

МЕТОДОЛОГІЯ ЗАСТОСУВАННЯ ІНТЕРВАЛЬНОЇ ОЦІНКИ ДЛЯ МІНІМІЗАЦІЇ РИЗИКУ КЛАСИФІКАТОРА ОПОРНО-ВЕКТОРНИХ МАШИН

Олександр Галкін

Національний технічний університет України "Київський політехнічний інститут", Київ, Україна

METHODOLOGY OF APPLICATION OF THE SPAN ESTIMATE FOR MINIMIZING OF THE RISK CLASSIFIER OF THE SUPPORT VECTOR MACHINES

Oleksandr Galkin

National Technical University of Ukraine "Kiev Polytechnic Institute", Kyiv, Ukraine

АНОТАЦІЯ. Досліджена методологія отримання обмежень та оцінок помилки процедури пропуску, що є об'єктивною оцінкою помилки узагальнення. Розглянуто випадок нероздільності та застосування обмежень на здатність узагальнення ОВМ. Дана точна оцінка процедури пропуску з використанням можливості заміни діаметра навчальних точок на інтервал опорних векторів після процедури пропуску.

Ключові слова: інтервал, опорний вектор, процедура пропуску.

АННОТАЦИЯ. Исследована методология получения ограничений и оценок ошибки процедуры пропуска, что является объективной оценкой ошибки обобщения. Рассмотрен случай неразделимости и применения ограничений на способность обобщения ОВМ. Дана точная оценка процедуры пропуска с использованием возможности замены диаметра учебных точек на интервал опорных векторов после процедуры пропуска.

Ключевые слова: интервал, опорный вектор, процедура пропуска.

SUMMARY. Purpose. Construction estimating or limiting core support-vector machines. **Methodology/approach.** A theoretical studies based on the known theories of vector sets. **Findings.** Methodology is investigated for obtaining bounds and estimates of the error of the leave-one-out procedure that is an objective estimate of generalization error. **Research limitations/implications.** After experimenting with the use of the smallest diameter of the sphere (which contains training points), the diameter of the smallest sphere (containing the support vectors) and the interval of support vectors using a set of data we were able to conduct an effective model selection model the optimal parameter values OVM to choose. Based on the database with separate data and database from noisy data and using the rule for interval prediction error test it was found that the prediction is very accurate and curves almost identical. **Originality/value.** The foundation laid for the development of new nuclear musculoskeletal vector machines.

Key words: span, support vector, leave-one-out procedure.

Вступ

Відомо, що ключовим фактором у продуктивності опорно-векторних машин є вибір ядра. Однак, на практиці, через складність налаштування параметрів, використовуються лише декілька різних типів ядер. Процедура налаштування параметрів ядра є надзвичайно важливою та може бути універсальним інструментом для вирішення різних завдань, таких як пошук правильної форми ядра, вибір ознак, знаходження відповідностей між помилкою та полем, виявлення викидів. В ідеалі, значення параметрів ядра повинно мінімізувати ризик істинного класифікатора ОВМ. На жаль, оскільки це значення є недоступним, для його знаходження необхідно будувати оцінки або обмеження.

Виклад основного матеріалу

Розглянемо спочатку випадок роздільних даних. Припустимо, що:

$$(x_1, y_1), \dots, (x_n, y_n)$$

є множиною опорних векторів, а $\alpha^0 = (\alpha_1^0, \dots, \alpha_n^0)$ є вектором множників Лагранжа для оптимальної гіперплощини.

Для будь-якого фіксованого опорного вектора x_p , визначимо множину Λ_p , як обмежену лінійну комбінацію точок $\{x_i\}_{i \neq p}$:

$$\Lambda_p = \sum_{i=1, i \neq p}^n \lambda_i x_i,$$

де $\sum_{i=1, i \neq p}^n \lambda_i x_i = 1$ та $\forall i \neq p, \alpha_i^0 + y_i y_p \alpha_p^0 \lambda_i \geq 0$.

Необхідно зауважити, що λ_i може бути менше 0.

Ми також визначаємо величину S_p , яка називається *інтервалом* [1] опорного вектора x_p та визначає відстань між x_p та наступною множиною:

$$S_p^2 = d^2(x_p, \Lambda_p) = \min_{x \in \Lambda_p} (x_p - x)^2. \quad (1)$$

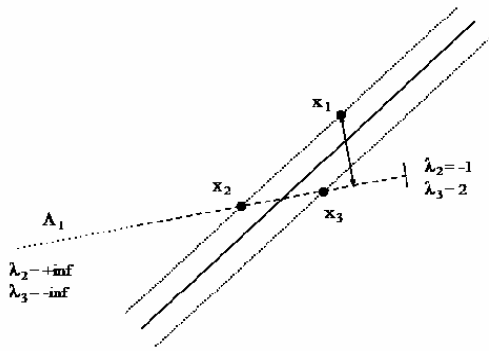


Рис. 1. Три опорних вектори з $\alpha_1 = \alpha_2 = \alpha_3 / 2$
 Fig. 1. Three support vectors with $\alpha_1 = \alpha_2 = \alpha_3 / 2$

Множина Λ_1 є напіввідкритою пунктирною лінією (рис.1):

$$\Lambda_1 = \{\lambda_2 x_2 + \lambda_3 x_3, \lambda_2 + \lambda_3 = 1, \lambda_2 \geq -1, \lambda_3 \leq 2\}.$$

Як показано на рисунку 2, може виникнути ситуація, коли $x_p \in \Lambda_p$.

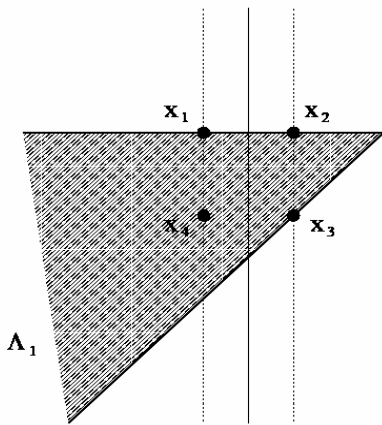


Рис. 2. $x_1 \in \Lambda_1$, тому $d(x_1, \Lambda_1) = 0$
 Fig. 2. $x_1 \in \Lambda_1$ so $d(x_1, \Lambda_1) = 0$

Множина Λ_1 обчислюється, використовуючи $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Максимальне значення S_p називається S -інтервалом:

$$S_p = d(x_p, \Lambda_p) = 0.$$

У цьому випадку точка x_p є лінійною комбінацією інших опорних векторів, а розширення розділової гіперплощини може не включати цієї точки. Крім того, функція

рішення буде незмінною після видалення цієї точки.

Інтуїтивно зрозуміло, що для невеликих $S_p = d(x_p, \Lambda_p)$, процедура пропуску має менше шансів зробити помилку на векторі x_p . Крім цього, ми доведемо (лема 2), що якщо $S_p < 1/(D\alpha_p^0)$ (D - діаметр найменшої сфери, що містить навчальні точки), тоді процедура пропуску класифікує x_p правильно.

Встановлюючи $\lambda_p = -1$, ми можемо записати S_p у вигляді:

$$S_p^2 = \min \left(\sum_{i=1}^n \lambda_i x_i \right)^2, \quad (2)$$

де $\lambda_p = -1$, $\sum_{i=1}^n \lambda_i = 0$ та $\alpha_i^0 + y_i y_p \alpha_p^0 \lambda_i \geq 0$.

Ми доведемо (лема 1), що $S_p \leq D_{sv}$, де D_{sv} є діаметром найменшої сфери, що містить всі опорні вектори. Отже:

$$S \leq D_{sv}. \quad (3)$$

Однак, залежно від взаємного розташування опорних векторів та $\alpha^0 = (\alpha_1^0, \dots, \alpha_n^0)$, значення інтервалу S може бути значно меншим, ніж діаметр D_{sv} опорних векторів. Дійсно, на прикладі рисунку 2, $d(x_1, \Lambda_1) = 0$, а враховуючи симетрію, $d(x_i, \Lambda_i) = 0$, для всіх i . Таким чином, у цьому прикладі $S = 0$.

Тепер ми узагальнимо поняття інтервалу для випадку нероздільних даних. У цьому випадку, ми будемо розрізняти дві категорії опорних векторів: опорні вектори, для яких:

$$0 < \alpha_i < C, i = 1, \dots, n^*$$

та опорні вектори, для яких:

$$\alpha_j = C, j = n^* + 1, \dots, n_{sv}.$$

Ми визначаємо інтервал опорних векторів, використовуючи опорні вектори першої категорії. Це означає, що розглядається значення $S_p = d(x_p, \Lambda_p)$, де

$$\Lambda_p = \sum_{i=1, i \neq p}^n \lambda_i x_i, \quad (4)$$

де $\sum_{i=1, i \neq p}^n \lambda_i = 1$, $\forall i \neq p$, $0 \leq \alpha_i^0 + y_i y_p \alpha_p^0 \lambda_i \leq C$.

Відмінність у визначенні інтервалу для випадків роздільності та нероздільності в тому, що у випадку нероздільності ми ігноруємо опорні вектори другої категорії та додаємо верхню границю C в обмеженнях на λ_i .

Отже, у випадку нероздільності [2], значення інтервалу опорних векторів залежить від значення C . Однак, не є очевидним те, що множина Λ_p є не порожньою. Це доведено в наступній лемі.

Лема 1. Як у випадку роздільності, так і у випадку нероздільності, множина Λ_p є не порожньою. Більш того, $S_p = d(x_p, \Lambda_p) \leq D_{sv}$ є діаметром найменшої сфери, що містить опорні вектори першої категорії.

Далі представлені нові обмеження на здатність узагальнення ОВМ. Ідея полягає в тому, що обмеження будуть залежати від інтервалу опорних векторів, які дають більш жорсткі обмеження в залежності від діаметру навчальних точок [3].

Перш за все, введемо наш фундаментальний результат:

Лема 2. Якщо в процедурі пропуску, опорний вектор x_p , що відповідає $0 < \alpha_p < C$, розпізнається невірно, тоді нерівність

$$\alpha_p^0 S_p \max(D, 1/\sqrt{C}) \geq 1$$

є справедливою.

Попередня лема приводить до наступної теореми для випадку роздільності:

Теорема 1. Припустимо, що ОВМ розділяє навчальні дані розміру n без помилок. Тоді, число помилок, допущених в процедурі пропуску, обмежено наступним чином:

$$\frac{SD}{M^2}.$$

Доведення. Припустимо, що за процедурою пропуску проведена невірна класифікація опорного вектора x_p . Тоді лема 2 дає

$\alpha_p^0 S_p D \geq 1$ (тут ми розглядаємо випадок роздільності та $C = \infty$) та нерівність

$$\alpha_p^0 SD \geq 1$$

є справедливою. Тепер підсумуємо ліву та праву частини цієї нерівності по всіх опорних векторах, де процедура пропуску робить помилку

$$\ell(x_1, y_1, \dots, x_n, y_n) \leq SD \sum_* \alpha_i^0 \leq SD \sum_{i=1}^n \alpha_i^0.$$

\sum_* вказує, що сума береться тільки по опорних векторах, де процедура пропуску робить помилку.

Дійсно, з умов Куна – Такера ($\alpha_i^0 [y_i (w_0 \cdot x_i + b_0) - 1] = 0$) та рівнянь:

$$\frac{\partial L(w, b, \alpha)}{\partial b} = \sum_{i=1}^n y_i \alpha_i = 0$$

$$\frac{\partial L(w, b, \alpha)}{\partial w} = w - \sum_{i=1}^n \alpha_i y_i x_i = 0,$$

ми маємо:

$$0 = \sum_{i=1}^n \alpha_i^0 [y_i (w_0 \cdot x_i + b_0) - 1] = w_0^2 - \sum_{i=1}^n \alpha_i^0. \quad (5)$$

Оскільки поле $M = 1/\|w_0\|$, ми отримуємо:

$$\ell(x_1, y_1, \dots, x_n, y_n) \leq \frac{SD}{M^2}.$$

Для випадку нероздільності, справедливою є наступна теорема.

Теорема 2. Кількість помилок, допущених процедурою пропуску, обмежена

$$S \max(D, 1/\sqrt{C}) \sum_{i=1}^n \alpha_i^0 + m,$$

де сума береться тільки за таких α_i , що відповідають опорним векторам першої категорії (для яких $0 < \alpha_i < C$), а m є кількістю опорних векторів другої категорії (для яких $\alpha_i = C$).

Доведення. Доведення цієї теореми аналогічно доведенню теореми 1. Ми розглядаємо всі опорні вектори другої категорії ($\alpha_i = C$) як помилку. Для першої категорії опорних векторів [4], ми оцінюємо число $\ell^*(x_1, y_1, \dots, x_n, y_n)$ помилок в процедурі

пропуску, використовуючи лему 2, як і в доведенні *теорему 1*.

Ми отримуємо:

$$\begin{aligned} \ell(x_1, y_1, \dots, x_n, y_n) &\leq \ell^*(x_1, y_1, \dots, x_n, y_n) + m \leq \\ &\leq S \max(D, 1/\sqrt{C}) \sum \alpha_i^* + m. \end{aligned}$$

Необхідно зауважити, що у випадку, коли $m = 0$ (випадок роздільності), рівність (5) є справедливою. У цьому випадку (за умови, що C є досить великим), обмеження, отримані в цих двох теоремах, збігаються.

Значення інтервалу S може бути набагато меншим, ніж діаметр. Крім цього, в *теоремах 1 і 2*, використовуючи нерівність (3), можна обмежити значення інтервалу S за діаметром найменшої сфери, що містить опорні вектори D_{SV} . Однак, відповідно до результатів експерименту, це може призвести до послаблення обмежень, оскільки інтервал може бути набагато меншим за діаметр.

Відповідно до доведення *леми 2*, діаметр навчальних точок D можна замінити на інтервал опорних векторів після процедури пропуску. Однак, оскільки множина опорних векторів після процедури пропуску є невідомою, ми обмежили цей невідомий інтервал за D . Тим не менш, це зауваження стало мотивацією проаналізувати випадок, коли множина опорних векторів залишається незмінною протягом процедури пропуску [5].

У цій ситуації ми маємо право замінити D на S в *лемі 2*, а також справедливою є наступна теорема.

Теорема 3. Якщо множини опорних векторів першої та другої категорії залишаться незмінними протягом процедури пропуску, тоді для будь-якого опорного вектора x_p , має місце наступна рівність:

$$y_p(f^0(x_p) - f^p(x_p)) = \alpha_p^0 S_p^2, \quad (6)$$

де f^0 та f^p є функціями рішення, заданими ОБМ під час тестування на всій навчальній множині та після того, як точка x_p була видалена.

Зауважимо, що при такому припущенні, вираз множини Λ_p (4), який визначає значення $S_p = d(x_p, \Lambda_p)$, може бути спрощений до

$$\Lambda_p = \sum_{i \neq p, 0 < \alpha_i^0 < C} \lambda_i \Phi(x_i), \quad (7)$$

де $\sum_{i \neq p} \lambda_i = 1$.

Теорема 3 призводить до кількох зауважень:

1. Припущення про те, що множина опорних векторів не змінюється під час процедури пропуску, може бути невірним. Тим не менш, пропорція точок, які порушують це припущення, як правило, є малою в порівнянні з кількістю опорних векторів. У цьому випадку, *теорема 3* забезпечує відповідне наближення результату процедури пропуску.

2. *Теорема 3* є більш потужною, ніж *лема 2* з трьох причин: термін $S_p \max(D, 1/\sqrt{C})$ стає S_p^2 , нерівність виявляється рівністю, а результат є справедливим для будь-якого опорного вектора.

3. З доведення *теорему 3* виходить, що якщо розглядати розширення $x^* = \sum \lambda_i x_i \in \Lambda_p$ точки мінімізації відстані між x_p та Λ_p , оптимальний розв'язок після видалення x_p стає $\alpha_i^p = \alpha_i^0 + y_i y_p \alpha_i^0 \lambda_i$. Це дає нову інтуїтивну інтерпретацію коефіцієнтів λ_i для визначення інтервалу: вони вимірюють зміни в α після видалення точки x_p з навчальної множини [6].

Нарешті, попередня теорема дозволяє обчислити число помилок, допущених процедурою пропуску:

Наслідок 1. Використовуючи припущення *теорему 3*, передбачення тестової помилки в процедурі пропуску визначається наступним чином:

$$\begin{aligned} t_n &= \frac{1}{n} \ell(x_1, y_1, \dots, x_n, y_n) = \\ &= \frac{1}{n} \text{Card}\{p : \alpha_p^0 S_p^2 \geq y_p f^0(x_p)\}. \end{aligned} \quad (8)$$

Як буде показано в експериментах, інтервальна оцінка (8) дає дуже точну оцінку

помилки пропуску. Однак, при стандартному виконанні, обчислення інтервалу S_p включає в себе мінімізацію та виконує $O(n^3)$ операцій. Крім того, загальна складність є $O(n^4)$, яка є такою ж, як складність явного виконання процедури пропуску.

Тим не менш, ще однією перевагою робити припущення, що множина опорних векторів є незмінною [7], є те, що обмеження нерівності, що входять у визначення (4) Λ_p , більше не потрібні для отримання спрощеного виразу (7). Як наслідок, можна обчислити інтервальну оцінку (8), використовуючи $O(n^3)$ операцій.

Як і раніше, ми припускаємо, що навчальні точки впорядковані таким чином, що $0 < \alpha_i^0 < C$ для $i = 1, \dots, n^*$, $\alpha_i^0 = C$ для $i = n^* + 1, \dots, n_{sv}$ та $\alpha_i^0 = 0$ для $i = n_{sv} + 1, \dots, n$.

Інтервал опорного вектора x_p визначається як відстань між точкою $\Phi(x_p)$ та множиною Λ_p , що визначається за (7). Тоді значення інтервалу можна записати у вигляді:

$$S_p^2 = \min_{\lambda} \max_{\mu} (\Phi(x_p) - \sum_{i \neq p}^{n^*} \lambda_i \Phi(x_i))^2 + 2\mu (\sum_{i \neq p}^{n^*} \lambda_i - 1).$$

Ми ввели множник Лагранжа μ для забезпечення дотримання обмежень $\sum \lambda_i = 1$.

Нехай K^* є $n^* \times n^*$ матрицею скалярних добутків між опорними векторами першої категорії та \tilde{K}_{sv} розширеною матрицею

$$\tilde{K}_{sv} = \begin{pmatrix} K^* & 1 \\ 1^T & 0 \end{pmatrix}.$$

Нарешті, нехай $\tilde{\lambda}$ буде розширеним вектором $(\lambda^T \mu)^T$.

Визначимо особливість характеру x_p :

Опорний вектор першої категорії
($0 < \alpha_p^0 < C, p \leq n^*$).

Значення інтервалу можна записати у вигляді:

$$S_p^2 = \min_{\lambda} \max_{\mu} (K(x_p, x_p) - 2v_p^T \tilde{\lambda} + \tilde{\lambda}^T H_p \tilde{\lambda}),$$

де H_p є підматрицею \tilde{K}_{sv} з видаленим рядком та стовпцем p , а v_p є p -м стовпцем \tilde{K}_{sv} .

Оскільки оптимальне значення $\tilde{\lambda} \in H_p^{-1} v_p$, ми маємо:

$$S_p^2 = K(x_p, x_p) - v_p^T H_p^{-1} v_p = 1 / (\tilde{K}_{sv}^{-1})_{pp}. \quad (9)$$

Остання рівність виходить з наступного перетворення:

$$\begin{pmatrix} A_1 & A^T \\ A & A_2 \end{pmatrix}^{-1} = \begin{pmatrix} B_1 & B^T \\ B & B_2 \end{pmatrix},$$

де $B_1 = (A_1 - AA_2^{-1}A^T)^{-1}$.

Опорний вектор другої категорії

$$(\alpha_p^0 = C, n^* < p \leq n_{sv}).$$

Аналогічне одержання призводить до модифікованої версії рівняння (9),

$$S_p^2 = K(x_p, x_p) - v_p^T \tilde{K}_{sv}^{-1} v_p,$$

де v_p є вектор розмірності $n^* + 1$, де i -й елемент є $K(x_p, x_i)$, а останній елемент дорівнюватиме 1.

Відсутність опорного вектора

$$(\alpha_p^0 = 0, p > n_{sv}).$$

У цьому випадку, відомо, що $f^p(x_p) = f^0(x_p)$. Це також є прямим наслідком *теорема 3*, оскільки $\alpha_p^0 = 0$. Обчислення інтервалу цієї точки не є обов'язковим, оскільки процедура пропуску не робить помилку на цій точці.

Замкнута форма, яку ми отримуємо, є особливо придатною, оскільки ми можемо обчислити значення інтервалу для кожного опорного вектора тільки шляхом інвертування матриці \tilde{K}_{sv} [9].

Попередні обмеження на здатність узагальнення опорних векторів включали діаметр найменшої сфери, що містить навчальні точки [4]. Ми показали в нерівності 3, що інтервал S є завжди меншим, ніж цей

діаметр, однак, для отримання підтвердження, були проведені декілька експериментів. Спочатку, ми порівнювали діаметр найменшої сфери (що містить навчальні точки), діаметр найменшої сфери (що містить опорні вектори), а також інтервал опорних векторів з використанням множини даних.

Множина даних складається з 7291 цифр розміром 16x16 з тестовою множиною на 2007 прикладів. Ми розділили навчальну множину на 23 підмножини, що містять 317 навчальних прикладів. Наша задача полягала в тому, щоб розділити цифри 0-4 від 5-9. Помилкові бар'єри на рисунку 3 є стандартними відхиленнями в 23 дослідженнях. Діаметри та інтервал на рисунку 3, побудовані для різних значень σ , а ширина використовуваного нами радіального ядра:

$$K(x, y) = e^{-\frac{|x-y|^2}{2\sigma^2}}$$

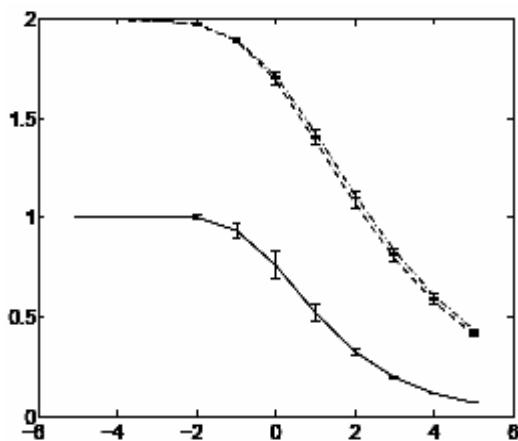


Рис. 3. Порівняння D , D_{SV} та S

Fig. 3. Comparison D , D_{SV} and S

У цьому прикладі, інтервал є до 6 разів меншим, ніж діаметр.

Тепер ми використовуємо інтервал для точного прогнозування тестової помилки. Це дозволить нам проводити вибір ефектної моделі, тобто вибирати оптимальні значення параметрів ОБМ. Важливо зауважити, що інтервал S визначається як максимум $S = \max_p S_p$, тому урахування різних значень S_p повинно забезпечити більш точну оцінку помилки узагальнення, ніж лише інтервал S . Таким чином, ми використали

правило інтервалу (8) для передбачення тестової помилки [9].

В наших експериментах були задіяні дві бази даних: база даних з роздільними даними, що була описана вище та база даних з наявністю шумів. Остання була поділена випадковим чином 100 разів на навчальну множину, що містить 200 прикладів та тестову множину, що містить 77 прикладів. Згідно рисунка 4а порівнюється тестова помилка та її передбачення, що задається правилом інтервалу (8) для різних значень ширини радіального ядра σ . На рисунку 4б виводяться ті ж функції для різних значень C з множини зашумлених даних. Для цього випадку, використовувалось радіальне ядро з $\sigma^2 = d = 9$.

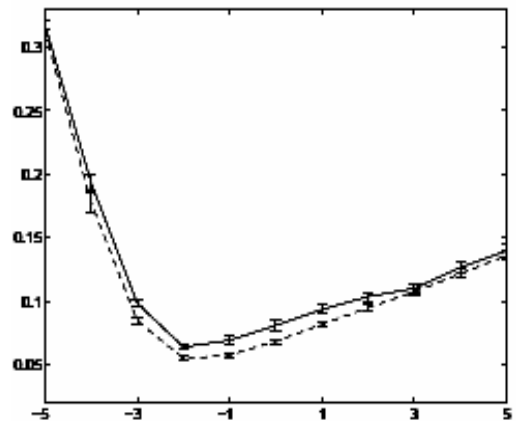


Рис. 4 а. Вибір σ в базі з роздільними даними
Fig. 4 a. Selecting σ in base with separate data

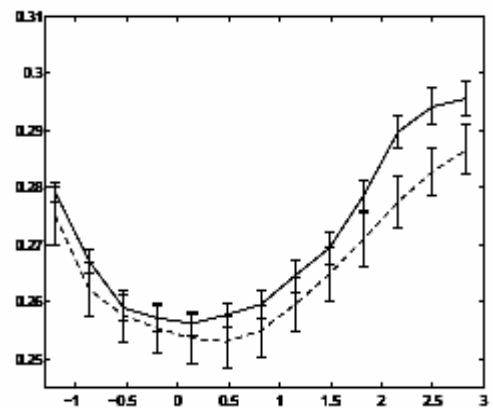


Рис. 4 б. Вибір C в базі з шумами
Fig. 4 b. Selecting C in base with noise

Відповідно до результатів, прогнозування є дуже точним, а криві - майже ідентичними.

Наостанок зауважимо, що складність обчислення інтервальної оцінки є $O(N_{sv}^3)$

(N_{SV} є числом опорних векторів, що має таку ж складність, як і стандартне навчання ОВМ).

Висновки

У даній статті досліджена методологія отримання обмежень та оцінок помилки процедури пропуску, що є об'єктивною оцінкою помилки узагальнення. Зокрема, детально проаналізована інтервальна оцінка для мінімізації ризику класифікатора ОВМ та узагальнено поняття інтервалу для випадку нероздільних даних. Крім цього, проаналізована точна оцінка процедури пропуску, де використана можливість заміни діаметра навчальних точок на інтервал опорних векторів після процедури пропуску.

Також у статті представлені результати практичного обчислення інтервальної оцінки. Провівши експерименти із використанням діаметра найменшої сфери (що містить навчальні точки), діаметра найменшої сфери (що містить опорні вектори), а також інтервалу опорних векторів з використанням множини даних, ми отримали можливість проводити вибір ефективної моделі, тобто вибирати оптимальні значення параметрів ОВМ.

Беручи за основу базу даних з роздільними даними та базу даних з зашумленими даними, а також використовуючи правило інтервалу для передбачення тестової помилки, було встановлено, що прогнозування є дуже точним, а криві майже ідентичними.

References

1. *Vapnik V., Chapelle O.*, 2000. Bounds on error expectation for support vector machines. *Neural Computation*, 12(9).
2. *Guyon I., Canu S., Boucheron S.*, 1997. Noise injection: Theoretical prospects. *Neural Computation*.
3. *Bousquet O., Elisseev A.*, 2002. Stability and generalization. *Journal of Machine Learning Research*, 2, 499-526.
4. *Vapnik V.*, 1995. *The Nature of Statistical Learning Theory*. Springer, New York.
5. *Kecman V.*, 2001. *Learning and Soft Computing – Support Vector Machines*. Neural Networks, Fuzzy Logic Systems, the MIT Press, Cambridge, MA.
6. *Mitchell T.*, 1997. *Machine Learning*. McGraw-Hill International.
7. *Suykens Johan A. K., Vandewalle, Joos P. L.*, 1999. Least squares support vector machine classifiers. *Neural Processing Letters*, vol. 9, no. 3, Jun.
8. *Breiman L., Friedman J.H., Olshen R.A., and Stone C.J.*, 1984. *Classification and Regression Trees*. Chapman & Hall.
9. *Campbell C., Ying Y.*, 2011. *Learning with Support Vector Machines*. Morgan and Claypool.
10. *Duda R., Hart P.*, 1973. *Pattern Classification and Scene Analysis*. John Wiley & Sons.

Надійшла до редакції
04.12.2012 р.

Затверджена до друку
10.12.2012 р.