INTERVAL PREDICTION BASED ON EXPERTS' STATEMENTS*

Gennadiy Lbov, Maxim Gerasimov

Abstract: In the work [1] we proposed an approach of forming a consensus of experts' statements in pattern recognition. In this paper, we present a method of aggregating sets of individual statements into a collective one for the case of forecasting of quantitative variable.

Keywords: interval prediction, distance between expert statements, consensus.

ACM Classification Keywords: I.2.6. Artificial Intelligence - knowledge acquisition.

Introduction

Let Γ be a population of elements or objects under investigation. By assumption, L experts give predictions of values of unknown quantitative feature Y for objects $a \in \Gamma$, being already aware of their description $X(a)$. We assume that $X(a) = (X_1(a),...,X_j(a),...,X_n(a))$, where the set X may simultaneously contain qualitative and quantitative features X_j, $j = \overline{1,n}$. Let D_j be the domain of the feature X_j, $j = \overline{1,n}$, D_y be the domain of the feature Y. The feature space is given by the product set $D = \prod_{j=1}^{n} D_j$.

In this paper, we consider statements S^i, $i = \overline{1,M}$; represented as sentences of type "if $X(a) \in E^i$, then $Y(a) \in G^i$", where $E^i = \prod_{j=1}^{n} E_j^i$, $E_j^i \subseteq D_j$, $E_j^i = [\alpha_j^i, \beta_j^i]$ if X_j is a quantitative feature, E_j^i is a finite subset of feature values if X_j is a nominal feature, $G^i = [y_1^i, y_2^i] \subseteq D_y$. By assumption, each statement S^i has its own weight w^i. Such a value is like a measure of "assurance".

Preliminary Analysis

We begin with some definitions.

Denote by $E^{i_1 i_2} := E^{i_1} \oplus E^{i_2} = \prod_{j=1}^{n} (E_j^{i_1} \oplus E_j^{i_2})$, where $E_j^{i_1} \oplus E_j^{i_2}$ is the *Cartesian join* of feature values $E_j^{i_1}$ and $E_j^{i_2}$ for feature X_j and is defined as follows. When X_j is a nominal feature, $E_j^{i_1} \oplus E_j^{i_2}$ is the union: $E_j^{i_1} \oplus E_j^{i_2} = E_j^{i_1} \bigcup E_j^{i_2}$. When X_j is a quantitative feature, $E_j^{i_1} \oplus E_j^{i_2}$ is a minimal closed interval such that $E_j^{i_1} \bigcup E_j^{i_2} \subseteq E_j^{i_1} \oplus E_j^{i_2}$ (see Fig. 1).

* The work was supported by the RFBR under Grant N07-01-00331a.

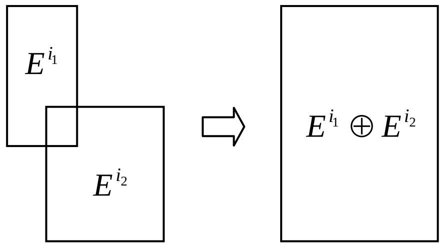

Fig. 1.

In the works [2, 3] we proposed a method to measure the distances between sets (e.g., E^1 and E^2) in heterogeneous feature space. Consider some modification of this method. By definition, put

$\rho(E^1,E^2)=\sum_{j=1}^{n}k_j\rho_j(E_j^1,E_j^2)$ or $\rho(E^1,E^2)=\sqrt{\sum_{j=1}^{n}k_j(\rho_j(E_j^1,E_j^2))^2}$, where $0\le k_j\le 1$, $\sum_{j=1}^{n}k_j=1$.

Values $\rho_j(E_j^1,E_j^2)$ are given by: $\rho_j(E_j^1,E_j^2)=\dfrac{|E_j^1\Delta E_j^2|}{|D_j|}$ if X_j is a nominal feature, $\rho_j(E_j^1,E_j^2)=\dfrac{r_j^{12}+\theta|E_j^1\Delta E_j^2|}{|D_j|}$ if X_j is a quantitative feature, where $r_j^{12}=\left|\dfrac{\alpha_j^1+\beta_j^1}{2}-\dfrac{\alpha_j^2+\beta_j^2}{2}\right|$. It can be proved that the triangle inequality is fulfilled if and only if $0\le\theta\le 1/2$.

The proposed measure ρ satisfies the requirements of distance there may be.

We first treat each expert's statements separately for rough analysis. Let us consider some special cases.

Case 1 ("coincidence"): $\max_j \max(\rho_j(E^{i_1},E^{i_1}\oplus E^{i_2}),\rho_j(E^{i_2},E^{i_1}\oplus E^{i_2}))<\delta$ and $\rho(G^{i_1},G^{i_2})<\varepsilon_1$, where δ, ε_1 are thresholds decided by the user, $i_1,i_2\in\{1,...,M\}$. In this case we unite statements S^{i_1} and S^{i_2} into resulting one: "if $X(a)\in E^{i_1}\oplus E^{i_2}$, then $Y(a)\in G^{i_1}\oplus G^{i_2}$".

Case 2 ("inclusion"): $\min(\max_j(\rho_j(E^{i_1},E^{i_1}\oplus E^{i_2})),\max_j(\rho_j(E^{i_2},E^{i_1}\oplus E^{i_2}))<\delta$ and $\rho(G^{i_1},G^{i_2})<\varepsilon_1$, where $i_1,i_2\in\{1,...,M\}$. In this case we unite statements S^{i_1} and S^{i_2} too: "if $X(a)\in E^{i_1}\oplus E^{i_2}$, then $Y(a)\in G^{i_1}\oplus G^{i_2}$".

Case 3 ("contradiction"): $\max_j \max(\rho_j(E^{i_1},E^{i_1}\oplus E^{i_2}),\rho_j(E^{i_2},E^{i_1}\oplus E^{i_2}))<\delta$ and $\rho(G^{i_1},G^{i_2})>\varepsilon_2$, where ε_2 is a threshold decided by the user, $i_1,i_2\in\{1,...,M\}$. In this case we exclude both statements S^{i_1} and S^{i_2} from the list of statements.

Consensus

Consider the list of l-th expert's statements after preliminary analysis $\Omega_1(l)=\{S^1(l),...,S^{m_l}(l)\}$. Denote by $\Omega_1=\bigcap_{l=1}^{L}\Omega_1(l)$, $M_1=|\Omega_1|$.

Determine values k_j from this reason: if far sets G^{i_1} and G^{i_2} corresponds to far sets $E_j^{i_1}$ and $E_j^{i_2}$, then the feature X_j is more "valuable" than another features, hence, value k_j is higher. We can use, for example, these values: $k_j = \frac{\tau_j}{\sum_{i=1}^{n} \tau_i}$, where $\tau_j = \sum_{u=1}^{M_1} \sum_{v=1}^{M_1} \rho(G^u, G^v) \rho_j(E_j^u, E_j^v)$, $j = \overline{1,n}$.

Denote by $r^{i_1 i_2} := d(E^{i_1 i_2}, E^{i_1} \bigcup E^{i_2})$.

The value $d(E,F)$ is defined as follows: $d(E,F) = \max\limits_{E' \subseteq E \setminus F} \min\limits_{j} \frac{k_j \,|\, E'_j \,|}{diam(E)}$, where E' is any subset such that its projection on subspace of quantitative features is a convex set (see Fig. 2), $diam(E) = \max\limits_{x,y \in E} \rho(x,y)$.

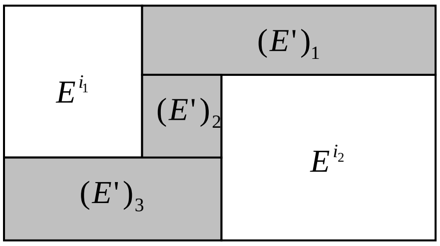

Fig. 2.

By definition, put $I_1 = \{\{1\},...,\{m_l\}\},...,$ $I_q = \{\{i_1,...,i_q\} \,|\, r^{i_u i_v} \le \delta$ and $\rho(G^{i_u}, G^{i_v}) < \varepsilon_1 \quad \forall u,v = \overline{1,q}\}$, where δ, ε_1 are thresholds decided by the user, $q = \overline{2,Q}$; $Q \le M_1$. Let us remark that the requirement $r^{i_u i_v} \le \delta$ is like a criterion of "insignificance" of the set $E^{uv} \setminus (E^{i_u} \bigcup E^{i_v})$. Notice that someone can use another value d to determine value r, for example:

$$d(E,F,G) = \max_{E' \subseteq E \setminus (F \bigcup G)} \frac{\min(diam(F \oplus E') - diam(F), diam(G \oplus E') - diam(G))}{diam(E)}.$$

Further, take any set $J_q = \{i_1,...,i_q\}$ of indices such that $J_q \in I_q$ and $\forall \Delta = \overline{1, Q-q} \quad J_q \not\subset J_{q+\Delta}$ $\forall J_{q+\Delta} \in I_{q+\Delta}$. Now, we can aggregate the statements S^{i_1}, …, S^{i_q} into the statement S^{J_q}:

$S^{J_q} =$"if $X(a) \in E^{J_q}$, then $Y(a) \in G^{J_q}$", where $E^{J_q} = E^{i_1} \oplus ... \oplus E^{i_q}$, $G^{J_q} = G^{i_1} \oplus ... \oplus G^{i_q}$.

By definition, put to the statement S^{J_q} the weight $w^{J_q} = \frac{\sum_{i \in J_q} c^{iJ_q} w^i}{\sum_{i \in J_q} c^{iJ_q}}$, where $c^{iJ_q} = 1 - \rho(E^i, E^{J_q})$.

The procedure of forming a consensus of single expert's statements consists in aggregating into statements S^{J_q} for all J_q under previous conditions, $q = \overline{1,Q}$.

Let us remark that if, for example, $k_1 < k_2$, then the sets E_1 and E_2 (see Fig. 3) are more suitable to be united (to be precise, the relative statements), then the sets F_1 and F_2 under the same another conditions.

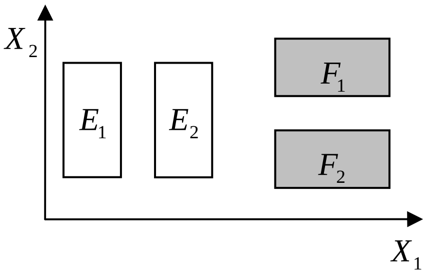

Fig. 3.

Note that we can consider another criterion of unification (instead of $r^{i_u i_v} \le \varepsilon$): aggregate statements S^{i_1}, ..., S^{i_q} into the statement S^{J_q} only if $w^{J_q} > \varepsilon'$, where ε' is a threshold decided by the user.

After coordinating each expert's statements separately, we can construct an agreement of several independent experts. The procedure is as above, except the weights: $w^{J_q} = \sum_{i \in J_q} c^{iJ_q} w^i$ (the more experts give similar statements, the more we trust in resulted statement).

Denote the list of statements after coordination by Ω_2, $M_2 := |\Omega_2|$.

Coordination

After constructing of a consensus of similar statements, we must form decision rule in the case of intersected non-similar statements. The procedure in such cases is as follows.

To each $h = \overline{2, M_2}$ consider statements $S^{(1)},...,S^{(h)} \in \Omega_2$ such that $\tilde{E}^h := E^{(1)} \cap ... \cap E^{(h)} \ne \varnothing$, where $E^{(i)}$ are related sets to statements $S^{(i)}$.

Denote $I(l) = \left\{ i \middle| S^i(l) \in \Omega_1(l), \quad E^i(l) \cap \tilde{E}^h \ne \varnothing \right\}$, where $E^i(l)$ are related sets to statements $S^i(l)$.

Consider related sets $G^i(l)$, where $l = \overline{1, L}$; $i \in I(l)$. Denote by $w^i(l)$ the weights of statements $S^i(l)$.

As above, unite sets $G^{(i_1)}(l_1)$,..., $G^{(i_q)}(l_q)$ if $\rho(G^{i_u}, G^{i_v}) < \varepsilon_1 \quad \forall u, v = \overline{1, q}$. Denote by $\tilde{G}^1$,..., $\tilde{G}^\lambda$,..., $\tilde{G}^\Lambda$ the sets $G^i(l)$ after procedure of unification. Consider the statements $\tilde{S}^\lambda$: "if $X(a) \in \tilde{E}^h$, then $Y(a) \in \tilde{G}^\lambda$".

In order to choose the best statement, we take into consideration these reasons:

1) similarities between sets $\tilde{E}^h$ and $E^i(l)$;
2) similarities between sets $\tilde{G}^\lambda$ and $G^i(l)$;
3) weights of statements $S^i(l)$;
4) we must distinguish cases when similar / contradictory statements produced by one or several experts.

We can use, for example, such values: $w^\lambda = \sum_{l=1}^{L} \frac{\sum_{i \in I(l)} (1 - \rho(G^{(i)}(l), \tilde{G}^{(\lambda)}))(1 - \rho(E^{(i)}(l), \tilde{E}^h)^2 w^i(l)}{\sum_{i \in I(l)} (1 - \rho(E^{(i)}(l), \tilde{E}^h)}$.

Denote by $\lambda^* := \arg\max_\lambda w^\lambda$.

Thus, we can make decision statement: $\tilde{S}^h =$ "if $X(a) \in \tilde{E}^h$, then $Y(a) \in \tilde{G}^{\lambda^*}$" with the weight $\tilde{w}^h := w^{\lambda^*} - \max_{\lambda \neq \lambda^*} w^{\lambda}$.

Denote the list of such statements by Ω_3.

Final decision rule is formed from statements in Ω_2 and Ω_3. Notice that we can range resulted statements in Ω_2 and Ω_3 by their weights and exclude "ignorable" statements from decision rule.

Conclusion

Suggested method of forming of united decision rule can be used for coordination of several experts statements, and different decision rules obtained from learning samples and/or time series.

Bibliography

[1] G.Lbov, M.Gerasimov. Constructing of a Consensus of Several Experts Statements. In: Proc. of XII Int. Conf. "Knowledge-Dialogue-Solution", 2006, pp. 193-195.

[2] G.S.Lbov, M.K.Gerasimov. Determining of distance between logical statements in forecasting problems. In: Artificial Intelligence, 2'2004 [in Russian]. Institute of Artificial Intelligence, Ukraine.

[3] G.S.Lbov, V.B.Berikov. Decision functions stability in pattern recognition and heterogeneous data analysis [in Russian]. Institute of Mathematics, Novosibirsk, 2005.

Authors' Information

Gennadiy Lbov - Institute of Mathematics, SB RAS, Koptyug St., bl.4, Novosibirsk, Novosibirsk State University, Russia; e-mail: lbov@math.nsc.ru

Maxim Gerasimov - Institute of Mathematics, SB RAS, Koptyug St., bl.4, Novosibirsk State University, Russia, e-mail: max_post@ngs.ru