in some type of political activity. This assumption is because these users are randomly chosen from a database of German political Tweets. On the other side, this bias is mildly mitigated in two steps. First, matrix A is created by a list of friends of all 50,000 random users and not only the friends of the labeled 8,146 users. Thus, the feature sets are from a bigger set of observations. Second, we added some randomness by removing some columns of matrix A in the final step.

Metric Learning and Label Propagation

We resampled 60 users per political party out of the 8,146 labeled users of A. We learned matrix $\mathcal{M}$ based on the 240 users. Next, we transformed the whole matrix A using $\mathcal{M}$ by applying

$$A_{\mathcal{M}} = A \times \mathcal{M}$$

Using the transformed $A_{\mathcal{M}}$, we made a 10-nearest neighbors graph using a Euclidean distance function to make graph G. Finally, we applied the label propagation algorithm on G that has 50,000 vertices, out of which, the labels of 240 are introduced to the algorithm. The labels of the other 49,760 are estimated using the label propagation algorithm.

Results

We performed the resampling and the computations 10 times to make sure the results are robust. For each trial, we applied a random forest classifier on the 240 training data as a

random forest	A (not transformed)	0.23
label propagation		0.20
random forest	$A_{\mathcal{M}}$ (transformed)	0.30
label propagation		0.62

Table 1: Average accuracy of the predictions over 10 resamples

benchmark result. We also applied the random forest classifier and the label propagation method on A directly to improve our understanding regarding how much the $LMNN$ metric learning method contributes to the accuracy of the results. Table 1 shows the average accuracy of the estimations on the remaining 8,146-240=7,906 labeled users with a known political orientation.

Referring to Table 1, we observe that the transformation increases the accuracy of the random forest classifier and the label propagation algorithm. We also observe that the combination of the metric learning algorithm and the label propagation method results to a much higher accuracy of estimation.

Discussion

In this paper, we proposed a new method to estimate the political orientation of Twitter users. Our method has many distinguishing features: The method requires few training observations, requires few learning features, is based on a multidimensional latent space, and is easily expendable to new users even if they have zero political activity on Twitter.

Based on Table 1, the high accuracy of the model is due to the transformation of the initial matrix using the function learned by the *LMNN* algorithm. The cost function of the *LMNN* algorithm has two parts. One part pulls the observations of the same class closer to each other, and the other part pushes the observations of different classes far apart. Additionally, since the *LMNN* algorithm is based on optimizing a k-nearest neighbor model on the training observations, the trained matrix $\mathcal{M}$ transforms the observations based on their relation to other observations in their vicinity and not the whole dataset. These characteristics have crucial implications reagarding the accuracy of our estimation.

As aforementioned, the initial matrix, A, has a special structural feature because it represents a homophilic social network, which means that users with similar political identity are assumed to demonstrate similar behavior on Twitter. Therefore, we expected that users with similar political identity would follow similar politicians, similar celebrities, similar sportsmen, and so forth.

When we apply the *LMNN* algorithm to this homophilic network, we accentuate the extant distinctive features formed due to the existing cognitive biases in self-categorization and group formation (Geschke, Lorenz, and Holtz 2018; Madsen, Bailey, and Pilditch 2018).

The matrix $\mathcal{M}$ learns different combinations of features that help distinguish normal Twitter users based on their political orientation. The matrix $\mathcal{M}$ also allows different combination of features for each class because it is based on a k-nearest neighbor algorithm that considers a bounded proximity of the users. Our model detects the political orientation of users with high accuracy, and by far outperforms other algorithms that have been applied to this task.

Due to the use of label propagation algorithm, this model can be later applied on any new user e to estimate her or his political orientation, as long as e is connected to the graph G. More generally, to predict the political orientation of user e, we must find a new set of users including e, forming a small graph g connected to the initial graph G.

This study provides valuable insights into the study of user behavior on online social networks. This study illustrates, that using mathematical algorithms that exhaust properties of social theories, we can improve the performance of models explaining human behavior. Furthermore, this study contradicts the general claim that a huge amount of data is required to make accurate predictions on social and political behavior. Finally, our method provides a novel technique to assign political partisanship, by having as input only the network of interpersonal connections.

References

Barberá, P. 2014. Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political Analysis* 23(1):76–91.

Barrett, R.; Berry, M. W.; Chan, T. F.; Demmel, J.; Donato, J.; Dongarra, J.; Eijkhout, V.; Pozo, R.; Romine, C.; and Van der Vorst, H. 1994. *Templates for the solution of linear systems: building blocks for iterative methods*, volume 43. Siam.

Cohen, R., and Ruths, D. 2013. Political orientation inference on twitter: It's not easy. *Proc. of ICWSM*.

Dokoohaki, N.; Zikou, F.; Gillblad, D.; and Matskin, M. 2015. Predicting swedish elections with twitter: A case for stochastic link structure analysis. In *Advances in Social Networks Analysis and Mining (ASONAM), 2015 IEEE/ACM International Conference on*, 1269–1276. IEEE.

Donkin, C.; Heathcote, B. R. A.; and Brown, S. D. 2015. Why is accurately labelling simple magnitudes so hard? a past, present and future look at simple perceptual judgment. *The Oxford Handbook of Computational and Mathematical Psychology* 121–141.

Gayo Avello, D.; Metaxas, P. T.; and Mustafaraj, E. 2011. Limits of electoral predictions using twitter. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence.

Geschke, D.; Lorenz, J.; and Holtz, P. 2018. The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *British Journal of Social Psychology*.

Golbeck, J., and Hansen, D. 2011. Computing political preference among twitter followers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1105–1108. ACM.

Groseclose, T., and Milyo, J. 2005. A measure of media bias. *The Quarterly Journal of Economics* 120(4):1191–1237.

Haselton, M. G.; Nettle, D.; and Murray, D. R. 2015. The evolution of cognitive bias. *The handbook of evolutionary psychology* 1–20.

Hegelich, S., and Shahrezaye, M. 2015. The communication behavior of german mps on twitter: Preaching to the converted and attacking opponents. *European Policy Analysis* 1(2):155–174.

Hornsey, M. J. 2008. Social identity theory and self-categorization theory: A historical review. *Social and Personality Psychology Compass* 2(1):204–222.

Kahneman, D., and Tversky, A. 1973. On the psychology of prediction. *Psychological review* 80(4):237.

Kenrick, D. T.; Neuberg, S. L.; Griskevicius, V.; Becker, D. V.; and Schaller, M. 2010. Goal-driven cognition and functional behavior: The fundamental-motives framework. *Current Directions in Psychological Science* 19(1):63–67.

Madsen, J. K.; Bailey, R. M.; and Pilditch, T. D. 2018. Large networks of rational agents form persistent echo chambers. *Scientific reports* 8(1):12391.

McCombs, M. 2014. *Setting the agenda: Mass media and public opinion*. John Wiley & Sons.

Papakyriakopoulos, O.; Hegelich, S.; Shahrezaye, M.; and Serrano, J. C. M. 2018. Social media and microtargeting: Political data processing and the consequences for germany. *Big Data & Society* 5(2):1 – 15.

Ridings, C. M., and Gefen, D. 2004. Virtual community attraction: Why people hang out online. *Journal of Computer-mediated communication* 10(1):JCMC10110.

Sharot, T.; Korn, C. W.; and Dolan, R. J. 2011. How unrealistic optimism is maintained in the face of reality. *Nature neuroscience* 14(11):1475.

Weinberger, K. Q., and Saul, L. K. 2009. Distance metric learning for large margin nearest neighbor classification. *J. Mach. Learn. Res.* 10:207–244.

Wong, F. M. F.; Tan, C. W.; Sen, S.; and Chiang, M. 2013. Quantifying political leaning from tweets and retweets. *ICWSM* 13:640–649.

Zhu, X., and Ghahramani, Z. 2002. Learning from labeled and unlabeled data with label propagation.

Zhu, X.; Ghahramani, Z.; and Lafferty, J. D. 2003. Semi-supervised learning using gaussian fields and harmonic functions. In *Proceedings of the 20th International conference on Machine learning (ICML-03)*, 912–919.