Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter

Eiji ARAMAKI
The University of Tokyo
JST PRESTO
Tokyo, Japan
eiji.aramaki@gmail.com

Sachiko MASKAWA
The University of Tokyo
Tokyo, Japan
sachiko.maskawa@gmail.com

Mizuki MORITA
National Institute of
Biomedical Innovation
Osaka, Japan
morita.mizuki@gmail.com

Abstract

With the recent rise in popularity and scale of social media, a growing need exists for systems that can extract useful information from huge amounts of data. We address the issue of detecting influenza epidemics. First, the proposed system extracts influenza related tweets using Twitter API. Then, only tweets that mention actual influenza patients are extracted by the support vector machine (SVM) based classifier. The experiment results demonstrate the feasibility of the proposed approach (0.89 correlation to the gold standard). Especially at the outbreak and early spread (early epidemic stage), the proposed method shows high correlation (0.97 correlation), which outperforms the state-of-the-art methods. This paper describes that Twitter texts reflect the real world, and that NLP techniques can be applied to extract only tweets that contain useful information.

1 Introduction

Twitter[1], a popular micro-blogging service, has received much attention recently. It is an online network used by millions of people around the world to stay connected to their friends, family members, and co-workers through their computers and mobile telephones (Milstein et al., 2010).

Nowadays, Twitter users have increased rapidly. Its community estimated as 120 million worldwide, posts more than 5.5 million messages (*tweets*) every day (reported by Twitter.com in March 2011). Twitter can potentially serve as a valuable information resource for various applications. Huberman et al. (2009) analyzed the relations among friends. Boyd et al. (2010) investigated commutation activity. Sakaki et al. (2010) addressed the detection of earthquakes. Among the numerous potential applications, this study addresses the issue of detecting influenza epidemics, which presents two outstanding advantages over current methods.

- **Large Scale**: More than a thousand messages include the word "*influenza*" each day (Nov. 2008 – Oct. 2009). Such a huge data volume dwarfs traditional surveillance resources.

- **Real-time**: Twitter enables real-time and direct surveillance. This characteristic is extremely suitable for influenza epidemic detection because early stage detection is important for influenza warnings.

Although Twitter based influenza warnings potentially offer the advantages noted above, it might also expose inaccurate or biased information from tweets like the following (brackets **[]** indicate the comments):

- *Headache? You might have **flu.*** **[Suspicions]**
- *The World Health Organization reports the avian influenza, or bird **flu**, epidemic has spread to nine Asian countries in the past few weeks.* **[General News]**

[1] http://twitter.com/

Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, pages 1568–1576,
Edinburgh, Scotland, UK, July 27–31, 2011.

- *Are you coming down with **influenza**?* **[Question]**

Although these tweets include mention of "*influenza*" or "*flu*", they do not indicate that an influenza patient is present nearby. We regard such messages (merely suspicions/questions, general news, etc.) as **negative influenza tweets**. We call others **positive influenza tweets**. In our experiments, 42% of all tweets that include "*influenza*" are negative influenza tweets. The huge volume of such negative tweets biases the results.

This paper presents a proposal of a machine-learning based classifier to filter out negative influenza tweets. First, we build an annotated corpus of pairs of a tweet and positive/negative labels. Then, a support vector machine (SVM) (Cortes and Vapnik, 1995) based sentence classifier extracts only positive influenza tweets from tweets. In the experiments, the results demonstrated the high correlation (0.89 of the correlation), which is equal performance to that of the state-of-the-art method.

The specified research point of this study is two-fold:

(1) This report describes that an SVM-based classifier can filter out the negative influenza tweets (*f*-measure=0.76).
(2) Experiments empirically demonstrate that the proposed method detects the influenza epidemics with high accuracy (correlation ratio=0.89): it outperforms the state-of-the-art method.

2 Influenza Epidemic Detection

The detection of influenza epidemics is a national mission in every country for two reasons.

(1) Anti-influenza drugs, which differ among influenza types, must be prepared before the epidemics.
(2) We can only slightly predict what type of influenza will spread in any given season.

This situation naturally demands the early detection of influenza epidemics. This section presents a description of previous methods of influenza epidemic detection.

2.1 Traditional Approaches

Most countries have their own influenza surveillance organization/center: the U.S. has the Centers for Disease Control and Prevention (CDC)[2], the E.U. has its European Influenza Surveillance Scheme (EISS), and Japan has its Infection Disease Surveillance Center (IDSC). Their surveillance systems fundamentally rely on both virology and clinical data. For example, the IDSC gathers influenza patient data from 5,000 clinics and releases summary reports. Such manual systems typically have a 1–2 week reporting lag. This time lag is sometimes pointed out as a major flaw.

2.2 Recent Approaches

In an attempt to provide earlier influenza detection, various new approaches are proposed each year.

Espino et al. (2003) described a telephone triage service, a public service, to give advice to users via telephone. They investigated the number of telephone calls and reported a significant correlation with influenza epidemics.

Magruder (2003) used the amount of over-the-counter drug sales. Because an influenza patient usually requires anti-influenza drugs, this approach is reasonable. However, in most countries, anti-influenza drugs are not available at the drug store (only hospitals provide such drugs).

The state-of-the-art approach is that proposed by Ginsberg et al. (2009). They used Google web search queries that correlate with an influenza epidemic. Their approach demonstrated high accuracy (average correlation ratio of 0.97; min=0.92; max=0.99)[3]. Several research groups have used similar approaches. Polgreen et al. (2008) used a Yahoo! query log. Hulth et al. (2009) used a query log of a Switzerland web search engine.

Although the above approaches use different information, they share the same approach, which is to observe patient actions directly. This approach was sufficient to obtain more numerous data than traditional services. Nevertheless, such information is unfortunately limited only to the service provider. For example, web search queries are available only for several companies: Google, Yahoo!, and Microsoft.

This paper examines Twitter data, which are widely available. Note that Paul and Dredze (2011) also propose a similar Twitter based approach. While they focus on a word distribution, this paper

[2] http://www.cdc.gov/flu/weekly/
[3] Their service is available at http://www.google.org/flutrends/ (Google Flu Trend).

employs a sentence classification (discrimination of negative influenza tweets).

3 Influenza Corpus

As described in Section 1, it is necessary to filter out **negative influenza tweets** to infer precise amounts of influenza epidemics. To do so, we constructed the influenza corpus (Section 3). Then, we trained the SVM-based classifier using the corpus (Section 4).

The corpus comprises pairs of sentences and a label (positive or negative). Several examples are presented in Table 1. This corpus was built using the following procedure.

3.1 Influenza Tweet

First, we collected 300 million tweets, starting from 2008 November to 2010 June, via Twitter API. Crawling results are presented in Figure 1. We extracted only influenza-related tweets using a simple word look-up of "*influenza*". This operation gave us 0.4 million tweets. We separated the data into two data groups.

Training Data are 5,000 tweets sent in November 2008. These were annotated by human annotators, and were then used for training.

Test Data are the other data. They were used in experiments of influenza epidemics detection. Because of the three dropout periods (Figure 1), the test data were separated into four periods (winter 2008, summer 2009, winter 2009, and summer 2010).

3.2 Positive–negative Annotation

To each tweet in the training dataset, a human annotator assigned one of two labels: positive or negative. In this labeling procedure, we regarded a tweet that meets the following two conditions as positive data.

Condition 1 (A Tweet person or Surrounding persons have Flu): one or more people who have influenza should exist *around* the tweet person. Here, we regard "*around*" as a distance in the same city. In cases in which the distance is unknown, we regard it as negative. Because of this annotation policy, the re-tweet type message is negative.

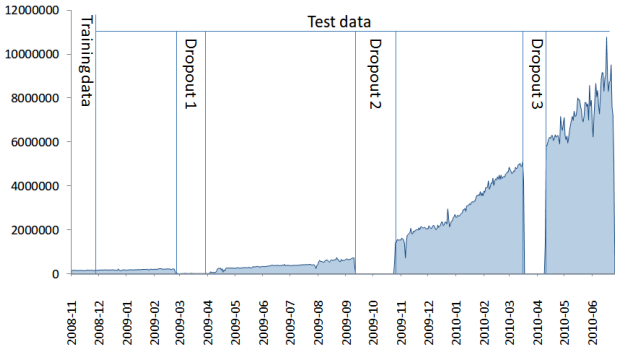

Figure 1: Twitter Data used in this Study.
The data include three dropout periods because the Twitter API specifications changed in those periods. The dropout periods were removed from evaluation in the experiments (Section 5).

Table 1: Corpus (Tweets with a Positive or Negative Label)

Positive(+1)/ Negative(-1)	Tweet
+1	A bad influenza is going around in our lab.
+1	I caught the flu. I was burning up.
+1	I think I'm coming down with the flu.
+1	It's the flu season. I had it and now he do es.
+1	Don't give me the flu. (Nearby people have the flu)
+1	My flu is worse than it was yesterday.
-1	In the normal flu season, 80 percent of deaths occur in people over 65 (Simply a fact)
-1	Influenza is now raging throughout Japan. (Too general.)
-1	His wife also contracted the bird flu, but has recovered. (Where is his wife?)
-1	You might have the flu. Has anyone around you had it? (Where are you?)
-1	Bird flu damage is spreading in Japan. (Too general.)

"+1" indicates a positive influenza tweet. "-1" indicates a negative influenza tweet. The case arc "()" indicates the reason for the positive or negative annotation.

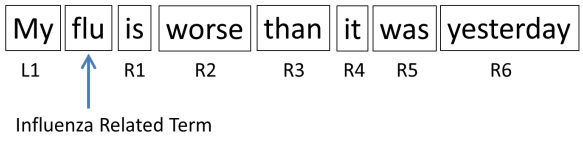

Figure 2: Feature Representation.
The word boundary is detected by a morph analyzer JUMAN[4].

[4] http://nlp.kuee.kyoto-u.ac.jp/nl-resource/juman.html

Condition 2 (Tense/Modality): The tense should be the present tense (current) or *recent past*. Here, we define the "*recent past*" as the prior 24 hour period (such as "*yesterday*"). The sentence should be affirmative (not interrogative and not subjunctive).

4 Influenza Positive–negative Classifier

Using the corpus (Section 3), we built a classifier that judges whether a given tweet is positive or negative. This task setting is similar to a sentence classification (such as spam e-mail filtering, sentiment analysis, and so on). We used a popular means for sentence classification, which is based on a machine learning classifier under the bag-of-words (BOW) representation (Figure 2). The parameters were investigated in preliminary experiments in terms of feature window size (Section 4.1) and machine-learning methods (Section 4.2). These preliminary experiments were conducted under the ten-fold cross variation manner using the training set.

4.1 Feature (window size)

Performance was dependent on the window size (the number of left/right side words). Figure 3 depicts the performance obtained using various window sizes. The best performance was scored at the BOTH=6 setting. Therefore, this window size was used for the following experiments. These results also indicated that entire sentences (BOTH=∞) are unsuitable for this task.

4.2 Machine Learning Method

We compared various machine-learning methods from two points of view: accuracy and time. The result, presented in Table 2, shows that SVM with a polynomial kernel showed feasibility from both viewpoints of accuracy and the training time.

5 Experiments

We assessed the detection performance using actual influenza reports provided by the Japanese IDSC.

5.1 Comparable Methods

We compared the various methods as follows:

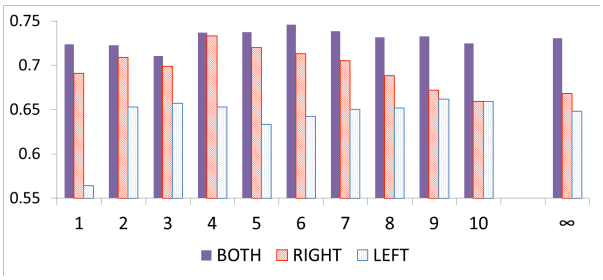

Figure 3: Window size and Accuracy (*F*-measure). RIGHT shows a method used only the right context. LEFT shows a method used only the left context. BOTH represents a method using both the right and left context. The number shows the window size. ∞ uses all words in each context direction.

Classifier	F-Measure	Training Time (sec)
AdaBoost (Freund 1996)	0.592	40.192
Bagging (Breiman 1996)	0.739	30.310
Decision Tree (Quinlan1993)	0.698	239.446
Logistic Regression	0.729	696.704
Naive Bayes	0. 741	**7.383**
Nearest Neighbor	0.695	22.441
Random Forest (Breiman 2001)	0.729	38.683
SVM (RBF kernel) (Cortes and Vapnik 1995)	0.738	92.723
SVM (polynomial kernel; d=2)	**0.756**	13.256

Table 2: Machine Learning Methods and Performance (*F*-measure and Training Time)

- **`TWEET-SVM`**: The proposed SVM-based method (window size = 6).
- **`TWEET-RAW`**: A simple frequency-based method. This approach outputs the relative frequency of word "*influenza*" appearing in Twitter.
- **`DRUG`**: The amounts of drug sales (sales of cold medicines). Statistics are provided by the Japanese Ministry of Health, Labor and Welfare.
- **`GOOGLE`**: Google flu trend detection (Japanese version). This method uses a query log of the Google search engine (Ginsberg et al., 2009)[5].

[5] http://www.google.org/flutrends/

5.2 Gold Standard and Test-Set

For gold standard data, we used data that are described in Section 2, as reported from IDSC. The report is released once a week. Therefore, the evaluation is done on a weekly basis.

We split the data into four seasons as follows:

- Season I: winter 2008,
- Season II: summer 2009,
- Season III: winter 2009,
- Season IV: summer 2010.

To investigate further detailed evaluations, we split the winters into two sub-seasons: **before the peak** and **after the peak**. We regard the peak point as the day with the highest number in that season. The statistics derived from the data are presented in Table 3.

Excessive News Period: In our experimental data, Season II and the earlier peak of Season III are special periods because news related to swine flu (H1N1 flu) is extremely hot in those seasons (Fig. 4). This paper calls them **Excessive News Periods**. We also investigated the results with and without the excessive news period.

Figure 4: A CNN news on "swine flu" in June 2009 (Season II in our experiment).

Experimental data include such excessive news periods.

5.3 Evaluation Metric

The evaluation metric is based on correlation (Pearson correlation) between the gold standard value and the estimated value.

5.4 Result

The results are presented in Table 4. In the non-excessive news period, the proposed method achieved the highest performance (0.890 correlation). This correlation is considerably higher than the query-based approach (**GOOGLE**), demonstrating the basic feasibility of the proposed approach. However, during the excessive news periods, the proposed method suffers from an avalanche of news, generating a news bias. This phenomenon is a remaining problem to be resolved in future studies.

6 Discussion

6.1 SVM-based Negative Filtering contributes to Performance

In most seasons, the proposed SVM approach (**TWEET-SVM**) shows higher correlation than the simple word lookup method (**TWEET-RAW**). The average improvement is 0.196 (max 0.56; min-0.009), which significantly boosts the correlation. This result demonstrates the basic feasibility of the proposed approach. In the future, more advantages attributable to the proposed approach can be obtained if the classification performance improves.

6.2 All Methods Suffer from News Bias in Excessive News Period

All methods expose the poor performance that prevails during the excessive news period (from Season II to Season III before the peak). Especially, tweet-based methods show dramatically reduced correlation, which indicates that Twitter is vulnerable to newswire bias.

One reason for that vulnerability is that Twitter is a kind of communication tool by which a tweet affects other people. Consequently, the possibility exists that a few tweets related to "flu" might spread widely, generating an explosive burst of influenza-related tweets. Future studies must address this burst phenomenon.

<table>
<tr><td colspan="5">All Season</td></tr>
<tr><td colspan="5">79 weeks (0.221)</td></tr>
<tr><td colspan="2">Season I

2008/11/9 - 2009/4/5</td><td>Season II

2009/4/12 - 2009/7/5</td><td colspan="2">Season III

2009/7/12 - 2010/2/14</td><td>Season IV

2010/2/21 - 2010/7/4</td></tr>
<tr><td colspan="2">22 weeks (0.423)</td><td rowspan="2">13 weeks
(0.553)</td><td colspan="2">26 weeks (0.388)</td><td rowspan="2">18 weeks
(0.468)</td></tr>
<tr><td>Before peak
2008/11/9-2009/1/25</td><td>After peak
2009/2/1-2009/4/5</td><td>Before peak
2009/7/12-2009/11/29</td><td>After peak
2009/12/6-2010/2/14</td></tr>
<tr><td>12weeks (0.576)</td><td>10 weeks
(0.632)</td><td></td><td>15 weeks
(0.514)</td><td>11 weeks
(0.602)</td><td></td></tr>
<tr><td colspan="2">Non-excessive news period</td><td colspan="2">Excessive news period</td><td colspan="2">Non-excessive news period</td></tr>
</table>

Table 3: Test-set Tracks and the number of data points (=weeks).
The number in the bracket indicates the statistical significance level.

		TWEET-RAW	TWEET-SVM (Proposed Method)	DRUG	GOOGLE
Excessive news period		0.001	0.060	**0.844**	**0.918**
Non- excessive news period		**0.831**	**0.890**	0.308	**0.847**
Season I		**0.683**	**0.816**	-0.208	**0.817**
	Before peak	**0.914**	**0.974**	-0.155	**0.962**
	After peak	**0.952**	**0.955**	0.557	**0.959**
Season II		-0.009	-0.018	0.406	0.232
Season III		0.382	**0.474**	**0.684**	**0.881**
	Before peak	0.390	0.474	**0.919**	**0.924**
	After peak	**0.960**	**0.944**	0.364	**0.936**
Season IV		0.391	**0.957**	0.130	**0.976**

Table 4: Results (Correlation Ratio).
The number in bold indicates the significance correlation (p=0.05). The number with underline indicates the highest value in each season.

6.3 Tweets have Advantages in Early Stage Detection

From practical viewpoints, the most important task is to detect influenza epidemics before the peak (early stage detection). Consequently, the correlation of the two seasons, Season I before the peak and Season III before the peak, presents the practical performance. Figure 5 portrays detailed results of all methods.

In Season I before the peak (Figure 5 Left), the proposed method (TWEET-SVM) shows the best performance among all methods.

In Season II before the peak (Figure 5 Right), all methods including the proposed method showed poor correlation because they are included in the excessive news periods. During that season, the newswires heavily reported the swine flu twice (April 2009 and May 2009). Because of this news, we can see two peaks in Twitter-based methods (TWEET-SVM and TWEET-RAW), which indicates that Twitter is more sensitive to the newswires.

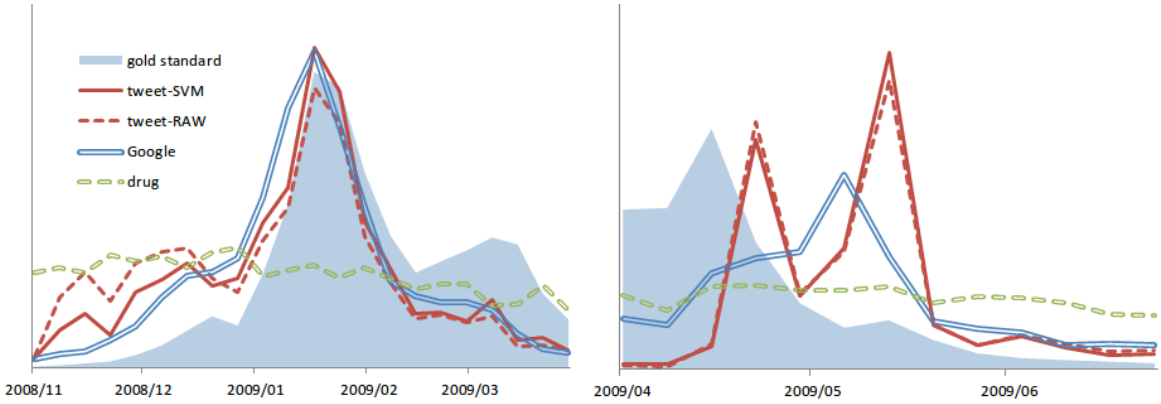

Figure 5: Predicted Values in Season I (Left) and Season II (Right):
the X-axis shows the date; the Y-axis shows the relative predicted value using each method.

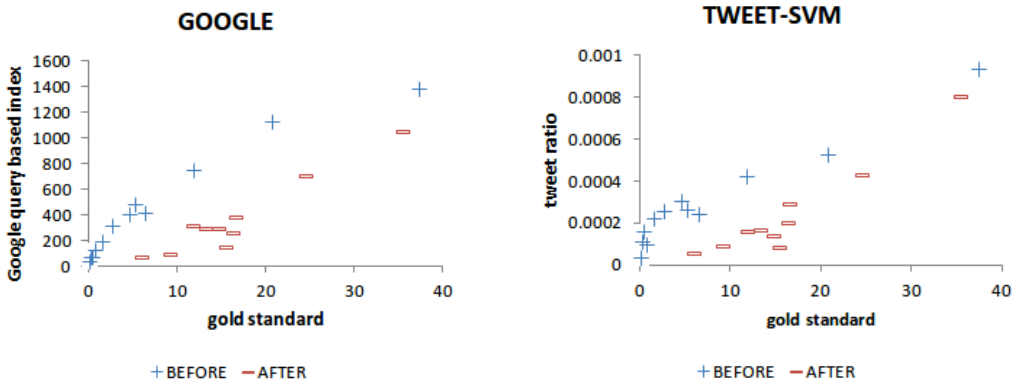

Figure 6: Patient Actions (Web Search Query and Tweet) is Sensitive before the Epidemic Peaks.
Distribution between the gold standard and Detected Values (Search Engine Query (Left) and Tweet (Right)): "+" denotes the distribution before the peak; "-" denotes the distribution after the peak.

6.4 Human Action is Sensitive before Epidemics

Figure 6 presents the distribution between the detected values (using **GOOGLE** and using **TWEET-SVM**) and the gold standard value (before the peak is shown by "+"; that after the peak is shown as "-"). Although the detected values fundamentally correlate with the gold standard, we can see different sensitivity before and after peak (The distribution before peak "+" is a higher value than after peak "-".).

Results show that human action, a web search (**GOOGLE**) and a tweet (**TWEET-SVM**), highly corresponds to the real influenza before the epidemic peaks, and vice versa. More acute detection is possible if we incorporate a model considering this aspect of human nature.

7 Related Works

The core technology of the proposed method is to classify whether the event is positive or negative. This task is similar to negation identification, which is a traditional topic, especially in medical fields. Therefore, we can find many previous studies of the topic in the relevant literature. An algorithm based approach, *NegEx* (Chapman et al., 2001), *Negfinder* (Mutalik et al., 2001), and *ConText* (Chapman et al., 2007), a machine learning based approach (Elkin et al., 2005; Huang and H.J. Lowe, 2007).

	Previous Negation (Syntactic)	This study: Negative Influenza (Semantic)
I caught a flu.	Positive sentence	Positive Influenza
I don't have the flu!	Negative sentence	Negative Influenza
I have enough flu drugs.	Positive sentence	Negative Influenza
I have not recovered from the flu.	Negative sentence	Positive Influenza

Table 5: Our target influenza negation (semantic) and previous negation (syntactic)

Although these approaches specifically examine the syntactic negation, this study detects the negative influenza, which is a specified semantic negation. Table 5 presents the difference between both negations. In general, the semantic operation is difficult in general. However, this paper revealed that the domain (influenza domain) specific semantic operation provides reasonable results.

Another aspect of this study is the target material, Twitter data, which have drawn much attention. Twitter can provide suitable material for many applications such as named entity recognition (NER) (Finin et al., 2010) and sentiment analysis (Barbosa and Feng, 2010). Although these studies specifically examine the fundamental NLP techniques, this study directly targets an NLP application that can contribute to our daily life.

8 Conclusion

This paper proposed a new Twitter-based influenza epidemics detection method, which relies on the Natural Language Processing (NLP). Our proposed method could successfully filter out the negative influenza tweets (*f*-measure=0.76), which are posted by the ones who did not actually catch the influenza. The experiments with the test data empirically demonstrate that the proposed method detects influenza epidemics with high correlation (correlation ratio=0.89), which outperforms the state-of-the-art Google method. This result shows that Twitter texts precisely reflect the real world, and that the NLP technique can extract the useful information from Twitter streams.

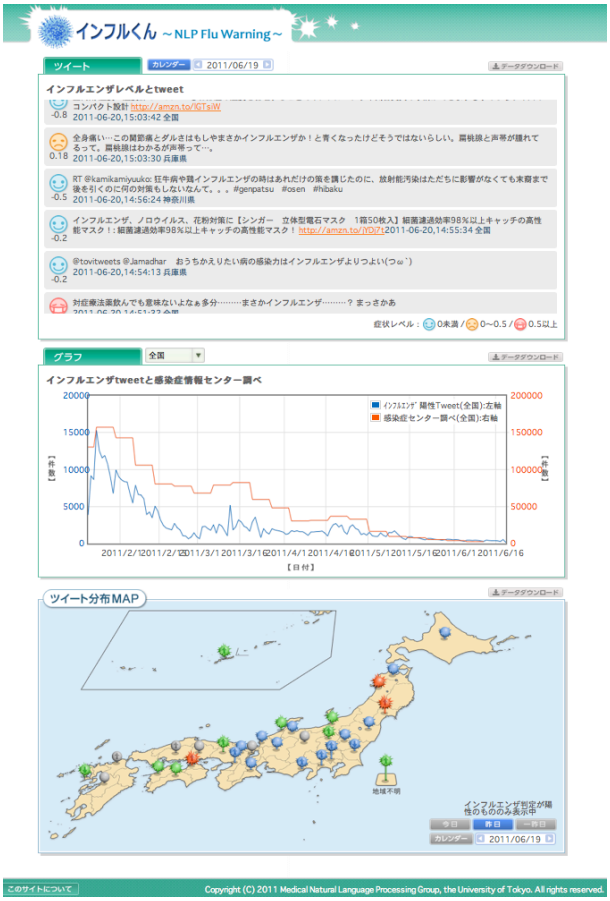

Figure 7: An influenza severance system "*INFLU kun*" using the proposed method is available at http://mednlp.jp/influ/.

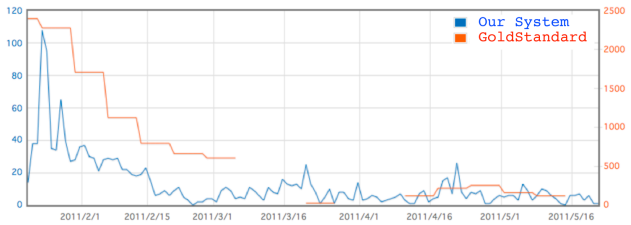

Figure 8: The Timeline of Influenza Epidemics in Fukushima. While the Infection Disease Surveillance Center (IDSC) sometimes stops (gold standard) due to the Great East Japan Earthquake, the proposed system could continue to work (Our System).

Available Resources

Corpus: The corpus of this study is provided at the http://mednlp.jp/~aramaki/KAZEMIRU/.

Web System: The web service is also released at http://mednlp.jp/influ/ (Figure 7 and Figure 8).

References

Barbosa, L. and J. Feng. 2010. Robust Sentiment Detection on Twitter from Biased and Noisy Data. In Proc. 23rd Intl. Conf. on Computational Linguistics (COLING).

Boyd, D., S. Golder, and G. Lotan. 2010. Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. In Proc. HICSS43.

Breiman L. Random Forests. 2001. Machine learning, 45(1): 5–32.

Breiman, L. Bagging predictors. 1996. Machine learning, 24(2):123–140.

Cortes C. and V. Vapnik. 1995. Support vector networks. In Machine Learning, pp. 273–297.

Chapman, W., W. Bridewell, P. Hanbury, G.F. Cooper, and B. Buchanan. 2001. A simple algorithm for identifying negated findings and diseases in discharge summaries. Journal of Biomedical Informatics, 5:301-310.

Chapman, W., J. Dowling, and D. Chu. 2007. ConText: An algorithm for identifying contextual features from clinical text. Biological, translational, and clinical language processing (BioNLP2007), pp. 81–88.

Elkin, P.L., S.H. Brown, B.A. Bauer, C.S. Husser, W. Carruth, L.R. Bergstrom, and D.L. Wahner-Roedler. 2005. A controlled trial of automated classification of negation from clinical notes. BMC Medical Informatics and Decision Making 5:13.

Espino, J., W. Hogan, and M. Wagner. 2003. Telephone triage: A timely data source for surveillance of influenza-like diseases. In Proc. of Annual Symposium of AMIA, pp. 215–219.

Finin, T., W. Murnane, A. Karandikar, N. Keller, J. Martineau, and M. Dredze. 2010. Annotating named entities in Twitter data with crowdsourcing. In Proc. NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk (CSLDAMT '10), pp. 80-88.

Freund, Y. and R. Schapire. 1996. Experiments with a new boosting algorithm. In Machine Learning Intl. Workshop, pp.148–156.

Ginsberg, J., M.H. Mohebbi, R.S. Patel, and L. Brammer. 2009. Detecting influenza epidemics using search engine query data, Nature Vol. 457 (19).

Huang, Y. and H.J. Lowe. 2007. A novel hybrid approach to automated negation detection in clinical radiology reports. Journal of the American Medical Informatics Association, 14(3):304-311.

Huberman, B. and D. R. F. Wu. 2009. Social networks that matter: Twitter under the microscope. First Monday, Vol. 14.

Hulth, A., G. Rydevik, and A. Linde. 2009. Web Queries as a Source for Syndromic Surveillance. PLoS ONE 4(2).

Johnson, HA., MM. Wagner, WR. Hogan, W. Chapman, RT. Olszewski, J. Dowling, and G. Barnas. 2004. Analysis of Web access logs for surveillance of influenza. Stud. Health Technol. Inform. 107(Pt 2):1202-1206.

Magruder, S. 2003. Evaluation of over-the-counter pharmaceutical sales as a possible early warning indicator of human disease. Johns Hopkins University APL Technical Digest 24:349–353.

Milstein, S., A. Chowdhury, G. Hochmuth, B. Lorica, and R. Magoulas. 2008. Twitter and the micro-messaging revolution: Communication, connections, and immediacy, 140 characters at a time. O'Reilly Media.

Mutalik, P.G., A. Deshpande, and P.M. Nadkarni. 2001. Use of general purpose negation detection to augment concept indexing of medical documents: A quantitative study using theUMLS. Journal of the American Medical Informatics Association, 8(6):598-609.

Paul, MJ. and M. Dredze. 2011. You Are What You Tweet: Analyzing Twitter for Public Health. In Proc. of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM).

Polgreen, PM., Y. Chen, D.M. Pennock, and F.D. Nelson. 2008. Using Internet Searches for Influenza Surveillance, Clinical Infectious Diseases Vol. 47 (11) pp. 1443-1448.

Quinlan. J. 1993. C4. 5: programs for machine learning. Morgan Kaufmann.

Sakaki, T., M. Okazaki, and Y. Matsuo. 2010. Earthquake shakes Twitter users: real-time event detection by social sensors, in Proc. of Conf. on World Wide Web (WWW).