Detection of Patients with Influenza Syndrome Using Machine-Learning Models Learned from Emergency Department Reports

Arturo López Pineda*, Fu-Chiang Tsui, Shyam Visweswaran and Gregory F. Cooper

University of Pittsburgh. Department of Biomedical Informatics, Pittsburgh, PA, USA

Objective

Compare 7 machine learning algorithms with an expert constructed Bayesian network on detection of patients with influenza syndrome.

Introduction

Early detection of influenza outbreaks is critical to public health officials. Case detection is the foundation for outbreak detection. Previous study by Elkin el al. demonstrated that using individual emergency department (ED) reports can better detect influenza cases than using chief complaints [1]. Our recent study using ED reports processed by Bayesian networks (using expert constructed network structure) showed high detection accuracy on detection of influenza cases [2].

Methods

The dataset used in this study includes 182 ED reports with confirmed PCR influenza tests (Jan 1, 2007-Dec 31, 2009) and 40853 ED reports as control cases from 8 EDs in UPMC (Jul 1, 2010-Aug 31, 2010). All ED reports were deidentified by De-ID software with IRB approval.

An NLP system, Topaz, was used to extract relevant findings and symptoms from the reports and encoded them with the UMLS concept unique identifier codes [2]. Two subsets were created: DS1-train (67% of cases) and DS1-test (remaining 33%).

The algorithms used for training the models are: Naïve Bayes Classifier, Efficient Bayesian Multivariate Classification (EBMC) [3], Bayesian Network with K2 algorithm, Logistic Regression (LR), Support Vector Machine (SVM), Artificial Neural Networks (ANN) and Random Forest (RF).

The predictive performance of each method was evaluated using the area under the receiver operator characteristic (AUROC) and the Hosmer-Lemeshow (HL) statistical significance testing, that describes the lack-of-fit of the model to the dataset.

Results

The evaluation results of all the models using DS1-test, including the AUROC, its confidence interval, p-value (between each algorithm and the expert) and the calibration with HL are shown in Table 1.

Conclusions

All models achieved high AUROC values. The pairwise comparison of p-values in Table 1 demonstrates that the AUROCs of all the machine-learning models and the expert model were not significantly different. Nevertheless, EBMC is the best fitted. The model created by EBMC is shown in Figure 1.

One limitation of the study is that the test dataset has low influenza prevalence, which may bias the detection algorithm performance. We are in the process of testing the algorithms using higher prevalence rate.

The same process could also be applied to other diseases to further research the generalizability of our method.

Predictive performance and Calibration

Algorithm	AUROC	95% CI	p-value	Calibration: HL
NaïveBayes	0.9988	(0.9983, 0.9994)	0.2342	4880.63
EBMC	0.9993	(0.9989, 0.9998)	0.2255	4.53
BN-K2	0.9994	(0.9990, 0.9998)	0.2228	1315.71
LR	0.9829	(0.9512, 1.0000)	0.8935	177.01
SVM	0.9996	(0.9993, 0.9999)	0.2189	12.30
RandForest	0.9995	(0.9993, 0.9998)	0.2201	16.30
A-NN	0.9991	(0.9986, 0.9997)	0.2300	275.81
Expert	0.9798	(0.9483, 1.0000)	1.0000	14374.67

Area under the ROC curve (AUROC) with 95% Confidence Interval; p-value relative to the Expert model; and Hosmer-Lemeshow calibration statistic
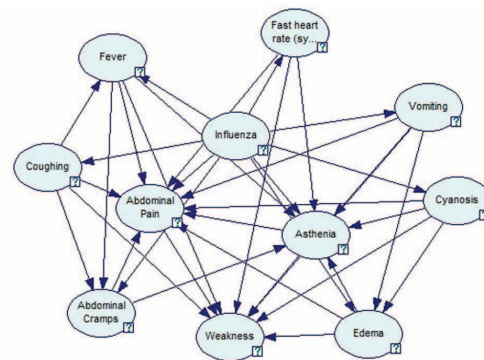
Influenza Syndrome model created using the EBMC algorithm

Keywords

influenza; machine-learning; ED reports

References

[1] Elkin, P. L., Froehling, D. A., Wahner-Roedler, D. L., Brown, S. H., & Bailey, K. R. (2012). Comparison of natural language processing biosurveillance methods for identifying influenza from encounter notes. Annals of Internal Medicine, 156(1 Pt 1), 11–18.

[2] Tsui, F.-C., Wagner, M., Cooper, G. F., Que, J., Harkema, H., Dowling, J., Sriburadej, T., et al. (2011). Probabilistic Case Detection for Disease Surveillance Using Data in Electronic MedicalRecords. Online Journal of Public Health Informatics, 1–17.

[3] Cooper, G. F., Hennings-Yeomans, P. P., & Barmada, M. M. (2010). An efficient bayesian method for predicting clinical outcomes from genome-wide data. AMIA 2010 Symposium Proceedings, 2010, 127–131.

*Arturo López Pineda
E-mail: arl68@pitt.edu