© April 2018 | IJIRT | Volume 4 Issue 11 | ISSN: 2349-6002

A novel multitask learning frame work for forecasting models of the locations

R.Mamatha[1], Mis.N.Vinayasree[2]
[1]*Student, KMM Institute of PG Studies, Tirupathi*
[2]*Asst.Professor, KMM Institute of PG Studies, Tirupathi*

Abstract- **Spatial occasion determining from online networking is conceivably amazingly helpful however experiences basic difficulties, for example, the dynamic examples of highlights (catchphrases) and geographic heterogeneity (e.g., spatial relationships, imbalanced examples, and unique populaces in various areas). Most existing methodologies (e.g., LASSO relapse, dynamic question extension, and burst identification) address a few, yet not all, of these difficulties. Here we propose a novel multi-errand learning system that intends to simultaneously address every one of the difficulties included. In particular, given a gathering of areas (e.g., urban communities), anticipating models are worked for all the areas all the while by extricating and using proper shared data that viably expands the example estimate for each area, in this way enhancing the estimating execution. The new model consolidates both static highlights got from a predefined vocabulary by area specialists and dynamic highlights produced from dynamic inquiry extension in a multi-errand include learning structure. Distinctive systems to adjust homogeneity and decent variety amongst static and dynamic terms are likewise explored. Furthermore, productive calculations in light of Iterative Group Hard Thresholding are created to accomplish proficient and successful model preparing and expectation. Broad trial assessments on Twitter information from common distress and flu episode datasets show the adequacy and productivity of our proposed approach.**

Index Terms- **Twitter, event detection, earthquake, LASSO.**

I. INTRODUCTION

Three technical challenges must be overcome when addressing this problem: 1) Dynamic features. The language used in micro blogs is highly informal, ungrammatical, and dynamic. Most existing methods treat fixed keywords as features, but expressions in tweets may dynamically evolve, rendering the use of fixed features and historical training data insufficient. Ideally, an event forecasting system must combine the judicious use of static (fixed) features with an awareness of subtle changes involving dynamic features. 2) Geographic heterogeneity. Existing models usually build a single predictive model for all the different locations. However, different cities have different characteristics, such as population, weather (e.g., humidity, temperature), and administrative structures (e.g., capital cities versus noncapital cities). As a result, it is difficult to impute basal levels of occurrence uniformly. Considering civil unrest as an example, finding 1000 tweets mentioning the keyword "protest" is not likely to be a strong indicator of an upcoming civil unrest event in a city with a population of a few million users but could be a strong signal in a much smaller city with a population of only 10,000. To consider the geographical heterogeneity, some works propose to establish the corresponding model for each different location separately. But because each model only utilizes the data of its corresponding location, the data scarcity problem (especially for non-large locations) is a serious challenge that degrades the model performance and generalization.

II. ALGORITHM

The FISTA algorithm performs well for convex problems. Even worse, they also involve discrete constraints, which make the problems particularly challenging to solve. Motivated by the success of the iterative hard thresholding algorithm for solving l0-regularized problems [7] and recent advances in non convex iterative shrinkage algorithms, we propose to employ the Iterative Group Hard Thresholding framework to solve both problems.

III. ALGORITHM FOR MODELS CMTFL-I AND CMTFL-II

We thus focus on Problemin the following discussion. The details are summarized in Algorithm 1. Here, data parallelism strategy is utilized to achieve the calculation of the gradient ∇f 0 (w i−1 j) in parallel for m different tasks. First, the variable H to store the array of gradients is defined. Then all of the tasks are evenly assigned onto multiple processors to calculate ∇f 0 (w i−1 j). After the calculation, the results from each processor are sent back to each Hj ∈ H. The detailed settings are specified in experiment section.

Algorithm 1 Algorithm for CMTFL-I and CMTFL-II

Require: $\boldsymbol{X}, \boldsymbol{Y}, \rho, \eta > 1$
Ensure: solution $\boldsymbol{W}$

```
Initialize W^0, η ← 1.
for i ← 1, 2, . . . do
  Initialize L
  for j ← 1 . . . m do in parallel
    H_j ← ∇f′(w_j^{i−1})
  end for
  repeat
    S^i ← W^i − (1/L)∇H
    W^i ← proj(S^i) (defined in Lemma 1)
    L ← ηL
  until line search criterion is satisfied
  if the objective stop criterion satisfied then
    return W^i
  end if
end for
```

. The key idea of IGHT is to first use the gradient information in the current iteration to provide the first-order approximation of the objective function, then apply the projection operators to ensure the next iteration satisfies the given constraints. Specifically, we use the combination of the linear approximation of the function f(W) at a given point W0 and a quadratic penalty term, and solve the following problem: min W f(W0) + h∇f(W0), W − W0 i + ρ 2 kW − W0 k 2 F ,

s.t. X j I(kw j Kk > 0) ≤ u,

X j I(kw j Dk > 0) ≤ v, (7)

where ρ is a positive constant that can be estimated by a line search scheme. By ignoring the constants and rearranging the terms in Problem, we obtain the following sub-problem:

min W 1 2 kW − Sk 2 2

s.t.X j I(kw j Kk > 0) ≤ u

X j I(kw j Dk > 0) ≤ v. (8)

where S = W0 − 1 c ∇f(W0). aims to find the optimal point satisfying the constraint set that is closest to fixed point S.

IV. ALGORITHM FOR MODEL CMTFL-III

Note that Problem encompasses an norm in the objective function similar to that in Problem and utilizes a l0-norm constraint similar to that in Problem (4). Accordingly, the solution to Problem (6) combines these notions from IGHT and FISTA. The details are summarized in Algorithm 2. Similar to Algorithm 1, data parallelism has been applied to different tasks in the loop in Line 4 and the loop in Line 10. The key idea of the algorithm for CMTFL-III is as follows. First, we denote

f (W) = Pm j f 0 (wj),

where f 0 (wj) = P t log(1 + exp(−Yj,τ (wj · Xj,t)) + ρ1 kwjk 2 F .

Applying a linear approximation, we get the first-order approximation to the original objective function in Problem 6, as shown in the following equation:

min W f W0 + ∇f W0 , W − W0 + ρ 2 W − W0 2 F + ρ0kWKk2,1

s.t. X j w j D 0 ≤ v,

where ρ is a positive constant that can be estimated using a line search scheme.

Algorithm 2 Algorithm for CMTFL-III

Require: $\boldsymbol{X}, \boldsymbol{Y}, \rho_0, \rho_1, \eta > 1$
Ensure: solution $\boldsymbol{W}$

```
1:  Initialize W^0, η ← 1.
2:  for i ← 1, 2, . . . do
3:    Initialize L, H
4:    for j ← 1 . . . m do in parallel
5:      H_j ← ∇f′(w_j^{i−1})
6:    end for
7:    repeat
8:      S ← W^{i−1} − (1/L)∇H
9:      W_D^i ← proj(S_D)
10:     for j ← 1 . . . d do in parallel
11:       [W_K^i]_j ← prox_{ℓ2,1}([S_K]_j)
12:     end for
13:     L ← ηL
14:   until line search criterion is satisfied
15:   W^i ← [W_K^i; W_D^i]
16:   if the objective stop criterion satisfied then
17:     return W^i
18:   end if
19: end for
```

V. CONCLUSION

We propose an arrangement of effective calculations in light of the IGHT that can foresee spatial occasions continuously. Our observational outcomes exhibit that we can successfully recognize common turmoil and flu episode occasions, outflanking existing techniques by a considerable edge on both exactness and review. Different contextual investigations are given to show the value of the proposed strategy in pragmatic applications.

REFERENCES

[1] M. Sarah, C. Abdur, H. Gregor, L. Ben, and M. Roger, "Twitter and the Micro-Messaging Revolution," technical report, O'Reilly Radar, 2008.

[2] A. Java, X. Song, T. Finin, and B. Tseng, "Why We Twitter: Understanding Microblogging Usage and Communities," Proc. Ninth WebKDD and First SNA-KDD Workshop Web Mining and Social Network Analysis (WebKDD/SNA-KDD '07), pp. 56-65, 2007.

[3] B. Huberman, D. Romero, and F. Wu, "Social Networks that Matter: Twitter under the Microscope," ArXiv E-Prints, http:// arxiv.org/abs/0812.1045, Dec. 2008.

[4] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, A Social Network or A News Media?" Proc. 19th Int'l Conf. World Wide Web (WWW '10), pp. 591-600, 2010.

[5] G.L. Danah Boyd and S. Golder, "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter," Proc. 43rd Hawaii Int'l Conf. System Sciences (HICSS-43), 2010.

[6] A. Tumasjan, T.O. Sprenger, P.G. Sandner, and I.M. Welpe, "Predicting Elections with Twitter: What 140 Characters Reveal About Political Sentiment," Proc. Fourth Int'l AAAI Conf. Weblogs and Social Media (ICWSM), 2010.

[7] P. Galagan, "Twitter as a Learning Tool. Really," ASTD Learning Circuits, p. 13, 2009.

[8] K. Borau, C. Ullrich, J. Feng, and R. Shen, "Microblogging for Language Learning: Using Twitter to Train Communicative and Cultural Competence," Proc. Eighth Int'l Conf. Advances in Web Based Learning (ICWL '09), pp. 78-87, 2009.

[9] J. Hightower and G. Borriello, "Location Systems for Ubiquitous Computing," Computer, vol. 34, no. 8, pp. 57-66, 2001.

[10] M. Weiser, "The Computer for the Twenty-First Century," Scientific Am., vol. 265, no. 3, pp. 94-104, 1991.