Inductive Modeling in Subjectivity/Sentiment Analysis (case study: dialog processing)

Mikhail Alexandrov[1], Xavier Blanco[1], Angels Catena[1], Natalia Ponomareva[2]

[1]*Autonomous University of Barcelona, Barcelona, 08193, Spain*
[2]*University of Wolverhampton, Stafford Street, Wolverhampton, WV1 1SB, UK*

dyner1950@mail.ru,{xavier.blanco,angels.catena}@uab.cat
nata.ponomareva@wlv.ac.uk

Abstract. *Subjectivity/sentiment analysis is an area of natural language processing, which aims to determine people opinions or sentiments with respect to some subject or event. This paper presents a methodology for constructing empirical formulae based on lexico-syntactic properties of people utterances in order to evaluate their politeness, satisfaction and competence. This methodology includes three steps: 1) Linguistic analysis, where a set of linguistic indicators (LIs) regarding each characteristic is selected; 2) Manual annotations, where documents are manually estimated by experts; 3) Inductive modeling using the Ivakhnenko method of model self-organization (IMMSO) in order to find an optimal model describing dependency between LIs and manual estimations. The suggested methodology is applied to a real set of dialogs between passengers and Directory Inquires in Barcelona railway station.*

Keywords

Subjectivity Analysis, Sentiment Analysis, Inductive Modeling

1 Introduction

Subjectivity/sentiment analysis (SSA) aims at automatic evaluation of personal characteristics of individuals on a basis of their utterances. Personal characteristics can refer to personal reactions concerning some issue, like satisfaction/dissatisfaction, like/dislike, interest, etc., or to some social and educational features, like politeness, competence, communicativeness, etc. Although these characteristics are determined by the personality of an individual, they also depend on a situation a person deals with and, therefore, their analysis can be useful not only from psychological point of view. For example, mining of these characteristics can help companies to reveal weak features of their products or services, estimate a level of product advertisement or accomplish other analysis in order to improve and consolidate the position of a company in the market. Nowadays SSA becomes a subject of consideration in International Conferences on Weblogging and Social Media [8].

In the paper we consider three personal characteristics: politeness, satisfaction and competence. As an automatic evaluation of these characteristics is quite difficult and time-consuming task, a semi-supervised learning based on a set of LIs and manual experts annotations was employed. Normally semi-supervised methods, like, for example, Mean Squared Regression, assume a certain type of functional dependency between features (in our case LIs) and manual annotations. The goal of our approach consists in automatic discovery of this dependency if at least a class of functions (polynomials, harmonic, etc.) is given. The best function is selected using IMMSO suggested by Ivakhnenko and developed by his colleagues [2,3,7]. As an example, in this paper we consider series of polynomials in order to find the one that better describes the character of dependency between LIs and manual estimations.

It should be said that we have already applied Inductive Modeling (IM) for politeness estimation [1]. As a difference with our previous work we use two external criteria (regularity and unbiasedness) instead of only one (regularity). Numerical estimation of politeness and satisfaction were also described in [5,6] where linearity between LIs and manual estimations was assumed. Results of IM carried out in this paper justify linear type of dependency.

Our corpus consists of dialogs between passengers and Directory Inquires of Barcelona railway station. In spite of the fact that the linguistic analysis was focused on the subject of railway inquires, the suggested methodology itself is domain independent and can be easily adopted to any similar types of problems.

2 Methodology description

The process of IM contains the following steps:

- Objects (dialogs) are parameterized on the basis of selected LIs. It can be polite words, expressions of approval, specialised vocabulary revealing competence, etc. The number of LIs occurrences is normalized (if it is necessary) regarding a dialog length.
- All objects (dialogs) are evaluated manually and these estimations can be considered as an experimental data. We use the scales [0,1] or [-1,1] with the step 0.5.
- The dialog set is divided into two parts: one for training the model and the other one for testing it. One additional independent part is used for estimation of model quality (standard error).

US: Good evening, *Could you* tell me the schedule of trains to Zaragoza for tomorrow?
DI: For tomorrow morning?
US: Yes
DI: There is one train at 7-30 and another at 8-30
US: And later?
DI: At 10-30
US: And till the noon?
DI: At 12
US: *Could you* tell me the schedule till 4 p.m. more or less?
DI: At 1-00 and at 3-30
US: 1-00 and 3-30
DI: hmm, hmm <SIMULTANEOUSLY>
US: And the next one?
DI: I will see, one moment. The next train leaves at 5-30
US: 5-30
DI: hmm, hmm < SIMULTANEOUSLY >
US: Well, and how much time does it take to arrive?
DI: 3 hours and a half
US: For all of them?
DI: Yes
US: Well, *could you* tell me the price?
DI: 3800 pesetas for a seat in the second class
US: Well, and what about a return ticket?
DI: The return ticket has a 20% of discount
US: Well, so, it is a little bit more than 6 thousands, no?
DI: Yes
US: Well, *thank you very much*
DI: Don't mention it, good bye

Fig. 1. Example of dialog with indicators of politeness

Figure 1 shows an example of a dialog we consider in the paper. For estimation of politeness we used (15+15+10) dialogs and for estimation of satisfaction and competence – (20+20+10) dialogs. The final models are constructed on a full data set.

To find the best polynomial model we use multicriteria decision making described in [4]: first, the vicinity of the global minimum for the criterion of regularity (K_r) is determined, and then, the model with the minimum value of the criterion of unbiasedness (K_u) is chosen.

3 Inductive modeling

3.1 Politeness evaluation

We consider the following series of polynomial models of increasing complexity for politeness estimation:

Model 0: $F(\boldsymbol{g},\boldsymbol{w},\boldsymbol{v}) = A_0$
Model 1: $F(\boldsymbol{g},\boldsymbol{w},\boldsymbol{v}) = C_0\boldsymbol{g} + B_{10}\boldsymbol{w} + B_{01}\boldsymbol{v}$
Model 2: $F(\boldsymbol{g},\boldsymbol{w},\boldsymbol{v}) = C_0\boldsymbol{g} + B_{10}\boldsymbol{w} + B_{01}\boldsymbol{v} + B_{11}\boldsymbol{vw}$
Model 3: $F(\boldsymbol{g},\boldsymbol{w},\boldsymbol{v}) = C_0\boldsymbol{g} + B_{10}\boldsymbol{w}^2 + B_{01}\boldsymbol{v}^2$
Model 4: $F(\boldsymbol{g},\boldsymbol{w},\boldsymbol{v}) = C_0\boldsymbol{g} + B_{11}\boldsymbol{vw} + B_{20}\boldsymbol{w}^2 + B_{02}\boldsymbol{v}^2$
Model 5: $F(\boldsymbol{g},\boldsymbol{w},\boldsymbol{v}) = C_0\boldsymbol{g} + B_{10}\boldsymbol{w} + B_{01}\boldsymbol{v} + B_{11}\boldsymbol{vw} + B_{20}\boldsymbol{w}^2 + B_{02}\boldsymbol{v}^2$

Linguistic indicators are: 1) $\boldsymbol{g}$ – first greeting (binary parameter 0/1); 2) $\boldsymbol{w}$ – polite words ('please', 'thank you', etc.); 3) $\boldsymbol{v}$ – polite verbal forms ('could you', 'would you', etc.). The last two parameters are normalized on a dialog length. The coefficients A_0,C_0, B_{ij} are unknown variables to be determined. The manual estimations belong to the interval [0,1], where 0 means an ordinary politeness and 1 means an excessive politeness.

Table 1 and Figure 2 show the values of the criteria K_r and K_u for different models.

Table 1. Criterion K_r and K_u (models of politeness)

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
K_r	0.51	**0.26**	0.44	0.76	1.65	0.36
K_u	0.02	**0.08**	0.28	0.55	1.18	0.25

Obviously, lineal Model-1 is the best model according to the multicriteria decision making. Its standard error is equal to 0.16. The final equation is:

$$F(\boldsymbol{g},\boldsymbol{w},\boldsymbol{v}) = 0.18\boldsymbol{g} + 3.29\boldsymbol{w} + 3.43\boldsymbol{v}$$

3.2 Satisfaction evaluation

We used the following set of polynomial models for satisfaction estimation:

Model 0: $F(\boldsymbol{b},\boldsymbol{f},\boldsymbol{q}) = A_0$
Model 1: $F(\boldsymbol{b},\boldsymbol{f},\boldsymbol{q}) = B_{100}\boldsymbol{b} + B_{010}\boldsymbol{f} + B_{001}\boldsymbol{q}$
Model 2: $F(\boldsymbol{b},\boldsymbol{f},\boldsymbol{q}) = B_{100}\boldsymbol{b} + B_{010}\boldsymbol{f} + B_{001}\boldsymbol{q} + B_{110}\boldsymbol{bf} + B_{101}\boldsymbol{bq} + B_{011}\boldsymbol{fq}$
Model 3: $F(\boldsymbol{b},\boldsymbol{f},\boldsymbol{q}) = B_{200}\boldsymbol{b}^2 + B_{020}\boldsymbol{f}^2 + B_{002}\boldsymbol{q}^2$
Model 4: $F(\boldsymbol{b},\boldsymbol{f},\boldsymbol{q}) = B_{110}\boldsymbol{bf} + B_{101}\boldsymbol{bq} + B_{011}\boldsymbol{fq} + B_{200}\boldsymbol{b}^2 + B_{020}\boldsymbol{f}^2 + B_{002}\boldsymbol{q}^2$
Model 5: $F(\boldsymbol{b},\boldsymbol{f},\boldsymbol{q}) = B_{100}\boldsymbol{b} + B_{010}\boldsymbol{f} + B_{001}\boldsymbol{q} + B_{110}\boldsymbol{bf} + B_{101}\boldsymbol{bq} + B_{011}\boldsymbol{fq} + B_{200}\boldsymbol{b}^2 + B_{020}\boldsymbol{f}^2 + B_{002}\boldsymbol{q}^2$

Linguistic indicators are: 1)-2) ***b*** and ***f*** – positive user feedback in a dialog body ('well', 'ok', 'all right', 'correct', 'splendid', etc) and in farewell respectively. They are presented separately as they have different level of importance. 3) ***q*** – question-answers, which convey dissatisfaction. The indicators ***b*** and ***q*** are normalized on a text length. The coefficients B_{ijk} are unknown variables to be determined. The manual estimations belong to the interval [-1,1], where -1 means strong dissatisfaction and 1 means full satisfaction. Table 2 and Figure 3 show the values of the criteria K_r and K_u for different models.

Table 2. Criterion K_r and K_u (models of satisfaction)

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
K_r	0.96	**0.69**	0.64	**0.69**	0.73	0.73
K_u	0.05	**0.15**	0.22	**0.10**	0.39	0.47

The results are not single-valued: the lowest values for both criteria give two models: Model-1 and Model-3. Their standard errors are equal to 0.35 and 0.38 respectively. The final equations are :

$$F(\boldsymbol{b},\boldsymbol{f},\boldsymbol{q}) = 0.18\boldsymbol{b} + 0.06\boldsymbol{f} - 1.11\boldsymbol{q} \qquad F(\boldsymbol{b},\boldsymbol{f},\boldsymbol{q}) = 0.20\boldsymbol{b}^2 + 0.006\boldsymbol{f}^2 - 1.78\boldsymbol{q}^2$$

3.3 Competence evaluation

For competence estimation we used the same series of polynomials as for satisfaction estimation although a set of LIs is different:

Model 0: $F(\boldsymbol{f},\boldsymbol{l},\boldsymbol{q}) = A_0$
Model 1: $F,\boldsymbol{f},\boldsymbol{l},\boldsymbol{q}) = B_{100}\boldsymbol{f} + B_{010}\boldsymbol{l} + B_{001}\boldsymbol{q}$
etc.

Linguistic indicators are: 1) ***f*** – level of competence in the first passenger question ('any train to…?' vs. 'regional express at night to…?', etc.); 2) ***l*** – specialised lexicon used in dialogs (name of trains and services); 3) ***q*** – question-answers, which refer to passenger questions that got either positive or negative answer. Table 3 and Figure 4 show the values of the criteria K_r and K_u for different models.

Table 3. Criterion K_r and K_u (models of competence)

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
K_r	0.93	**0.26**	0.27	0.86	0.66	0.28
K_u	0.04	**0.16**	0.23	0.39	0.78	0.23

The Model-1 obtained the best result for both criteria. Its standard error is equal to 0.26. The final equation is:

$$F(\boldsymbol{f},\boldsymbol{l},\boldsymbol{q}) = 0.52\boldsymbol{f} + 0.19\boldsymbol{l} + 0.16\boldsymbol{q}$$

Finally Figure 5 shows the values of criteria for lineal models used in the paper.

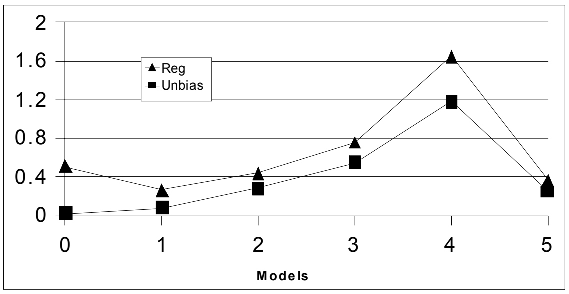

Fig.2. Criteria for politeness evaluation

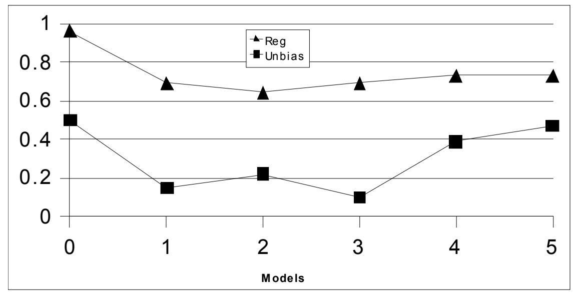

Fig.3. Criteria for satisfaction evaluation

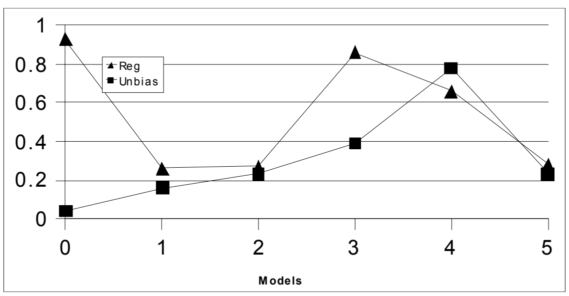

Fig.4. Criteria for competence evaluation

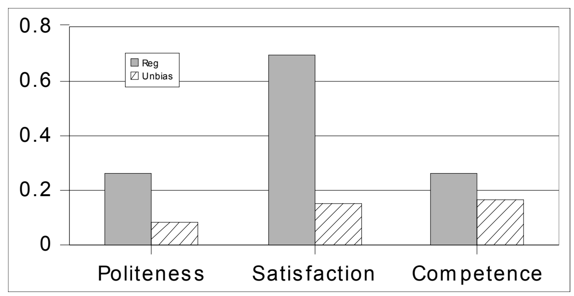

Fig.5. Criteria for lineal models

4 Conclusions

This paper presents the methodology for automatic estimation of various personal characteristics, such as politeness, satisfaction and competence. The obtained errors are less than the steps of the manual dialog estimation. The suggested methodology can be used for other applications of subjectivity/sentiment analysis.

References

[1] Alexandrov, M., Blanco, X., Ponomareva, N., Rosso, P. (2007): Constructing Empirical Models for Automatic Dialog Parameterization. In: *"Text, Speech and Dialogue",* Springer, LNCS, N_4629, pp. 455-462

[2] Ivakhnenko, A.: Inductive Method of Model Self-Organization of Complex Systems. *Kiev, Naukova Dumka Publ.*1982 (rus.)

[3] Ivakhnenko, A., Stepashko, V.: Noise Immunity of Modeling. *Kiev, Naukova Dumka Publ.*1985 (rus.)

[4] Ivakhnenko, A., Savchenko, E.: Investigation of Efficiency of Additional Determination Method of the Model Choice in the Modelling Problems with Application of GMDH Algorithms. *Journ. of Inform. and Automat. Sciences, Begell House Inc*, 2008, Vol. 40, N_3, pp. 4-22

[5] Ponomareva, N., Blanco, X. (2008): Regression Model for Politeness Estimation Trained on Examples. In *"NooJ Applications"*, Proc. of the Intern. Conf. NooJ-2007. Cambridge Scholars Publishing, UK, 2008, 7 pp.

[6] Ponomareva, N., Catena, A.: Linguistic Models Construction and Analysis for Satisfaction Estimation. In: *"Text, Speech, Dialog"*, Springer, LNCS, 2009 (to be published)

[7] Stepashko, V.: Method of Critical Variances as an Analytical Tool of the Inductive Modeling Theory. *Journ. of Inform. and Automat. Sciences, Begell House Inc*, 2008, Vol. 40, N_3, pp. 47-58

[8] Webloging conferences: www.icwsm.org/2008/; www.icwsm.org/2009/; www.lsi.us.es/womsa09/