

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ”

МИТНИК ОЛЕГ ЮРІЙОВИЧ

УДК 004.855:681.518

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ СИНТЕЗУ РОБАСТНИХ
НЕЙРОНЕЧІТКИХ МОДЕЛЕЙ СТОХАСТИЧНИХ ПРОЦЕСІВ

05.13.06. — Інформаційні технології

АВТОРЕФЕРАТ
дисертації на здобуття наукового ступеня
кандидата технічних наук

Київ – 2008

Дисертацією є рукопис.

Робота виконана в навчально-науковому комплексі “Інститут прикладного системного аналізу” Національного технічного університету України “Київський політехнічний інститут”.

Науковий керівник: доктор технічних наук, професор
Бідюк Петро Іванович,
Навчально-науковий комплекс “Інститут прикладного системного аналізу” Національного технічного університету України “Київський політехнічний інститут”,
провідний науковий співробітник відділу математичних методів системного аналізу.

Офіційні опоненти: доктор фізико-математичних наук, професор
Крак Юрій Васильович,
Київський національний університет імені Тараса Шевченка,
професор кафедри моделювання складних систем;

доктор технічних наук, професор
Томашевський Валентин Миколайович,
Національний технічний університет України “Київський політехнічний інститут”,
професор кафедри автоматизованих систем обробки інформації та управління.

Захист відбудеться 2008 р. о годині на засіданні спеціалізованої вченої ради Д26.002.03 в Національному технічному університеті України “Київський політехнічний інститут” за адресою: 03056, Київ, просп. Перемоги 37, корп. 35, ауд. 006.

З дисертацією можна ознайомитись у бібліотеці Національного технічного університету України “КПІ”.

Автореферат розісланий 2008р.

Вчений секретар
спеціалізованої вченої ради
д.т.н., професор

Новіков О.М.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Прогнозування поведінки технічних процесів успішно виконується на основі диференційних рівнянь. Тоді як прогнозування економічних, екологічних, соціальних процесів або так зване інтелектуальне прогнозування виконується на основі правил, оскільки досвідчені менеджери приймають ефективні рішення на рівні практичних міркувань. Для побудови інтелектуального прогнозу доцільно використовувати експертні системи на основі нечітких баз знань. Такі бази знань використовують апарат нечіткої логіки, сила якого полягає у здатності створювати кількісне представлення для лінгвістичних змінних, а також ефективно відображати залежності між цими змінними у вигляді нечітких правил. Однак нечіткі бази знань, які побудовані виключно на основі експертної інформації, є суб'єктивними і повністю залежать від кваліфікації експертів. Тому необхідно підкріплювати висновки і рішення експертів технічним аналізом процесів. Для цього, після побудови апріорної моделі процесу у формі нечіткої бази знань, проводять навчання цієї моделі на емпіричних даних. Здатність навчатися на даних притаманна нейроматематичним технологіям. Тож гібридизація нечітких та нейроматематичних технологій в так звані нейронечіткі технології є важливим інструментом інтелектуального прогнозування. Нейронечіткі технології дозволяють перетворювати апріорні нечіткі правила в аналітичні моделі, навчати ці моделі, і перетворювати їх знову в нечіткі правила, які обґрунтовані технічним аналізом. Проблема навчання моделей процесів з людським фактором це проблема відновлення стохастичних залежностей. Тож дисертаційна робота присвячена дослідженню некоректних за Адамаром задач відновлення стохастичних залежностей на основі емпіричних даних, побудові нейронечітких моделей за допомогою інтелектуальних інформаційних технологій, та практичному застосуванню нових моделей для синтезу нечітких баз знань та прогнозування стохастичних процесів.

Можна виділити три основні групи інформаційних технологій синтезу моделей стохастичних процесів та інтелектуального аналізу даних. З точки зору теорії суб'єктивної ймовірності — байєсівські технології. З точки зору теорії статистичного навчання — технологія опорних векторів. З точки зору теорії адаптивного моделювання — нейронечіткі технології. Сучасні байєсівські технології і так званий байєсівський індуктивний висновок ґрунтується на теорії суб'єктивної ймовірності. Історія становлення байєсівської школи, яка розглядає ймовірність як ступінь впевненості по відношенню до випадкової події, пов'язана з іменами Л.Дж.Севіджа (L.J.Savage), Дж.Кейнса, Г.Джефрі, Ф.Рамсея, А.Вальда. Подальші дослідження Е.Джейнса (E.Jaynes) в цьому напрямку стосуються проблеми “об'єктивного” вибору апріорних ймовірностей байєсівського висновку на основі принципу максимальної ентропії. Байєсівський підхід до ідентифікації систем знайшов своє продовження в роботах С.Гала (S.Gull), де вперше була сформульована концепція байєсівського під-

твердження параметрів апріорних розподілів для задач регресійного аналізу. Потім Д.Маккей (D.Maskau) використав і розвинув байєсівський принцип підтвердження з метою регуляризації штучних нейронних мереж. В наш час байєсівські технології активно використовуються пошуковими системами мережі Інтернет та системами фільтрації електронної пошти.

Початком принципово іншого підходу до аналізу процесів стала статистична теорія навчання В.Н.Вапника, яка намагається пояснити процес навчання зі статистичної точки зору, вивчає проблеми індуктивного висновку та прогнозування. В рамках цієї теорії сформульований принцип мінімізації структурного ризику, який став передумовою створення машини опорних векторів для задач класифікації. Метод опорних векторів був поширений А.Смолою (A.Smola) та Б.Шолькопфом (B.Schölkopf) на задачі побудови робастних регресійних моделей і розроблена так звана регресія опорних векторів (РОВ). Подальші дослідження пов'язані в основному з оптимальним вибором гіперпараметрів РОВ, які визначають характеристики шуму. Так, використавши байєсівську індукцію Маккея для знаходження гіперпараметрів, М.Лоу (M.Law) та Дж.Квок (J.Kwok) започаткували байєсівську РОВ. Питання про вибір ефективного характеристичного простору для РОВ залишається відкритим.

Нейротехнології беруть свій початок з фундаментальної роботи В.МакКаллока і В.Піттса в області моделювання нервової системи. Саме з цієї роботи почався розвиток штучного інтелекту та нейронних мереж, який пов'язаний з іменами В.М.Глушкова, О.А.Павлова, О.Г.Івахненко, Н.Вінера, Ф.Розенблата, Д.Румельхарта та Дж.Хопфілда. Новим етапом розвитку штучного інтелекту стала концепція нечіткої логіки, запропонована Л.Заде та розвинута Е.Мамдані, Т.Такагі, М.Сугено для адаптивних нечітких контролерів. Зв'язок між нейронними моделями і нечіткою логікою вперше було встановлено М.Брауном (M.Brown) і К.Харрісом (C.Harris) в так званих нейронечітких адаптивних мережах, які володіють лінгвістичною прозорістю нечіткої логіки в поєднанні з аналітичною зручністю нейронних моделей. Типовими представниками таких нейронечітких архітектур стали система адаптивного нечіткого виводу (ANFIS) Р.Джанга (R.Jang) та адаптивний сплайновий алгоритм моделювання (ASMOD) Т.Кавлі (T.Kavli). Оскільки нейронечіткі мережі потерпають від прокляття вимірності, то подальші дослідження спрямовані на вирішення саме цієї проблеми. Зокрема, К.Харріс і К.Хонг (X.Hong) запропонували нейронечіткий алгоритм моделювання поліноміальної складності, де базисні поліноми Бернштейна виступають в ролі функцій належності.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконана в ННК "Інститут прикладного системного аналізу" НТУУ "КПІ" відповідно до плану науково-дослідних робіт: "Розробка якісно-кількісного методу ситуаційного аналізу на основі байєсівських мереж" (№ ДР 0103U000528), "Розробка та впровадження автоматизованої інформаційної системи підтримки прийняття рішень в управлінні проектами" (№ ДР 0102U000245).

Мета і завдання дослідження. Підвищити якість прогнозування стохастичних процесів за допомогою нових інформаційних технологій та методів розв’язання некоректних задач відновлення залежностей. А саме:

1. Створити нову інформаційну технологію синтезу нечітких баз знань пониженої складності на основі нейронечітких моделей.
2. За допомогою технології опорних векторів забезпечити робастність нейронечітких моделей для описання стохастичних процесів з домінуючою випадковою складовою.
3. Вивчити недоліки існуючих методів регуляризації некоректних задач відновлення залежностей та отримати новий удосконалений критерій адекватності нейронечітких моделей.
4. Розробити новий метод побудови нейронечітких моделей, який забезпечить високоякісне прогнозування стохастичних процесів. Метод повинен враховувати апіорну інформацію про структуру моделі, працювати на коротких вибірках і оцінювати похибки прогнозу.
5. На основі розроблених методів створити нову інформаційну систему обробки статистичних даних та застосувати її до розв’язку реальних задач прогнозування.

Об’єкт дослідження — стохастичні процеси.

Предмет дослідження — нові інформаційні технології та методи розв’язання некоректних задач відновлення стохастичних залежностей.

Методика досліджень ґрунтується на основі теорії ймовірності, статистичного оцінювання і навчання, оптимального планування, нечітких множин, математичного програмування, чисельних методів. Програмну реалізацію запропонованих методів та алгоритмів виконано в середовищах MATLAB 6.1, JDK 1.5.0.4. Для порівняння отриманих результатів використано пакети “NeuroShell 2”, “GMDH Modeler 0.9.37” та “NeuroSolutions 5”.

Наукова новизна одержаних результатів. В дисертаційній роботі отримані наступні нові результати:

1. Розвинуто метод побудови нейронечітких моделей у формі Бернштейна. Для визначення барицентричних координат новий прискорений метод використовує швидке обернене відображення Кастельжо. Вперше запропоновано еволюційний метод побудови робастних нейронечітких моделей у формі Бернштейна, де для визначення барицентричних координат використовується оптимальне обернене відображення Кастельжо.
2. Створена нейронечітка інформаційна технологія синтезу нечітких баз знань квадратичної складності на основі так званих збалансованих нейронечітких моделей у формі Бернштейна.

3. На основі наближення маргінальної правдоподібності методом Лапласа отримано новий критерій адекватності байєсівської регресії опорних векторів, який, на відміну від відомих критеріїв, не порушує робастності моделей та має просту аналітичну форму.
4. Розроблено новий індуктивний метод побудови збалансованих робастних нейронечітких моделей у формі Бернштейна (ПРІАМ) на основі байєсівської регресії опорних векторів у характеристичному просторі поліноміальних функцій Без'є-Бернштейна, який відрізняється від аналогів підвищеною якістю прогнозування стохастичних процесів з домінуючою випадковою складовою і можливістю задавати апіорну структуру моделі.
5. Створено нову інформаційну систему обробки статистичних даних, математичний апарат якої реалізує запропоновані методи. Унікальність системи полягає в її здатності оперувати в онлайн режимі (в мережі Інтернет).

Практичне значення одержаних результатів полягає в тому, що запропоновані автором інформаційні технології та методи можуть бути використані в системах підтримки прийняття рішень при математичному моделюванні і прогнозуванні стохастичних процесів з домінуючою випадковою складовою. Зокрема отримані результати впроваджені в Севастопольську гідрометеорологічну обсерваторію МНС України.

Особистий внесок здобувача. Всі наукові результати, викладені в дисертації, отримані автором самостійно. В написаних у співавторстві роботах здобувачеві належать: швидке обернене відображення Кастельжо [1]; еволюційний метод визначення положення максимумів першої похідної спектру відбиття листків рослинності [2, 3]; новий математичний метод аналізу даних гідробіологічного моніторингу [6].

Апробація результатів дисертації. Основні результати дисертаційної роботи доповідалися і обговорювалися на міжнародних наукових конференціях “Системний аналіз та інформаційні технології” (Київ, 2003-2004 рр.), “Стан та перспективи розвитку новітніх науково-освітніх комп’ютерних технологій” (Миколаїв, 2003 р.), “Інтелектуальні системи прийняття рішень та прикладні аспекти інформаційних технологій” (Євпаторія, 2005 р.), “Моделювання та дослідження стійкості динамічних систем” (Київ, 2005, 2007 рр.), “Интеллектуальный анализ информации” (Київ, 2006 р.), на одинадцятій міжнародній науковій конференції імені академіка М.Кравчука (Київ, 2006 р.).

Публікації. Основні положення дисертації викладені в 14 друкованих роботах, з яких 5 у провідних фахових виданнях.

Структура та обсяг дисертації. Дисертація складається із вступу, чотирьох розділів, висновків, списку літератури з 212 джерел на 24 сторінках, чотирьох додатків. Загальний обсяг роботи становить 183 сторінки, з яких 134 сторінки основного тексту.

ОСНОВНИЙ ЗМІСТ РОБОТИ

Узагальнена задача відновлення стохастичної залежності формально поставлена у наступному вигляді. Нехай задані дані спостережень $\mathcal{D} = \{(y_j, \mathbf{x}_j) : j = 1, \dots, N\}$, де $\mathbf{x} = \text{col}(x^1, \dots, x^n) \in \mathcal{X} \subseteq \mathbb{R}^n$, $y \in \mathcal{Y} \subseteq \mathbb{R}$. Висувається гіпотеза про існування стохастичної залежності, відповідно до якої кожному вектору \mathbf{x} ставиться у відповідність число y , отримане за допомогою випадкового випробування за законом $P(y|\mathbf{x})$. Задача відновлення стохастичної залежності полягає у відшуванні умовної щільності розподілу $P(y|\mathbf{x})$ на основі скінченної вибірки даних \mathcal{D} . Дана обернена задача некоректно поставлена, оскільки має більш ніж один розв'язок, і зводиться до пошуку щільності апостеріорного прогнозного розподілу $P(y|\mathbf{x}, \mathcal{D})$, який є наближенням реальної стохастичної залежності. Наскільки близьким буде таке наближення залежить від повноти даних спостережень. Процес побудови $P(y|\mathbf{x}, \mathcal{D})$, або його умовного математичного сподівання, тобто функції регресії $y(\mathbf{x}) = \mathbb{E}y_x = \int yP(y|\mathbf{x}, \mathcal{D})dy$, будемо називати емпіричним моделюванням.

В першому розділі розглянуті основні проблеми емпіричного моделювання, серед яких некоректність постановки самої задачі навчання за даними спостережень і слабка структурованість представлення моделей. Значна увага приділяється механізму регуляризації задачі навчання як засобу для розкриття невизначеності відношення зміщення-варіації моделі. Розглянутий розклад аналізу варіацій як засіб забезпечення структурованості представлення моделі. Наведені типові приклади використання цього розкладу. Перший розділ закінчується постановкою задач дисертаційного дослідження.

Другий розділ присвячений дослідженню нейронечітких технологій. Введене формальне означення нейронечітких моделей відповідно до якого нейронечітка модель в канонічній формі має вигляд:

$$f(\mathbf{x}) = \sum_i w_i \mu_{A^i}(\mathbf{x}), \quad \sum_i \mu_{A^i}(\mathbf{x}) = 1,$$

де μ_{A^i} — дійсні невід'ємні функції належності змінної \mathbf{x} , A^i — лінгвістичні терми визначені на \mathcal{X} . Як відомо, основним результатом для нейронечітких моделей є наступне твердження. Дефазифікований методом середнього значення результат нечіткого логічного висновку системи Мамдані, в якій функції нечіткої логіки визначені алгебраїчними операторами добутку та суми, відповідає виходу нейронечіткої моделі. При цьому, якщо для представлення входу та виходу нечіткої системи Мамдані використовуються відповідно дійсні невід'ємні функції належності $\mu_{A^i}(\mathbf{x})$ та $\mu_{B^j}(y)$, де A^i , B^j — лінгвістичні терми визначені відповідно на \mathcal{X} та \mathcal{Y} , і $\sum_i \mu_{A^i}(\mathbf{x}) = 1$, а також $\forall j : \int \mu_{B^j}(y)dy = \text{const}$, тоді

$$w_i = \sum_j c_{ij} y_j^c, \quad 1 = \sum_j c_{ij}, \quad 0 \leq c_{ij}, \quad y_j^c = \frac{\int \mu_{B^j}(y) y dy}{\int \mu_{B^j}(y) dy}, \quad (1)$$

де c_{ij} — вага правила: якщо $\mathbf{x} \in A^i$ тоді $y \in B^j$. Тож задача генерації нечітких правил на основі аналітичної нейронечіткої моделі зводиться до відшукування ймовірностей c_{ij} за заданими коефіцієнтами w_i . Для цього необхідно розв'язати системи лінійних рівнянь з обмеженнями (1).

В загальному випадку функції належності визначають не для всього вектору входу \mathbf{x} , а для кожної вхідної змінної x^k окремо. Показано, що в цьому випадку нечітка база знань має експонентну складність відносно розмірності вектору входу. З метою зменшення кількості нечітких правил і спрощення самих правил досліджені нейронечіткі моделі у формі Бернштейна, які гарантують поліноміальну складність нечіткої бази знань. Для представлення регресії $y(\mathbf{x})$ модель у формі Бернштейна має вигляд:

$$f(\mathbf{x}) = b + \sum_{k=1}^n B_k^d(x^k) + \sum_{p=1}^{n-1} \sum_{q=p+1}^n B_{pq}^d(x^p, x^q), \quad (2)$$

де поліноми у формі Бернштейна B_k^d , B_{pq}^d однієї та двох змінних відповідно визначаються як лінійні комбінації базисних поліномів Бернштейна порядку d від барицентричних координат s та $\mathbf{u} = \text{col}(u, v)$:

$$B_k^d(x^k) = \sum_{j=0}^d w_j^k \phi_j^d(s(x^k)), \quad B_{pq}^d(x^p, x^q) = \sum_{i+r+t=d} w_{irt}^{pq} \phi_{irt}^d(\mathbf{u}(x^p, x^q)).$$

Базисні поліноми Бернштейна однієї та двох змінних мають вигляд:

$$\phi_j^d(s) = \binom{d}{j} s^j (1-s)^{d-j}, \quad \phi_{irt}^d(\mathbf{u}) = \binom{d}{i, r, t} u^i v^r (1-u-v)^t.$$

Барицентричні координати s та \mathbf{u} обчислюються за допомогою оберненого відображення Кастельжо $\Psi_k(\Lambda_k) : x^k \mapsto s(x^k) \in [0; 1]$, та $\Psi_{pq}(\Lambda_{pq}) : (x^p, x^q) \mapsto \mathbf{u}(x^p, x^q) \in \Delta\{u \geq 0, v \geq 0, u+v \leq 1\}$ на основі систем контрольних точок $\Lambda_k = \{a_j \in \mathbb{R} : j = 0, \dots, d\}$, $\Lambda_{pq} = \{\mathbf{a}_{ijk} \in \mathbb{R}^2 : i+j+k = d\}$. Обернене відображення Кастельжо використовує чисельний метод зворотного розповсюдження помилки для розв'язання рівнянь:

$$x^k = \sum_{j=0}^d a_j \phi_j^d(s), \quad \text{col}(x^p, x^q) = \sum_{i+j+k=d} \mathbf{a}_{ijk} \phi_{ijk}^d(\mathbf{u}).$$

Нейронечітку модель у формі Бернштейна можна представити як лінійну в параметрах модель $\mathcal{M}(\mathbf{x}, \mathbf{w}) = b + \mathbf{w} \cdot \Phi(\mathbf{x})$ в евклідовому просторі \mathcal{W} із скалярним добутком $\mathbf{x} \cdot \mathbf{z} = \sum_i x_i z_i$. Простір \mathcal{W} будемо називати простором поліноміальних функцій Без'є-Бернштейна. Вектор параметрів $\mathbf{w} = \text{col}(\dots, w_j^k, \dots, w_{irt}^{pq}, \dots) \in \mathcal{W}$. Вектор-функція $\Phi : \mathcal{X} \rightarrow \mathcal{W}$ складається з функцій належності однієї $\{\phi_i^d(\cdot)\}$ та двох $\{\phi_{irt}^d(\cdot, \cdot)\}$ змінних $\Phi : \mathbf{x} \mapsto \text{col}(\dots, \phi_j^d(x^k), \dots, \phi_{irt}^d(x^p, x^q), \dots)$.

Для нейронечітких моделей у формі Бернштейна нечіткі правила генеруються незалежно на основі окремих нейронечітких підмоделей у канонічній формі. Кожна нейронечітка підмодель визначається поліномом у формі Бернштейна. При цьому функції належності це базисні поліноми Бернштейна від барицентричних координат.

В результаті дослідження нейронечітких моделей у формі Бернштейна виявлені такі недоліки. Повільна збіжність оберненого відображення Кастельжо на основі чисельного методу зворотного розповсюдження помилки. Конфігурація контрольних точок не визначена однозначно. Необґрунтована можливість незалежного використання нечітких правил, які сформовані на основі окремих підмоделей. Функції належності у формі базисних поліномів Бернштейна високих порядків сильно перекриваються.

В роботі розроблено прискорений метод навчання нейронечітких моделей у формі Бернштейна. Основою методу є швидке обернене відображення Кастельжо, зміст якого полягає в наступному. Нехай для довільної множини точок $\{\mathbf{z}\} \subset \mathbb{R}^2$ задана система контрольних точок $\Lambda = \{\mathbf{a}_{ijk} \in \mathbb{R}^2 : i + j + k = d\}$ порядку d така, що задовольняє умовам:

1. Опукла оболонка $\text{co}\{\mathbf{a}_{ijk}\} \supset \{\mathbf{z}\}$.
2. Множина внутрішніх точок $\text{int co}\{\mathbf{a}_{ijk}\} \neq \emptyset$.
3. Виконуються наступні рівності $2\mathbf{a}_{i+1j+1k} = \mathbf{a}_{ij+2k} + \mathbf{a}_{i+2jk}$,
 $2\mathbf{a}_{ij+1k+1} = \mathbf{a}_{ij+2k} + \mathbf{a}_{ijk+2}$, $2\mathbf{a}_{i+1jk+1} = \mathbf{a}_{i+2jk} + \mathbf{a}_{ijk+2}$, $i + j + k + 2 = d$.

Тоді $\Psi(\Lambda) : \mathbf{z} \mapsto \mathbf{u}(\mathbf{z}) = (\mathbf{a}_{0d0} - \mathbf{a}_{00d}, \mathbf{a}_{d00} - \mathbf{a}_{00d})^{-1} (\mathbf{z} - \mathbf{a}_{00d})$, і при цьому $\mathbf{u} \in \Delta\{u \geq 0, v \geq 0, u + v \leq 1\}$. Аналогічний результат отримано також для оберненого відображення Кастельжо однієї змінної. Таким чином виведена чітка схема розміщення контрольних точок, яка забезпечує більш швидке (в 3–4 рази) навчання нейронечітких моделей у формі Бернштейна.

Досліджено також питання про оптимальне, за заданим критерієм, розміщення контрольних точок. Відповідно введено ряд означень оптимальних систем контрольних точок. Зокрема А-оптимальною системою контрольних точок будемо називати таку, що мінімізує слід коваріаційної матриці оцінок коефіцієнтів \mathbf{w} : $\Lambda_{\text{А-опт}} = \arg \min \text{tr cov}(\hat{\mathbf{w}}(\Lambda))$. Як відомо, А-оптимальність мінімізує середню дисперсію оцінок коефіцієнтів, що має зміст суми квадратів головних півосей еліпсоїду розсіювання оцінок. Оскільки А-оптимальність покращує обумовленість інформаційної матриці, то підвищується робастність побудованої моделі. Для розв'язання такої оптимізаційної задачі запропоновано використовувати еволюційні методи, зокрема генетичний алгоритм.

Для обґрунтування незалежного використання нейронечітких підмоделей вводяться так звані збалансовані нейронечіткі моделі у вигляді:

$$f(\mathbf{x}) = b + \sum_{k=1}^n g_k(x^k) + \sum_{p=1}^{n-1} \sum_{q=p+1}^n g_{pq}(x^p, x^q),$$

$$g_k(x^k) = \sum_i w_i^k \mu_{A_k^i}(x^k), \quad g_{pq}(x^p, x^q) = \sum_j w_j^{pq} \mu_{A_{pq}^j}(x^p, x^q),$$

де математичне сподівання функцій g_k , g_{pq} за умови багатовимірної рівномірної розподілу випадкової величини \mathbf{x} на своїй області визначення дорівнює нулю. При цьому нейронечіткі підмоделі визначаються як $f_k = b + g_k$ та $f_{pq} = b + g_{pq}$. Величина b називається зміщенням нейронечіткої збалансованої моделі і характеризує середнє значення виходу. Таким чином, за відсутності знань щодо значення змінної x^k логічно припустити (за принципом максимальної ентропії), що змінна x^k з однаковою ймовірністю може приймати будь-яке значення зі своєї області визначення. При цьому вплив цієї змінної на вихід моделі після розкриття невизначеності за допомогою математичного сподівання буде нульовим. Тож збалансованість дозволяє генерувати нечіткі правила на основі окремої підмоделі незалежно від інших підмоделей. Визначена достатня умова збалансованості нейронечіткої моделі:

$$\forall i : \mathbb{E} [\mu_{A_k^i}(x^k)] = \text{const}_k, \quad \forall j : \mathbb{E} [\mu_{A_{pq}^j}(x^p, x^q)] = \text{const}_{pq}, \quad \mathbf{x} \sim \mathcal{U}(\mathcal{X}),$$

а також сума коефіцієнтів при функціях належності рівна нулю: $\sum_i w_i^k = 0$, $\sum_j w_j^{pq} = 0$. Показано, що нейронечіткі моделі у формі Бернштейна можна звести до збалансованих і отримати збалансовані нейронечіткі моделі у формі Бернштейна. При цьому обернене відображення Кастельжо повинно бути побудоване так, щоб рівномірний розподіл змінної \mathbf{x} індукував рівномірний розподіл відповідних барицентричних координат. Швидке обернене відображення Кастельжо задовольняє цій умові.

В третьому розділі викладені основи РОВ. Розглянута байєсівська РОВ, яка будується наступним чином. Нехай систематична складова стохастичної залежності може бути представлена деякою параметричною моделлю \mathcal{M} із простору моделей \mathcal{H} . Нехай задано простір моделей \mathcal{H} . Тоді, відповідно до байєсівської теорії прийняття рішень, в якості оптимальної вибирається модель з найбільшою апостеріорною ймовірністю $P(\mathcal{M}|\mathcal{D})$. Припускаючи рівномірний розподіл моделей в просторі \mathcal{H} , моделі ранжуються за величиною їх маргінальної правдоподібності $P(\mathcal{D}|\mathcal{M})$ (підтвердження), яка виражає здатність моделі \mathcal{M} генерувати дані \mathcal{D} . Маргінальна правдоподібність визначається як:

$$P(\mathcal{D}|\mathcal{M}) = \int_{\mathcal{W}} P(\mathcal{D}|\mathbf{w}, \mathcal{M}) F(d\mathbf{w}|\mathcal{M}), \quad (3)$$

де F — апіорна ймовірнісна міра на \mathcal{W} . Ймовірність $P(\mathcal{D}|\mathbf{w}, \mathcal{M})$ виражає здатність моделі \mathcal{M} з заданими параметрами \mathbf{w} генерувати дані \mathcal{D} і називається правдоподібністю. Оскільки підтвердження (3) може бути обчислена аналітично тільки в найпростіших випадках, наприклад у випадку нормального шуму, то застосовують різноманітні методи наближення. Серед них метод Монте-Карло, розповсюдження сподівань, варіаційні методи, метод Лапласа.

Для того щоб отримати точкову оцінку параметрів моделі, байєсівська теорія прийняття рішень мінімізує умовний байєсівський ризик:

$$R_{\mathcal{L}}(\hat{\mathbf{w}}) = \mathbb{E}[\mathcal{L}(\hat{\mathbf{w}}, \mathbf{w})] = \int_{\mathcal{W}} \mathcal{L}(\hat{\mathbf{w}}, \mathbf{w}) F(d\mathbf{w}|\mathcal{D}, \mathcal{M}), \quad \mathcal{L}(\hat{\mathbf{w}}, \mathbf{w}) = \lim_{t \rightarrow 0} |\hat{\mathbf{w}} - \mathbf{w}|^t,$$

де \mathcal{L} — такий функціонал ризику, при якому оптимальна байєсівська оцінка параметрів моделі це мода апостеріорного розподілу параметрів (МАП оцінка):

$$\mathbf{w}_{\text{мап}} = \arg \min_{\mathbf{w} \in \mathcal{W}} R_{\mathcal{L}}(\mathbf{w}) = \arg \max_{\mathbf{w} \in \mathcal{W}} P(\mathbf{w}|\mathcal{D}, \mathcal{M}).$$

Байєсівська РОВ будується за наступних припущень. Випадкова складова стохастичної залежності вважається адитивним шумом і $y_j = \mathcal{M}(\mathbf{x}_j, \mathbf{w}) + \delta_j$, де δ_j — незалежні і однаково розподілені за законом $P(\delta|\mathbf{w}, \mathcal{M})$ випадкові величини. Априорна щільність розподілу параметру \mathbf{w} моделі \mathcal{M} згідно з принципом максимальної ентропії задана у вигляді багатовимірного нормального розподілу з нульовим математичним сподіванням і одиничною матрицею коваріації: $P(\mathbf{w}|\mathcal{M}) = \mathcal{N}(\mathbf{w}|\mathbf{0}, \mathbf{I})$. Моделі розглядаються як лінійні в параметрах в деякому характеристичному просторі \mathcal{W} : $\mathcal{M}(\mathbf{w}, \mathbf{x}) = b + \mathbf{w} \cdot \Phi(\mathbf{x})$, $\Phi: \mathcal{X} \rightarrow \mathcal{W}$. Гомоскедастична модель шуму задана у вигляді:

$$P(\delta_j|\mathbf{w}, \mathcal{M}) = \frac{\beta}{2(1 + \epsilon\beta)} \exp(-\beta|\delta_j|_{\epsilon}), \quad (4)$$

де $|\cdot|_{\epsilon}$ це ϵ -нечутлива функція втрат Вапника. $|\delta|_{\epsilon} = 0$ для $|\delta| \leq \epsilon$, $|\delta|_{\epsilon} = |\delta| - \epsilon$ для $|\delta| > \epsilon$. Така функція втрат забезпечує розрідженість представлення і робастність моделей РОВ. Параметри ϵ і β називають гіперпараметрами.

Для заданої моделі \mathcal{M} МАП оцінка байєсівської РОВ породжує задачу мінімізації регуляризованого ризику:

$$R_{\text{рег}}(\mathbf{w}) = \beta \sum_{j=1}^N |\delta_j|_{\epsilon} + \frac{1}{2} \|\mathbf{w}\|^2 \longrightarrow \min_{\mathbf{w}},$$

яка зводиться до типової задачі опуклого квадратичного програмування з розв'язком у вигляді:

$$f_{\text{мап}}(\mathbf{x}) = \sum_j (\alpha_j - \alpha_j^*) \Phi(\mathbf{x}_j) \cdot \Phi(\mathbf{x}) + b_{\text{мап}}, \quad \mathbf{w}_{\text{мап}} = \sum_j (\alpha_j - \alpha_j^*) \Phi(\mathbf{x}_j),$$

$$b_{\text{мап}} = \text{середнє}_j \left\{ \begin{array}{l} y_j - \mathbf{w} \cdot \Phi(\mathbf{x}_j) - \epsilon, \quad \alpha_j \in (0, \beta) \\ y_j - \mathbf{w} \cdot \Phi(\mathbf{x}_j) + \epsilon, \quad \alpha_j^* \in (0, \beta) \end{array} \right\},$$

де α_j, α_j^* — множники Лагранжа, які задовольняють $\sum_{j=1}^N (\alpha_j - \alpha_j^*) = 0$, $\alpha_j, \alpha_j^* \in [0; \beta]$. Таким чином параметри і сама модель представлені як лінійна комбінація елементів розрідженої множини навчальної вибірки в характеристичному просторі, для яких один з множників α_j або α_j^* ненульовий. Такі елементи називають опорними векторами. Опорні вектори лежать поза ϵ -смугою нечутливості або на її верхній чи нижній межах.

Більш складною проблемою є обчислення підтвердження моделі. Асимптотично точну оцінку можна отримати за допомогою методу Монте-Карло. Відносно точну оцінку дає метод розповсюдження сподівань. В роботі використано метод Лапласа, який хоч і є неточним в порівнянні з вище згаданими методами, проте потребує значно менші обчислювальні витрати. Аналіз показує, що цей метод систематично переоцінює підтвердження і відтак не може бути використаний для його точного підрахунку. Але, оскільки метод Лапласа переоцінює підтвердження систематично для всіх моделей, він може бути використаний для швидкого порівняння адекватності моделей. Метод Лапласа дає оцінку:

$$-\ln P(D|\mathcal{M}) \approx R_{\text{рег}}(\mathbf{w}_{\text{мап}}) - N \ln \frac{\beta}{2(1 + \epsilon\beta)} + \frac{1}{2} \ln \det \mathbf{H}_{\text{мап}}, \quad (5)$$

де $\mathbf{H}_{\text{мап}} = \nabla_{\mathbf{w}}^2 R_{\text{рег}}(\mathbf{w}_{\text{мап}})$ — гессіан регуляризованого ризику в точці, яка є модою апостеріорного розподілу. При цьому основною проблемою є недиференційовність функції втрат в критичних точках, які лежать на верхній та нижній межах ϵ -смуги. Тому, як правило, застосовують різноманітні згладжені аналоги. Але вони порушують розрідженість та робастність РОВ і призводять до складного байєсівського висновку. В роботі запропоновано використовувати локально згладжену в нескінченно малому околі критичних точок функцію втрат і доведено, що для будь-якої ϵ -нечутливої смуги РОВ існує як завгодно мале відхилення $\Delta\epsilon$ таке, що для ϵ_1 -нечутливої смуги ($\epsilon_1 = \epsilon - \Delta\epsilon$) гессіан регуляризованого ризику $\mathbf{H}_{\text{мап}} = \mathbf{I}$ і при цьому вектори зберігають властивість опорності. Як наслідок, отримано новий критерій байєсівського підтвердження (КБП) адекватності моделей РОВ:

$$-\ln P(\mathcal{D}|\mathcal{M}) \approx \text{КБП}(\mathcal{M}, \epsilon, \beta) = R_{\text{рег}}(\mathbf{w}_{\text{мап}}) - N \ln \frac{\beta}{2(1 + \epsilon\beta)}. \quad (6)$$

Задача максимізації підтвердження зводиться до задачі мінімізації КБП. Оскільки КБП залежить від гіперпараметрів, то в межах тієї самої моделі можна знайти оптимальні значення гіперпараметрів, які мінімізують (6). Цю задачу розв'язано рефлексивним методом Ньютона, де на кожному кроці будується РОВ і знаходиться $\mathbf{w}_{\text{мап}}$.

Основні переваги байєсівської РОВ полягають в наступному. Розклад опорних векторів не залежить від розмірності вхідного простору. Задача квадратичного програмування дає єдиний розв'язок. Модель РОВ робастна. Байєсівський підхід дозволяє робити оцінку для довірчих інтервалів. КБП забезпечує високу швидкість перебору моделей. Серед недоліків слід відзначити неточне обчислення підтвердження. Далі в роботі досліджені питання вибору простору моделей і характеристичного простору. З метою поєднання переваг байєсівської РОВ із зручністю нейронечіткого підходу Харріса в якості характеристичного простору використано простір поліноміальних функцій Без'є-Бернштейна, де для обчислення барицентричних координат використовується швидке обернене відображення Кастельжо.

Введено поняття конфігурації характеристичного простору поліноміальних функцій Без'є-Бернштейна у вигляді верхньотрикутної матриці \mathbf{C} розмірності $n \times n$, де кожен елемент $c_{i \geq j} \geq 0$ відображає ступінь впливу фактору $x_{i=j}$ або пари факторів x_i, x_j на вихідну змінну y і визначає порядок базисних поліномів Бернштейна для відповідної поліноміальної функції. Характеристичне відображення у цьому випадку матиме вигляд: $\Phi(\mathbf{C}) : \mathbf{x} \mapsto \text{col} \left(\dots, \phi_j^{c_k}(x^k), \dots, \phi_{irt}^{c_{pq}}(x^p, x^q), \dots \right)$. Це відображення індукує простір моделей у формі Бернштейна $\mathcal{H} = \text{span} \left\{ \dots, \phi_j^{c_k}, \dots, \phi_{irt}^{c_{pq}}, \dots \right\}$, які узагальнюють (2):

$$\mathcal{H} \ni \mathcal{M}(\mathbf{w}, \mathbf{x}) = b + \sum_{k=1}^n B_k^{c_k}(x^k) + \sum_{p=1}^{n-1} \sum_{q=p+1}^n B_{pq}^{c_{pq}}(x^p, x^q), \quad (7)$$

Для загальності міркувань покладемо $B^0(\cdot) \equiv 0$, що фактично означає відсутність впливу відповідного фактору або пари факторів на вихід. Доведено, що РОВ в характеристичному просторі поліноміальних функцій Без'є-Бернштейна індукує збалансовані нейронечіткі моделі у формі Бернштейна, які можна використовувати для побудови нечітких правил на основі окремих підмоделей: $f_k(x^k) = b + B_k^{c_k}(x^k)$, $f_{pq}(x^p, x^q) = b + B_{pq}^{c_{pq}}(x^p, x^q)$.

Основним результатом третього розділу є індуктивний метод побудови збалансованих нейронечітких моделей у формі Бернштейна, який визначає порядок перебору моделей в просторі \mathcal{H} . Схема алгоритму ПРИАМ (Поліноміальної РОВ Індуктивний Алгоритм Моделювання), який реалізує даний метод, має вигляд:

Задані: дані спостережень \mathcal{D} , рівень збіжності $\mu > 0$, початкова модель $\mathcal{M}^{(0)} = \mathcal{H}(\mathbf{C}^{(0)})$, яка відповідає апріорним сподіванням.

Результат: субоптимальна модель \mathcal{M}_{opt} .

Алгоритм починає роботу з

нормування вхідного та вихідного просторів; ітератор $t \leftarrow 0$;

повторювати

$$\mathcal{M}_{\text{opt}} \leftarrow \mathcal{M}^{(t)};$$

генерація множини моделей кандидатів: $\left\{ \mathcal{M}_{ij}^{(t+1)} = \mathcal{H}(\mathbf{C}_{ij}^{(t+1)}) \right\}_{ij}$,

де $\mathbf{C}_{ij}^{(t+1)} = \mathbf{C}^{(t)} \pm \mathbf{1}_{ij}$, $1 \leq i \leq j \leq n$, $\mathbf{1}_{ij}$ — нульова матриця з одиницею в рядку i в стовпчику j , що відповідає мінімальній зміні конфігурації;

для кожної моделі кандидата $\mathcal{M}_{ij}^{(t+1)}$ виконати

обчислення КБП $\left(\mathcal{M}_{ij}^{(t+1)} \right)$, мінімізація (6) за гіперпараметрами;

кінець перебору моделей кандидатів;

вибір моделі з найменшим значенням КБП:

$$\mathcal{M}^{(t+1)} = \arg \left(\text{КБП}^{(t+1)} = \min \text{КБП} \left(\mathcal{M}_{ij}^{(t+1)} \right) \right);$$

$t \leftarrow t + 1$;

поки не виконається критерій зупинки: $\text{КБП}^{(t)} + \mu > \text{КБП}^{(t-1)}$;

кінець роботи алгоритму.

Якщо визначити розмірність задачі навчання розмірністю вхідного простору n і об'ємом вибірки даних спостережень N , то складність алгоритму ПРИАМ складається з витрат на перебір моделей кандидатів і витрат на розв'язок задачі квадратичного програмування методом активних обмежень і становить $\mathcal{O}(n^2 N^3)$. Збіжність ПРИАМ забезпечується глобальною і квадратичною збіжністю рефлексивного методу Ньютона для мінімізації КБП за гіперпараметрами а також, як легко показати в силу оцінки знизу $\text{КБП} > -N \ln(\beta_{\max}/2)$ для $\beta_{\max} < \infty$, лінійною збіжністю пошуку субоптимальної моделі.

Перша частина четвертого розділу присвячена розробці архітектури нової інформаційної онлайн системи обробки статистичних даних, яка використовує розроблені в роботі нові алгоритми та методи. Так, інформаційна система представлена на трьох рівнях: концептуальному, логічному та фізичному. На концептуальному рівні визначені інформаційні процеси, процедури та потоки (рис. 1а). Відповідно користувач може як поповнювати базу вибірок новими вибірками так і завантажувати та обробляти вибірки алгоритмами з бази алгоритмів. Хостинг провайдер підтримує сайт та сервери, на яких розміщуються бази вибірок та алгоритмів. На логічному рівні описана взаємодія формалізованих моделей інформаційних процесів. На фізичному рівні створені підсистеми, які реалізують процеси обробки, накопичення даних та представлення знань. Схема носіїв цих підсистем зображена на рис. 1б. Користувач зі свого терміналу з браузером, який підтримує JRE версії не нижче 1.4, має можливість через Інтернет завантажити java-аплети, вибірки даних та будувати прогнозуючі моделі. Аплети взаємодіють з базою вибірок через JDBC. Java-аплети представляють собою написані на мові Java алгоритми: нейронечіткий алгоритм Харріса — В-В.NF та ПРИАМ — PRIAMUS. Структура аплетів та послідовність взаємодії користувача з аплетами зображена на рис. 2. Оскільки реалізовані механізми багатопоточності, то відповідно користувач має можливість отримати список задач (вибірок) та алгоритмів, вибрати необхідний алгоритм та задачу і завантажити паралельну сесію. В цій сесії користувач налаштовує параметри задачі та алгоритму, будує модель та отримує детальні результати прогнозу на основі цієї моделі.

В другій частині четвертого розділу наведені результати практичного застосування реально створеного прототипу такої інформаційної системи для відомих штучних та реальних еталонних моделей, для спрощеної моделі макроекономіки України (ІСЦ, РСГ), для гідрометеорологічної моделі вітрових хвиль (СГМО), для гідробіологічної моделі біомаси макрозообентосу (ІБПМ). Для кожної з моделей визначена оптимальна конфігурація характеристичного простору, побудований прогноз і довірчі інтервали, виконано порівняльний аналіз з такими провідними в області прогнозування алгоритмами як метод групового врахування аргументів (МГВА), нечіткий МГВА (НМГВА), рекурентні нейронні мережі (РНМ) та ANFIS. Також наведені представлення у вигляді

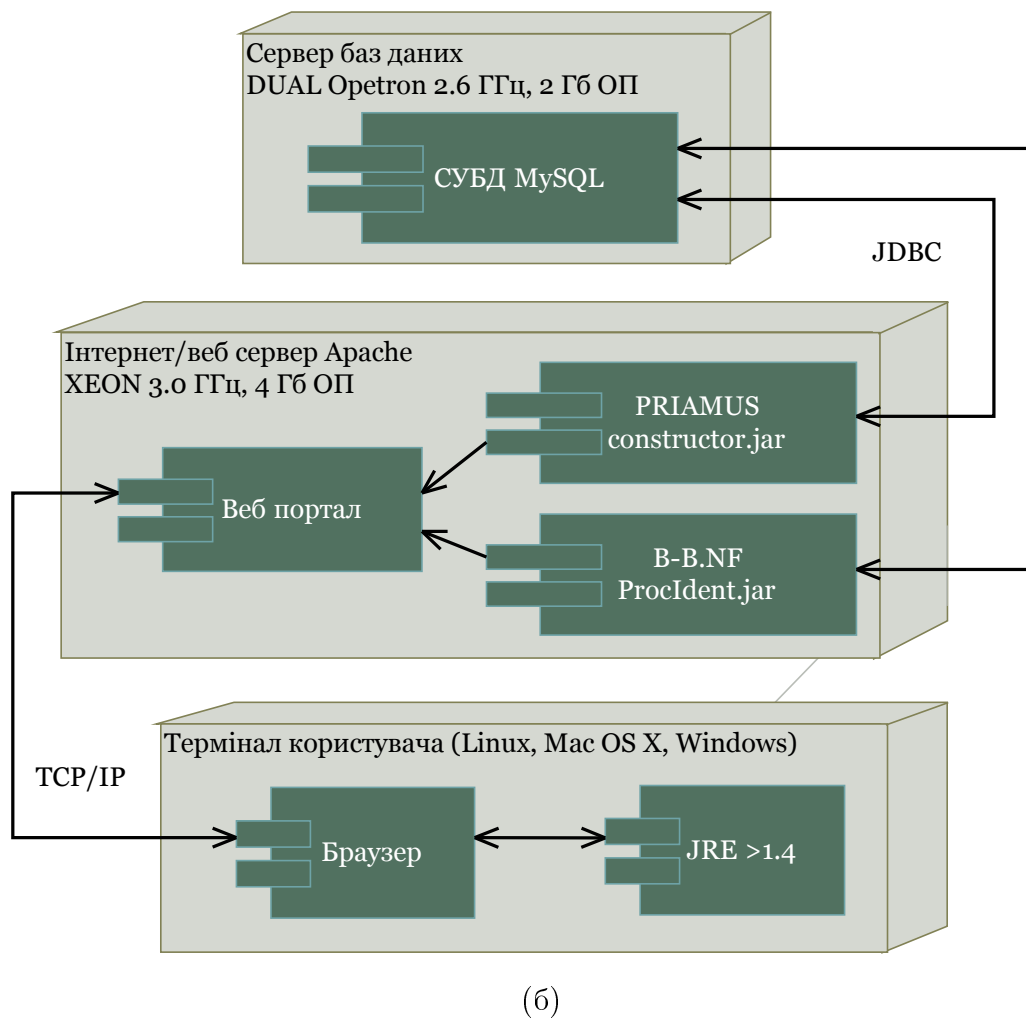
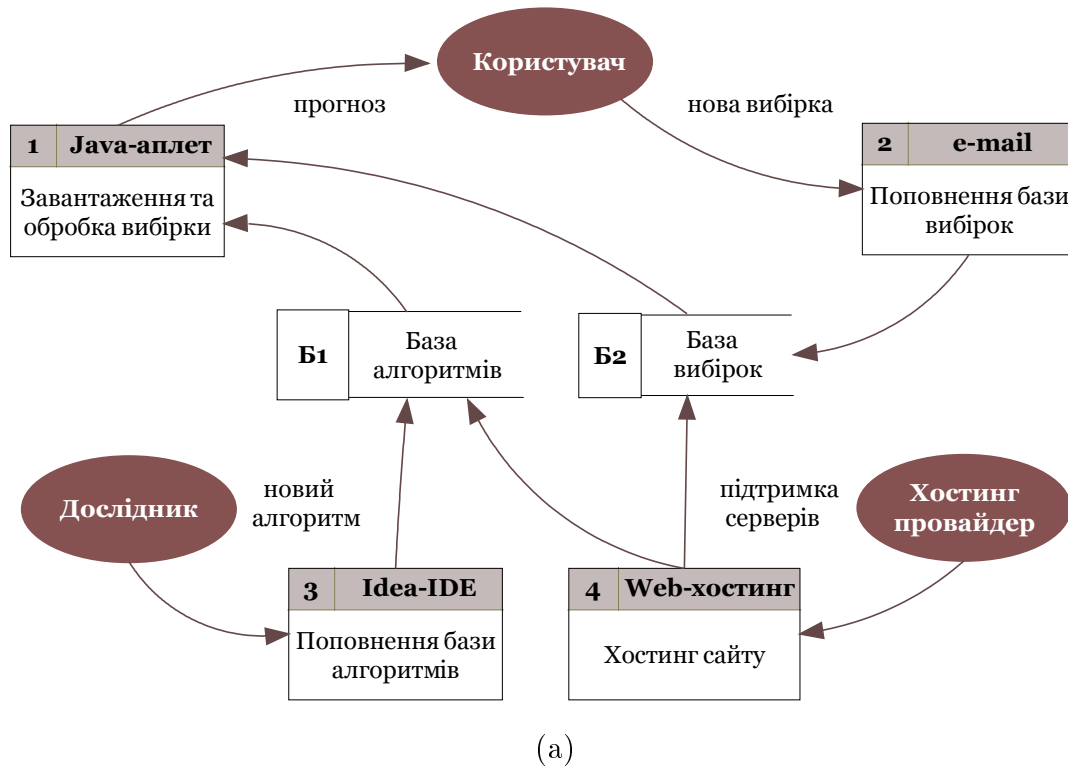


Рис. 1. (а) Діаграма інформаційних потоків. (б) Діаграма топології інформаційної системи.

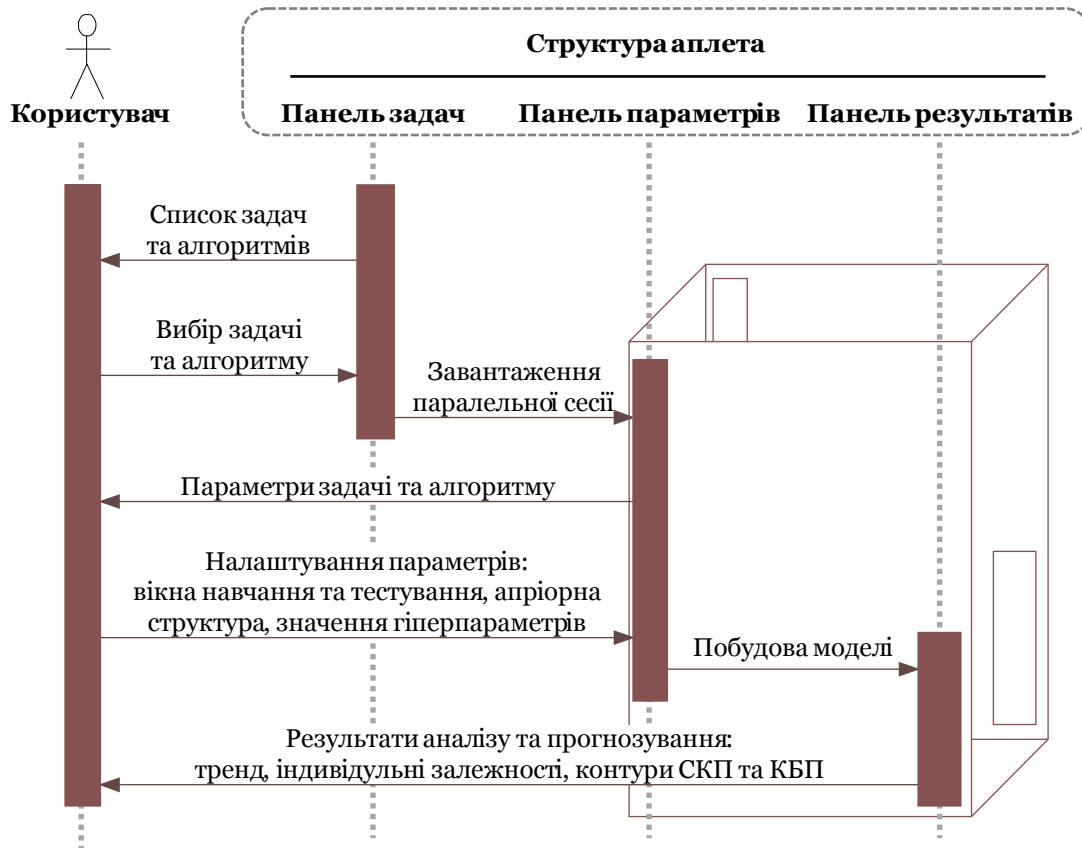


Рис. 2. Діаграма взаємодії користувача з аплетами.

розкладу опорних векторів, ітерації байєсівського виводу оптимальних значень гіперпараметрів, залишки навчальної та прогнозовної вибірок, залежності КБП та середньо квадратичної помилки в нормованому просторі (СКП_н) від ширини смуги нечутливості. Наведений в табл. 1 порівняльний аналіз підтверджує високу ефективність розробленого автором індуктивного методу. Для перевірки

Таблиця 1

Порівняння СКП_н прогнозу для різних методів

Моделі	Longley	Filippelli	Friedman	AMPG	ІСЦ	РСГ	СГМО	ІБІМ
$N_{\text{повна}}(N)$	16(11)	82(33)	1000(500)	392(300)	24(14)	24(14)	166(100)	50(25)
n	6	1	10	4	4	2	3	2
$F_{N-n, N-n}^{(80\%)}$	2.2	1.3	1.0	1.0	1.8	1.7	1.1	1.4
ПРІАМ	0.017	0.004	0.009	0.022	0.003	0.027	0.033	0.076
МГВА	0.051	0.016	0.029	0.024	0.130	0.042	0.025	0.124
НМГВА	0.001	0.039	0.037	0.026	0.008	0.098	0.043	0.053
РНМ	0.018	0.012	0.008	0.028	0.090	0.101	0.033	0.083
ANFIS	0.002	0.001	0.009	0.027	0.003	0.051	0.031	0.110

значимості переваги одного методу над іншим необхідно порівняти відношення СКП_н цих методів з наведеним табличним значенням критерію Фішера. Так, було визначено, що в середньому точність прогнозу ПРІАМ вища за точність МГВА на 52%, РНМ — на 40%, НМГВА — на 26%, ANFIS — на 10%.

Прототип інформаційної системи впроваджено в Севастопольську гідроме-

теорологічну обсерваторію для прогнозування висоти хвиль в Севастопольській бухті. Для побудови моделі використані дані щоденних спостережень за параметрами вітрових хвиль (висота, період, довжина, напрям) та показниками вітру (напрямок, швидкість, тривалість дії) з 1960 по 2006 рік. Наприклад, за допомогою ПРИАМ на основі навчальної вибірки в 100 точок (за 2006 рік) отримана оптимальна модель у формі Бернштейна в нормованому просторі:

$$f(\mathbf{x}) = 0.47 + 0.09\phi_0^1(x^1) - 0.09\phi_1^1(x^1) - \\ - 0.04\phi_0^1(x^2) + 0.04\phi_1^1(x^2) - 0.35\phi_0^1(x^3) + 0.35\phi_1^1(x^3),$$

де x^1 — синус кута напрямку вітру від напрямку на північ за годинниковою стрілкою, x^2 — косинус кута напрямку, x^3 — швидкість вітру. Для генерації нечітких правил випишемо нейронечіткі підмоделі в канонічній формі:

$$f_1(x^1) = 0.56\phi_0^1(x^1) + 0.38\phi_1^1(x^1), \quad f_2(x^2) = 0.43\phi_0^1(x^2) + 0.51\phi_1^1(x^2), \\ f_3(x^3) = 0.12\phi_0^1(x^3) + 0.82\phi_1^1(x^3).$$

Функції належності для змінних входу визначаються через базисні поліноми Бернштейна першого порядку ϕ_0^1, ϕ_1^1 . Вони відповідають нечітким змінним: західний та східний, південний та північний, слабкий та сильний вітер. Визначимо функції належності для висоти хвилі через п'ять В-сплайнів першого порядку побудовані на вузлах $\{-0.25; 0; 0.25; 0.5; 0.75; 1; 1.25\}$, які відповідають нечітким змінним: штиль, слабе хвилювання, помірне хвилювання, сильне хвилювання, шторм з центрами в точках $\{0; 0.25; 0.5; 0.75; 1\}$. Ці точки описують висоту хвилі від 0.1 до 2.4 метра. Розв'язавши по дві системи (1) для кожної з трьох підмоделей отримаємо нечітку базу знань з шести простих правил:

- | | |
|---|--|
| [| Правило 1: якщо вітер західний тоді хвилювання помірне (0.76) або сильне (0.24) |
| | Правило 2: якщо вітер східний тоді хвилювання слабе (0.48) або помірне (0.52) |
| | Правило 3: якщо вітер південний тоді хвилювання слабе (0.28) або помірне (0.72) |
| | Правило 4: якщо вітер північний тоді хвилювання помірне (0.96) або сильне (0.04) |
| | Правило 5: якщо вітер слабкий тоді штиль (0.52) або слабе хвилювання (0.48) |
| | Правило 6: якщо вітер сильний тоді сильне хвилювання (0.72) або шторм (0.28) |

На відміну від моделей, які згенеровані за допомогою МГВА, РНМ чи навіть ANFIS, цей набір правил є абсолютно прозорим для експертів. Експерт легко може визначити ступінь впливу того чи іншого фактору. Можна також визначити адекватність моделі реальним представленням експерта про процес. Так, наприклад, очевидно, що північний та західний вітер створюють більші хвилі, ніж південний та східний. Це збігається з уявленнями експертів про процеси в бухті, оскільки вихід у відкрите море з бухти має північно-західний напрям.

В додатках зведені детальні результати побудови прогнозів за допомогою ПРИАМ та їх порівняльний аналіз із результатами отриманими за допомогою

МГВА, НМГВА та РНМ. Також додатки містять скріншоти роботи розробленого прототипу інформаційної системи та акти впровадження результатів дисертаційної роботи.

ВИСНОВКИ

1. Проведено аналіз методів прогнозування стохастичних процесів та обґрунтована ефективність застосування нечітких баз знань (НБЗ) побудованих на основі нейронечітких технологій. Встановлено, що основним недоліком побудованих НБЗ є їх експонентна складність, а для синтезу НБЗ пониженої складності доцільно використовувати нейронечіткі моделі у формі Бернштейна.

2. Створена нейронечітка інформаційна технологія синтезу НБЗ квадратичної складності на основі збалансованих нейронечітких моделей у формі Бернштейна, де для обґрунтування декомпозиції моделі у формі Бернштейна на незалежні підмоделі використовується принцип максимуму ентропії.

3. Розвинуто прискорений метод побудови нейронечітких моделей у формі Бернштейна, який для визначення барицентричних координат використовує швидке обернене відображення Кастельжо (ОВК). Швидкодія побудови моделей при цьому зросла в 3–4 рази. Для синтезу робастних нейронечітких моделей у формі Бернштейна вперше запропоновано використовувати еволюційний метод на основі оптимального ОВК.

4. Показано, що байєсівська регресія опорних векторів (БРОВ) дозволяє будувати моделі стохастичних процесів з домінуючою випадковою складовою. Головними перевагами БРОВ є робастність оцінок її параметрів, розв'язання задачі опуклого квадратичного програмування, побудова довірчих інтервалів, здатність прогнозувати на коротких вибірках. Це дозволяє робити прогноз в умовах обмеженої кількості даних при сильній мультиколінеарності і шумі.

5. Для оцінки адекватності моделей БРОВ отримано новий критерій (критерій байєсівського підтвердження — КБП), який, на відміну від відомих критеріїв, не порушує робастності моделей та має просту аналітичну форму. Ефективність критерію підтверджена аналізом залежності похибки прогнозу від значення КБП на модельних прикладах.

6. На основі БРОВ в характеристичному просторі поліноміальних функцій Без'є-Бернштейна розроблено новий індуктивний метод побудови збалансованих робастних нейронечітких моделей у формі Бернштейна (ПРІАМ), який відрізняється від аналогів підвищеною якістю прогнозування стохастичних процесів. Метод враховує апріорну структуру моделі і працює на коротких вибірках. В середньому точність прогнозу ПРІАМ вища за точність МГВА на 52%, РНМ — на 40%, НМГВА — на 26%, ANFIS — на 10%.

7. Розроблена архітектура нової інформаційної системи обробки даних спостережень, яка реалізує запропоновані методи статистичного аналізу, та створено її прототип. Унікальність системи полягає в її здатності оперувати

в онлайн режимі. Впровадження прототипу інформаційної системи в Севастопольській гідрометеорологічній обсерваторії дало можливість побудувати ефективні прогнозуючі моделі, які використовуються для підтримки прийняття рішень прогнозистами з метою уточнення прогнозу параметрів хвиль.

СПИСОК ПУБЛІКАЦІЙ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Митник О. Ю., Бідюк П. І. Обернене відображення Кастельжо в нечітких нейронних моделях // Системні дослідження та інформаційні технології. — 2004. — № 2. — С. 24–34.
2. Бідюк П. І., Митник О. Ю. Застосування генетичного алгоритму в задачі оцінювання вмісту хлорофілу в рослинності // Наукові вісті НТУУ “КПІ”. — 2004. — № 4. — С. 65–70.
3. Бідюк П. І., Литвиненко В. І., Митник О. Ю. Порівняльний аналіз двох методів дистанційного оцінювання стану рослинності // Інформаційні технології і системи. — 2004. — Т. 7, № 1. — С. 108–116.
4. Митник О. Ю. Регуляризація поліноміальних функцій Без’є-Бернштейна і її застосування в задачі ідентифікації нелінійних систем // Системні дослідження та інформаційні технології. — 2006. — № 1. — С. 94–105.
5. Мытник О. Ю. Построение байесовской регрессии опорных векторов в характеристическом пространстве полиномиальных функций Безье-Бернштейна // Кибернетика и системный анализ. — 2007. — № 4. — С. 179–188.
6. Мытник О. Ю., Алёмов С. В. Использование байесовской модели регрессии опорных векторов для анализа данных гидробиологического мониторинга // Экология моря. — 2007. — Т. 73. — С. 64–69.
7. Митник О. Ю. Поліноміальні функції Без’є-Бернштейна в нечітких нейронних моделях // Системний аналіз та інформаційні технології : тези V міжнар. наук.-практ. конф. студентів, аспірантів та молодих вчених, 1–3 лип. 2003 р. — К. : НТУУ “КПІ”, 2003. — С. 80–81.
8. Оцінювання хлорофілу в рослинності на основі спектрального аналізу / П. І. Бідюк, І. В. Баклан, О. Ю. Митник, В. І. Литвиненко // Стан та перспективи розвитку новітніх науково-освітніх комп’ютерних технологій : матеріали наук.-практ. конф., 14–15 лист. 2003 р. — Миколаїв, 2003. — С. 62–65.
9. Митник О. Ю. Еволюційний метод побудови оптимального оберненого відображення Кастельжо // Системний аналіз та інформаційні технології : тези VI міжнар. наук.-практ. конф. студентів, аспірантів та молодих вчених, 1–3 лип. 2004 р. — К. : НТУУ “КПІ”, 2004. — С. 66–68.
10. Митник О. Ю., Бідюк П. І. Нейронечіткий підхід до ідентифікації нелінійних динамічних систем // Інтелектуальні системи прийняття рішень та прикладні аспекти інформаційних технологій : матеріали міжнар. наук. конф., 18–21 трав. 2005 р., Євпаторія. Т. 5. — Херсон : Вид-во Херсонського морського ін-ту, 2005. — С. 58–60.

11. Митник О. Ю. Нейронечіткий підхід до ідентифікації нелінійних динамічних систем // Моделювання та дослідження стійкості динамічних систем : тези міжнар. конф., 23–25 трав. 2005 р. — К. : Вісник КНУ, 2005. — С. 204.
12. Митник О. Ю. Байєсівський умовивід регресії опорних векторів // матеріали XI міжнар. наук. конф. імені академіка М. Кравчука, 18–20 трав. 2006 р. — К. : ТОВ “Задруга”, 2006. — С. 729.
13. Мытник О. Ю. Локальное сглаживание ϵ -нечувствительной функции потерь в байесовской модели регрессии опорных векторов // Интеллектуальный анализ информации : сборник трудов VI междунар. науч. конф., 16–19 мая, 2006 г. — К., 2006. — С. 189–198.
14. Митник О. Ю. Швидкий метод наближення маргінальної правдоподібності для оцінки адекватності моделей нелінійних стохастичних процесів // Моделювання та дослідження стійкості динамічних систем : тези міжнар. конф., 22–25 трав. 2007 р. — К. : Вісник КНУ, 2007. — С. 70.

АНОТАЦІЇ

Митник О.Ю. Інформаційні технології синтезу робастних нейронечітких моделей стохастичних процесів. — Рукопис.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 — Інформаційні технології. — Національний технічний університет України “Київський політехнічний інститут”, Київ, 2008.

Робота присвячена розробці нових нейронечітких інформаційних технологій синтезу нечітких баз знань пониженої складності для прогнозування стохастичних процесів. Введені збалансовані нейронечіткі моделі у формі Бернштейна, які генерують систему спрощених нечітких правил квадратичної складності. Досліджені методи навчання моделей стохастичних процесів як некоректної оберненої задачі відновлення стохастичних залежностей за даними спостережень. Розроблено індуктивний метод побудови збалансованих робастних нейронечітких моделей у формі Бернштейна на основі байєсівської регресії опорних векторів в характеристичному просторі поліноміальних функцій Без’є-Бернштейна (ПРІАМ). Розроблена архітектура та створено прототип інформаційної онлайн системи обробки статистичних даних, яка реалізує запропонований метод. Проведено порівняльний аналіз ПРІАМ з нечітким методом групового врахування аргументів та рекурентними нейронними мережами на реальних та штучних еталонних моделях, на економічних, метеорологічних, екологічних моделях. Результати експериментів доводять ефективність розробленого методу.

Ключові слова: стохастичні процеси, байєсівське навчання, робастні нейронечіткі моделі, регресія опорних векторів.

Мытник О.Ю. Информационные технологии синтеза робастных нейронечётких моделей стохастических процессов. — Рукопись.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.06 — Информационные технологии. — Национальный технический университет Украины “Киевский политехнический институт”, Киев, 2008.

Работа посвящена разработке новых нейронечётких информационных технологий синтеза нечётких баз знаний пониженной сложности для прогнозирования стохастических процессов. Введены сбалансированные нейронечёткие модели в форме Бернштейна, которые генерируют систему упрощённых нечётких правил квадратичной сложности. Исследованы методы обучения моделей стохастических процессов как некорректной обратной задачи восстановления стохастических зависимостей по данным наблюдений. Разработан индуктивный метод построения сбалансированных робастных нейронечётких моделей в форме Бернштейна на основе байесовской регрессии опорных векторов в характеристическом пространстве полиномиальных функций Безье-Бернштейна (ПРИАМ). Разработана архитектура и построен прототип информационной онлайн системы обработки статистических данных, которая реализует предложенный метод. Проведен сравнительный анализ ПРИАМ с нечётким методом группового учета аргументов и рекуррентными нейронными сетями на реальных и искусственных эталонных моделях, на экономических, метеорологических, экологических моделях. Результаты экспериментов подтверждают эффективность разработанного метода.

Ключевые слова: стохастические процессы, байесовское обучение, робастные нейронечёткие модели, регрессия опорных векторов.

Mytnyk O.Yu. Information technologies for robust neurofuzzy stochastic process model fusion. — Manuscript.

Thesis for a candidate of technical sciences degree by the speciality 05.13.06 — Information technologies. — National technical university of Ukraine “Kiev polytechnic institute”, Kiev, 2008.

In this paper we investigate the approaches to forecasting of stochastic processes based on observed data as an ill-posed stochastic dependence recovery problem. In the first chapter the regularization methods are studied to pose the problem well. The model transparency is also at the center of attention. It is well known that neurofuzzy modeling is an appealing resource for transparent knowledge representation. The fact that behavior of neurofuzzy models can be described naturally as a series of linguistic human-readable rules makes them suitable for expert and decision support systems.

The generalized neurofuzzy network of C.Harris is considered in the second chapter as the state-of-the-art approach to construct the polynomial complexity transparent models. Such drawbacks as uncertainty of the knots selection, and slowness of the inverse de Casteljau procedure using back propagation rule are noted.

The fast inverse de Casteljau mapping based on uniform knots layout is suggested. Thus we develop an algorithm for the fast training of the neurofuzzy models in Bernstein form based on analytical calculation of barycentric coordinates. In order to support independent generation of fuzzy rules based on separate neurofuzzy submodels we introduce the notion of the balanced neurofuzzy models. It has been shown that neurofuzzy models in Bernstein form can be balanced.

The last data mining contests have proved the advantage of Bayesian and support vector methods. Consequently the third chapter is devoted to study the Bayesian framework for support vector regression (SVR), where the main complexity is computation of evidence. Since the evidence is analytically tractable only in the simplest cases, usually the different approximation techniques are applied. Among them expectation propagation (EP), variational methods, Laplace's method (LM), Markov chain Monte Carlo (MCMC). Although LM is proved recently to be inaccurate comparing to EP and MCMC we derive Bayesian Evidence Criterion (BEC) for the fast model comparison, which drastically reduces the computational costs. Whereas BEC based on LM overestimates evidence and thus can't be used for direct evidence computation, at the same time BEC succeeds in model comparison since it overestimates evidence systematically for all models. In order to overcome the curse of smoothness of the ϵ -insensitive loss we propose to use the locally smoothed loss function, which preserves model sparseness and robustness. Bayesian framework for SVR has the following advantages. Support vector expansion is independent from input space dimension and takes away from curse of dimensionality for reconstruction problem. A unique solution is found after training as solution of a quadratic programming problem. SVR model possesses robustness and sparseness. Bayesian framework allows to make error bars estimation. BEC supplies us with extra fast model search.

Finally we present an inductive method to build balanced robust neurofuzzy model in Bernstein form based on Bayesian SVR in feature space spanned by Bézier-Bernstein polynomial functions (PRIAM — Polynomial Regression Inductive Algorithm Modeling) for stochastic dependence recovery problem. It combines the precedence of Bayesian inference, robustness of the support vector approach and transparency of the high end neurofuzzy modeling. Dual model conception allows PRIAM both be competitive with modern machine learning algorithms and be convenient for knowledge representation in expert systems. The complexity and convergence properties of PRIAM have been analysed.

We conduct experiments on synthetic reference data sets as well as on real world economic, ecologic, meteorologic models and compare PRIAM forecasts with results of group method of data handling (GMDH), fuzzy GMDH, recurrent neural networks and ANFIS. Our experiments show that PRIAM outperforms the methods listed above having parsimony model construction logic.

Keywords: stochastic processes, Bayesian learning, robust neurofuzzy models, support vector regression.